AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the accepted manuscript version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Hounyo, U. (2017). Bootstrapping integrated covariance matrix estimators in noisy jump–diffusion models with non-synchronous trading. *Journal of Econometrics*, 197(1), 130–152. <https://doi.org/10.1016/j.jeconom.2016.11.002>

Publication metadata

Title:	Bootstrapping integrated covariance matrix estimators in noisy jump–diffusion models with non-synchronous trading
Author(s):	Hounyo, U.
Journal:	<i>Journal of Econometrics</i>
DOI/Link:	10.1016/j.jeconom.2016.11.002
Document version:	Accepted manuscript (post-print)
Document license:	<i>[If the document is published under a Creative Commons, enter link to the license here]</i>

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

If the document is published under a Creative Commons license, this applies instead of the general rights.

Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading*

Ulrich Hounyo[†]

Department of Economics and Business Economics, Aarhus University,
CREATES, Aarhus University,

November 15, 2016

Abstract

We propose a bootstrap method for estimating the distribution (and functionals of it such as the variance) of various integrated covariance matrix estimators. In particular, we first adapt the wild blocks of blocks bootstrap method suggested for the pre-averaged realized volatility estimator to a general class of estimators of integrated covolatility. We then show the first-order asymptotic validity of this method in the multivariate context with a potential presence of jumps, dependent microstructure noise, irregularly spaced and non-synchronous data. Our results justify using the bootstrap to estimate the covariance matrix of a broad class of covolatility estimators. The bootstrap variance estimator is positive semi-definite by construction, an appealing feature that is not always shared by existing variance estimators of the integrated covariance estimator. As an application of our results, we also consider the bootstrap for regression coefficients. We show that the wild blocks of blocks bootstrap, appropriately centered, is able to mimic both the dependence and heterogeneity of the scores. We provide a proof of construction of bootstrap percentile and percentile- t intervals as well as variance estimates in this context. This contrasts with the traditional pairs bootstrap which is not able to mimic the score heterogeneity even in the simple case where no microstructure noise is present. Our Monte Carlo simulations show that the

*I would like to thank Jean Jacod for many useful comments and discussions. I would also like to thank Yacine Aït-Sahalia (the editor in charge), the associate editor as well as two anonymous referees for their comments and suggestions, which greatly improved the paper. I acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, as well as support from the Oxford-Man Institute of Quantitative Finance.

[†]Department of Economics and Business Economics, Aarhus University, 8210 Aarhus V., Denmark. Email: uhounyo@econ.au.dk.

wild blocks of blocks bootstrap improves the finite sample properties of the alternative approach based on the Gaussian approximation. We illustrate its practical use on high-frequency equity data.

JEL Classification: C15, C22, C58

Keywords: High-frequency data, market microstructure noise, non-synchronous data, jumps, realized measures, integrated covariance, wild bootstrap, block bootstrap.

1 Introduction

The covariation between asset returns is indispensable for risk management, portfolio selection, hedging and pricing of derivatives, etc. Presently, the availability of high-frequency financial intraday data such as stock prices or currencies allows us to accurately estimate the integrated covariance. An early popular estimator is realized covariance matrix, computed as the sum of outer product of vectors of high-frequency returns. The underlying idea is to use quadratic covariation as an ex-post covariance measure, whose increments can be studied to learn about the dependence of asset returns over a given period (see e.g. Andersen et al. (2003) and Barndorff-Nielsen et al. (2004)). An important characteristic of high frequency financial data is the presence of market microstructure effects: prices are observed with contamination errors (the so-called noise) due to the presence of bid-ask bounce effects, rounding errors, etc., which contribute to a discrepancy between the latent efficient price process and the price observed by the econometrician (e.g. Hasbrouck (2007)). In a univariate setting, market microstructure noise makes the standard realized volatility estimator biased and inconsistent. This has motivated the development of alternative estimators. Currently, there are four main univariate approaches to restore the consistency of realized volatility estimator, namely linear combination of realized volatilities obtained by subsampling (Zhang et al. (2005), and Zhang (2006)), kernel-based autocovariance adjustments (Barndorff-Nielsen et al. (2008)), the pre-averaging method (Podolskij and Vetter (2009), and Jacod et al. (2009)), and the maximum likelihood-based approach (Aït-Sahalia, Mykland and Zhang (2005) and Xiu (2010)).

In a multivariate setting, matters are further complicated with the distinctive feature of multivariate financial data: the phenomenon of non-synchronous trading, i.e. the prices of two assets are often not observed at the same time, leading to the well-known Epps effect, highlighted by Epps (1979). These factors create a further level of challenge to the problem of integrated covariance matrix estimation. The most prominent estimators of integrated covolatility that are consistent under non-synchronous observed data and contaminated by market microstructure noise include but are not limited to, the pre-averaged Hayashi-Yoshida estimator studied by Christensen et al. (2010), the multivariate realized kernel estimator of Barndorff-Nielsen et al. (2011), the flat-top realized kernel by Varneskov (2016), the two-scales covariance estimator of Zhang (2011), the generalized multi-scale covariance estimator of Bibinger (2011), the maximum likelihood based-estimator of Aït-Sahalia, Fan and Xiu (2010), Liu and

Tang (2014), Shephard and Xiu (2014), Corsi, Peluso and Audrino (2015), the Fourier based estimator of covariances of Park and Linton (2012), and the local method of moments estimator of Bibinger et al. (2014).

Despite the fact that these statistics are measured over large samples, their finite sample distributions are not necessarily well approximated by their asymptotic mixed normal distribution. Indeed, Zhang et al. (2011) showed in the univariate case that the asymptotic normal approximation is often inaccurate for the subsampling realized volatility estimator of Zhang et al. (2005), whose finite sample distribution is skewed and heavy tailed. They proposed Edgeworth corrections for this estimator as a way to improve upon the standard normal approximation. Similarly, Bandi and Russell (2011) discussed the limitations of asymptotic approximations in the context of realized kernels and proposed a finite sample procedure. As an alternative tool of inference in this context, Gonçalves and Meddahi (2009) introduced bootstrap methods for the realized volatility under no market microstructure noise, whereas Gonçalves et al. (2014) and Hounyo et al. (2016) extend the work of Gonçalves and Meddahi (2009) by allowing market microstructure effects.

In this paper, we focus on the class of estimators of integrated covolatility that can be written as the sum of miniature realized covolatility measure. Examples of potential estimators of integrated covolatility in this class include the realized covariance matrix, the cumulative covariance estimator developed in Hayashi and Yoshida (2005), the truncation-based estimators of integrated covariance of Mancini and Gobbi (2012), and some noise-robust estimators listed above (pre-averaging, realized kernel, two and multi-scale based covariance estimators), among others.

The main contribution of this paper is to propose a general bootstrap method for estimating the distribution as well as the variance of integrated covariance matrix estimators. The bootstrap technique employed here is related to previous work in the univariate case, in particular, the wild blocks of blocks bootstrap suggested in Hounyo et al. (2016) for the pre-averaging estimator. To handle both the dependence and heterogeneity of pre-averaged returns (most often in the form of heteroskedasticity), Hounyo et al. (2016) propose to combine the wild bootstrap with the blocks of blocks bootstrap. This procedure relies on the fact that the heteroskedasticity can be handled elegantly by use of the wild bootstrap, and a block-based bootstrap can be used to treat the serial correlation in the data. The current article draws ideas from this paper, but here we are faced with two additional challenges at the same time. We have to extend their univariate wild blocks of blocks bootstrap method to the multivariate case, but we also need to adapt this method for a broad class of covolatility estimators (not only for the pre-averaging based-estimator). The univariate method cannot be applied directly in this general context. This generalization faces the additional complexity of possibly having to deal with jumps, various types of noise, irregularly spaced and non-synchronous data. In particular, in a multivariate setting we first adapt the wild blocks of blocks bootstrap method studied by Hounyo et al. (2016) to a general class of statistics. Next, we give a set of high level conditions such that any bootstrap method of this form is asymptotically valid when estimating the

distribution as well as the variance of integrated covariance matrix estimator. We then verify these high-level conditions for an estimator of integrated covolatility which allow for a potential presence of jumps, dependent microstructure noise, irregularly spaced and non-synchronous data. Moreover, and as a by-product of the analysis, we establish a feasible asymptotic normality of the pre-averaged truncated Hayashi-Yoshida estimator under some mild conditions allowing the presence of infinite activity jumps processes with finite variations. It will be shown that under some regularity conditions, the pre-averaged truncated Hayashi-Yoshida estimator have the same convergence rate and asymptotic variance as those of Christensen's et al. (2013) pre-averaged Hayashi-Yoshida-type estimator. The main idea of construction of this estimator comes from Wang, Liu and Liu (2013), where they showed its consistency, but did not provide a central limit theorem. Our result enables us to construct confidence intervals of the estimator. The bootstrap variance estimator is positive semi-definite by construction, an appealing feature that is not always shared by existing variance estimators of the integrated covariance estimator. We also consider the multivariate realized kernels estimator and illustrate the application of our bootstrap method. We would like to emphasize that the generality and simplicity are the key advantages of the proposed bootstrap approach.

As an application of our results, we also consider the bootstrap for realized regression coefficients. We show that the wild blocks of blocks bootstrap, appropriately centered, is able to mimic both the dependence and heterogeneity of the scores, thus justifying the construction of bootstrap percentile and percentile- t intervals as well as asymptotic variance estimates in this context. This contrasts with the traditional pairs bootstrap analysed in Dovonon et al. (2013), which is not able to mimic the score heterogeneity even in the simple case where microstructure noise is absent and log-price process follows a continuous semimartingale, regularly spaced and synchronous. Our Monte Carlo simulations suggest that the wild blocks of blocks bootstrap method can improve upon the first-order asymptotic theory in finite samples. Although the wild blocks of blocks bootstrap that we propose here requires the choice of an additional tuning parameter (the block size), we follow Hounyo et al. (2016) and use an empirical procedure to select the block size that performs well in our simulations.

The remainder of this paper is organized as follows. In the next section, we provide the framework and introduce the general class of statistics of interest. In Section 3, after introducing the bootstrap method, we give a set of high level conditions such that any bootstrap method of this form is asymptotically valid when estimating the distribution as well as the asymptotic variance matrix of integrated covariance matrix estimator. Section 4 illustrates the bootstrap method and verifies these high level conditions for two estimators of integrated covolatility. In Section 5, we present the Monte Carlo results, while an empirical illustration is conducted in Section 6. Section 7 concludes. Two appendices are provided. Appendix A contains the tables with simulation and empirical results whereas Appendix B is a mathematical appendix providing the proofs.

2 General framework

2.1 Setup

It is well-known in finance that, under the no-arbitrage assumption, price processes must follow a semimartingale (see, e.g. Delbaen and Schachermayer (1994)). We consider a d -dimensional latent efficient log-price process $X_t = (X_t^{(1)}, \dots, X_t^{(d)})'$ defined on a probability space $(\Omega^{(0)}, \mathcal{F}^{(0)}, P^{(0)})$ equipped with a filtration $(\mathcal{F}_t^{(0)})_{t \geq 0}$. We model X as an Itô semimartingale process defined by the equation

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int \kappa(\delta(s, z)) (\mu - \nu)(ds, dz) + \int_0^t \int \kappa'(\delta(s, z)) \mu(ds, dz), \quad (1)$$

where $a = (a_t)_{t \geq 0}$ is a d -dimensional predictable locally bounded drift vector, $W = (W_t)_{t \geq 0}$ is d -dimensional Brownian motion and $\sigma = (\sigma_t)_{t \geq 0}$ is an adapted càdlàg $d \times d$ locally bounded process such that $\Sigma_t = \sigma_t \sigma_t'$ is the spot covariance matrix of X at time t . Whereas μ is a d -dimensional Poisson random measure on $\mathbb{R}_+ \times E$, with $(E, \mathcal{E})$ an auxiliary measurable space, on the space $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)})_{t \geq 0}, P^{(0)})$ and the predictable compensator (or intensity measure) of μ is $\nu(ds, dz) = ds \otimes \lambda(dz)$ for some given finite or σ -finite measure λ on $(E, \mathcal{E})$, δ is a d -dimensional predictable function on $\Omega^{(0)} \times \mathbb{R}_+ \times E$. Moreover, κ is a continuous truncation function on $\mathbb{R}^d$, that is a function from $\mathbb{R}^d$ into itself with compact support and $\kappa(x) = x$ on a neighbourhood of zero, and we set $\kappa'(x) = x - \kappa(x)$ to separate the martingale part of small jumps and the large jumps. Note that a , σ and δ should be such that the integrals in (1) make sense (see, e.g. Jacod and Shiryaev (2003) for a precise definition of the last two integrals).

In the special case where X is continuous, it has the form

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s. \quad (2)$$

It has long been recognized that asset prices do not always evolve continuously over a given time interval (e.g. Huang and Tauchen (2005), Barndorff-Nielsen and Shephard (2006)). Thus, in this paper we will mainly focus on (1). Under (1), the quadratic (co)variation of X is given by

$$\begin{aligned} [X]_t &= \int_0^t \Sigma_s ds + \sum_{s \leq t} (\Delta X_s) (\Delta X_s)' \\ &\equiv \Gamma_t + JC_t, \end{aligned}$$

where $\Delta X_s = X_s - X_{s-}$, $X_{s-} = \lim_{t \rightarrow s, t < s} X_t$. Thus $[X]_t$ is the sum of Γ_t (the integrated covolatility) and JC_t (the sum of products of simultaneous jumps (called co-jumps)). For empirical applications, one may be concerned with the behavior of Γ_t and JC_t in isolation making interesting to decompose the two sources of covariability in the price process. In this paper, our parameter of interest is integrated covariance matrix Γ_t . Without loss of generality, we let $t = 1$ (which we think of as a given day), omit the index t and define $\Gamma \equiv \Gamma_1 = \int_0^1 \Sigma_s ds$.

The presence of market frictions such as price discreteness, rounding errors, bid-ask spreads, gradual response of prices to block trades, etc, prevent us from observing the efficient price process X . Instead, we observe a noisy price process $Y = (Y^{(1)}, \dots, Y^{(d)})'$, given by

$$Y_t = X_t + \epsilon_t,$$

where ϵ_t represents the noise term that collects all the market microstructure effects. These prices are observed irregularly and non-synchronously over the interval $[0, 1]$. In particular, for all $k = 1, \dots, d$, we observed the component process $(Y^{(k)})$ at time points t_i^k for $i = 0, \dots, n_k$, given by

$$Y_{t_i^k}^k = X_{t_i^k}^k + \epsilon_{t_i^k}^k,$$

from which we compute n_k intraday returns defined as,

$$\Delta Y_{t_i^k}^k \equiv Y_{t_i^k}^k - Y_{t_{i-1}^k}^k, \quad i = 1, \dots, n_k, \quad (3)$$

with $0 = t_0^k < \dots < t_{n_k}^k = 1$ being partitions of the interval $[0, 1]$, which satisfies $\max_{1 \leq i \leq n_k} |t_i^k - t_{i-1}^k| \rightarrow 0$ as $n_k \rightarrow \infty$ for all $1 \leq k \leq d$.

In order to make both X and Y measurable with respect to the same kind of filtration, we define a new probability space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, P)$, which accommodates both processes. To this end, we follow Jacod et al. (2009) and assume one has a second space $(\Omega^{(1)}, (\mathcal{F}_t^{(1)})_{t \geq 0}, P^{(1)})$, where $\Omega^{(1)}$ denotes $\mathbb{R}^{[0,1]}$ and $\mathcal{F}^{(1)}$ the product Borel- σ -field on $\Omega^{(1)}$. Next, for any $t \in [0, 1]$, we define $Q_t(\omega^{(0)}, dy)$ to be the probability measure on $\mathbb{R}$, which corresponds to the transition from $X_t(\omega^{(0)})$ to the observed process Y_t . In the case of i.i.d. noise, this transition kernel is rather simple, but it becomes more pronounced in a general framework. $P^1(\omega^{(0)}, d\omega^{(1)})$ denotes the product measure $\otimes_{t \in [0,1]} Q_t(\omega^{(0)}, \cdot)$. The filtered probability space $(\Omega, (\mathcal{F}_t)_{t \in [0,1]}, P)$ on which the process Y lives is then defined with $\Omega = \Omega^{(0)} \times \Omega^{(1)}$, $\mathcal{F} = \mathcal{F}^{(0)} \times \mathcal{F}^{(1)}$, $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s^{(0)} \times \mathcal{F}_s^{(1)}$, and $P(d\omega^{(0)}, d\omega^{(1)}) = P^0(\omega^{(0)}) P^1(\omega^{(0)}, d\omega^{(1)})$.

2.2 Statistics of interest

The statistics of interest in this paper can be written as smooth functions of $\hat{\Gamma}^n \equiv (\hat{\Gamma}_{kl}^n)_{1 \leq k, l \leq d}$ where $\hat{\Gamma}^n$ is a consistent estimator of the integrated covariance matrix Γ , such that a central limit theorem (CLT) holds. We have, as $n \rightarrow \infty$,

$$\tau_n (\hat{\Gamma}^n - \Gamma) \rightarrow^{st} MN(\mathcal{B}, V), \quad (4)$$

where n denotes the sample size, $\tau_n = n^{\delta_1}$ with $\delta_1 \in (0, 1)$ is a known rate of convergence, $\rightarrow^{st} MN$ denotes stable convergence to a mixed Gaussian distribution (see Jacod and Shiryaev (2003, Ch. 8, Sect. 5c) for the definition and properties of stable convergence) and $V = (V_{kl,k'l'})_{1 \leq k, k', l, l' \leq d}$ is a $d \times d \times d \times d$ array, whose generic element $V_{kl,k'l'}$ corresponding to the asymptotic covariance between $\tau_n \hat{\Gamma}_{kl}^n$ and $\tau_n \hat{\Gamma}_{k'l'}^n$; $\mathcal{B}$ is a $d \times d$ array representing the asymptotic bias. In particular, we focus on the

class of estimators of Γ which can be written as

$$\hat{\Gamma}^n = \sum_{\alpha=1}^{J_n} \mathcal{Z}^n(\alpha) + \tilde{b}^n + \tilde{e}^n, \quad (5)$$

or equivalently using the individual entries of $\hat{\Gamma}^n$, $\mathcal{Z}^n(\alpha) \equiv (\mathcal{Z}_{kl}^n(\alpha))_{1 \leq k, l \leq d}$, $\tilde{b}^n \equiv (\tilde{b}_{kl}^n)_{1 \leq k, l \leq d}$, and $\tilde{e}^n \equiv (\tilde{e}_{kl}^n)_{1 \leq k, l \leq d}$, we have

$$\hat{\Gamma}_{kl}^n = \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha) + \tilde{b}_{kl}^n + \tilde{e}_{kl}^n, \quad (6)$$

where $J_n = \left\lfloor \frac{n}{b_n} \right\rfloor$, with $\lfloor \cdot \rfloor$ the integer part function and b_n is a sequence of integers such that

$$b_n \propto n^{\delta_2}, \quad (7)$$

where $\delta_2 \in (0, 1)$. $\tilde{b}^n$ can be interpreted as a bias-corrected estimator, which does not contribute to the asymptotic variance of the statistic of interest. This means that $\tau_n \hat{\Gamma}^n$ and $\tau_n \left(\sum_{\alpha=1}^{J_n} \mathcal{Z}^n(\alpha) + \tilde{e}^n \right)$ have the same asymptotic variance. $\tilde{e}^n$ collects border terms (arising near 0 and 1) due to end-effects which are estimator-specific. Usually, the following results also hold, as $n \rightarrow \infty$,

$$\tau_n (\tilde{b}^n - \tilde{b}) \rightarrow^P 0 \text{ and } \tau_n \left(\sum_{\alpha=1}^{J_n} \mathcal{Z}^n(\alpha) + \tilde{e}^n - \Gamma + \tilde{b} \right) \rightarrow^{st} MN(\mathcal{B}, V), \quad (8)$$

where $\tilde{b} = p \lim_{n \rightarrow \infty} \tilde{b}^n$. The border terms are asymptotically unbiased but not necessarily asymptotically mixed normal. In the simple case where no bias-correction is needed (i.e. $\tilde{b}_{kl}^n = 0$), for each $\alpha = 1, \dots, J_n$, the statistic $\mathcal{Z}_{kl}^n(\alpha)$ is essentially the same quantity as $\hat{\Gamma}_{kl}^n$, with the difference that it is computed only over time points t_i^k from the smaller interval $B_n(\alpha) = \left[\frac{(\alpha-1)b_n}{n}, \frac{\alpha b_n}{n} \right)$, whereas $\hat{\Gamma}_{kl}^n$ is computed over the whole interval $[0, 1]$. Thus in this case, $\mathcal{Z}^n(\alpha)$ is a miniature realized measure, which can help to get information about $\int_{\frac{(\alpha-1)b_n}{n}}^{\frac{\alpha b_n}{n}} \Sigma_s ds$. Similarly, when $\tilde{b}_{kl}^n \neq 0$, $\mathcal{Z}_{kl}^n(\alpha)$ is the analogue of $\sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha)$, but computed over time points t_i^k from $B_n(\alpha)$. The main advantage of writing $\hat{\Gamma}^n$ as in (6) is that it provides a unified bootstrap theory to dealing with a broad class of estimators of Γ . As we show in the next section, as long as this is possible and under some other regularity conditions, the wild blocks of blocks bootstrap method studied by Hounyo et al. (2016) applies now to the statistics $\mathcal{Z}_{kl}^n(\alpha)$ is first-order valid. Examples of potential estimators of integrated covolatility that can be written as (6) are listed in the introduction.

The exact expression of the conditional asymptotic variance V may be rather complicated and can involve substantially more complex quantities than the original parameter of interest Γ , see e.g. the important recent work of Mykland and Zhang (2014). One of our contributions is to justify the use of the bootstrap to estimate V . Let $V^n = \left(V_{kl, k'l'}^n \right)_{1 \leq k, k', l, l' \leq d}$ denote a consistent estimator of V , then together with the CLT result (4) we have that

$$T^n \equiv \left(\tilde{V}^n \right)^{-1/2} S^n \rightarrow^{st} N(\text{vec}(\mathcal{B}), I_{d^2}), \quad (9)$$

where $S^n = \tau_n \left(\text{vec} \left(\hat{\Gamma}^n \right) - \text{vec}(\Gamma) \right)$, vec is the vectorization operator that stacks columns of a matrix below one another, I_{d^2} is a d^2 -dimensional identity matrix and $\tilde{V}^n = \left(\tilde{V}_{kl}^n \right)_{1 \leq k, l \leq d^2}$ is a $d^2 \times d^2$ matrix, whose generic element $\tilde{V}_{kl}^n$ is given by

$$\tilde{V}_{kl}^n = V_{k-d\lfloor(k-1)/d\rfloor, \lfloor(k-1)/d\rfloor+1, l-d\lfloor(l-1)/d\rfloor, \lfloor(l-1)/d\rfloor+1}^n, \quad 1 \leq k, l \leq d^2.$$

This result can be applied in order to compute confidence region for some functionals of Γ that are important in practice, such as covariance, regression coefficient and correlation estimates. In particular, the asymptotic variance estimates for standard measures of dependence between two asset returns such as the realized covariance, the realized regression and the realized correlation coefficients are obtained by the delta method, whose finite sample properties are often poor. This motivates the bootstrap as an alternative method of inference in these contexts. The next section details how the bootstrap methodology can be used for these purposes in our general setup, which accommodates the potential presence of jumps, microstructure noise, irregularly spaced and non-synchronous trading.

3 The wild blocks of blocks bootstrap

3.1 Main results

Our aim in this section is to extend the wild blocks of blocks bootstrap method proposed by Hounyo et al. (2016) to the multivariate context allowing for the presence of jumps, noise, irregularly spaced and non-synchronous data. In particular, we propose a bootstrap method that can be used to consistently estimate the distribution of $\tau_n \left(h \left(\text{vec} \left(\hat{\Gamma}^n \right) \right) - h \left(\text{vec}(\Gamma) \right) \right)$, where $h : \mathbb{R}^{d^2} \rightarrow \mathbb{R}$ denotes a real valued function with continuous derivatives. This justifies for instance, the construction of bootstrap percentile (bootstrap unstudentized statistic) confidence intervals for covariance, regression and correlation. The bootstrap percentile intervals are easier to implement as they do not require an explicit estimator of the variance which is hard to compute in our context. However, in a setting where an estimator of the variance of $\hat{\Gamma}^n$ is available and relatively easy to compute, the construction of bootstrap percentile- t (bootstrap studentized statistic) confidence intervals may be useful. Then, we also show how and to what extent the wild blocks of blocks bootstrap can be used to construct bootstrap percentile- t for some standard measures of dependence between two asset returns.

Gonçalves and Meddahi (2009) proposed the wild bootstrap method for the realized volatility in the absence of market microstructure noise and Gonçalves et al. (2014) extend their work by allowing for the latter, see also Hounyo and Veliyev (2016). In particular, they focus on the pre-averaged realized volatility estimator proposed by Podolskij and Vetter (2009). In their ideal setting, pre-averaged returns are non-overlapping, implying that they are asymptotically uncorrelated as $n \rightarrow \infty$, but possibly heteroskedastic due to stochastic volatility, thus motivating the use of a wild bootstrap method.

When pre-averaged returns are overlapping, they are strongly dependent. This implies that the wild bootstrap is no longer valid when applied to pre-averaged returns. Instead, a block bootstrap method applied to the pre-averaged returns would seem appropriate. This amounts to a “blocks of blocks” bootstrap, as proposed by Politis and Romano (1992) and further studied by Bühlmann and Künsch (1995) (see also Künsch (1989)). Nevertheless, as Hounyo et al. (2016) show in the univariate case, such a bootstrap scheme is only consistent when volatility is constant. They argue that squared pre-averaged returns are heterogeneously distributed (in particular, their mean and variance are time-varying) and this creates a bias term in the blocks of blocks bootstrap variance estimator when volatility is stochastic. To avoid this problem, Hounyo et al. (2016) propose to combine the wild bootstrap with the blocks of blocks bootstrap. Here, we generalize their bootstrap method to the class of estimators of integrated covolatility, which can be written as in (6).

The general multivariate wild blocks of blocks bootstrap pseudo-data is given by

$$\mathcal{Z}_{kl}^{n*}(\alpha) = \begin{cases} \mathcal{Z}_{kl}^n(\alpha + 1) + (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha + 1))\eta_\alpha, & \text{if } \alpha = 1, \dots, J_n - 1 \\ \mathcal{Z}_{kl}^n(\alpha), & \text{if } \alpha = J_n, \end{cases} \quad (10)$$

where the external random variable η_α is an i.i.d. random variable independent of the data and whose moments are given by $\mu_q^* \equiv E^*(|\eta_\alpha|^q)$. As usual in the bootstrap literature, P^* (E^* and Var^*) denotes the probability measure (expected value and variance) induced by the bootstrap resampling, conditional on a realization of the original time series. In addition, for a sequence of bootstrap statistics Z_n^* , we write $Z_n^* = o_{P^*}(1)$ in probability, or $Z_n^* \rightarrow^{P^*} 0$, as $n \rightarrow \infty$, in probability, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{n \rightarrow \infty} P[P^*(|Z_n^*| > \delta) > \varepsilon] = 0$. Similarly, we write $Z_n^* = O_{P^*}(1)$ as $n \rightarrow \infty$, in probability if for all $\varepsilon > 0$ there exists a $M_\varepsilon < \infty$ such that $\lim_{n \rightarrow \infty} P[P^*(|Z_n^*| > M_\varepsilon) > \varepsilon] = 0$. Finally, we write $Z_n^* \rightarrow^{d*} Z$ as $n \rightarrow \infty$, in probability, if conditional on the sample, Z_n^* weakly converges to Z under P^* , for all samples contained in a set with probability P converging to one.

The bootstrap analogue of (6) is defined by

$$\hat{\Gamma}_{kl}^{n*} = \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^{n*}(\alpha), \quad (11)$$

and $\hat{\Gamma}^{n*} \equiv \left(\hat{\Gamma}_{kl}^{n*} \right)_{1 \leq k, l \leq d}$. Let $V_{kl, k'l'}^{n*} \equiv Cov^*(\tau_n \hat{\Gamma}_{kl}^{n*}, \tau_n \hat{\Gamma}_{k'l'}^{n*})$ denote the wild blocks of blocks bootstrap covariance between $\tau_n \hat{\Gamma}_{kl}^{n*}$ and $\tau_n \hat{\Gamma}_{k'l'}^{n*}$ based on an external random variables $\eta_\alpha \sim$ i.i.d. with mean $E^*(\eta_\alpha)$ and variance $Var^*(\eta_\alpha)$ and V^{n*} a $d \times d \times d \times d$ array, whose generic element is $V_{kl, k'l'}^{n*}$ such that (10) holds. Note that although $\hat{\Gamma}_{kl}^n$ contains a bias correction term (when $\tilde{b}_{kl}^n \neq 0$), we do not consider bias correction in the bootstrap world, even in the case where $\tilde{b}_{kl}^n \neq 0$. This is because the bias correction term $\tilde{b}^n$ by definition does not affect the asymptotic variance of $\hat{\Gamma}^n$. As long as the bootstrap method is able to consistently estimate this variance, no bias correction is needed in the bootstrap world. Since we can always center the bootstrap statistic $\hat{\Gamma}_{kl}^{n*}$ at its own theoretical mean $E^*(\hat{\Gamma}_{kl}^{n*})$ without affecting the bootstrap variance. For example, the bias correction term $\tilde{b}_{kl}^n$ for pre-

averaged realized covolatility estimator is crucially dependent of the noise assumption (see Hautsch and Podolskij (2013)) whereas the bootstrap estimator is robust regardless. Note also that by definition of $\tilde{e}^n$, hard edge-effects may influence the asymptotic variance V , because we did not necessarily impose that $\tau_n \tilde{e}^n \xrightarrow{P} 0$. Then in presence of hard end-effects, the absence of the bootstrap analogue of $\tilde{e}^n$ in $\hat{\Gamma}^{n*}$ may affect the bootstrap variance V^{n*} to consistently estimate the asymptotic variance V . However, even in such extreme situations, we can still achieve an asymptotically valid bootstrap for the studentized statistic T^n given by (9) (but not necessarily for S^n). The key aspect is that we studentize $\hat{\Gamma}^{n*}$ with $\hat{V}^{n*}$ an appropriate consistent estimator of the bootstrap variance V^{n*} based on bootstrap observations (cf. (15) below for the exact definition of $\hat{V}^{n*}$), implying that the asymptotic variance of the bootstrap t -statistic is I_{d^2} .

Our bootstrap method can be seen as a generalization of the wild blocks of blocks bootstrap method of Hounyo et al. (2016) to the general context described by (6). In particular, here we resample the statistics $\mathcal{Z}_{kl}^n(\alpha)$, which may be a block sum of functions of $\left(\Delta Y_{t_i^k}^k, \Delta Y_{t_j^l}^l\right)$ (see Section 4 for an example of statistics $\mathcal{Z}_{kl}^n(\alpha)$). As in the univariate case, to preserve the weak dependence, we divide the interval $[0, 1]$ into J_n non-overlapping sub-interval of length $\frac{b_n}{n}$ and generate the bootstrap observations within a given sub-interval $B_n(\alpha) = \left[\frac{(\alpha-1)b_n}{n}, \frac{\alpha b_n}{n}\right)$ using the same external random variable η_α . This preserves the dependence within each sub-interval. Also, as mentioned in Hounyo et al. (2016), we show that by centering around $\mathcal{Z}_{kl}^n(\alpha + 1)$ instead of $J_n^{-1} \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha)$ (as in the plain wild bootstrap method of Wu (1986) and Liu (1988)) yields an asymptotically valid bootstrap method for $\hat{\Gamma}_{kl}^n$. This is not necessary the case for the naive application of the original wild bootstrap of Liu (1988), which generates bootstrap observations $\mathcal{Z}_{kl}^{n*}(\alpha)$ as

$$\mathcal{Z}_{kl}^{n*}(\alpha) = J_n^{-1} \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha) - \left(\mathcal{Z}_{kl}^n(\alpha) - J_n^{-1} \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha) \right) \eta_\alpha, \quad \alpha = 1, \dots, J_n, \quad (12)$$

where η_α is i.i.d. $(0, 1)$. As we show in this paper, the new wild blocks of blocks bootstrap preserves the mean heterogeneity property of the statistics $\mathcal{Z}_{kl}^n(\alpha)$ even when volatility is stochastic, in our multivariate setting that allows for jumps, noise, irregularly spaced and non-synchronous data. The following result gives the bootstrap moments of $\left(\hat{\Gamma}_{kl}^{n*}, \hat{\Gamma}_{k'l'}^{n*}\right)$. In order to state our results, let $V_{kl,k'l'}^{n*} \equiv Cov^*\left(\tau_n \hat{\Gamma}_{kl}^{n*}, \tau_n \hat{\Gamma}_{k'l'}^{n*}\right)$ denote the wild blocks of blocks bootstrap covariance between $\tau_n \hat{\Gamma}_{kl}^{n*}$ and $\tau_n \hat{\Gamma}_{k'l'}^{n*}$ based on an external random variables $\eta_\alpha \sim$ i.i.d. with mean $E^*(\eta_\alpha)$ and variance $Var^*(\eta_\alpha)$, and V^{n*} a $d \times d \times d \times d$ array, whose generic element is $V_{kl,k'l'}^{n*}$ such that (10) holds.

Lemma 3.1. *Given (10) and (11), we have*

a)

$$\begin{aligned} E^* \left(\widehat{\Gamma}_{kl}^{n*} \right) &= \sum_{\alpha=1}^{J_n-1} \mathcal{Z}_{kl}^n(\alpha+1) + \mathcal{Z}_{kl}^n(J_n) \\ &+ \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) E^*(\eta_\alpha), \end{aligned}$$

in particular, if $E^*(\eta_\alpha) = 1$, we have that $E^*(\mathcal{Z}_{kl}^{n*}) = \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha) = \widehat{\Gamma}_{kl}^n - \widetilde{b}_{kl}^n - \widetilde{e}_{kl}^n$.

b)

$$V_{kl,k'l'}^{n*} = 2Var^*(\eta) \underbrace{\frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1))}_{\equiv V_{kl,k'l'}^n}.$$

Part a) of Lemma 3.1 states that in the case where $\widetilde{b}_{kl}^n = \widetilde{e}_{kl}^n = 0$,¹ if we let $E^*(\eta_\alpha) = 1$ then $\widehat{\Gamma}_{kl}^{n*}$ is an unbiased estimator of the integrated covariance Γ_{kl} . Part b) shows that the bootstrap covariance of $\tau_n \widehat{\Gamma}_{kl}^{n*}$ and $\tau_n \widehat{\Gamma}_{k'l'}^{n*}$ depends on the variance of the external random variable η , as well as the statistic $V_{kl,k'l'}^n$ which is based on a "local estimation" of the covariance of $\mathcal{Z}_{kl}^n$ and $\mathcal{Z}_{k'l'}^n$. It follows then that a sufficient condition for the bootstrap to provide a consistent estimator of the conditional asymptotic variance V is that $Var^*(\eta) = \frac{1}{2}$, and the sequence of $\mathcal{Z}_{kl}^n(\alpha)$, $\alpha = 1, \dots, J_n$, is such that $V_{kl,k'l'}^n \rightarrow^P V_{kl,k'l'}$, as $n \rightarrow \infty$. Next, we provide a set of high level conditions that allow us to derive the first-order asymptotic validity of the bootstrap method. Note that this is a high level condition that does not depend on specifying whether the process X is a continuous martingale or observed with error or not. However, for some estimators, it might hold only with some restrictions.

Condition A

A.1. The choice of the external random variable η is such that $Var^*(\eta) = \frac{1}{2}$, and as $n \rightarrow \infty$

$$V_{kl,k'l'}^n = \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1)) \rightarrow^P V_{kl,k'l'}.$$

A.2. $\left(\frac{n}{b_n}\right)^{1+\varepsilon} \sum_{\alpha=1}^{J_n} |\mathcal{Z}_{kl}^n(\alpha)|^{2+\varepsilon} = O_P(1)$, for some $\varepsilon > 0$, as $n \rightarrow \infty$.

A.3. For the same $\varepsilon > 2$ as in A.2., it holds that, as $n \rightarrow \infty$

$$\frac{b_n}{n} = o\left(\tau_n^{-\frac{2+\varepsilon}{1+\varepsilon}}\right).$$

A.1. requires that the choice of the external random variable η as well as the statistic $\mathcal{Z}_{kl}^n(\alpha)$ are such that the bootstrap variance V^{n*} yields a consistent estimator of the asymptotic variance V . This condition is very general and do not impose any structure on $\mathcal{Z}_{kl}^n(\alpha)$. We could replace A.1. by a

¹An example of such estimator is the realized covariance matrix absent microstructure noise.

condition on the sequence of $\mathcal{Z}_{kl}^n(\alpha)$, $\alpha = 1, \dots, J_n$ such that they are conditionally independent, a moment condition on $\mathcal{Z}_{kl}^n(\alpha)$ ($E \left(|\mathcal{Z}_{kl}^n(\alpha)|^{2+\varepsilon} |\mathcal{F}_{(\alpha-1)b_n}^n \right) < \infty$, for some $\varepsilon > 0$) and more importantly the following homogeneity condition on the means

$$M_n \equiv \tau_n^2 \sum_{\alpha=1}^{J_n-1} (\mu_{kl}^n(\alpha) - \mu_{kl}^n(\alpha+1)) (\mu_{k'l'}^n(\alpha) - \mu_{k'l'}^n(\alpha+1)) \rightarrow^P 0, \quad (13)$$

where $\mu_{kl}^n(\alpha) = E \left(\mathcal{Z}_{kl}^n(\alpha) | \mathcal{F}_{(\alpha-1)b_n}^n \right)$. This means homogeneity condition is suitable for financial high frequency data, in particular for estimators of integrated covolatility. This is not necessary the case of a naive application of the original wild bootstrap of Liu (1988), which will require in our context to verify the following condition

$$M_n^L \equiv \tau_n^2 \sum_{\alpha=1}^{J_n} \left(\mu_{kl}^n(\alpha) - J_n^{-1} \sum_{\alpha=1}^{J_n} \mu_{kl}^n(\alpha) \right) \left(\mu_{k'l'}^n(\alpha) - J_n^{-1} \sum_{\alpha=1}^{J_n} \mu_{k'l'}^n(\alpha) \right) \rightarrow^P 0. \quad (14)$$

In the context of time series, see e.g. Liu (1988) and Gonçalves and White (2002) (cf. Assumption 2.2) for similar restriction of the heterogeneity on the means. It is easy to see that in our setting, the homogeneity condition defined in (14) does not hold even in the very simple univariate stochastic volatility model without noise, where we also rule out drift, leverage effect, jumps and we suppose that prices are observed at equidistant date. In particular, in this case (for simplicity) we can let $J_n = n$ and consider as statistic of interest the realized volatility estimator defined by $\hat{\Gamma}^n = \sum_{\alpha=1}^{J_n} \mathcal{Z}^n(\alpha) = \sum_{\alpha=1}^n \left(\Delta Y_{\frac{\alpha}{n}} \right)^2$. We can show that

$$\begin{aligned} M_n^L &= n \sum_{\alpha=1}^n \left(\int_{\frac{\alpha-1}{n}}^{\frac{\alpha}{n}} \sigma_s^2 ds - n^{-1} \sum_{\alpha=1}^n \int_{\frac{\alpha-1}{n}}^{\frac{\alpha}{n}} \sigma_s^2 ds \right)^2 = n \sum_{\alpha=1}^n \left(\int_{\frac{\alpha-1}{n}}^{\frac{\alpha}{n}} \sigma_s^2 ds \right)^2 - \left(\int_0^1 \sigma_s^2 ds \right)^2 \\ &\rightarrow^P \int_0^1 \sigma_s^4 ds - \left(\int_0^1 \sigma_s^2 ds \right)^2, \end{aligned}$$

which is not equal to zero (one exception is when the volatility is constant). Whereas for the new bootstrap method, the mean homogeneity condition requires that

$$M_n = n \sum_{i=1}^{n-1} \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma_s^2 ds - \int_{\frac{i}{n}}^{\frac{i+1}{n}} \sigma_s^2 ds \right)^2 \rightarrow^P 0.$$

In contrast to Liu's condition, we can show that under some regularity conditions (Riemann integrability of σ), we always have $M_n \rightarrow^P 0$, even if the volatility is stochastic. This explains the new centering suggested in (10). See Section 4, for more general stochastic volatility model.

Condition A.2. and A.3. are conditions used to show that a CLT holds for $\tau_n \left(\hat{\Gamma}^{n*} - E^* \left(\hat{\Gamma}^{n*} \right) \right)$ in the bootstrap world. Part A.2. is a Lyapounov type condition that drives the asymptotic normality of $\sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha)$, whereas part A.3. restricts the choice of the block size b_n , such that the CLT holds. Note that, when the sequence of $\mathcal{Z}_{kl}^n(\alpha)$, $\alpha = 1, \dots, J_n$ can be shown to be conditionally independent

by letting $b_n = 1$, in this case we will simply use $b_n = 1$, i.e. $J_n = n$.

Under this high level condition, we can prove the following results. Theorem 3.1 is the main result of our paper, and its proof is postponed to the Appendix. We need to introduce some notation. For $1 \leq k, k', l, l' \leq d$, let

$$\check{V}_{kl,k'l'}^{n*} = \left(2 \frac{\text{Var}^*(\eta)}{E^*(\eta^2)} \right) \cdot \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^{n*}(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^{n*}(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1)).$$

We also let

$$\begin{aligned} \hat{V}_{kl}^{n*} &= \check{V}_{k-d\lfloor(k-1)/d\rfloor, [(k-1)/d]+1, l-d\lfloor(l-1)/d\rfloor, [(l-1)/d]+1}^{n*}, \quad 1 \leq k, l \leq d^2, \\ \hat{V}^{n*} &= \left(\hat{V}_{kl}^{n*} \right)_{1 \leq k, l \leq d^2}, \text{ a } d^2 \times d^2 \text{ matrix.} \end{aligned} \quad (15)$$

$$\begin{aligned} S^n &= \tau_n \left(\text{vec}(\hat{\Gamma}^n) - \text{vec}(\Gamma) \right), \\ S^{n*} &= \tau_n \left(\text{vec}(\hat{\Gamma}^{n*}) - E^* \left(\text{vec}(\hat{\Gamma}^{n*}) \right) \right), \\ T^{n*} &= \left(\hat{V}^{n*} \right)^{-1/2} S^{n*}. \end{aligned} \quad (16)$$

Theorem 3.1. *Suppose that $\hat{\Gamma}^n$ can be written as in (5).*

a) *Under Condition A.1., as $n \rightarrow \infty$,*

$$V_{kl,k'l'}^{n*} \xrightarrow{P} V_{kl,k'l'}, \text{ so that } V^{n*} \xrightarrow{P} V.$$

b) *Assume that (4) holds. Under Conditions A.1., A.2. and A.3., if for some $\varepsilon > 0$, $E^* |\eta_\alpha|^{2+\varepsilon} \leq \Delta < \infty$, then as $n \rightarrow \infty$,*

$$\sup_{x \in \mathbb{R}^{d^2}} |P^*(S^{n*} \leq x) - P(S^n - \text{vec}(\mathcal{B}) \leq x)| \xrightarrow{P} 0.$$

c) *Assume that (9) holds. Under Conditions A.2., A.3., and in addition assume that as $n \rightarrow \infty$,*

$$\left(\hat{V}^{n*} \right)^{-1} \tilde{V}^{n*} = \left(\tilde{V}^{n*} \right)^{-1} \hat{V}^{n*} \xrightarrow{P^*} I_{d^2}, \quad (17)$$

in probability-P, then as $n \rightarrow \infty$,

$$\sup_{x \in \mathbb{R}^{d^2}} |P^*(T^{n*} \leq x) - P(T^n - \text{vec}(\mathcal{B}) \leq x)| \xrightarrow{P} 0,$$

where T^n is given by (9) and T^{n} is defined in (16).*

Part a) of Theorem 3.1 shows that the bootstrap variance estimator is consistent for the asymptotic variance V according to Condition A. Part b) provides a theoretical justification for using the wild blocks of blocks bootstrap unstudentized statistic S^{n*} to consistently estimate the entire distribution of $\hat{\Gamma}^n$. Whereas part c) justifies the first-order validity of the bootstrap when applied to the studentized statistic T^n . In the following, let $\tilde{V}^{n*}$ denote the wild blocks of blocks bootstrap variance of

$\tau_n \text{vec}(\hat{\Gamma}^{n*})$, we have $\tilde{V}^{n*} = \left(\tilde{V}_{kl}^{n*}\right)_{1 \leq k, l \leq d^2}$ which is a $d^2 \times d^2$ matrix with generic element $\tilde{V}_{kl}^{n*}$ given by

$$\tilde{V}_{kl}^{n*} = V_{k-d\lfloor(k-1)/d\rfloor, \lfloor(k-1)/d\rfloor+1, l-d\lfloor(l-1)/d\rfloor, \lfloor(l-1)/d\rfloor+1}^{n*}, \quad (18)$$

with $1 \leq k, l \leq d^2$.

Remark 1. Here, assume that the external random variable η has a normal distribution such that $\text{Var}^*(\eta) = \frac{1}{2}$. Given (10), (11) and Lemma 3.1, it follows that the conditional distribution of $\text{vec}(\hat{\Gamma}^{n*})$ is exactly Gaussian in finite samples. In particular, $S^{n*} = \tau_n \left(\text{vec}(\hat{\Gamma}^{n*}) - E^* \left(\text{vec}(\hat{\Gamma}^{n*}) \right) \right) \sim N \left(0, \tilde{V}^{n*} \right)$. Thus, the distribution of S^{n*} is completely characterized by $\tilde{V}^{n*}$. Then, it can be shown that result in part b) of Theorem 3.2 holds without imposing Conditions A.2 and A.3. In the following, for simplicity, let's consider the 2-dimesional case. In particular, consider as parameter of interest the integrated covariance Γ_{12} ; we have $\text{vec}(\hat{\Gamma}^{n*}) = \left(\hat{\Gamma}_{11}^{n*} \quad \hat{\Gamma}_{21}^{n*} \quad \hat{\Gamma}_{12}^{n*} \quad \hat{\Gamma}_{22}^{n*} \right)'$, and the bootstrap 95% symmetric confidence interval for Γ_{12} suggested by part b) of Theorem 3.2 is simply

$$\left(\hat{\Gamma}_{12}^n - \frac{1.96}{\tau_n} \sqrt{\tilde{V}_{33}^{n*}}, \hat{\Gamma}_{12}^n + \frac{1.96}{\tau_n} \sqrt{\tilde{V}_{33}^{n*}} \right), \quad (19)$$

where

$$\tilde{V}_{33}^{n*} = V_{12,12}^{n*} = \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{12}^n(\alpha) - \mathcal{Z}_{12}^n(\alpha+1))^2.$$

In this case, we do not need to resample bootstrap pseudo-data as in (10) before computing a bootstrap unstudentized (percentile) confidence interval. An interpretation of this result in part b) of Theorem 3.2 is that the mixed normal asymptotic distribution of S^n can be estimated by the (non-asymptotic) mixed normal distribution of S^{n*} . We would like to emphasize that it is no longer necessary the case when η is not normally distributed. In addition, note that $T^{n*} = \left(\hat{V}^{n*}\right)^{-1/2} S^{n*}$ is not necessary Gaussian in finite samples even if η is normally distributed, since $\hat{V}^{n*}$ is random conditional on the data. For this reason, even in the case where η is normally distributed, in order to compute bootstrap studentized (percentile-t) intervals, we cannot avoid the resampling procedure given by (10).

Remark 2. Note that, without the condition $\text{Var}^*(\eta) = \frac{1}{2}$, result in part c) of Theorem 3.2 holds. Since given the definitions of $\check{V}_{kl,k'l'}^{n*}$ and T^{n*} , the studentized statistic T^{n*} is invariant to multiplication of η by a constant.

Allowing η to be non-normal is important to expect corrected higher-order departures from normality. It is well known that the finite sample distribution of S^n might differ from the limiting Gaussian distribution (which amounts to the bootstrap distribution of S^{n*} when η is normal), so that the approximation error $\sup_{x \in \mathbb{R}^{d^2}} |P^*(S^{n*} \leq x) - P(S^n - \text{vec}(\mathcal{B}) \leq x)|$ might be very large. The finite sample distribution of S^n might be heavily skewed, which cannot be well approximated by using the asymptotic normal distribution. When η is normally distributed, the bootstrap skewness of S^{n*} is exactly

equal to zero since the conditional distribution of S^{n*} is exactly normal which is not necessarily the case of S^n .

To gain further insight in the following, let us consider once again the simplified univariate stochastic volatility model without noise, no jumps, no drift term, no leverage effect, with prices observed at equidistant date and where the statistic of interest is the realized volatility. In such a model, we have $S^n = \tau_n \left(\sum_{i=1}^n \left(\Delta Y_{\frac{i}{n}} \right)^2 - \int_0^1 \sigma_s^2 ds \right)$ and the skewness is $E(S^n)^3 = \frac{1}{\tau_n} 8n^2 \sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma_s^2 ds \right)^3 \neq 0$, with $\tau_n = \sqrt{n}$ (see e.g. Gonçalves and Meddahi (2009) (cf. Lemma S.3)). Therefore, in this simple model, generate η to be normal will prevent the wild blocks of blocks bootstrap to accurately match the skewness of S^n up to $o(\tau_n^{-1})$, which is a first step to showing that the bootstrap can offer an asymptotic refinement. This is in line with the existing results in the bootstrap realized volatility context, where the wild bootstrap approach proposed by Gonçalves and Meddahi (2009) (cf. p 294) based on a standard normal external random variable is not a good choice (although it is first-order asymptotically valid) for inference on integrated volatility because it does not replicate the higher order cumulants of realized volatility (even in the toy model where $\Delta Y_{\frac{i}{n}} \sim N\left(0, \int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma_s^2 ds\right)$, despite the normality properties of $\Delta Y_{\frac{i}{n}}$).

Recently, Hounyo (2013) proposed the local Gaussian bootstrap method which exploits the local Gaussianity and the local constancy of volatility of high-frequency returns over blocks of consecutive M_n observations to generate a bootstrap approximation of realized volatility. Specifically, the local Gaussian bootstrap high-frequency returns are obtained by drawing a random draw from a normal distribution with mean zero and variance given by the realized volatility computed over the corresponding block. Hounyo (2013) shows that when the block size $M_n \rightarrow \infty$, the local Gaussian bootstrap method is able to provide asymptotic refinement. It is worth emphasizing that in our context, the fact that η with normal distribution is not an optimal choice for the wild blocks of blocks bootstrap is not in contrast to the local Gaussian bootstrap main idea in Hounyo (2013) for realized volatility (in the absence of noise). The reason is twofold. The first reason is that the blocking technique used in Hounyo (2013) is not motivated by a desire to cope with serial correlation in high-frequency returns (as it is the case of the wild blocks of blocks bootstrap), but rather by the need to exploit the source of efficiency related to block based estimators, see e.g. Mykland and Zhang (2009). The second and main reason is that in absence of noise the local Gaussian bootstrap method in Hounyo (2013) resamples the raw returns, but not the squared raw returns or sum of squared raw returns as it is the case of the wild blocks of blocks bootstrap. Heuristically, in the context of the resampling method in the current paper, when the statistics of interest is the realized volatility the application of a similar idea as in Hounyo (2013) should not involve a normal η , but rather suggest η with distribution given by a weighted sum of chi squared. The optimal choice of η is beyond the scope of this paper and will be left for future research.

The statistics of interest in this paper can be written as smooth functions of $\hat{\Gamma}^n$. The following theorem proves that the wild blocks of blocks bootstrap is first-order asymptotically valid when applied

to smooth functions of the vectorized of $\hat{\Gamma}^n$. We consider a general class of nonlinear transformations that satisfy the following assumption. Throughout we let ∇h , ($d \times 1$ vector-valued function) denote the gradient of h .

Assumption H: Let $E_\Gamma \equiv \lim_{n \rightarrow \infty} E^* \left(\hat{\Gamma}^{n*} \right)$. The function $h : \mathbb{R}^{d^2} \rightarrow \mathbb{R}$ is continuously differentiable with $\nabla h(\text{vec}(\Gamma))$ and $\nabla h(\text{vec}(E_\Gamma))$ are non-zero for any sample path of Γ .

The statistics of interest are defined as

$$S_h^n = \tau_n \left(h \left(\text{vec} \left(\hat{\Gamma}^n \right) \right) - h \left(\text{vec}(\Gamma) \right) \right), \quad (20)$$

$$T_h^n = \left(\tilde{V}_h^n \right)^{-1/2} S_h^n, \quad (21)$$

where

$$\tilde{V}_h^n \equiv \nabla' h \left(\text{vec} \left(\hat{\Gamma}^n \right) \right) \tilde{V}^n \nabla h \left(\text{vec} \left(\hat{\Gamma}^n \right) \right). \quad (22)$$

The wild blocks of blocks bootstrap version of S_h^n and T_h^n are given by

$$S_h^{n*} = \tau_n \left(h \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) - h \left(E^* \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) \right) \right) \quad (23)$$

and

$$T_h^{n*} = \left(\hat{V}_h^{n*} \right)^{-1/2} \left[\tau_n \left(h \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) - h \left(E^* \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) \right) \right) \right], \quad (24)$$

respectively, where $\hat{V}_h^{n*} = \nabla' h \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) \hat{V}^{n*} \nabla h \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right)$. In the following, let

$$V_h^{n*} \equiv \nabla' h \left(E^* \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) \right) \tilde{V}^{n*} \nabla h \left(E^* \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right) \right)$$

denote the wild blocks of blocks bootstrap variance of $\tau_n h \left(\text{vec} \left(\hat{\Gamma}^{n*} \right) \right)$, where $\tilde{V}^{n*} = \left(\tilde{V}_{kl}^{n*} \right)_{1 \leq k, l \leq d^2}$ with $\tilde{V}_{kl}^{n*}$ is defined by (18). The next theorem establishes the first-order asymptotic validity of the bootstrap for some smooth functions of the vectorized of $\hat{\Gamma}^n$.

Theorem 3.2. Suppose that $\hat{\Gamma}^n$ can be written as in (5) and Assumption H hold.

- a) Under Condition A.1., as $n \rightarrow \infty$, $V_h^{n*} \rightarrow^P V_h \equiv \lim_{n \rightarrow \infty} \text{Var} \left(\tau_n h \left(\text{vec} \left(\hat{\Gamma}^n \right) \right) \right)$.
- b) Assume that (4) holds. Under Conditions A.1., A.2. and A.3., if for some $\varepsilon > 0$, $E^* |\eta_\alpha|^{2+\varepsilon} \leq \Delta < \infty$, then as $n \rightarrow \infty$,

$$\sup_{x \in \mathbb{R}} |P^* (S_h^{n*} \leq x) - P (S_h^n - h(\text{vec}(\mathcal{B})) \leq x)| \rightarrow^P 0.$$

- c) Assume that (9) holds. Under Conditions A.2., A.3., and in addition suppose that as $n \rightarrow \infty$, (17) holds and for some $\varepsilon > 0$, $E^* |\eta_\alpha|^{2+\varepsilon} \leq \Delta < \infty$, then

$$\sup_{x \in \mathbb{R}} |P^* (T_h^{n*} \leq x) - P (T_h^n - h(\text{vec}(\mathcal{B})) \leq x)| \rightarrow^P 0.$$

Parts b) and c) of Theorem 3.2 justify constructing bootstrap unstudentized (percentile) and studentized (percentile- t) intervals, respectively, for smooth functions of Γ (i.e., $h(\text{vec}(\Gamma))$). Specifically, a $100(1 - \alpha)\%$ symmetric bootstrap percentile interval for $h(\text{vec}(\Gamma))$ based on the wild blocks of blocks bootstrap is given by

$$IC_{perc, 1-\alpha}^* = \left(h(\text{vec}(\hat{\Gamma}^n)) - \frac{p_{1-\alpha}^*}{\tau_n}, h(\text{vec}(\hat{\Gamma}^n)) + \frac{p_{1-\alpha}^*}{\tau_n} \right), \quad (25)$$

where $p_{1-\alpha}^*$ is the $1 - \alpha$ quantile of the bootstrap distribution of $|S_h^{n*}|$. Similarly, a $100(1 - \alpha)\%$ symmetric bootstrap percentile- t interval for $h(\text{vec}(\Gamma))$ is given by

$$IC_{perc-t, 1-\alpha}^* = \left(h(\text{vec}(\hat{\Gamma}^n)) - \frac{q_{1-\alpha}^*}{\tau_n} \sqrt{\tilde{V}_h^n}, h(\text{vec}(\hat{\Gamma}^n)) + \frac{q_{1-\alpha}^*}{\tau_n} \sqrt{\tilde{V}_h^n} \right), \quad (26)$$

where $\tilde{V}_h^n$ is defined by (22), $q_{1-\alpha}^*$ is the $(1 - \alpha)$ -quantile of the bootstrap distribution of $|T_h^{n*}|$. See the next section cf. equation (30) for specific examples of $h(\text{vec}(\hat{\Gamma}^n))$ as well as $\tilde{V}_h^n$.

Remark 3. *In contrast with comments in Remark 1, when h is nonlinear, even if η is normally distributed, resampling bootstrap observations as in (10) become important. In particular, given result in part b) of Theorem 3.2, we recommend to use the exact distribution of S_h^{n*} instead of the normal approximation. Thus, one can avoid some approximation errors (due to the delta method) in finite samples.*

3.2 The bootstrap for realized covariation measures

In this section we show how we can apply Theorem 3.2 in order to prove first-order asymptotic validity of the bootstrap for some functionals of the matrix $\hat{\Gamma}^n$ that are important in practice. The focus will be on realized covariance, realized regression and realized correlation coefficients. For the k th and l th asset, these quantities are given by

$$\hat{\Gamma}_{kl}^n, \hat{\beta}_{lk}^n = \frac{\hat{\Gamma}_{kl}^n}{\hat{\Gamma}_{kk}^n} \text{ and } \hat{\rho}_{lk}^n = \frac{\hat{\Gamma}_{kl}^n}{\sqrt{\hat{\Gamma}_{kk}^n \hat{\Gamma}_{ll}^n}},$$

which under certain conditions consistently estimate

$$\Gamma_{kl} = \int_0^1 \Sigma_{kl}(s) ds, \beta_{lk} = \frac{\Gamma_{kl}}{\Gamma_{kk}} \text{ and } \rho_{lk} = \frac{\Gamma_{kl}}{\sqrt{\Gamma_{kk} \Gamma_{ll}}},$$

respectively. For each of these measures, the non-studentized statistics analogue of (20) are given by

$$S_{\Gamma_{kl}}^n \equiv \tau_n (\hat{\Gamma}_{kl}^n - \Gamma_{kl}), S_{\beta_{lk}}^n \equiv \tau_n (\hat{\beta}_{lk}^n - \beta_{lk}), \text{ and } S_{\rho_{lk}}^n \equiv \tau_n (\hat{\rho}_{lk}^n - \rho_{lk}),$$

respectively. Similarly, the corresponding bootstrap percentile statistics analogue of (23) for $\hat{\Gamma}_{kl}$, $\hat{\beta}_{lk}^n$ and $\hat{\rho}_{lk}^n$ are given by

$$S_{\Gamma_{kl}}^{n*} \equiv \tau_n \left(\hat{\Gamma}_{kl}^{n*} - E^* \left(\hat{\Gamma}_{kl}^{n*} \right) \right), \quad S_{\beta_{lk}}^{n*} \equiv \tau_n \left(\hat{\beta}_{lk}^{n*} - \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)} \right), \quad \text{and} \quad S_{\rho_{lk}}^{n*} \equiv \tau_n \left(\hat{\rho}_{lk}^{n*} - \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{\sqrt{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)} \sqrt{E^* \left(\hat{\Gamma}_{ll}^{n*} \right)}} \right),$$

respectively, where $\hat{\Gamma}_{kl}^{n*}$ is defined in (11), $\hat{\beta}_{lk}^{n*} = \frac{\hat{\Gamma}_{kl}^{n*}}{\hat{\Gamma}_{kk}^{n*}}$ and $\hat{\rho}_{lk}^{n*} = \frac{\hat{\Gamma}_{kl}^{n*}}{\sqrt{\hat{\Gamma}_{kk}^{n*}} \sqrt{\hat{\Gamma}_{ll}^{n*}}}$. According to part a) of Theorem 3.2, we can use the wild blocks of blocks bootstrap variance of $S_{\Gamma_{kl}}^{n*}$, $S_{\beta_{lk}}^{n*}$ and $S_{\rho_{lk}}^{n*}$ to consistently estimate the variance of $S_{\Gamma_{kl}}^n$, $S_{\beta_{lk}}^n$ and $S_{\rho_{lk}}^n$, respectively. In particular, for the realized covariance measure, a consistent estimator of $V_{\Gamma_{kl}} = \lim_{n \rightarrow \infty} \text{Var} \left(\tau_n \hat{\Gamma}_{kl}^n \right)$ based on the bootstrap method is given by

$$V_{\Gamma_{kl}}^{n*} = V_{kl,kl}^{n*} = \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} \left(\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1) \right)^2. \quad (27)$$

Similarly, for the realized regression,

$$V_{\beta_{lk}}^{n*} = \left(\hat{\Gamma}_{kk}^n \right)^{-2} \hat{g}_{\beta_{lk}} \rightarrow^P V_{\beta_{lk}} \equiv \lim_{n \rightarrow \infty} \text{Var} \left(\tau_n \hat{\beta}_{lk} \right), \quad (28)$$

where $\hat{g}_{\beta_{lk}} = \left(1, -\frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)} \right) B_{\beta_{lk}}^{n*} \left(1, -\frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)} \right)'$ with $B_{\beta_{lk}}^{n*} = \begin{pmatrix} V_{kl,kl}^{n*} & V_{kl,kk}^{n*} \\ \bullet & V_{kk,kk}^{n*} \end{pmatrix}$.

For the realized correlation, the bootstrap estimator of $V_{\rho_{lk}} = \lim_{n \rightarrow \infty} \text{Var} \left(\tau_n \hat{\rho}_{lk}^n \right)$ is given by

$$V_{\rho_{lk}}^{n*} = \left(\hat{\Gamma}_{kk}^n \hat{\Gamma}_{ll}^n \right)^{-1} \hat{g}_{\rho_{lk}}, \quad (29)$$

where $\hat{g}_{\rho_{lk}}$ is defined by $\hat{g}_{\rho_{lk}} = \left(-\frac{1}{2} \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)}, 1, -\frac{1}{2} \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{ll}^{n*} \right)} \right) B_{\rho_{lk}}^{n*} \left(-\frac{1}{2} \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{kk}^{n*} \right)}, 1, -\frac{1}{2} \frac{E^* \left(\hat{\Gamma}_{kl}^{n*} \right)}{E^* \left(\hat{\Gamma}_{ll}^{n*} \right)} \right)'$

with $B_{\rho_{lk}}^{n*} = \begin{pmatrix} V_{kk,kk}^{n*} & V_{kk,kl}^{n*} & V_{kk,ll}^{n*} \\ \bullet & V_{kl,kl}^{n*} & V_{kl,ll}^{n*} \\ \bullet & \bullet & V_{ll,ll}^{n*} \end{pmatrix}$. Note that all the required terms are easy to compute (see

Lemma 3.1), so it is rather simple to implement the bootstrap variance estimator of the variance of $V_{\Gamma_{kl}}$, $V_{\beta_{lk}}$ and $V_{\rho_{lk}}$.

The studentized statistics analogue of (21) are given by

$$T_{\Gamma_{kl}}^n \equiv \left(\tilde{V}_{\Gamma_{kl}}^n \right)^{-1/2} S_{\Gamma_{kl}}^n, \quad T_{\beta_{lk}}^n \equiv \left(\tilde{V}_{\beta_{lk}}^n \right)^{-1/2} S_{\beta_{lk}}^n, \quad \text{and} \quad T_{\rho_{lk}}^n \equiv \left(\tilde{V}_{\rho_{lk}}^n \right)^{-1/2} S_{\rho_{lk}}^n,$$

where $\tilde{V}_{\Gamma_{kl}}^n$, $\tilde{V}_{\beta_{lk}}^n$ and $\tilde{V}_{\rho_{lk}}^n$ are any consistent estimator of $V_{\Gamma_{kl}}$, $V_{\beta_{lk}}$ and $V_{\rho_{lk}}$, respectively. According to part a) of Theorem 3.2, one can use

$$\tilde{V}_{\Gamma_{kl}}^n = V_{\Gamma_{kl}}^{n*}, \quad \tilde{V}_{\beta_{lk}}^n = V_{\beta_{lk}}^{n*}, \quad \text{and} \quad \tilde{V}_{\rho_{lk}}^n = V_{\rho_{lk}}^{n*}, \quad (30)$$

where $V_{\Gamma_{kl}}^{n*}$, $V_{\beta_{lk}}^{n*}$ and $V_{\rho_{lk}}^{n*}$ are given by (27), (28) and (29), respectively. Similarly, the corresponding

bootstrap percentile- t statistics (analogue of (24) for $\hat{\Gamma}_{kl}$, $\hat{\beta}_{lk}^n$ and $\hat{\rho}_{lk}^n$ are given by

$$T_{\Gamma_{kl}}^{n*} \equiv \left(\hat{V}_{\Gamma_{kl}}^{n*}\right)^{-1/2} S_{\Gamma_{kl}}^{n*}, \quad T_{\beta_{lk}}^{n*} \equiv \left(\hat{V}_{\beta_{lk}}^{n*}\right)^{-1/2} S_{\beta_{lk}}^{n*}, \quad \text{and} \quad T_{\rho_{lk}}^{n*} \equiv \left(\hat{V}_{\rho_{lk}}^{n*}\right)^{-1/2} S_{\rho_{lk}}^{n*},$$

where $\hat{V}_{\Gamma_{kl}}^{n*}$, $\hat{V}_{\beta_{lk}}^{n*}$ and $\hat{V}_{\rho_{lk}}^{n*}$ are consistent estimator of $V_{\Gamma_{kl}}^{n*}$, $V_{\beta_{lk}}^{n*}$ and $V_{\rho_{lk}}^{n*}$, respectively. More specifically, we will use:

$$\hat{V}_{\Gamma_{kl}}^{n*} = \check{V}_{kl,kl}^{n*} = \left(2 \frac{\text{Var}^*(\eta)}{E^*(\eta^2)}\right) \cdot \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^{n*}(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1))^2, \quad (31)$$

for the realized covariance measure.

$$\hat{V}_{\beta_{lk}}^{n*} = \left(\hat{\Gamma}_{kk}^{n*}\right)^{-2} \hat{g}_{\beta_{lk}}^*, \quad (32)$$

where $\hat{g}_{\beta_{lk}}^* = \left(1, -\frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{kk}^*}\right) \hat{B}_{\beta_{lk}}^{n*} \left(1, -\frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{ll}^*}\right)'$ with $\hat{B}_{\beta_{lk}}^{n*} = \begin{pmatrix} \hat{V}_{kl,kl}^{n*} & \hat{V}_{kl,kk}^{n*} \\ \bullet & \hat{V}_{kk,kk}^{n*} \end{pmatrix}$ for the realized regression. Whereas, for the realized correlation,

$$\hat{V}_{\rho_{lk}}^{n*} = \left(\hat{\Gamma}_{kk}^{n*} \hat{\Gamma}_{ll}^{n*}\right)^{-1} \hat{g}_{\rho_{lk}}^*, \quad (33)$$

where $\hat{g}_{\rho_{lk}}^*$ is defined by $\hat{g}_{\rho_{lk}}^* = \left(-\frac{1}{2} \frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{kk}^*}, 1, -\frac{1}{2} \frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{ll}^*}\right) \hat{B}_{\rho_{lk}}^{n*} \left(-\frac{1}{2} \frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{kk}^*}, 1, -\frac{1}{2} \frac{\hat{\Gamma}_{kl}^*}{\hat{\Gamma}_{ll}^*}\right)'$ with $\hat{B}_{\rho_{lk}}^{n*} = \begin{pmatrix} \hat{V}_{kk,kk}^{n*} & \hat{V}_{kk,kl}^{n*} & \hat{V}_{kk,ll}^{n*} \\ \bullet & \hat{V}_{kl,kl}^{n*} & \hat{V}_{kl,ll}^{n*} \\ \bullet & \bullet & \hat{V}_{ll,ll}^{n*} \end{pmatrix}$.

4 Illustration of the bootstrap scheme

The general results presented so far for a multivariate diffusion model with a potential presence of jumps, noise, irregularly spaced and non-synchronous data are stated quite compactly. Hence, it is helpful to focus on a particular case in order to enhance intuition. In this section, we show in details how our bootstrap scheme can be applied for the class of statistics called the *pre-averaged truncated Hayashi-Yoshida estimator*. We then verify the high level Condition A for this estimator of integrated covolatility. Next, we consider the multivariate realized kernels estimator and illustrate the application of our bootstrap method. In order to discuss these results, let us first introduce the assumptions on the sampling scheme. Some of the assumptions made here are specific for the pre-averaging estimator, and others may be considered when using a different estimator. We follow Christensen et al. (2013) and assume that the observation times $t_i^k, i = 0, \dots, n_k, k = 1, \dots, d$ satisfy the following conditions:

Assumption 1 - Sampling scheme

- (a) **(Time transformation)** t_i^k 's are transformations of an equidistant grid, i.e. there exist strictly monotonic (deterministic) functions $f_k : [0, 1] \rightarrow [0, 1]$ in $C^1([0, 1])$ with non-zero right and left

derivative in 0 and 1, respectively, and with $f_k(0) = 0$, $f_k(1) = 1$ such that

$$t_i^k = f_k^{-1}(i/n_k), \quad i = 0, \dots, n_k, \quad k = 1, \dots, d.$$

(b) (Boundedness of f'_k) There exists a natural number $M > 0$ such that

$$M^{-1} < \sup_{x \in [0,1]} |f'_k(x)| < M, \quad k = 1, \dots, d.$$

(c) (Comparable number of observations) Set $n = \sum_{k=1}^d n_k$. It holds that

$$\frac{n_k}{n} \rightarrow m_k \in (0, 1], \quad k = 1, \dots, d.$$

(d) (Joint grid points) The grids $(t_i^k), (t_j^l)$ ($1 \leq k, l \leq d$) have n_{kl} common points which are denoted by $(t_p^{kl})_{1 \leq p \leq n_{kl}}$. They have the representation $t_p^{kl} = f_{kl}^{-1}(p/n_{kl}) \rightarrow m_{kl} \in [0, 1]$, where the functions f_{kl} satisfy the same assumptions as f_k in (a) and (b).

Assumption 1 amounts to Assumption T in Christensen et al. (2013). As they explain, condition (a) makes the explicit computation of the asymptotic covariance matrix of the pre-averaged Hayashi-Yoshida estimator possible. Condition (c) implies that the observation numbers n_k have the same order. Condition (b) means that the points of the l th grid do not lie dense between any two successive points of the k th grid, i.e. the number of points t_j^l that lie in the interval $[t_{i-1}^k, t_i^k]$ is uniformly bounded by a constant for all $1 \leq k, l \leq d$. When these last two conditions (similar number of observations and uniform boundedness of the number of points t_j^l that belong to $[t_{i-1}^k, t_i^k]$) are fulfilled we say that the sampling schemes are comparable. See for instance Lemma 6.1 of Christensen et al. (2013) where conditions (b) and (c) imply that the amount of time points t_i^k contained in all sub-interval $[a, b]$ of $[0, 1]$ is of the same order as in the equidistant case for all k . Finally, condition (d) means that the number of common points can be negligible compared to n (if $m_{kl} = 0$) or it can be of order n (if $m_{kl} > 0$).

We assume that ϵ_t is m -dependent in tick time and that ϵ_t is independent of X_t . Assumption 2 below collects these assumptions.

Assumption 2 - Noise component.

(a) The noise component ϵ_t is m -dependent in tick time, which means that for $t_i^k \leq t_j^l$ the random variables $\epsilon_{t_i^k}^k$ and $\epsilon_{t_j^l}^l$ are independent if $\|t_i^k - t_j^l\| > m$ with

$$\|t_i^k - t_j^l\| = \min \left(j - \max \left\{ z \mid t_z^l \leq t_i^k \right\}, \min \left\{ z \mid t_z^k \geq t_j^l \right\} - i \right)$$

and similarly for $t_j^l < t_i^k$.

(b) $E(\epsilon_t) = 0$, and $E(\epsilon_t \epsilon_t') = \Psi \in \mathbb{R}^{d \times d}$, and the marginal law Q of ϵ has finite eight moments.

(c) ϵ_t is independent from the latent log-price X_t .

Note that this assumption is specific for the pre-averaging estimator, and can be called to question at very high frequencies. See, e.g. Hansen and Lunde (2006), Voev and Lunde (2007) and Diebold and Strasser (2012) for further discussion of this assumption. For instance, for the pre-averaged covolatility estimator, we could allow for dependence between X and ϵ , at the cost of slowing down the speed at which this estimator converges to the true integrated covariation (see Christensen et al. (2010), Section 3.4 for details). We could also consider a general noise model, allowing for both exogenous and endogenous components with polynomially decaying autocovariances as in Varneskov (2016) for realized kernel-based estimators and Koike (2015) for the pre-averaged truncated Hayashi-Yoshida estimator.

We rule out jumps in σ_t , formally, we make the following assumption.

Assumption 3 - Volatility

σ_t is locally bounded away from zero and is a continuous semimartingale.

This assumption is common in the realized volatility literature (e.g. equation (3) of Barndorff-Nielsen et al. (2008); Assumption 2 of Mykland and Zhang (2009) or equation (3) of Gonçalves and Meddahi (2009)). Assumption 3 can be relaxed (see Assumption H1 of Barndorff-Nielsen et al. (2006) for a weaker assumption on σ).

We follow Aït-Sahalia and Jacod (2014) and impose the following structural assumption on the characteristics of X .

Assumption 4 - Activity of jumps

For any $r \in [0, 1)$, there is a sequence of stopping time $(\tilde{\tau}_n)$ increasing to ∞ , and a deterministic nonnegative function $\tilde{\gamma}_n$ such that $\int_{\mathbb{R}^d} \tilde{\gamma}_n(z)^r \lambda(dz) < \infty$ and that $\|\delta(\omega, t, z)\| \wedge 1 \leq \tilde{\gamma}_n(z)$, for all (ω, t, z) with $t \leq \tilde{\tau}_n(\omega)$.

This assumption amounts to Assumption 1 of Aït-Sahalia and Xiu (2016). It restricts the activity of the jumps in the prices. As explained, e.g. by Todorov and Bollerslev (2010), intuitively, if the small jumps are too frequent, they become statistically indistinguishable from the diffusive part of the process, rendering the decomposition meaningless.

4.1 Pre-averaged truncated Hayashi-Yoshida estimator

Let us introduce the estimator of interest $\hat{\Gamma}_{kl}^n$. This estimator combines three techniques: the Hayashi-Yoshida method (to deal with the nonsynchronicity of the observation times), the pre-averaging method (to remove the noise) and the thresholding approach proposed by Mancini (2001) (to exclude the jumps).

The pre-averaging approach proposed by Podolskij and Vetter (2009), studied by Jacod et al. (2009) and further extended to the multivariate context by Christensen et al. (2010) and Christensen et al. (2013) is one way to lessen the influence of the noise and help us to get information about Γ .

To describe this technique, let k_n be a sequence of integers, which defines the window length over which the pre-averaging of returns is performed. In particular, suppose

$$\frac{k_n}{\sqrt{n}} = \theta + o\left(n^{-1/4}\right), \quad (34)$$

for some $\theta > 0$. Similarly, let g be a weighting function on $[0, 1]$ such that $g(0) = g(1) = 0$, $\int_0^1 g(s)^2 ds > 0$, and assume g is continuous and piecewise continuously differentiable with a piecewise Lipschitz derivative g' . An example of a function that satisfies these restrictions is $g(x) = (x \wedge (1 - x))$.

For all $k = 1, \dots, d$, $i = 0, \dots, n_k - k_n + 1$, the pre-averaged returns in tick time $\bar{Y}_{t_i^k}^k$ are obtained by computing the weighted sum of all consecutive returns performed in (3) over each block of size k_n

$$\bar{Y}_{t_i^k}^k = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \Delta Y_{t_{i+j}^k}^k. \quad (35)$$

Note that under Assumption 1, n , n_k and n_l are of the same order and that n controls the universal pre-averaging window k_n . Based on the pre-averaged returns $\bar{Y}_{t_i^k}^k$, Wang, Liu and Liu (2013) generalize Christensen's et al. (2010) pre-averaged Hayashi-Yoshida-type estimator for the integrated covariance Γ_{kl} between assets k and l by allowing jumps. See also in absence of noise, Mancini and Gobbi's (2012) Hayashi-Yoshida-type jumps robust estimator of Γ_{kl} and the related work of Jacod and Todorov (2009). The idea of Wang, Liu and Liu (2013) is to select only some of the cross variations $\bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l$ in order to estimate $\Gamma_{kl} = \int_0^1 \Sigma_{kl}(s) ds$, and precisely the ones for which there is an intersection between the time intervals $(t_i^k, t_{i+k_n}^k]$ and $(t_j^l, t_{j+k_n}^l]$ in a way such that, we exclude at the same time all terms containing jumps over a given threshold. To construct the estimator, we follow e.g. Jacod and Protter (2014) and define the truncation level as

$$v_n^i = \alpha^{(i)} u_n^\varpi, \text{ for some } \alpha^{(i)} > 0, \varpi \in \left(0, \frac{1}{2}\right), i = k, l, \quad (36)$$

where $u_n = k_n/n$, and k_n is given by (34). Let $A_{ij}^{kl} = \{(i, j) : (t_i^k, t_{i+k_n}^k] \cap (t_j^l, t_{j+k_n}^l] \neq \emptyset\}$ and $C_{ij}^{kl} = \{(i, j) : \{i : |\bar{Y}_{t_i^k}^k| < v_n^k\} \neq \emptyset \text{ and } \{j : |\bar{Y}_{t_j^l}^l| < v_n^l\} \neq \emptyset\}$. Then, the pre-averaged truncated Hayashi-Yoshida estimator is defined as

$$\hat{\Gamma}_{kl}^n = \frac{1}{(\psi k_n)^2} \sum_{i=0}^{n_k - k_n + 1} \sum_{j=0}^{n_l - k_n + 1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}} 1_{C_{ij}^{kl}}, \quad (37)$$

where k_n satisfies (34), $\psi = \int_0^1 g(s) ds$ and $1_{\{\cdot\}}$ is the indicator function discarding pre-averaged returns that do not overlap in time and exceeding a predetermined threshold values. For the simple function $g(x) = (x \wedge (1 - x))$, $\psi = 1/4$. This estimator has the profound advantage that it does not throw away information that is typically lost using a synchronization procedure. Wang, Liu and Liu (2013) under some regularity conditions showed the consistency of $\hat{\Gamma}_{kl}^n$ given by (37), but did not provide the complete asymptotic theory (i.e., a CLT result). However, before being able to apply our general

bootstrap results derived in Section 3, an asymptotic mixed normal result needs to be established for $\hat{\Gamma}_{kl}^n$ given by (37). Next, we show a feasible CLT result for the pre-averaged truncated Hayashi-Yoshida estimator.

Theorem 4.1. *Suppose (1) and $\left(\frac{1}{2(2-r)} \vee \frac{2\delta_2}{4-r}\right) \leq \varpi < 1/2$, for some $\delta_2 \in (\frac{1}{2}, \frac{2}{3})$. Furthermore suppose that $\hat{\Gamma}_{kl}^n$ is given by (37). Under Assumptions 1, 2, 3 and 4, as $n \rightarrow \infty$,*

a) (9) holds true, where $\tau_n = n^{1/4}$, $\hat{\Gamma}^n = \left(\hat{\Gamma}_{kl}^n\right)_{1 \leq k, l \leq d}$ with, $\mathcal{B} \equiv (\mathcal{B}_{kl})_{1 \leq k, l \leq d}$ where $\mathcal{B}_{kl} = 0$,

$$\tilde{V}_{kl}^n = V_{k-d\lfloor(k-1)/d\rfloor, \lfloor(k-1)/d\rfloor+1, l-d\lfloor(l-1)/d\rfloor, \lfloor(l-1)/d\rfloor+1}^n, \quad 1 \leq k, l \leq d^2,$$

such that

$$V_{kl, k'l'}^n = \frac{\sqrt{n}}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1)),$$

where

$$\mathcal{Z}_{kl}^n(\alpha) = \frac{1}{(\psi k_n)^2} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l - k_n + 1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}} 1_{C_{ij}^{kl}}, \quad (38)$$

$$B_n(\alpha) = \left[\frac{(\alpha-1)b_n}{n}, \frac{\alpha b_n}{n} \right), \quad J_n = \left\lfloor \frac{n}{b_n} \right\rfloor, \quad \text{with } b_n \text{ follows (7) such that } 1/2 < \delta_2 < 2/3.$$

b) In addition if Assumption H holds, as $n \rightarrow \infty$,

$$T_h^n \rightarrow^{st} N(0, 1), \quad (39)$$

where T_h^n is given by (21).

We can use this feasible asymptotic distribution result to build confidence intervals of (transforms of) integrated covariance matrix. Note that, this alternative method of inference does not require any resampling of one's data. For simplicity, let's consider the case of the integrated covariance Γ_{kl} . A two-sided feasible $100(1 - \alpha)\%$ level interval for Γ_{kl} is given by:

$$IC_{Feas, 1-\alpha} = \left(\hat{\Gamma}_{kl}^n - \frac{z_{1-\alpha/2}}{\tau_n} \sqrt{\tilde{V}_h^n}, \hat{\Gamma}_{kl}^n + \frac{z_{1-\alpha/2}}{\tau_n} \sqrt{\tilde{V}_h^n} \right), \quad (40)$$

where more generally $\tilde{V}_h^n$ is defined by (22), whereas in the simple case of Γ_{kl} ,

$$\tilde{V}_h^n = V_{kl, kl}^n = \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1))^2,$$

with $\mathcal{Z}_{kl}^n(\alpha)$ given by (38). $z_{1-\alpha/2}$ is such that $\Phi(z_{1-\alpha/2}) = 1 - \alpha/2$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For instance, $z_{0.975} = 1.96$ when $\alpha = 0.05$.

Note that the asymptotic variance of $\hat{\Gamma}^n$ in Theorem 4.1 is the same one as in Theorem 3.4 of Christensen et al. (2013) (in the continuous case). The exact expression of $V_{kl, k'l'}$ is rather complicated, and the reader can find it in Christensen et al. (2013).

Remark 4. *One could prove that under the same conditions from Theorem 4.1 without using Assumption 3 that (4) holds true, where $\tau_n = n^{1/4}$, $\hat{\Gamma}^n = (\hat{\Gamma}_{kl}^n)_{1 \leq k, l \leq d}$ such that $\hat{\Gamma}_{kl}^n$ is given by (37), i.e., the infeasible version of result in part a) of Theorem 4.1). But this result is not used in practice. Assumption 3 is also used by Christensen et al. (2013) to derive the consistency result of the variance estimator of their estimator of Γ_{kl} in the continuous case (cf. equation (4.2) in Christensen et al. (2013)). By imposing this assumption, we are able to build on their results when proving our result. In particular, Assumption 3 plays an essential role in the proof of the estimation of the variance $V_{kl,k'v}^n$. In order for the statistics $\mathcal{Z}_{kl}^n(\alpha)$ and $\mathcal{Z}_{kl}^n(\alpha + 1)$ to estimate the same quantity up to an error sufficiently small, one usually postulates that σ_t is a continuous Itô semimartingale itself.*

Recently, Koike (2015) uses an additional synchronization step based on refresh times for the construction of the pre-averaged truncated Hayashi-Yoshida estimator. As he argued, this additional step enables him to develop fully an asymptotic theory for the estimator in a general setting, which allow for instance a kind of endogenous noise. Nevertheless, even in Koike's (2015) framework the general bootstrap scheme described in (10) can be applied, but for sake of brevity and to ease the exposition, in this paper we will describe the bootstrap method to $\hat{\Gamma}_{kl}^n$ given by (37). In addition, by avoiding an additional synchronization step as in Koike (2015), and imposing Assumption 1 (the sampling scheme as in Christensen et al. (2013)), we are able to build on results in Christensen et al. (2013) when proving our CLT result for $\hat{\Gamma}_{kl}^n$ given by (37). A full exploration of the validity of our high level conditions for Koike's (2015) estimator of Γ_{kl} is left for future research.

Given the results of Section 3, and Theorem 4.1, the asymptotic validity of our bootstrap method for $\hat{\Gamma}_{kl}^n$ given by (37) can be established for non-studentized as well as studentized statistics by verifying the appropriate conditions. Our next result shows that this is the case.

Theorem 4.2. *Suppose all conditions from Theorem 4.1 hold true. Furthermore, suppose that for some $\varepsilon > 0$, $E^* |\eta_\alpha|^{4+\varepsilon} \leq \Delta < \infty$. Then as $n \rightarrow \infty$, Conditions A and (16) are verified and the conclusions of Theorem 3.2 hold.*

This result extends the work of Hounyo (2013), where a local Gaussian bootstrap method have been proposed for inference on integrated covolatility under no jumps by allowing for the latter. It also provides an alternative to the recent general local Gaussian bootstrap method introduced by Dovonon et al. (2015) for jump tests.

4.2 Multivariate realized kernels estimator

In this section, we do not intend to use the same general framework as in Section 4.1 (sampling scheme, activity jumps). Rather our aim here is to focus on a different related and competing covolatility estimator other than the pre-averaged truncated Hayashi-Yoshida estimator and then illustrate the bootstrap method to enhance intuition. Specifically, we illustrate the application of our bootstrap

scheme in context of realized kernel based estimators. The equivalence between realized kernel and pre-averaging based estimator is well-established in the literature, see e.g. Jacod et al (2009) and Christensen et al. (2010). The existence of such relation between both estimators is the key aspect behind the validity of our bootstrap method. In the univariate setting, Jacod et al (2009) show that apart from border terms, i.e. terms close to 0 and 1, the pre-averaging estimator of integrated volatility given by

$$\hat{\Gamma}_{kk}^n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \left(\bar{Y}_{t_i^k}^k \right)^2 - \underbrace{\frac{\psi_1}{2n\theta^2\psi_2} \sum_{i=1}^n \left(\Delta Y_{t_i^k}^k \right)^2}_{\text{bias correction term}},$$

where $\psi_1 = \int_0^1 g'(u)^2 du$ and $\psi_2 = \int_0^1 g(u)^2 du$ coincides with the one-lag "flat-top" realized kernel estimator in Barndorff-Nielsen et al. (2008) using kernel weights

$$k(s) = \psi_2^{-1} \int_s^1 g(u) g(u-s) du, \quad (41)$$

where $g(u)$ is defined as in Section 4.1. In particular, when we choose the bandwidth of the realized kernel estimator equal to the size of the pre-averaging window k_n , the realized kernel and pre-averaging based-estimators have the same asymptotic distribution. Next, we show that similar results hold in multivariate setting.² To fix ideas, let consider synchronous data in the following. Denote the number of synchronized returns by N regardless of the sampling scheme. We follow Barndorff-Nielsen et al. (2011) and we average m prices at the very beginning and end of the day. More specifically, let $n, m_n \in \mathbb{N}^+$, with $n-1+2m_n = N$, $m_n \propto n^{\delta_3}$ with the jittering rate $\delta_3 \in (0, 1)$, and redefine the sample as

$$Y_{t'_0} = \frac{1}{m_n} \sum_{i=1}^{m_n} Y_{t_{i-1}}, \quad Y_{t'_n} = \frac{1}{m_n} \sum_{i=1}^{m_n} Y_{t_{N-m+i}} \text{ and } Y_{t'_i} = Y_{t_{m+i-1}}, \text{ for } i = 1, \dots, n-1.$$

According to equation (1) of Barndorff-Nielsen et al. (2011) (see also equation (5) of Varneskov (2016)), the multivariate realized kernel can be rewritten as

$$\hat{\Gamma}^n = \sum_{i=1}^n k(0) y_i y'_i + \sum_{i=1}^{n-1} \sum_{h=1}^{n-i} k\left(\frac{h}{H_n}\right) (y_i y'_{i+h} + y_{i+h} y'_i), \quad (42)$$

where the high frequency vector returns $y_i \equiv Y_{t'_i} - Y_{t'_{i-1}}$, $i = 1, \dots, n$ and $k: \mathbb{R} \rightarrow \mathbb{R}$ is a non-stochastic weight function. That is characterised by:

Assumption K. (i) $k(0) = 1, k'(0) = 0$; (ii) k is twice differentiable with continuous derivatives; (iii) define $k^{0,0} = \int_0^\infty k(x)^2 dx$, $k^{1,1} = \int_0^\infty k'(x)^2 dx$, and $k^{2,2} = \int_0^\infty k''(x)^2 dx$, then $k^{0,0}, k^{1,1}, k^{2,2} < \infty$; (iv)

²Note that the univariate version of the realized kernel estimator of Barndorff-Nielsen et al. (2011) is not identical to the univariate realized kernel estimator of Barndorff-Nielsen et al. (2008) (for more details see Section 6.1 of Barndorff-Nielsen et al. (2011)).

$$\int_{-\infty}^{\infty} k(x) \exp(ix\lambda) dx \geq 0 \text{ for all } \lambda \in \mathbb{R}.$$

Condition (iv) guarantees $\hat{\Gamma}^n$ to be positive semi-definite. The choice of kernel function impacts the asymptotic distribution of the realized kernels. Therefore may influence the bootstrap approximation of the limiting distribution of $\hat{\Gamma}^n$. In the following development in addition to Assumption K, we assume that $k(x)$ is such that $k(x) = 0$ for $x > 1$. An example of such kernel is the Parzen kernel which is advocated by Barndorff-Nielsen et al. (2011). It is elementary (though somewhat tedious) to show that the realized kernel estimator defined by (42) can be written as

$$\begin{aligned} \hat{\Gamma}^n = & \sum_{i=2}^{n-H_n-m_n+2} y_i y'_i + \sum_{i=2}^{n-H_n-m_n+1} \sum_{h=1}^{H_n} k\left(\frac{h}{H_n}\right) (y_i y'_{i+h} + y_{i+h} y'_i), \\ & + \sum_{i=n-H_n-m_n+3}^{n-1} y_i y'_i + \sum_{i=n-H_n-m_n+2}^{n-2} \sum_{h=1}^{\min(H_n, n-i-1)} k\left(\frac{h}{H_n}\right) (y_i y'_{i+h} + y_{i+h} y'_i) \\ & + \underbrace{(y_1 y'_1 + y_n y'_n) + \sum_{h=1}^{H_n} k\left(\frac{h}{H_n}\right) (y_1 y'_{1+h} + y_{1+h} y'_1 + y_{n-h} y'_n + y_n y'_{n-h})}_{\text{border terms } \tilde{e}^n}. \end{aligned}$$

One can control the effects of jittered end-points on the asymptotic distribution of the realized kernel estimator $\hat{\Gamma}^n$ by choosing the parameters m_n and H_n appropriately. In particular, we have (see e.g. Lemma B.3 of Varneskov (2016))

$$\tilde{e}^n = O_P\left(\frac{m_n}{n}\right) + O_P\left(\frac{(H_n m_n)^{1/2}}{n}\right) + O_P\left(\frac{1}{m_n}\right) + O_P\left(\left(\frac{H_n}{m_n n}\right)^{1/2}\right) + O_P\left(\frac{m_n}{(H_n n)^{1/2}}\right) + O_P\left(\frac{1}{n^{1/2}}\right).$$

Consider a flat-top realized kernel using kernel weights given by (41), and let $\psi_2^{H_n} = \frac{1}{H_n} \sum_{i=1}^{H_n} g^2\left(\frac{i}{H_n}\right)$. It is easy to see that a such kernel verifies Assumption K and is such that $k(x) = 0$ for $x > 1$. Hence, we can rewrite $\hat{\Gamma}^n$ as in (5), where the main term driving the asymptotic distribution of $\hat{\Gamma}^n$ is

$$\begin{aligned} \sum_{\alpha=1}^{J_n} \mathcal{Z}^n(\alpha) &= \frac{1}{\psi_2^{H_n} H_n} \sum_{i=2}^{n-H_n-m_n+2} \sum_{h=1}^{H_n} g^2\left(\frac{h}{H_n}\right) y_i y'_i \\ &\quad + \frac{1}{\psi_2^{H_n} H_n} \sum_{i=2}^{n-H_n-m_n+1} \sum_{h=1}^{H_n-1} \sum_{l=h+1}^{H_n} g\left(\frac{l}{H_n}\right) g\left(\frac{l-h}{H_n}\right) (y_i y'_{i+h} + y_{i+h} y'_i) \\ &= \frac{1}{\psi_2^{H_n} H_n} \sum_{i=m_n+1}^{n-H_n+1} \bar{Y}_{t_i} (\bar{Y}_{t_i})'. \end{aligned} \tag{43}$$

We are now ready to state the bootstrap first-order asymptotic validity result for $\hat{\Gamma}^n$ given by (42).

Theorem 4.3. *Suppose (2), assume that $\hat{\Gamma}^n$ is given by (42) such that (41) holds. Suppose that we observe Y at regular time points. Let Assumptions 2, 3, and 4 be fulfilled.*

a) *If $H_n = \theta n^{3/5}$ for some $\theta > 0$, and $m_n^{-1} = o\left(\sqrt{H_n/n}\right) = o\left(n^{-1/5}\right)$, then (4) holds true, where*

$\tau_n \propto \sqrt{n/H_n}$, $\hat{\Gamma}^n = \left(\hat{\Gamma}_{kl}^n\right)_{1 \leq k, l \leq d}$ with, $\mathcal{B} = \theta^{-2} |k''(0)| \bar{\Psi}$, and $V = 4\theta k^{0,0} \int_0^1 \Sigma_s \oplus \Sigma_s ds$, where $\bar{\Psi}$ is the average long run variance of ϵ .

b) In addition if Assumption H holds, and for some $\varepsilon > 0$, $E^* |\eta_\alpha|^{4+\varepsilon} \leq \Delta < \infty$, then as $n \rightarrow \infty$, Conditions A is verified and the conclusions of Theorem 3.2 hold, where $J_n = \left\lfloor \frac{n-H_n-m_n+1}{b_n} \right\rfloor \propto \frac{n}{b_n}$, with b_n follows (7) such that $1/2 < \delta_2 < 2/3$, $\tau_n = n^{1/5}$ and for $1 \leq \alpha \leq J_n$

$$\mathcal{Z}_{kl}^n(\alpha) = \frac{1}{\psi_2^{H_n} H_n} \sum_{i=1}^{b_n} \bar{Y}_{t_{m_n+i} + \frac{(\alpha-1)}{J_n}}^k \bar{Y}_{t_{m_n+i} + \frac{(\alpha-1)}{J_n}}^l. \quad (44)$$

Part (a) of Theorem 4.3 is well known in the literature (see e.g. Theorem 3 of Barndorff-Nielsen et al. (2011)). This result is only given here for completeness. We emphasize that the paper by Barndorff-Nielsen et al. (2011) goes much further in developing the multivariate realized kernel estimation technology, including non-synchronous trading and allowing certain types of measurement error (such as endogenous noise). Furthermore, their results are extended in Varneskov (2016), who also suggests a class of kernels that are $n^{1/4}$ -consistent and efficient and also accomodate α -mixing dependent exogenous noise.

Given that the realized kernel estimator is very similar to that of the multi-scale estimator of Zhang (2006), cf. Bibinger and Mykland (2016), we conjectured that similar analysis as for kernel-based estimators holds for the multi-scale based estimators, but we have not investigated this. We left the full exploration of this for future research. Due to the relation between realized kernel and pre-averaging based estimator, in the simulation study below, we will not treat separately realized kernel estimator.

5 Monte Carlo results

In this section, we assess by Monte Carlo simulation the accuracy of the feasible asymptotic theory based on the limiting result in Theorem 4.1, which is inspired by the approach in Wang, Liu and Liu (2013) (and combines the methods in Christensen's et al. (2013) and Jing, Liu and Kong (2014)). We find that this approach leads to important coverage probability distortions when returns are not sampled too frequently. We also compare the finite sample performance of this approach with the wild blocks of blocks bootstrap method. The design of our Monte Carlo study is roughly identical to that used by Christensen et al. (2010) and Barndorff-Nielsen et al. (2011) with some minor differences. In particular, in addition to the continuous semimartingale component, we add a jump process. About the microstructure noise, in addition to the case of i.i.d. noise, we also look at the case of autocorrelated noise. Here we briefly describe the Monte Carlo design we use.

To simulate log-prices we consider the following bivariate stochastic volatility model

$$dX_t^{(i)} = dX_t^{c(i)} + dJ_t^{(i)}, \text{ for } i = 1, 2,$$

where $X_t^{c(1)}$ and $X_t^{c(2)}$ are continuous semimartingales, $J_t^{(1)}$ and $J_t^{(2)}$ are Lévy processes. The bivariate factor stochastic volatility model of Barndorff-Nielsen et al. (2011) is used to generate the continuous semimartingales $X_t^{c(1)}$ and $X_t^{c(2)}$:

$$dX_t^{c(i)} = a^{(i)}dt + \rho^{(i)}\sigma_t^{(i)}dB_t^{(i)} + \sqrt{1 - [\rho^{(i)}]^2}\sigma_t^{(i)}dW_t, \text{ for } i = 1, 2,$$

where $B^{(i)}$ and W are independent Brownian motions. In this model, the term $\rho^{(i)}\sigma_t^{(i)}dB_t^{(i)}$ is an idiosyncratic component, while $\sqrt{1 - [\rho^{(i)}]^2}\sigma_t^{(i)}dW_t$ is a common factor.

The spot volatility is modeled as $\sigma_t^{(i)} = \exp\left(\beta_0^{(i)} + \beta_1^{(i)}\varrho_t^{(i)}\right)$ with an Ornstein-Uhlenbeck specification for $\varrho_t^{(i)} : d\varrho_t^{(i)} = \alpha^{(i)}\varrho_t^{(i)}dt + dB_t^{(i)}$. This implies that there is perfect correlation between the innovations of $\rho^{(i)}\sigma_t^{(i)}dB_t^{(i)}$ and $\sigma_t^{(i)}$, while it is $\rho^{(i)}$ between the increments of $X_t^{c(i)}$ and $\varrho_t^{(i)}$. Note that, the magnitude of correlation between the two underlying price processes $X_t^{c(1)}$ and $X_t^{c(2)}$ is $R = \sqrt{1 - [\rho^{(1)}]^2}\sqrt{1 - [\rho^{(2)}]^2}$. Finally, we set

$$J_t^{(1)} = L_t^{(1)}, \text{ and } J_t^{(2)} = RL_t^{(1)} + \sqrt{1 - R^2}L_t^{(2)},$$

where $L^{(1)}$ and $L^{(2)}$ are two independent tempered-stable processes (or CGMY processes) with Lévy measure given by

$$\nu^{(i)}(x) = c^{(i)} \frac{\exp(-d^{(i)-}|x|)}{|x|^{1+b^{(i)}}} 1_{\{x < 0\}} + c^{(i)} \frac{\exp(-d^{(i)+}x)}{x^{1+b^{(i)}}} 1_{\{x > 0\}}, \quad i = 1, 2,$$

where $b^{(i)} \in (0, 1)$ to ensure finite variation.

For the continuous semimartingales $X_t^{c(i)}$, the reported results are based on the following configuration of parameters for both processes: $(a^{(i)}, \beta_0^{(i)}, \beta_1^{(i)}, \rho^{(i)}, \alpha^{(i)}) = (0.03, -5/16, 1/8, -1/40, -0.3)$, so that $\beta_0^{(i)} = [\beta_1^{(i)}]^2 / [2\alpha^{(i)}]$. We note that this particular choice of parameters also means that the volatility process has been normalized, in the sense that $E\left(\int_0^1 [\sigma^{(i)}(s)ds]^2\right) = 1$. We use the same parameter values for the jump process as in Aït-Sahalia and Xiu (2016). In particular, for $i = 1, 2$, we set $b^{(i)} = 0.5$, $d^{(i)+} = 3$, $d^{(i)-} = 5$, and $c^{(i)}$ are calibrated such that the percentage of quadratic variation contributed by jumps in $X^{(i)}$ are 10%. This reflects the empirical results in Huang and Tauchen (2005).

We simulate data for the unit interval $[0, 1]$, and normalize one second to be $1/23400$, so that $[0, 1]$ represent 6.5 hours worth of trading, which is then further decomposed into $N = 23,400$ subintervals of equal length $1/N$. In constructing noisy prices $Y^{(i)}$, we first generate a complete high frequency record of N equidistant observations of the efficient price $X^{(i)}$ using a standard Euler scheme. We initialize the spot volatility $\sigma_t^{(i)}$ at the start of each interval by drawing the initial values for the $\varrho_t^{(i)}$ processes from its stationary distribution, i.e. $\varrho_0^{(i)} \sim N\left(0, [2\alpha^{(i)}]^{-1}\right)$. The size of the market microstructure noise is an important parameter. We follow Barndorff-Nielsen et al. (2011) and model the noise magnitude as $\xi^2 = \omega^2 / \sqrt{\int_0^1 \sigma_s^4 ds}$. We fix ξ^2 equal to 0, 0.001 and 0.01 (which covers scenarios with no noise through low-to-high levels of noise) and let $\omega^2 = \xi^2 \sqrt{\int_0^1 \sigma_s^4 ds}$. This means that the variance of the

noise process increases with the level of volatility of the efficient price $X^{(i)}$, as documented by Bandi and Russell (2006). These values are motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of the Dow Jones Industrial Average, see also Aït-Sahalia and Yu (2009). We follow Kalnina (2011) and add autocorrelated microstructure noise simulated as an $MA(1)$ process (for a given frequency of the observations):

$$\epsilon_{\frac{j}{n}}^{(i)} = u_{\frac{j-1}{n}}^{(i)} + \gamma u_{\frac{j}{n}}^{(i)}, \text{ where } u^{(i)} | \{\sigma, X\} \stackrel{i.i.d.}{\sim} N\left(0, \frac{\omega^2}{1+\gamma^2}\right),$$

so that $Var(\epsilon^{(i)}) = \omega^2$. The observed process is then given by $Y^{(i)} = X^{(i)} + \epsilon^{(i)}$. Three different values of γ are considered, $\gamma = 0$, $\gamma = -0.5$ and $\gamma = -0.9$ (which covers scenarios of i.i.d. noise, moderate and high level of correlation of noise). We follow Christensen et al. (2010) and use the conservative choice of k_n ($\theta = 1$, implying that $k_n = \sqrt{n}$). We also follow the literature and use the weight function $g(x) = (x \wedge (1-x))$ to compute the pre-averaged returns. In order to reduce finite sample biases associated with Riemann integrals, we replace in (37), $\psi = \int_0^1 g(s) ds$ by its Riemann approximation

$$\text{given by } \psi_n = \frac{1}{k_n} \sum_{i=0}^{k_n} g\left(\frac{i}{k_n}\right).$$

Finally, we extract irregular, non-synchronous data from the complete high-frequency record using Poisson process sampling to generate actual observation times, $\{t_j^{(i)}\}$. In particular, we consider two independent Poisson processes with intensity parameter $\lambda = (\lambda_1, \lambda_2)$. Here λ_i denotes the average waiting time (in seconds) for new data from process $Y^{(i)}$, so that an average day will have N/λ_i observations of $Y^{(i)}$, $i = 1, 2$. We vary λ_1 through $(3, 10, 60)$ to capture the influence of liquidity on the performance of the pre-averaged multivariate volatility estimator and we set $\lambda_2 = 2\lambda_1$ such that on average $Y^{(2)}$ refreshes at half the pace of $Y^{(1)}$.

We determine the truncation level parameters $\{\alpha^{(i)}, \varpi^{(i)}\}$ for individual processes. For $i = 1, 2$, we follow Podolskij and Ziggel (2010) and use $\varpi^{(i)}/2 = 0.17$, and choose $\alpha^{(i)}$ such that

$$\alpha^{(i)} = 2.3 \sqrt{BV^{(i)}(1, 1) + \frac{\psi_1}{\theta} NV^{(i)}}, \quad (45)$$

where

$$BV^{(i)}(1, 1) = \frac{\pi}{2} n^{-\frac{1}{2}} \sum_{j=0}^{n_i - 2k_n + 1} \left| \bar{Y}_{t_j^i}^i \right| \left| \bar{Y}_{t_{j+k_n}^i}^i \right|, \text{ and } NV^{(i)} = \frac{\sum_{j=1}^{n_i} \left(\Delta Y_{t_j^i}^i \right)^2}{2n_i}.$$

The intuition behind this choice is rather simple. In absence of jumps on the interval $[t_j^i, t_{j+k_n}^i]$ and under some regularity conditions, $\bar{Y}_{t_j^i}^i$ is asymptotically distributed as $n^{-1/4} N\left(0, \theta \psi_2 \left(\sigma_{t_j^i}^{(i)}\right)^2 + \frac{\psi_1}{\theta} \omega^2\right)$. In (45) the quantity 2.3 is approximately the 99%-quantile of the standard normal distribution. In addition, under some appropriate conditions (see e.g. Podolskij and Vetter (2009)) we have that

$$BV^{(i)}(1, 1) + \frac{\psi_1}{\theta} NV^{(i)} \rightarrow^P \int_0^1 \left(\theta \psi_2 \left(\sigma_s^{(i)} \right)^2 + \frac{\psi_1}{\theta} \omega^2 \right) ds.$$

For more details on this choice, see Section 5 in Podolskij and Ziggel (2010).

Table 1 gives the actual coverage probability rates of 95% confidence intervals of the three co-variation measures (integrated covariance, integrated correlation and integrated regression coefficients) computed over 10,000 replications. Results based on the asymptotic normal distribution and the wild blocks of blocks bootstrap method are included under the label CLT and WB, respectively.

In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications. We consider the bootstrap percentile method computed at the 95% level. To generate the bootstrap data we use two different external random variables.

WB1 $\eta_\alpha \sim \text{i.i.d. } N(1, 1/2)$.

WB2 A two point distribution $\eta_\alpha \sim \text{i.i.d.}$ such that:

$$\eta_\alpha = \begin{cases} \left(\frac{1}{2}\right)^{1/2} \left(\frac{1+2\sqrt{2}+\sqrt{5}}{2}\right), & \text{with prob } p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \left(\frac{1}{2}\right)^{1/2} \left(\frac{1+2\sqrt{2}-\sqrt{5}}{2}\right), & \text{with prob } 1-p = \frac{\sqrt{5}+1}{2\sqrt{5}} \end{cases},$$

for which $E^*(\eta_\alpha) = 1$ and $Var^*(\eta_\alpha) = 1/2$.

Note that both choices of η_α are asymptotically valid when used to construct bootstrap percentile as well as percentile- t intervals. WB2 is related to a certain external random variable with a two point distribution used in Gonçalves et al. (2014) cf. Section 3 also WB2 in their notation. But we have centered and scaled appropriately such that $E^*(\eta_\alpha) = 1$ and $Var^*(\eta_\alpha) = 1/2$. Based on simulations results, Gonçalves et al. (2014) advocated the use of WB2. The choice of the bootstrap block size is critical. We follow Politis, Romano and Wolf (1999) and Hounyo et al. (2016) and use the Minimum Volatility Method to choose the bootstrap block (for further details see Hounyo et al. (2016)).

For the three covariation measures, all intervals tend to undercover. The degree of undercoverage is especially large, when the average arrival times of trades is not too frequent. Results are not very sensitive to the noise magnitude nor to the level of correlation.

Starting with unstudentized statistics. The results show that bootstrap percentile intervals based on WB1 are close to the CLT-based intervals. Whereas for WB2, it is the case only when the average arrival times of trades is very frequent, see e.g. when $\lambda = (3, 6)$. The gains associated with the wild blocks of blocks bootstrap method based on WB2 can be quite substantial, especially for larger values of λ_1 and λ_2 (long average waiting time for new data from process $Y^{(1)}$ and $Y^{(2)}$), when distortions of the CLT-based intervals are larger. For instance, when $\gamma = -0.5$ (moderate level of correlation of noise), $\xi^2 = 0.01$ (high level of noise), and $\lambda = (60, 120)$ (illiquid assets), for the regression coefficient, the coverage rate for a symmetric bootstrap percentile interval based on WB2 is equal to 85.55%, whereas it is equal to 81.18% for the feasible asymptotic theory based on the limiting result in Theorem 4.1. The gains are especially important for the correlation coefficient, when the asymptotic theory-based intervals does worst. The bootstrap percentile interval based on WB2 has a rate of 89.80%, whereas

the CLT-based interval has a rate of 80.10%. For the covariance, these numbers are equal to 85.33% and 81.09%, for the percentile bootstrap based on WB2 and the CLT-based intervals, respectively. When the average arrival times of trades become frequent, all methods have coverage rates closer to the desired level. For instance, for the CLT-based intervals, when $\gamma = -0.9$ (high level of correlation of noise), $\xi^2 = 0.001$ (low level of noise), and $\lambda = (3, 6)$ (liquid assets), a two-sided 95% confidence interval for the covariance measure between the two assets has coverage rate equal to 92.13%, whereas it is equal to 92.00% for the regression coefficient. These numbers increase to 93.95% and 93.91% for the bootstrap percentile-based intervals using WB2. The bootstrap performance is quite remarkable for the correlation coefficient where it essentially removes all finite sample bias associated with the first-order asymptotic theory based on the limiting result in Theorem 4.1.

For percentile- t intervals, results based on both external random variables WB1 and WB2 are close, but slightly different. Overall, a comparison between bootstrap percentile and percentile- t confidence intervals shows that results based on bootstrap studentized statistics outperform those based on bootstrap unstudentized statistics.

In summary, the results in Table 1 show that the performance of the asymptotic theory-based intervals, the bootstrap percentile and percentile- t intervals in terms of coverage rate crucially depends on the average arrival times of trades. In fact for non-frequent arrival times of trade, the asymptotic normal approximation is often inaccurate and leads to important coverage distortions. In addition, results suggest that the choice of the external random variable η is important when the average arrival times of trades is not too frequent. Formal study of optimal choice of the external random variable in this general context is an interesting research question which we will consider elsewhere. In all cases, the bootstrap percentile- t intervals outperform the the CLT-based intervals.

6 Empirical application

To illustrate some empirical features of the wild blocks of blocks bootstrap theory developed above, we analyse high-frequency assets prices for four assets. In the analysis we focus on the realized beta estimator based on pre-averaged returns. In particular, we compare the empirical properties of the bootstrap to the feasible asymptotic procedure based on the limiting result in Theorem 4.1. The data is the collection of trades recorded on the NYSE in July 2013, taken from the TAQ database through the Wharton Research Data Services (WRDS) system. This results in 22 distinct trading days. We picked 3 equities at random from the S&P 500 constituents list as of July 1, 2014. They are Microsoft Co. (listed under the ticker symbol (MSFT)), Boeing Co. (BA) and WPX Energy Inc. (CPWR). We then added a 4th element, namely the S&P 500 Depository Receipt (ticker symbol SPY). The SPY is an exchange-traded fund that tracks the large-cap segment of the US stock market. As such, it can be viewed as generating market-wide index returns. For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. until 4 p.m. Eastern Standard Time.

Our procedure for cleaning the data is identical to that used by Barndorff-Nielsen et al. (2011) (for further details see this paper). Table 2 reports some summary statistics of the data (before and after cleaning). As can be seen, these equities display varying degrees of liquidity with MSFT and SPY being the most liquid, while CPWR is the least liquid.

Figure 1: Summary statistics of raw and pre-averaged SPY trade data over regular exchange opening days in July 2013.

To implement the pre-averaged returns in tick time as given in (35), we select the tuning parameter θ by following the conservative rule ($\theta = 1$, implying that $k_n = \sqrt{n}$). We tune the truncation level parameters $\{\alpha^{(i)}, \varpi^{(i)}\}$ as in our Monte Carlo simulation, in particular, $\varpi^{(i)} = 0.34$, and $\alpha^{(i)}$ is given by (45). For the bootstrap as well as the CLT-based approach, to choose the block size b_n , we follow Politis, Romano and Wolf (1999) and use the minimum volatility method (see Appendix A of Hounyo

Figure 2: 95% Confidence Intervals (CI's) for the daily the pre-averaged truncated Hayashi-Yoshida estimator-based beta estimates, for each regular exchange opening days for BA, CPWR and MSFT in July 2013, calculated using the feasible asymptotic theory based on the limiting result in Theorem 4.1 (CI's with bars), and the wild blocks of blocks bootstrap method (CI's with lines). The pre-averaged truncated Hayashi-Yoshida estimator-based beta estimate is the middle of all CI's by construction. Days on the x -axis.

et al. (2016) for details).

We start by analysing the high frequency data. Figure 1 shows time series, autocorrelation and histogram of raw returns as well as of pre-averaged returns for SPY. We observe a pronounced serial correlation in raw returns and in pre-averaged returns. In particular, for raw returns the first autocorrelation is large and negative. This is typical of noisy data and unlikely to arise from a Brownian semimartingale. Note that, the strong autocorrelation observed for pre-averaged returns in Panel D of Figure 1 is due to the fact that we have considered overlapping pre-averaged returns, which rely on many common raw returns. This has nothing to do with the fact that raw returns are possibly noisy. In fact, the correlogram (not reported here) of non-overlapping pre-averaged returns shows that the latter are almost uncorrelated (even for the first lag). The effect of pre-averaging is nicely illustrated by comparing Panel E and F of Figure 1. It appears that pre-averaging helps to reduce price discreteness effect observed in raw returns. At the same time, return distribution is now much closer to being Gaussian. These results are not surprising, it confirms theoretical properties of pre-averaged returns. In particular, under mild conditions on the dynamics of the price process we have that $n^{1/4}\bar{Y}_{\frac{\alpha}{n}}^k|\mathcal{F}_{(\frac{\alpha-1}{n})b_n}^n \stackrel{a}{\sim} N\left(0, \theta\psi_2\left(\sigma_{\frac{\alpha}{n}}^{(k)}\right)^2 + \frac{\psi_1}{\theta}\omega^2\right)$. Similar patterns (not reported here) are observed for MSFT, BA and CPWR.

We now turn to the realized beta for MSFT, BA and CPWR. We consider bootstrap percentile- t intervals, computed at the 95% level, where η_α is generated using WB2. The results are displayed in Figure 2 in terms of daily 95% symmetric confidence intervals for the latent realized beta. Two types of intervals are presented: our proposed wild blocks of blocks bootstrap method and the feasible asymptotic theory based on the limiting result in Theorem 4.1. The pre-averaged truncated Hayashi-Yoshida estimator-based beta estimate is in the center of both confidence intervals by construction. In fact, similar series of confidence intervals for beta was also graphed by Dovonon et al. (2013) in their Figures 1 and 2, except that they used daily log-returns to calculate estimated betas (based on realized covariance) over intervals of one quarter. The emphasis of their paper was to illustrate the usefulness of the bootstrap as a method of inference on beta in a context, where the mechanics of trading is perfect so that there is no market microstructure effects and prices are observed synchronously. In Figure 2, beta is estimated using full record transaction prices. For all stocks considered in the present study, the width of confidence intervals (the bootstrap and the asymptotic theory-based) varies through time. Also, there are a lot of variability in the daily estimate of beta, but all of them lie in the positive region. This means that, these stocks move in the same direction as the market.

As illustrated below, a closer analysis of Figure 2 show that these common patterns observed for MSFT, BA and CPWR hide different empirical features which allow us to gain valuable insights into the empirical performance of the wild blocks of blocks bootstrap method. For MSFT: the most liquid stock after SPY considered in our analysis, a comparison of the bootstrap intervals with the intervals based on the feasible asymptotic approach suggests that the two types of interval tend to be quite similar. In contrast to MSFT, for the less liquid stock considered here, i.e. CPWR, in most of the

cases the confidence intervals for daily beta based on the bootstrap method are usually wider than the confidence intervals using the feasible asymptotic theory. For BA, there is no evidence about the relative empirical performance of the bootstrap and the asymptotic theory-based. These observations lead us to conclude that the degree of liquidity of assets, specifically the non-trading of MSFT, BA or CPWR versus SPY influences the width of confidence intervals, although the conclusion might change for other data sets. Note that, as our Monte Carlo simulations showed, the asymptotic theory-based approach typically have undercoverage problems whereas the bootstrap intervals have coverage rates closer to the desired level. Therefore, if the goal is to control the coverage probability, shorter intervals are not necessarily better.

7 Conclusion

This paper proposes the bootstrap as a method of inference for integrated covariance matrix. We show that the wild blocks of blocks bootstrap studied by Hounyo et al. (2016) can be used to simultaneously handle the presence of dependence, jumps, heterogeneity, irregularly spaced and non-synchronous trading properties of high-frequency data. This combination of properties is unique in the bootstrap literature, so it is worthwhile exploring this bootstrap method in some detail. The percentile bootstrap method is particularly useful because it circumvents the need for an explicit estimator of the asymptotic variance, which has proved difficult in our context.

We provide a set of conditions under which this method is asymptotically valid to first-order. We then verify these conditions for two estimators of integrated covolatility: the pre-averaged truncated Hayashi-Yoshida estimator and the multivariate realized kernel estimator. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improves the finite sample properties of the alternative approach based on the Gaussian approximation (first-order asymptotic theory). Furthermore, an empirical illustration highlights the usefulness of our approach as an alternative method of inference for realized covariation measures and its applicability to real high-frequency data. In future work, we plan to study the higher-order accuracies of this bootstrap method. Another important extension is to provide a theoretical optimal choice of the block size b_n for confidence interval construction.

Appendix A

Tables 1 reports the actual coverage rates for the feasible asymptotic theory based on the limiting result in Theorem 4.1 and for our bootstrap methods using the optimal block size by minimizing confidence interval volatility. In Table 2 we provide some descriptive statistics and liquidity measures for BA, CPWR, MSFT and SPY recorded on the NYSE in July 2013.

Table 1. Coverage rates of nominal 95% intervals

	$\gamma = 0$ (i.i.d. noise)						$\gamma = -0.5$						$\gamma = -0.9$					
	Percentile			Percentile- t			Percentile			Percentile- t			Percentile			Percentile- t		
	CLT	WB1	WB2	WB1	WB2	WB2	CLT	WB1	WB2	WB1	WB2	WB2	CLT	WB1	WB2	WB1	WB2	WB2
Covariance																		
$\xi^2 = 0$																		
$\lambda = (3, 6)$	92.13	92.13	94.01	96.20	95.16		92.17	92.17	94.04	96.25	95.18		92.12	92.12	93.93	95.83	94.98	
$\lambda = (10, 20)$	88.89	88.89	91.85	94.82	94.29		88.89	88.89	91.82	94.80	94.25		88.89	88.89	91.80	94.74	94.19	
$\lambda = (60, 120)$	81.16	81.16	85.31	88.63	87.77		81.08	81.08	85.29	88.63	87.77		81.12	81.12	85.25	88.51	87.66	
$\xi^2 = 0.001$																		
$\lambda = (3, 6)$	92.15	92.15	94.02	96.19	95.17		92.18	92.18	94.06	96.27	95.21		92.13	92.13	93.95	95.84	94.96	
$\lambda = (10, 20)$	88.90	88.90	91.87	94.84	94.31		88.91	88.91	91.84	94.81	94.25		88.90	88.90	91.81	94.75	94.21	
$\lambda = (60, 120)$	81.17	81.17	85.32	88.65	87.78		81.10	81.10	85.33	88.65	87.79		81.13	81.13	85.26	88.53	87.68	
$\xi^2 = 0.01$																		
$\lambda = (3, 6)$	92.16	92.16	94.04	96.22	95.17		92.19	92.19	94.06	96.26	95.19		92.12	92.12	93.95	95.84	94.99	
$\lambda = (10, 20)$	88.92	88.92	91.86	94.85	94.34		88.90	88.90	91.83	94.82	94.26		88.90	88.90	91.84	94.75	94.20	
$\lambda = (60, 120)$	81.16	81.16	85.31	88.64	87.78		81.09	81.09	85.33	88.62	87.81		81.15	81.15	85.26	88.52	87.67	
Regression																		
$\xi^2 = 0$																		
$\lambda = (3, 6)$	92.06	91.97	93.90	96.10	95.06		92.01	91.98	93.90	96.10	95.06		91.68	91.60	93.94	96.06	95.02	
$\lambda = (10, 20)$	88.94	88.99	91.88	94.85	94.31		88.92	88.97	91.82	94.79	94.26		88.61	88.68	91.74	94.73	94.20	
$\lambda = (60, 120)$	81.20	81.35	85.35	88.67	87.82		81.21	81.34	85.54	88.87	88.00		80.37	80.47	85.42	88.75	87.88	
$\xi^2 = 0.001$																		
$\lambda = (3, 6)$	92.03	91.97	93.91	96.12	95.08		91.92	91.92	93.92	96.13	95.08		91.71	91.65	93.94	96.07	95.03	
$\lambda = (10, 20)$	88.92	89.01	91.89	94.86	94.33		88.89	88.89	91.83	94.80	94.27		88.63	88.67	91.74	94.74	94.21	
$\lambda = (60, 120)$	81.20	81.32	85.37	88.62	87.85		81.18	81.18	85.55	88.88	88.02		80.40	80.51	85.43	88.76	87.90	
$\xi^2 = 0.01$																		
$\lambda = (3, 6)$	92.03	91.98	93.92	96.12	95.08		91.92	91.87	93.92	96.13	95.07		91.71	91.66	93.94	96.06	95.03	
$\lambda = (10, 20)$	88.94	88.99	91.89	94.87	94.34		88.89	88.94	91.83	94.81	94.28		88.64	88.68	91.73	94.74	94.23	
$\lambda = (60, 120)$	81.23	81.36	85.38	88.62	87.86		81.18	81.36	85.55	88.88	88.02		80.41	80.50	85.42	88.76	87.92	
Correlation																		
$\xi^2 = 0$																		
$\lambda = (3, 6)$	92.18	92.04	93.95	96.78	95.10		91.96	91.83	93.97	96.70	95.04		92.04	92.00	93.91	96.40	94.86	
$\lambda = (10, 20)$	87.96	88.17	93.28	94.95	94.31		87.85	87.99	93.23	94.90	94.27		87.78	87.96	93.22	94.82	94.17	
$\lambda = (60, 120)$	80.03	80.23	89.85	91.56	91.12		80.01	80.11	89.78	91.47	91.08		79.97	80.12	89.68	91.40	91.02	
$\xi^2 = 0.001$																		
$\lambda = (3, 6)$	92.20	92.18	93.96	96.79	95.12		92.00	92.00	93.98	96.69	95.03		92.06	92.01	93.92	96.41	94.87	
$\lambda = (10, 20)$	87.97	88.17	93.29	94.96	94.31		87.98	87.98	93.24	94.89	94.28		87.81	87.97	93.23	94.82	94.18	
$\lambda = (60, 120)$	80.06	80.25	89.86	91.57	91.13		80.08	80.12	89.80	91.47	91.09		79.98	80.15	89.69	91.39	91.03	
$\xi^2 = 0.01$																		
$\lambda = (3, 6)$	92.22	92.19	93.97	96.78	95.10		92.24	92.24	93.99	96.70	95.04		92.04	92.00	93.92	96.42	94.89	
$\lambda = (10, 20)$	87.98	88.18	93.30	94.97	94.30		88.01	88.01	93.24	94.90	94.29		87.82	87.98	93.23	94.83	94.19	
$\lambda = (60, 120)$	80.09	80.31	89.86	91.58	91.13		80.10	80.14	89.80	91.47	91.10		79.98	80.14	89.70	91.38	91.04	

Notes: CLT-intervals based on the Normal; WB1 wild blocks of blocks bootstrap based on the external random variable WB1; WB2 wild blocks of blocks bootstrap based on the external random variable WB2. Only in the case for covariance, WB1 percentile intervals are computed using the normal approximation, i.e., we proceed as in (19), otherwise percentile intervals are based on (25); Percentile- t intervals are based on (26). 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 2. Descriptive statistics and number of data before and after filtering.

Stock	BA	CPWR	MSFT	SPY
Raw trades	783,150	155,413	3,160,226	5,557,249
Corrected/Abnormal/Zeros	10	26	36	12
Time aggregation	645,249	125,242	2,889,825	5,191,067
# Trades	137,891	30,145	270,365	366,170
Intensity	6,268	1,370	12,289	16,644

Note. This table reports some descriptive statistics and liquidity measures for the selection of stocks included in our empirical application. Raw trades is the total number of data available from these exchanges during the trading session, while # trades is the total sample remaining after filtering the data. Intensity is the average number of data per day.

Appendix B

In the following and throughout the appendix, K denotes a constant, which may change from line to line and from (in)equality to (in)equality.

Proof of Lemma 3.1 Part a). Given (10) and (11), result follows directly since we can write

$$\begin{aligned}
E^* \left(\hat{\Gamma}_{kl}^{n*} \right) &= \sum_{\alpha=1}^{J_n} E^* \left(\mathcal{Z}_{kl}^{n*}(\alpha) \right) \\
&= \sum_{\alpha=1}^{J_n-1} E^* \left(\mathcal{Z}_{kl}^{n*}(\alpha) \right) + E^* \left(\mathcal{Z}_{kl}^{n*}(J_n) \right) \\
&= \sum_{\alpha=1}^{J_n-1} [\mathcal{Z}_{kl}^n(\alpha+1) + (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) E^*(\eta_\alpha)] + \mathcal{Z}_{kl}^n(J_n).
\end{aligned}$$

Then, under the condition $E^*(\eta_\alpha) = 1$, we have that

$$\begin{aligned}
E^* \left(\hat{\Gamma}_{kl}^{n*} \right) &= \sum_{\alpha=1}^{J_n} \mathcal{Z}_{kl}^n(\alpha) \\
&= \hat{\Gamma}_{kl}^n + \tilde{b}_{kl}^n.
\end{aligned}$$

Proof of Lemma 3.1 Part b). Given the definition of $V_{kl,k'l'}^{n*}$, equations (10) and (11) we have that

$$\begin{aligned}
V_{kl,k'l'}^{n*} &= \tau_n^2 E^* ((Z_{kl}^{n*} - E^*(Z_{kl}^{n*})) (Z_{k'l'}^{n*} - E^*(Z_{k'l'}^{n*}))) \\
&= \tau_n^2 \sum_{\alpha=1}^{J_n-1} \sum_{\alpha'=1}^{J_n-1} E^* \left((Z_{kl}^{n*}(\alpha) - E^*(Z_{kl}^{n*}(\alpha))) (Z_{k'l'}^{n*}(\alpha') - E^*(Z_{k'l'}^{n*}(\alpha'))) \right) \\
&= \tau_n^2 \sum_{\alpha=1}^{J_n-1} \sum_{\alpha'=1}^{J_n-1} ((Z_{kl}^n(\alpha) - Z_{kl}^n(\alpha+1))) (Z_{k'l'}^n(\alpha') - Z_{k'l'}^n(\alpha'+1)) Cov^*(\eta_\alpha, \eta_{\alpha'}).
\end{aligned}$$

Using the fact that $\eta_\alpha \sim \text{i.i.d.}$, result follows, then we get

$$\begin{aligned}
V_{kl,k'l'}^{n*} &= 2Var^*(\eta) \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (Z_{kl}^n(\alpha) - Z_{kl}^n(\alpha+1)) (Z_{k'l'}^n(\alpha) - Z_{k'l'}^n(\alpha+1)) \\
&= 2Var^*(\eta) V_{kl,k'l'}^n.
\end{aligned}$$

Proof of Theorem 3.1 Part a). Result follows directly given part b) of Lemma 3.1 and Condition A1.

Proof of Theorem 3.1 Part b). Let $\hat{\Gamma}_{kl}^{n*}(\alpha) \equiv (Z_{kl}^{n*}(\alpha))_{1 \leq k, l \leq d}$, where $Z_{kl}^{n*}(\alpha)$ is defined in (10), and let $x_\alpha^* \equiv vec(\hat{\Gamma}_{kl}^{n*}(\alpha))$. We have that $S^{n*} \equiv n^{1/4} (vec(\hat{\Gamma}^{n*}) - E^*(vec(\hat{\Gamma}^{n*}))) = \tau_n \sum_{\alpha=1}^{J_n} (x_\alpha^* - E^*(x_\alpha^*))$. The proof follows from showing that for any $\lambda \in \mathbb{R}^{d^2}$ such that $\lambda' \lambda = 1$, $\sup_{x \in \mathbb{R}} |P^*(\sum_{\alpha=1}^{J_n} \tilde{x}_\alpha^* \leq x) - \Phi(x/(\lambda' \tilde{V} \lambda))| \xrightarrow{P} 0$, where $\tilde{x}_\alpha^* = \tau_n \lambda' (x_\alpha^* - E^*(x_\alpha^*))$, and $\tilde{V} = (\tilde{V}_{kl})_{1 \leq k, l \leq d^2}$ is a $d^2 \times d^2$ matrix, whose generic element $\tilde{V}_{kl}$ is given by

$$\tilde{V}_{kl} = V_{k-d[(k-1)/d], [(k-1)/d]+1, l-d[(l-1)/d], [(l-1)/d]+1},$$

with $1 \leq k, l \leq d^2$. Clearly, $E^*(\sum_{\alpha=1}^{J_n} \tilde{x}_\alpha^*) = 0$ and $Var^*(\sum_{\alpha=1}^{J_n} \tilde{x}_\alpha^*) = \lambda' \tilde{V} \lambda \xrightarrow{P} \lambda' \tilde{V} \lambda$ by part a). Thus, by Katz's (1963) Berry-Essen Bound, for some small $\varepsilon > 0$ and some constant $K > 0$ which changes from line to line, $\sup_{x \in \mathbb{R}} |P^*(\sum_{\alpha=1}^{J_n} \tilde{x}_\alpha^* \leq x) - \Phi(x/(\lambda' \tilde{V} \lambda))| \leq K \sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon}$. Next, we show that $\sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon} = o_p(1)$. We have that

$$\begin{aligned}
\sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon} &= \sum_{\alpha=1}^{J_n} E^* |\tau_n \lambda' (x_\alpha^* - E^*(x_\alpha^*))|^{2+\varepsilon} \\
&\leq 2^{2+\varepsilon} \tau_n^{2+\varepsilon} \sum_{\alpha=1}^{J_n} E^* |\lambda' x_\alpha^*|^{2+\varepsilon} \\
&\leq 2^{2+\varepsilon} \tau_n^{2+\varepsilon} \sum_{\alpha=1}^{J_n} E^* |x_\alpha^*|^{2+\varepsilon} \\
&\leq K \tau_n^{2+\varepsilon} E^* |\eta_1|^{2+\varepsilon} \sum_{\alpha=1}^{J_n} |x_\alpha|^{2+\varepsilon},
\end{aligned}$$

where the first inequality follows from the C_r and the Jensen inequalities; the second inequality uses the Cauchy-Schwarz inequality and the fact that $\lambda' \lambda = 1$; and the third inequality follows from the C_r

and the Jensen inequalities. We let $|z|^2 = (z'z)$ for any vector z . It follows that

$$\begin{aligned}
\sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon} &\leq K \tau_n^{2+\varepsilon} E^* |\eta_1|^{2+\varepsilon} \sum_{\alpha=1}^{J_n} |x_\alpha|^{2(1+\varepsilon/2)} \\
&\leq K \tau_n^{2+\varepsilon} E^* |\eta_1|^{2+\varepsilon} \sum_{\alpha=1}^{J_n} \left(\sum_{k=1}^d \sum_{l=1}^d (\mathcal{Z}_{kl}^n(\alpha))^2 \right)^{1+\varepsilon/2} \\
&\leq \underbrace{K E^* |\eta_1|^{2+\varepsilon}}_{=O(1)} \underbrace{\tau_n^{2+\varepsilon} \left(\frac{b_n}{n} \right)^{1+\varepsilon}}_{=o(1)} \underbrace{\left(\frac{n}{b_n} \right)^{1+\varepsilon} \sum_{\alpha=1}^{J_n} |\mathcal{Z}_{kl}^n(\alpha)|^{2+\varepsilon}}_{=O_P(1)} = o_P(1), \tag{46}
\end{aligned}$$

where consistency follows since for any $\varepsilon > 0$, $E^* |\eta_\alpha|^{2+\varepsilon} \leq \Delta < \infty$, and by using Conditions A.2. and A.3.

Proof of Theorem 3.1 Part c). Given that $T^n \rightarrow^{st} N(\text{vec}(\mathcal{B}), I_{d^2})$, it suffices to show that $T^{n*} \rightarrow^{d*} N(0, I_{d^2})$ in probability- P . Let

$$H^{n*} = (\tilde{V}^{n*})^{-1/2} S^{n*},$$

and note that, we can write

$$T^{n*} = (\hat{V}^{n*})^{-1/2} (\tilde{V}^{n*})^{1/2} H^{n*},$$

where S^{n*} , $\hat{V}^{n*}$ and $\tilde{V}^{n*}$ are defined in the main text. Given (17) and the fact that $\hat{V}^{n*}$ and $\tilde{V}^{n*}$ are non singular in large samples with probability approaching one we have

$$(\hat{V}^{n*})^{-1} \tilde{V}^{n*} = (\tilde{V}^{n*})^{-1} \hat{V}^{n*} \rightarrow^{P*} I_{d^2}.$$

Hence, to obtain the desired CLT for T^{n*} , it suffices to show that

$$H^{n*} \rightarrow^{d*} N(0, I_{d^2}). \tag{47}$$

To prove (47), we proceed as in the proof of part b) of Theorem 3.1. Similarly, the proof follows from showing that for any $\lambda \in \mathbb{R}^{d^2}$ such that $\lambda' \lambda = 1$, $\sup_{x \in \mathbb{R}} |P^*(\sum_{\alpha=1}^{J_n} \tilde{y}_\alpha^* \leq x) - \Phi(x)| \xrightarrow{P} 0$, where $\tilde{y}_\alpha^* \equiv (\lambda' \tilde{V}^{n*} \lambda)^{-1/2} \tilde{x}_\alpha^* = (\lambda' \tilde{V}^{n*} \lambda)^{-1/2} \tau_n \lambda' (x_\alpha^* - E^*(x_\alpha^*))$, and $\tilde{V}^{n*} = (\tilde{V}_{kl}^{n*})_{1 \leq k, l \leq d^2}$ is a $d^2 \times d^2$ matrix whose generic element $\tilde{V}_{kl}^{n*}$ is defined by (18). Clearly, $E^*(\sum_{\alpha=1}^{J_n} \tilde{y}_\alpha^*) = 0$ and $\text{Var}^*(\sum_{\alpha=1}^{J_n} \tilde{y}_\alpha^*) = 1$ by part a). Thus, by Katz's (1963) Berry-Essen Bound, for some small $\varepsilon > 0$ and some constant $K > 0$, $\sup_{x \in \mathbb{R}} |P^*(\sum_{\alpha=1}^{J_n} \tilde{y}_\alpha^* \leq x) - \Phi(x)| \leq K \sum_{\alpha=1}^{J_n} E^* |\tilde{y}_\alpha^*|^{2+\varepsilon}$. Next, we show that $\sum_{\alpha=1}^{J_n} E^* |\tilde{y}_\alpha^*|^{2+\varepsilon} = o_P(1)$. We have that

$$\begin{aligned}
\sum_{\alpha=1}^{J_n} E^* |\tilde{y}_\alpha^*|^{2+\varepsilon} &= |\lambda' \tilde{V}^{n*} \lambda|^{-(2+\varepsilon)/2} \sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon} \\
&\leq \left| \min_{1 \leq i \leq d^2} \varphi_i(\tilde{V}^{n*}) \right|^{-(2+\varepsilon)/2} |\lambda' \lambda|^{-(2+\varepsilon)/2} \sum_{\alpha=1}^{J_n} E^* |\tilde{x}_\alpha^*|^{2+\varepsilon} = o_P(1),
\end{aligned}$$

where $\varphi_i(\tilde{V}^{n*})$ denotes the i -th eigenvalue of $\tilde{V}^{n*}$, results follows given that $\tilde{V}^{n*}$ is non singular in large samples with probability approaching one, $\lambda'\lambda = 1$ and by using (46).

Proof of Theorem 3.2. Parts a) and b). Since S^n converges stably in distribution to $N((\text{vec}(\mathcal{B})), \tilde{V})$, by an application of the delta method (see Podolskij and Vetter (2010, Proposition 2.5(iii))),

$$S_h^n \rightarrow^{st} N\left(h(\text{vec}(\mathcal{B})), \nabla' h\left(\text{vec}\left(\int_0^1 \Sigma_s ds\right)\right) \tilde{V} \nabla h\left(\text{vec}\left(\int_0^1 \Sigma_s ds\right)\right)\right).$$

Similarly, by a mean value expansion, and conditionally on the original sample,

$$S_h^{n*} = \tau_n \nabla' h\left(E^*\left(\text{vec}(\hat{\Gamma}^{n*})\right)\right) \left(\text{vec}(\hat{\Gamma}^{n*}) - E^*\left(\text{vec}(\hat{\Gamma}^{n*})\right)\right) + o_{P^*}(1),$$

since $\hat{\Gamma}_{kl}^{n*} - E^*\left(\hat{\Gamma}_{kl}^{n*}\right) \rightarrow^{P^*} 0$ in probability. It follows that

$$S_h^{n*} \rightarrow^{st} N\left(0, \nabla' h\left(\text{vec}\left(\int_0^1 \Sigma_s ds\right)\right) \tilde{V} \nabla h\left(\text{vec}\left(\int_0^1 \Sigma_s ds\right)\right)\right)$$

in probability, given Theorem 3.1. The result follows from Polya's theorem (see, e.g. Serfling (1980)), given that the normal distribution is continuous.

Proof of Theorem 3.2. Part c). Follows similarly as the proof of parts b) of Theorem 3.2.

Proofs of results in Section 4

As in Jacod and Protter (2014), we assume in the following that the processes a, σ, δ and X are bounded processes satisfying (1). As Jacod et al. (2010) explain, this assumption simplifies the mathematical derivations without loss of generality (by a standard localization procedure detailed e.g. in Jacod (2008)). Formally, we derive our results under the following assumption.

Assumption 5. X satisfies equation (1) with a and σ adapted càdlàg processes such that a, σ, δ and X are bounded processes. In particular, for some constant K and nonnegative deterministic function $\tilde{\gamma}$, we have

$$\|a_t(\omega)\| \leq K, \|\sigma_t(\omega)\| \leq K, \|X_t(\omega)\| \leq K, \|\delta(\omega, t, z)\| \leq \tilde{\gamma}(z) \leq K, \int_{\mathbb{R}^d} \tilde{\gamma}(z)^r \lambda(dz) \leq K.$$

Notation

We define the continuous part of X by X' and discontinuous martingale part by X'' , that is

$$X'_t = X_0 + \int_0^t a'_s ds + \int_0^t \sigma_s dW_s, \quad X''_t = X_t - X'_t \quad (48)$$

where $a'_s = a_s - \int_0^t \int \kappa(\delta(s, z)) \nu(ds, dz)$. Then, observe that we can write

$$Y_t = Y'_t + Y''_t, \quad (49)$$

where $Y'_t = X'_t + \epsilon_t$ and $Y''_t = X''_t$. Let

$$PHY(Y)_{kl}^n = \frac{1}{(\psi k_n)^2} \sum_{i=0}^{n_k-k_n+1} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}}, \quad (50)$$

where $\psi = \int_0^1 g(s) ds$, and $A_{ij}^{kl} = \{(i, j) : (t_i^k, t_{i+k_n}^k] \cap (t_j^l, t_{j+k_n}^l] \neq \emptyset\}$. We also let $PTHY(Y)_{kl}^n = \hat{\Gamma}_{kl}^n$ denote the pre-averaged truncated Hayashi-Yoshida estimator as defined by (37). Finally, since below we will rely on different estimator $\hat{\Gamma}^n$ of the integrated covariance matrix Γ , we will write $\tilde{V}^n(\hat{\Gamma}^n)$ and $V^n(\hat{\Gamma}^n)$ instead of simply $\tilde{V}^n(\hat{\Gamma}^n)$ and $V^n(\hat{\Gamma}^n)$, respectively, to precise that the variance estimator is rely on $\hat{\Gamma}^n$.

Lemma 7.1. *Suppose (1) and Assumptions 1-5 hold. Furthermore suppose that $\hat{\Gamma}_{kl}^n$ is given by*

$$\hat{\Gamma}_{kl}^n = PHY(Y')_{kl}^n = \frac{1}{(\psi k_n)^2} \sum_{i=0}^{n_k-k_n+1} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}}.$$

Then, we have

$$V_{kl,k'l'}^n = \frac{\tau_n^2}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1)) \rightarrow^P V_{kl,k'l'}.$$

where $\tau_n = n^{1/4}$,

$$\mathcal{Z}_{kl}^n(\alpha) = \frac{1}{(\psi k_n)^2} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}},$$

$B_n(\alpha) = \left[\frac{(\alpha-1)b_n}{n}, \frac{\alpha b_n}{n} \right)$, $J_n = \left\lfloor \frac{n}{b_n} \right\rfloor$, with b_n follows (7) such that $1/2 < \delta_2 < 2/3$ and $V_{kl,k'l'}$ is the same as in Theorem 3.4 of Christensen et al. (2013).

Proof of Lemma 7.1. Given the definitions of $V_{kl,k'l'}^n$ and $\mathcal{Z}_{kl}^n(\alpha)$, after adding and subtracting appropriately, we get that

$$\begin{aligned} V_{kl,k'l'}^n &= \frac{\sqrt{n}}{2} \left(\sum_{\alpha=1}^{J_n-1} (2\mathcal{Z}_{kl}^n(\alpha) \mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha) \mathcal{Z}_{k'l'}^n(\alpha+1) - \mathcal{Z}_{kl}^n(\alpha+1) \mathcal{Z}_{k'l'}^n(\alpha)) \right) \\ &\quad + \frac{\sqrt{n}}{2} (\mathcal{Z}_{kl}^n(J_n) \mathcal{Z}_{k'l'}^n(J_n) - \mathcal{Z}_{kl}^n(1) \mathcal{Z}_{k'l'}^n(1)) \\ &\equiv L_{kl,k'l'}^n + R_{kl,k'l'}^n. \end{aligned}$$

where the remainder term is

$$\begin{aligned} R_{kl,k'l'}^n &= \frac{\sqrt{n}}{2} (\mathcal{Z}_{kl}^n(J_n) \mathcal{Z}_{k'l'}^n(J_n) - \mathcal{Z}_{kl}^n(1) \mathcal{Z}_{k'l'}^n(1)) \\ &= O_P\left(n^{-\frac{3}{2}} b_n^2\right) = O_P\left(\left(\frac{b_n}{n^{3/4}}\right)^2\right) = o_P(1), \end{aligned}$$

so long as $\delta_2 < 3/4$, where we used the definitions of $\mathcal{Z}_{kl}^n(\alpha) = \frac{1}{(\psi k_n)^2} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}}$, the Cauchy-Schwartz inequality, the fact that for some $q > 0$, $E\left(\left|\bar{Y}_{t_i^k}^k\right|^q\right) \leq K n^{-q/4}$ uniformly in i .

Proof of Theorem 4.1 Part a). Under the stated assumptions, given the definitions of $PHY(Y')^n$, $PTHY(Y)$ and the central limit theorem of Christensen et al. (2013, equation (4.6) with $b = 1$) in conjunction with Christensen et al. (2013, Remark 4.5) and a stable Cramer-Wold theorem, we have that, as $n \rightarrow \infty$,

$$\left(\tilde{V}^n(PHY(Y')^n)\right)^{-1/2} n^{1/4} (vec(PHY(Y')^n) - vec(\Gamma)) \rightarrow^d N(0, I_{d^2}).$$

Thus, to obtain the desired result, we need to prove that

$$n^{1/4} (vec(PTHY(Y)^n) - vec(PTHY(Y')^n)) \rightarrow^P 0, \text{ and } \tilde{V}^n(PTHY(Y)^n) - \tilde{V}^n(PTHY(Y')^n) \rightarrow^P 0.$$

This amounts to show that

$$n^{1/4} (PTHY(Y)_{kl}^n - PTHY(Y')_{kl}^n) \rightarrow^P 0, \text{ and} \quad (51)$$

$$V_{kl,k'l'}^n(PTHY(Y)^n) - V_{kl,k'l'}^n(PTHY(Y')^n) \rightarrow^P 0. \quad (52)$$

To show (51), let $F_u(x) = x^k x^l 1_{C^{kl}}$, where $C^{kl} = \{|x^k| < u^l\} \cap \{|x^l| < u^l\}$ and $u = u^k \vee u^l$. Following the same argument on page 385 in Jacod and Protter (2014), we can show that for $w_n = (v_n^k \vee v_n^l) / \sqrt{u_n}$ with $u_n = k_n/n$, that

$$|F_{w_n}(x+y) - F_{w_n}(x)| \leq w_n^{\frac{-2}{1-2\omega}} \|x\|^{2+\frac{2}{1-2\omega}} + \left((1 + \|x\|^2) (\|y\| \wedge 1 + \|y\|^2 \wedge w_n^2)\right).$$

Next, let $x^a = \bar{Y}_{t_a^k}^{la} / \sqrt{u_n}$, $y^a = \bar{Y}_{t_a^k}^{na} / \sqrt{u_n}$, with $a = k, l$. According to (16.4.9) in Jacod and Protter (2014) in conjunction with results in part 3 in the proof of Lemma 16.4.5 of Jacod and Protter (2014), for some $q > 0$, we have

$$E \|x\|^q \leq K, \quad E(\|y\| \wedge 1) \leq K u_n^{1-r/2} \phi_n \text{ and } E(\|y\|^2 \wedge w_n^2) \leq K u_n^{\omega(2-r)} \phi_n, \quad (53)$$

where $\phi_n \rightarrow 0$, as $n \rightarrow 0$, see also e.g., the proof of Theorem 2 of Ait-Sahalia and Xiu (2016) for similar results. Thus, using the above inequalities, together with the fact that by definition

$$n^{1/4} (PTHY(Y)_{kl}^n - PTHY(Y')_{kl}^n) = \frac{n^{1/4}}{(\psi k_n)^2} \sum_{i=0}^{n_k-k_n+1} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} \left(\bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l - \bar{Y}_{t_i^k}^{lk} \bar{Y}_{t_j^l}^{ll} \right) 1_{C_{ij}^{kl}},$$

it follows that

$$\begin{aligned} \frac{n^{1/4}}{(\psi k_n)^2} \underbrace{\sum_{i=0}^{n_k-k_n+1} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} E \left| \left(\bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l - \bar{Y}_{t_i^k}^{lk} \bar{Y}_{t_j^l}^{ll} \right) 1_{C_{ij}^{kl}} \right|}_{\leq K n k_n} &\leq K n^{1/4} n^{1/2} \cdot \underbrace{u_n}_{=n/k_n} \left(u_n + u_n^{1-r/2} \phi_n + u_n^{\omega(2-r)} \phi_n \right) \\ &\leq K \left(n^{-1/4} + n^{\frac{(r-1)}{4}} \phi_n + n^{\frac{1-2\omega(2-r)}{4}} \phi_n \right) \\ &\rightarrow 0, \end{aligned}$$

which conclude the proof of (51).

To show (52), remark that by definition of $V_{kl,k'l'}^n(PTHY(Y)^n)$ and $V_{kl,k'l'}^n(PTHY(Y')^n)$, we can

write

$$\begin{aligned}
& V_{kl,k'l'}^n (PTHY (Y)^n) - V_{kl,k'l'}^n (PTHY (Y')^n) \\
&= \frac{\sqrt{n}}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n (\alpha) \mathcal{Z}_{k'l'}^n (\alpha) - \mathcal{Z}_{kl}^m (\alpha) \mathcal{Z}_{k'l'}^m (\alpha)) \\
&\quad - \frac{\sqrt{n}}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n (\alpha) \mathcal{Z}_{k'l'}^n (\alpha+1) - \mathcal{Z}_{kl}^m (\alpha) \mathcal{Z}_{k'l'}^m (\alpha+1)) \\
&\quad - \frac{\sqrt{n}}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n (\alpha+1) \mathcal{Z}_{k'l'}^n (\alpha) - \mathcal{Z}_{kl}^m (\alpha+1) \mathcal{Z}_{k'l'}^m (\alpha)) \\
&\quad + \frac{\sqrt{n}}{2} \sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n (\alpha+1) \mathcal{Z}_{k'l'}^n (\alpha+1) - \mathcal{Z}_{kl}^m (\alpha+1) \mathcal{Z}_{k'l'}^m (\alpha+1)) \tag{54}
\end{aligned}$$

where

$$\mathcal{Z}_{kl}^n (\alpha) = \frac{1}{(\psi k_n)^2} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{A_{ij}^{kl}} 1_{C_{ij}^{kl}}, \text{ and } \mathcal{Z}_{kl}^m (\alpha) = \frac{1}{(\psi k_n)^2} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} \bar{Y}_{t_i^k}^{l_k} \bar{Y}_{t_j^l}^{l_l} 1_{A_{ij}^{kl}} 1_{C_{ij}^{kl}}.$$

Let

$$\pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} = \left(\bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l \bar{Y}_{t_{i'}^{k'}}^{k'} \bar{Y}_{t_{j'}^{l'}}^{l'} - \bar{Y}_{t_i^k}^{l_k} \bar{Y}_{t_j^l}^{l_l} \bar{Y}_{t_{i'}^{k'}}^{l_{k'}} \bar{Y}_{t_{j'}^{l'}}^{l_{l'}} \right) 1_{C_{ij}^{kl}} 1_{C_{i'j'}^{k'l'}},$$

then (54) becomes

$$\begin{aligned}
& V_{kl,k'l'}^n (PTHY (Y)^n) - V_{kl,k'l'}^n (PTHY (Y')^n) \\
&= \frac{\sqrt{n}}{2 (\psi k_n)^4} \sum_{\alpha=1}^{J_n-1} \left(\sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} \sum_{t_{i'}^{k'} \in B_n(\alpha)} \sum_{j'=0}^{n_{l'}-k_n+1} 1_{A_{i'j'}^{k'l'}} \right) \pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} \\
&\quad - \frac{\sqrt{n}}{2 (\psi k_n)^4} \sum_{\alpha=1}^{J_n-1} \left(\sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} \sum_{t_{i'}^{k'} \in B_n(\alpha+1)} \sum_{j'=0}^{n_{l'}-k_n+1} 1_{A_{i'j'}^{k'l'}} \right) \pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} \\
&\quad - \frac{\sqrt{n}}{2 (\psi k_n)^4} \left(\sum_{\alpha=1}^{J_n-1} \sum_{t_i^k \in B_n(\alpha+1)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} \sum_{t_{i'}^{k'} \in B_n(\alpha)} \sum_{j'=0}^{n_{l'}-k_n+1} 1_{A_{i'j'}^{k'l'}} \right) \pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} \\
&\quad + \frac{\sqrt{n}}{2 (\psi k_n)^4} \sum_{\alpha=1}^{J_n-1} \left(\sum_{t_i^k \in B_n(\alpha+1)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} \sum_{t_{i'}^{k'} \in B_n(\alpha+1)} \sum_{j'=0}^{n_{l'}-k_n+1} 1_{A_{i'j'}^{k'l'}} \right) \pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} \\
&\equiv V_{kl,k'l'}^{(1)n} (Y, Y') + V_{kl,k'l'}^{(2)n} (Y, Y') + V_{kl,k'l'}^{(3)n} (Y, Y') + V_{kl,k'l'}^{(4)n} (Y, Y').
\end{aligned}$$

It follows that the claim staded in (52) can be reduced to

$$V_{kl,k'l'}^{(1)n} (Y, Y') \xrightarrow{P} 0, \quad (55)$$

$$V_{kl,k'l'}^{(2)n} (Y, Y') \xrightarrow{P} 0, \quad (56)$$

$$V_{kl,k'l'}^{(3)n} (Y, Y') \xrightarrow{P} 0 \text{ and} \quad (57)$$

$$V_{kl,k'l'}^{(4)n} (Y, Y') \xrightarrow{P} 0. \quad (58)$$

The results in (55), (56), (57) and (58) can be proven exactly in the same way. In fact, it is sufficient to show (55). Thus in the following we show only (55). To this end, we let $F_u(x) = x^k x^l x^{k'} x^{l'} 1_{C^{kl}} 1_{C^{k'l'}}$, where $C^{kl} = \{|x^k| < u^l\} \cap \{|x^l| < u^l\}$ and $u = \max(u^k, u^l, u^{k'}, u^{l'})$. Following the same argument on page 385 in Jacod and Protter (2014), we can now show that for $w_n = \max(v_n^k, v_n^l, v_n^{k'}, v_n^{l'}) / \sqrt{u_n}$ with $u_n = k_n/n$, that

$$|F_{w_n}(x+y) - F_{w_n}(x)| \leq w_n^{\frac{-2}{1-2\omega}} \|x\|^{4+\frac{2}{1-2\omega}} + \left((1 + \|x\|^4) (\|y\| \wedge 1 + \|y\|^4 \wedge w_n^4) \right).$$

Next, let $x^a = \bar{Y}_{t_a^k}^a / \sqrt{u_n}$, $y^a = \bar{Y}_{t_a^k}^a / \sqrt{u_n}$, with $a = k, l, k', l'$. Thus, using (53) and the inequality $(\|y\| \wedge w_n)^p \leq w_n^{p-m} (\|y\| \wedge w_n)^m$ for $0 < m < p$, we have

$$E \left(\|y\|^4 \wedge w_n^4 \right) \leq K w_n^2 E (\|y\| \wedge w_n)^2 \leq K u_n^{\omega(4-r)-1} \phi_n, \quad (59)$$

where $\phi_n \rightarrow 0$, as $n \rightarrow 0$. Therefore, we obtain

$$\begin{aligned} & \frac{\sqrt{n}}{2(\psi k_n)^4} \sum_{\alpha=1}^{J_n-1} \left(\underbrace{\sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}}}_{\leq K b_n k_n} \underbrace{\sum_{t_{i'}^{k'} \in B_n(\alpha)} \sum_{j'=0}^{n_{l'}-k_n+1} 1_{A_{i'j'}^{k'l'}}}_{\leq K b_n k_n} \right) \underbrace{E \left| \pi_{t_i^k, t_j^l, t_{i'}^{k'}, t_{j'}^{l'}} \mathcal{F}_{\frac{(\alpha-1)b_n}{n}}^n \right|}_{\leq K u_n^2 \left(\underbrace{u_n}_{=n/k_n} + u_n^{1-r/2} \phi_n + u_n^{\omega(4-r)-1} \phi_n \right)} \\ & \leq K n^{\delta_2 - \frac{1}{2}} \left(n^{-\frac{1}{2}} + n^{\frac{r-2}{4}} \phi_n + n^{\frac{-\omega(4-r)-2}{2}} \phi_n \right) \\ & \leq K \left(n^{\delta_2-1} + n^{\frac{4\delta_2-4+r}{4}} \phi_n + n^{\frac{2\delta_2-3-\omega(4-r)}{2}} \phi_n \right) \rightarrow 0, \end{aligned}$$

which conclude the proof of (52).

Proof of Theorem 4.1 Part b). The result follows by the application of the delta method.

Proof of Theorem 4.2. The validity of Condition A.1. is a consequence of Lemma 7.1 in conjunction with result in (52). Condition A.2. also holds because under our assumptions we have $\sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} \leq$

Kb_nk_n . Then, for some $q > 1$, by using C-r inequality, follows by Cauchy-Schwartz inequality, we have

$$\begin{aligned}
E |\mathcal{Z}_{kl}^n(\alpha)|^q &\leq K (b_nk_n)^{q-1} \frac{1}{k_n^{2q}} \sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}} E \left| \bar{Y}_{t_i^k}^k \bar{Y}_{t_j^l}^l 1_{C_{ij}^{kl}} \right|^q \\
&\leq K (b_nk_n)^{q-1} \frac{1}{k_n^{2q}} \underbrace{\sum_{t_i^k \in B_n(\alpha)} \sum_{j=0}^{n_l-k_n+1} 1_{A_{ij}^{kl}}}_{\leq b_nk_n} \left(E \left| \bar{Y}_{t_i^k}^k 1_{\left\{ \left| \bar{Y}_{t_i^k}^k \right| < v_n^k \right\}} \right|^{2q} \right)^{1/2} \left(E \left| \bar{Y}_{t_j^l}^l 1_{\left\{ \left| \bar{Y}_{t_j^l}^l \right| < v_n^l \right\}} \right|^{2q} \right)^{1/2} \\
&\leq K (b_nk_n)^q \frac{1}{k_n^{2q}} E \left| \bar{Y}_{t_i^k}^k 1_{\left\{ \left| \bar{Y}_{t_i^k}^k \right| < v_n^k \right\}} \right|^{2q} \\
&\leq K \left(\frac{b_n}{k_n^2} \right)^q, \tag{60}
\end{aligned}$$

where the last inequality follows since for $\left(\frac{1}{2(2-r)} \vee \frac{2\delta_2}{4-r} \right) \leq \varpi < 1/2$, for some $\delta_2 \in \left(\frac{1}{2}, \frac{2}{3} \right)$, $E \left| \bar{Y}_{t_i^k}^k 1_{\left\{ \left| \bar{Y}_{t_i^k}^k \right| < v_n^k \right\}} \right|^{2q} \leq \frac{K}{k_n^q}$. Condition A.3. follows since for any $\varepsilon > 2$ and $1/2 < \delta_2 < 2/3$ we have that $\frac{2+\varepsilon}{4} + (\delta_2 - 1)(1 + \varepsilon) < 0$. This conclude the proof of Condition A. Thus results in parts a) and b) of Theorem 3.2 holds for $\hat{\Gamma}_{kl}^n$ given by (37). Finally, in order to claim that the conclusions in part (c) of Theorem 3.2 also holds for $\hat{\Gamma}_{kl}^n$ given by (37), we have to verify that (17) is also satisfy. In particular, we show that (1) $Bias^* \left(\check{V}_{kl,k'l'}^{n*} \right) = 0$, and (2) $Var^* \left(\check{V}_{kl,k'l'}^{n*} \right) \xrightarrow{P} 0$. It is easy to verify that (1) holds by the definition of $\check{V}_{kl,k'l'}^{n*}$ and $V_{kl,k'l'}^{n*}$. To prove (2), note that

$$\begin{aligned}
Var^* \left(\check{V}_{kl,k'l'}^{n*} \right) &= E^* \left(\check{V}_{kl,k'l'}^{n*} - V_{kl,k'l'}^{n*} \right)^2 \\
&= \left(\frac{Var^*(\eta) \tau_n^2}{E^*(\eta^2)} \right)^2 E^* \left[\sum_{\alpha=1}^{J_n-1} (\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1)) (\eta_\alpha^2 - E^*(\eta^2)) \right]^2 \\
&= \left(\frac{Var^*(\eta) \tau_n^2}{E^*(\eta^2)} \right)^2 E^* (\eta^2 - E^*(\eta^2))^2 \sum_{\alpha=1}^{J_n-1} [(\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1))]^2 \\
&\leq K \tau_n^4 \sum_{\alpha=1}^{J_n-1} [(\mathcal{Z}_{kl}^n(\alpha) - \mathcal{Z}_{kl}^n(\alpha+1)) (\mathcal{Z}_{k'l'}^n(\alpha) - \mathcal{Z}_{k'l'}^n(\alpha+1))]^2 \\
&\leq Kn \sum_{\alpha=1}^{J_n-1} \left[(\mathcal{Z}_{kl}^n(\alpha) \mathcal{Z}_{k'l'}^n(\alpha))^2 + (\mathcal{Z}_{kl}^n(\alpha) \mathcal{Z}_{k'l'}^n(\alpha+1))^2 \right] \\
&\quad + Kn \sum_{\alpha=1}^{J_n-1} \left[(\mathcal{Z}_{kl}^n(\alpha+1) \mathcal{Z}_{k'l'}^n(\alpha))^2 + (\mathcal{Z}_{kl}^n(\alpha+1) \mathcal{Z}_{k'l'}^n(\alpha+1))^2 \right] \\
&\equiv K \left(V_{kl,k'l'}^{(1)n} + V_{kl,k'l'}^{(2)n} + V_{kl,k'l'}^{(3)n} + V_{kl,k'l'}^{(4)n} \right).
\end{aligned}$$

It follows that $Var^* \left(\check{V}_{kl,k'l'}^{n*} \right) \xrightarrow{P} 0$ can be reduced to $V_{kl,k'l'}^{(i)n} \xrightarrow{P} 0$, for $i = 1, \dots, 4$. In fact, it is sufficient to show for $i = 1$, since results for $i = 1, \dots, 4$ can be proven exactly in the same way. Thus in the following we show only, that $V_{kl,k'l'}^{(1)n} \xrightarrow{P} 0$. We utilize the fact that convergence in L_1 implies convergence

in probability. In particular, we show that $E \left| V_{kl,k'l'}^{(1)n} \right| \rightarrow 0$. By using the triangular inequality follows by Cauchy-Schwartz inequality, we have

$$\begin{aligned}
E \left| V_{kl,k'l'}^{(1)n} \right| &= Kn E \left| \sum_{\alpha=1}^{J_n-1} E \left(\mathcal{Z}_{kl}^n(\alpha) \mathcal{Z}_{k'l'}^n(\alpha) \right)^2 \right| \\
&\leq Kn \sum_{\alpha=1}^{J_n-1} E \left(|\mathcal{Z}_{kl}^n(\alpha)|^2 |\mathcal{Z}_{k'l'}^n(\alpha)|^2 \right) \\
&\leq Kn \sum_{\alpha=1}^{J_n-1} \underbrace{E \left(|\mathcal{Z}_{kl}^n(\alpha)|^4 \right)^{1/2} E \left(|\mathcal{Z}_{k'l'}^n(\alpha)|^4 \right)^{1/2}}_{\leq K \left(\frac{b_n}{k_n} \right)^2} \\
&\leq K \frac{b_n^3}{n^2} \rightarrow 0,
\end{aligned}$$

where the last inequality follows given (38) and the fact that $J_n \propto \frac{n}{b_n}$ with b_n follows (7) such that $1/2 < \delta_2 < 2/3$.

Proof of Theorem 4.3. Part (a) follows under our assumptions by using Theorem 3 of Barndorff-Nielsen et al. (2011). The proof of part (b) is achieved by using the same arguments alike the ones presented in the proof of Theorem 4.2. To see this, let

$$\begin{aligned}
\phi_1(s) &= \int_s^1 g'(u) g'(u-s) du, \quad \phi_2(s) = \int_s^1 g(u) g(u-s) du, \quad \Phi_{ij} = \int_0^1 \phi_i(s) \phi_j(s) ds, \\
\text{and for } i &= 1, 2, \quad \psi_i = \phi_i(0).
\end{aligned}$$

Next, note that under assumptions of Theorem 4.3, $|k''(0)| = \psi_1/\psi_2$ and $k^{0,0} = \Phi_{22}$, which follows from the identity:

$$k''(s) = \frac{1}{\psi_2} \int_s^1 g(u) g''(u-s) du,$$

and integration by parts (recall that $g(0) = g(1) = 0$). Therefore we can write $\mathcal{B} = \theta^{-2} |k''(0)| \bar{\Psi} = \theta^{-2} \psi_1/\psi_2 \bar{\Psi}$, and $V = 4\theta k^{0,0} \int_0^1 \Sigma_s \oplus \Sigma_s ds = 4\theta \Phi_{22} \int_0^1 \Sigma_s \oplus \Sigma_s ds$. In other words, under our assume conditions, the asymptotic distribution of the realized kernel estimator given by (42) coincides with those of the modulated realized covariance estimator based on pre-averaged returns given by $\frac{1}{\psi_2 H_n} \sum_{i=0}^{n-H_n+1} \bar{Y}_{t_i} (\bar{Y}_{t_i})'$ (see e.g. Theorem 4 of Christensen et al. (2010)). Recall In particular, the validity of Condition A.1. follows the same line as the proof of Lemma 7.1. Condition A.2. also follows because under our assumptions, for some $q > 0$ we have $E(|\bar{Y}_{t_i}^k|^q) \leq K \left(\sqrt{\frac{H_n}{n}} \right)^q$ uniformly in i and similarly we have $\mathcal{Z}_{kl}^n(\alpha) = O_P \left(\frac{b_n}{n} \right)$ uniformly in α . Condition A.3. follows since for any $\varepsilon > 2$ and $1/2 < \delta_2 < 2/3$ we have that $\frac{2+\varepsilon}{4} + (\delta_2 - 1)(1 + \varepsilon) < 0$. This conclude the proof of Condition A. The proof of (17) also follows similarly and therefore we omit the details.

References

- [1] Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys (2003). "Modeling and forecasting realized volatility," *Econometrica*, 71, 529-626.
- [2] Aït-Sahalia, Y., J. Fan, and D. Xiu (2010). "High frequency covariance estimates with noisy and asynchronous financial data," *Journal of the American Statistical Association*, 18, 351-416.
- [3] Aït-Sahalia, P.A. Mykland, and L. Zhang (2005). "How often to sample a continuous-time process in the presence of market microstructure noise," *Review of Financial Studies*, 105, 1504-1517.
- [4] Aït-Sahalia, Y., and Xiu, D. (2016). "Increased Correlation Among Asset Classes: Are Volatility or Jumps to Blame, or Both?" *Journal of Econometrics*, Forthcoming.
- [5] Aït-Sahalia, Y., and D. Yu (2009). "High frequency market microstructure noise estimates and liquidity measures," *Annals of Applied Statistics*, 3, 422-457.
- [6] Bandi, F., and J. Russell (2006). "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79(3), 655-692.
- [7] Bandi, F., and J. Russell (2011). "Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations," *Journal of Econometrics*, 160, 145-159.
- [8] Barndorff-Nielsen, O., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard, (2006). "A central limit theorem for realised power and bipower variations of continuous semimartingales." In Y. Kabanov, R. Lipster, and J. Stoyanov (Eds.), *From Stochastic Analysis to Mathematical Finance*, Festschrift for Albert Shiryaev, 33 - 68. Springer.
- [9] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008). "Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise," *Econometrica*, 76, 1481-1536.
- [10] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2011). "Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," *Journal of Econometrics*, 162, 149-169.
- [11] Barndorff-Nielsen, O. and N. Shephard, (2004). "Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics," *Econometrica*, 72, 885-925.
- [12] Barndorff-Nielsen, O. and N. Shephard, (2006). "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, 4, 1-30.
- [13] Bibinger, M. (2011). "Efficient covariance estimation for asynchronous noisy high-frequency data," *Scandinavian Journal of Statistics*, 38, 23-45.

- [14] Bibinger, M., N. Hautsch, P. Malec, and M. Reis (2014). “Estimating the quadratic covariation matrix from noisy observations: local method of moments and efficiency,” *Annals of Statistics*, 42(4), 1312–1346.
- [15] Bibinger, M. and P.A. Mykland (2016). “Inference for multi-dimensional high-frequency data with an application to conditional independence testing,” *Scandinavian Journal of Statistics*, Forthcoming.
- [16] Bühlmann, P. and H. R. Künsch (1995). “The blockwise bootstrap for general parameters of a stationary time series,” *Scandinavian Journal of Statistics*, 22(1), 35-54.
- [17] Christensen, K., S. Kinnebrock, and M. Podolskij (2010). “Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data,” *Journal of Econometrics*, 159, 116-133.
- [18] Christensen, K., M. Podolskij, and M. Vetter (2013). “On covariation estimation for multivariate continuous Ito semimartingales with noise in non-synchronous observation schemes,” *Journal of Multivariate Analysis*, 120, 59-84.
- [19] Corsi, F., S. Peluso, and F. Audrino (2014). “Missing asynchronicity: a Kalman-EM approach to multivariate realized covariance estimation,” *Journal of Applied Econometrics*, 30(3), 377-397.
- [20] Delbaen, F., and W. Schachermayer (1994). “A general version of the fundamental theorem of asset pricing,” *Mathematische Annalen*, 300, 463-520.
- [21] Diebold, F.X. and Strasser, G.H. (2013). “On the correlation structure of microstructure noise: a financial economic approach,” *Review of Economics Studies*, 80, 1304-1337.
- [22] Dovonon, P., Gonçalves, S., Hounyo, U. and N. Meddahi (2015). “Bootstrapping high-frequency jump tests,” manuscript.
- [23] Dovonon, P., Gonçalves, S. and N. Meddahi, (2013). “Bootstrapping realized multivariate volatility measures,” *Journal of Econometrics*, 172, 49-65.
- [24] Epps, T. W., (1979). “Comovements in stock prices in the very short run,” *Journal of the American Statistical Association*, 74(366), 291-298.
- [25] Gonçalves, S. and N. Meddahi (2009). “Bootstrapping realized volatility,” *Econometrica*, 77(1), 283-306.
- [26] Gonçalves, S., Hounyo, U. and N. Meddahi (2014). “Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns,” *Journal of Financial Econometrics*, 12(4), 679-707.
- [27] Gonçalves, S., and H. White (2002). “The bootstrap of the mean for dependent heterogeneous arrays,” *Econometric Theory*, 18, 1367-1384.
- [28] Hansen, P.R. and A. Lunde (2006). “Realized variance and market microstructure noise,” *Journal of Business and Economics Statistics*, 24, 127-161.

- [29] Hasbrouck, J. (2007). “Empirical Market Microstructure,” *Oxford University Press*.
- [30] Hautsch N., and M. Podolskij (2013). “Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, Implementation, and Empirical Evidence,” *Journal of Business and Economic Statistics*, 31(2), 165-183.
- [31] Hayashi, T., and N. Yoshida (2005). “On covariance estimation of non-synchronously observed diffusion processes,” *Bernoulli*, 11, 359-379.
- [32] Hounyo, U. (2013). “Bootstrapping realized volatility and realized beta under a local Gaussianity assumption,” Research paper 2013-30, CREATES, Aarhus University.
- [33] Hounyo, U. , Gonçalves, S., and N. Meddahi (2016). “Bootstrapping pre-averaged realized volatility under market microstructure noise,” *Econometric Theory*, Forthcoming.
- [34] Hounyo, U., and B. Veliev (2016). “Validity of Edgeworth expansions for realized volatility estimators,” *Econometrics Journal*, 19(1), 1-32.
- [35] Huang, X. and G. Tauchen (2005). “The relative contribution of jumps to total price variance,” *Journal of Financial Econometrics*, 3(4), 456-499.
- [36] Jacod, J. (2008). “Asymptotic properties of realized power variations and related functionals of semimartingales,” *Stochastic Processes and Their Applications*, 118, 517-559.
- [37] Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2009). “Microstructure noise in the continuous case: the pre-averaging approach,” *Stochastic Processes and Their Applications*, 119, 2249-2276.
- [38] Jacod, J., M. Podolskij, and M. Vetter (2010). “Limit theorems for moving averages of discretized processes plus noise,” *Annals of Statistics* 38, 1478-1545.
- [39] Jacod, J. and P. Protter (2014). “Discretization of Processes,” *Springer-Verlag*.
- [40] Jacod, J. and A.N. Shiryaev (2003). “Limit Theorems for Stochastic Processes,” 2nd ed. *Springer-Verlag, Berlin*.
- [41] Jacod, J. and V. Todorov (2009). “Testing for Common Arrivals of Jumps for Discretely Observed Multi-dimensional Processes,” *Annals of Statistics*, 37, 1792-1838.
- [42] Jing, B., Liu, Z., and X. Kong (2014). “On the estimation of integrated volatility with jumps and microstructure noise,” *Journal of Business and Economic Statistics* 32(3), 457-467.
- [43] Kalnina, I. (2011). “Subsampling high frequency data,” *Journal of Econometrics*, 161(2), 262-283.
- [44] Katz, M.L., (1963). “Note on the Berry-Esseen theorem,” *Annals of Mathematical Statistics*, 34, 1107-1108.

- [45] Koike, Y. (2015). "Estimation of integrated covariances in the simultaneous presence of nonsynchronicity, microstructure noise and jumps," *Econometric Theory*, Forthcoming
- [46] Künsch, H.R. (1989). "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics* 17, 1217-1241.
- [47] Liu, R.Y. (1988). "Bootstrap procedure under some non-i.i.d. models," *Annals of Statistics* 16, 1696-1708.
- [48] Liu, C. and C. Y. Tang (2014). "A quasi-maximum likelihood approach to covariance matrix with high frequency data," *Journal of Econometrics*. *Forthcoming*.
- [49] Mancini, C. (2001). "Disentangling the jumps of the diffusion in a geometric jumping Brownian motion," *Giornale dell'Istituto Italiano degli Attuari* 64, 19-47.
- [50] Mancini, C. and F. Gobbi (2012). "Identifying the Brownian covariation from the co-jumps given discrete observations," *Econometric Theory*, 28, 249-273.
- [51] Mykland, P.A. and L. Zhang (2009). "Inference for continuous semimartingales observed at high frequency," *Econometrica*, 77, 1403-1455.
- [52] Mykland, P.A. and L. Zhang (2014). "Assessment of uncertainty in high frequency data: The observed asymptotic variance," *Working paper, University of Chicago*.
- [53] Park, S., and O. Linton (2012). "Estimating the quadratic covariation matrix for an asynchronously observed continuous time signal masked by additive noise," *FMG Discussion Papers* 703.
- [54] Podolskij, M., and M. Vetter (2009). "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15(3), 634-658.
- [55] Podolskij, M., and M. Vetter (2010). "Understanding limit theorems for semimartingales: a short survey," *Statistica Neerlandica*, 64, 329-351.
- [56] Podolskij, M., and D. Ziggel (2010). "New tests for jumps in semimartingale models," *Statistical Inference for Stochastic Processes* 13(1), 15-41.
- [57] Politis, D. N. and Romano, J. P. (1992). "A general resampling scheme for triangular arrays of α -mixing random variables," *Annals of Statistics*, 20, 1985-2007.
- [58] Politis, D.N., Romano, J.P., Wolf, M. (1999). "Subsampling," Springer-Verlag, New York.
- [59] Serfling, R.J., (1980). "Approximation theorems of mathematical statistics," Wiley, New York.
- [60] Shephard, N. Xiu, D. (2014). "Econometric analysis of multivariate realised QML: estimation of the covariation of equity prices under asynchronous trading," *Working paper No. 12-14, The University of Chicago Booth School of Business*.

- [61] Todorov, V. and T. Bollerslev, (2010). “Jumps and betas: a new theoretical framework for disentangling and estimating systematic risks,” *Journal of Econometrics*, 157, 220-235.
- [62] Varneskov, R. T. (2016). “Flat-top realized kernel estimation of quadratic covariation with non-synchronous and noisy asset prices,” *Journal of Business and Economic Statistics*, 34(1), 1-22.
- [63] Voev, V. and A. Lunde, (2007). “Integrated covariance estimation using high-frequency data in the presence of noise,” *Journal of Financial Econometrics*, 5, 68-104.
- [64] Wang, K., J. Liu, and Z. Liu (2013). “Disentangling the effect of jumps on systematic risk using a new estimator of integrated co-volatility,” *Journal of Banking and Finance* 37(5), 1777-1786.
- [65] Wu, C.F.J., (1986). “Jackknife, bootstrap and other resampling methods in regression analysis,” *Annals of Statistics* 14, 1261-1295.
- [66] Xiu, D. (2010). “Quasi-maximum likelihood estimation of volatility with high frequency data,” *Journal of Econometrics*, 159, 235-250.
- [67] Zhang, L. (2006). “Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach,” *Bernoulli*, 12, 1019-1043.
- [68] Zhang, L. (2011). “Estimating covariation: Epps effect, microstructure noise,” *Journal of Econometrics*, 160, 33-47.
- [69] Zhang, L, P.A. Mykland, and Y. Aït-Sahalia (2005). “A tale of two time-scales: determining integrated volatility with noisy high frequency data,” *Journal of the American Statistical Association*, 100, 1394-1411.
- [70] Zhang, L., P.A. Mykland, and Y. Aït-Sahalia (2011). “Edgeworth expansions for realized volatility and related estimators,” *Journal of Econometrics*, 160, 190-203.