UNIVERSITÀ DEGLI STUDI FIRENZE
DISEI
DIPARTIMENTO DI SCIENZE PER L'ECONOMIA E L'IMPRESA

Working Papers
Quantitative Methods for Social Sciences

Volatility of volatility estimation: central limit theorems for the Fourier transform estimator and empirical study of the daily time series stylized facts

Giulia Livieri, Maria Elvira Mancino, Stefano Marmi, Giacomo Toscano

Working Paper N. 01/2022

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

Volatility of volatility estimation: central limit theorems for the Fourier transform estimator and empirical study of the daily time series stylized facts

Giulia Livieri[*], Maria Elvira Mancino[†], Stefano Marmi[‡], Giacomo Toscano[§]

January 27, 2022

[*]Scuola Normale Superiore, Pisa, Italy
[†]University of Firenze, Dept. of Economics and management, Firenze, Italy
[‡]Scuola Normale Superiore, Pisa, Italy
[§]University of Firenze, Dept. of Economics and management, Firenze, Italy (corresponding author)

Abstract

We study the asymptotic normality of two feasible estimators of the integrated volatility of volatility based on the Fourier methodology, which does not require the pre-estimation of the spot volatility. We show that the bias-corrected estimator reaches the optimal rate $n^{1/4}$, while the estimator without bias-correction has a slower convergence rate and a smaller asymptotic variance. Additionally, we provide simulation results that support the theoretical asymptotic distribution of the rate-efficient estimator and show the accuracy of the latter in comparison with a rate-optimal estimator based on the pre-estimation of the spot volatility. Finally, we reconstruct the daily volatility of volatility of the S&P500 and EUROSTOXX50 indices over long samples via the rate-optimal Fourier estimator and provide novel insight into the existence of stylized facts about its dynamics.

Keywords: volatility of volatility, non-parametric estimation, central limit theorem, stochastic volatility, Fourier analysis.

JEL Classification: C14, C58.

1 Introduction

In the last decades, different stochastic volatility models have been proposed to describe the evolution of asset prices, motivated by empirical studies on the patterns of volatilities in financial time series. Further, the availability of high-frequency data has given impulse to devise statistical techniques aimed at the efficient estimation of model parameters in the stochastic volatility framework, e.g., the leverage and the volatility of volatility processes. The estimation of these model parameters is rather complicated, the main difficulties being due to the fact that some factors are unobservable. In particular, the estimation of the volatility of volatility is a challenging task, because a pre-estimation of the spot volatility is typically required as a first step, due to the latency of the volatility process.

Unlike the case of the integrated volatility, the non-parametric estimation of the integrated volatility of volatility is a relatively recent topic. [Barndorff-Nielsen and Veraart, 2009] propose a new class of stochastic volatility of volatility models, with an extra source of randomness, and show that the volatility of volatility can be estimated non-parametrically by means of the quadratic variation of the preliminarily estimated squared volatility process, which they name *pre-estimated spot variance based realized variance*. [Vetter, 2015] proposes an estimator of the integrated volatility of volatility which is also based on increments of the pre-estimated spot volatility process and attains the optimal convergence rate in the absence of noise. The common feature of these estimators is that they first reconstruct the unobservable volatility path via some consistent estimator thereof and then compute the volatility of volatility using the estimated paths as a proxy of the corresponding unknown paths. The issue of estimating the volatility of volatility in the presence of jumps is studied in [Cuchiero and Teichmann, 2015]: first, the authors combine jump robust estimators of the integrated variance and the Fourier-Fejer inversion formula to get an estimator of the instantaneous volatility path; secondly, they use again jump robust estimators of the integrated volatility, in which they plug the estimated path of the volatility process, to obtain an estimator of the volatility of volatility. In the same spirit of [Barndorff-Nielsen and Veraart, 2009, Vetter, 2015], [Li et al., 2021] also propose an estimator of the integrated volatility of volatility by means of a pre-

estimation of the spot volatility, but, in order to extend the study to the case when the observed price process contains jumps and microstructure noise, the authors adopt a threshold pre-averaging estimator of the volatility, following [Jing et al., 2014].

In this paper, we focus on the estimation of the integrated volatility of volatility via the Fourier estimation method by [Malliavin and Mancino, 2002], which does not require the pre-estimation of the spot volatility. An early application of the Fourier methodology to identify the parameters (volatility of volatility and leverage) of stochastic volatility models has been proposed by [Barucci and Mancino, 2010], where the authors prove a consistency result for the estimator of both the integrated leverage and volatility of volatility in the absence of noise. In the presence of microstructure noise, [Sanfelici et al., 2015] study the finite-sample properties of the Fourier estimator of the volatility of volatility introduced in [Barucci and Mancino, 2010] and show its asymptotic unbiasedness. However, the convergence rate of the estimator is not established, not even in the absence of microstructure noise contamination.

In the present paper we fill this gap. Specifically, after proving that the Fourier estimator of the volatility of volatility by [Sanfelici et al., 2015] has a sub-optimal rate of convergence, we define its bias-corrected version and prove that it reaches the optimal convergence rate $n^{1/4}$. We also show that the non-corrected estimator with slower rate of convergence displays a smaller asymptotic error variance. Further, we provide feasible versions of the two CLT's that exploit the product formula for the Fourier coefficients of the volatility of volatility and the fourth power of volatility. The same property of the Fourier coefficients is used in [Livieri et al., 2019] for the estimation of the quarticity.

The asymptotic results are supported by a simulation exercise, where we also compare the finite-sample performance of the rate-efficient Fourier estimator with that of the rate-efficient realized estimator based on the pre-estimation of the spot volatility by [Ait-Sahalia and Jacod, 2014]. The comparative study suggests that the Fourier estimator works quite well on the daily horizon, while the performance of the realized estimator appears to be not satisfactory. This feature may be related to the fact that, differently from the other volatility of volatility estimators, which rely on the pre-estimation of the instantaneous volatility path via a numerical differentiation, the Fourier approach relies only on the reconstruction of integrated quantities, i.e., the Fourier coefficients of the volatility. As it was early observed in [Malliavin and Mancino, 2002], this is a peculiarity of the Fourier estimator that renders the proposed method easily implementable and computationally stable.

Finally, we present an empirical exercise where the Fourier estimator is applied to obtain the daily time series of the volatility of volatility of the S&P500 and EUROSTOXX50 indices over, resp., the periods May 1, 2007 - August 6, 2021 and June 29, 2005 - May 28, 2021. As a result, we obtain some novel insight into the empirical regularities that characterize the daily dynamics of the volatility of volatility, which - to the best of our knowledge - up to now had been scarcely explored in the literature. Specifically, we find that the daily volatility of volatility of both the indices spikes in correspondence of periods of financial turmoil (e.g., during the financial crisis of 2008 and the outbreak of the COVID pandemic in 2020). Additionally, we also find that it is usually positively (resp., negatively) correlated with the volatility (resp., the asset return), but appears to be less persistent than the volatility. Finally, we observe that its empirical distribution is satisfactorily approximated by a log-normal distribution in years characterized

by higher financial stability, as it is the case for the volatility.

This novel insight appears to be valuable in view of the relevance of the volatility of volatility for scholars and practitioners. Indeed, on the one hand, market operators regularly "trade" the volatility of many financial asset classes via quoted and O.T.C. volatility derivatives (e.g., variance swaps, VIX futures and VIX options), hence the importance of the availability of accurate estimates of the volatility of the "traded" volatility. On the other hand, the need for efficient estimates of the volatility of volatility arises also in a number of technical tasks, e.g., the calibration of stochastic volatility of volatility models ([Barndorff-Nielsen and Veraart, 2009], [Sanfelici et al., 2015]), the estimation of the leverage coefficient ([Kalnina and Xiu, 2017], [Ait-Sahalia et al., 2017]), the inference of future returns (Bollerslev et al. (2009)) and spot volatilities ([Mykland and Zhang, 2009]). Furthermore, [Bandi et al., 2020] have recently provided empirical support to the dependence between the volatility of volatility of equity assets and structural sources of risk related to firms' characteristics.

The paper is organized as follows. Section 2 contains the assumptions and definitions. Section 3 states the central limit theorems, which are supported by the simulation study in Section 4. Finally, Section 5 contains the empirical results and Section 6 concludes. The proofs are given in Appendix A, while Appendix B contains some auxiliary lemmas on the Fejer and Dirichlet kernels.

2 Volatility of Volatility estimators: definition and assumptions

This section presents the general non-parametric stochastic volatility model which will be considered throughout the paper and defines two estimators of the integrated volatility of volatility based on the Fourier estimation method introduced in [Malliavin and Mancino, 2002, Malliavin and Mancino, 2009]. The class considered includes most of the continuous stochastic volatility models commonly used in high-frequency finance and is assumed (to cite one among many others) in [Ait-Sahalia and Jacod, 2014], chp.8.3.

We make the following assumptions.

(A.I) The log-price process p and the variance process v are continuous Itô semimartingales **on** $[0,T]$ satisfying the stochastic differential equations

$$\begin{cases} dp(t) = \sigma(t)\, dW_t + b(t)\, dt \\ dv(t) = \gamma(t) dZ_t + \beta(t) dt \end{cases}$$

where $v := \sigma^2$, while W and Z are Brownian motions on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying the usual conditions, possibly correlated (in this regard, note that it is not restrictive to assume a constant correlation ρ).

(A.II) The processes σ, b, γ and β are continuous adapted stochastic processes defined on the same probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$, such that for any $p \geq 1$

$$E[\int_0^T \sigma^p(t)dt] < \infty \ , \ E[\int_0^T b^p(t)dt] < \infty \ , \ E[\int_0^T \gamma^p(t)dt] < \infty \ , \ E[\int_0^T a^p(t)dt] < \infty.$$

The processes are specified in such a way that the spot volatility and volatility of volatility, resp. σ and γ, are a.s. positive.

(A.III) The process γ is a continuous Itô semimartingale, whose drift and diffusion processes are continuous adapted stochastic processes defined on the same probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$.

The assumptions (A.I)-(A.II)-(A.III) are standard in the non-parametric setting and are considered, e.g., in [Barndorff-Nielsen and Veraart, 2009, Ait-Sahalia and Jacod, 2014, Cuchiero and Teichmann, 2015, Vetter, 2015, Li et al., 2021].

By changing of the origin of time and scaling the unit of time, one can always modify the time window $[0,T]$ to $[0,2\pi]$. Suppose that the asset log-price p is observed at discrete, irregularly-spaced points in time on the grid $\{0 = t_{0,n} \leq \dots t_{i,n} \dots \leq t_{n,n} = 2\pi\}$. For simplicity, we omit the second index n. Denote $\rho(n) := \max_{0\leq h\leq n-1} |t_{h+1} - t_h|$ and suppose that $\rho(n) \to 0$ as $n \to \infty$.

Consider the following interpolation formula

$$p_n(t) := \sum_{i=0}^{n-1} p(t_i) I_{[t_i, t_{i+1}[}(t).$$

For any integer k, $|k| \leq 2N$, the discretized version of the Fourier coefficient $c_k(dp)$ is denoted by

$$c_k(dp_n) := \frac{1}{2\pi} \sum_{i=0}^{n-1} e^{-\mathrm{i}kt_i} \left(p(t_{i+1}) - p(t_i)\right), \tag{1}$$

where the symbol i is the imaginary unit $\sqrt{-1}$. Further, for any $|k| \leq N$, define the convolution formula

$$c_k(v_{n,N}) := \frac{2\pi}{2N+1} \sum_{|s|\leq N} c_s(dp_n) c_{k-s}(dp_n). \tag{2}$$

In [Malliavin and Mancino, 2009] it is proved that (2) is a consistent estimator of the k-th Fourier coefficient of the volatility process[1] and in [Barucci and Mancino, 2010, Sanfelici et al., 2015] it is shown that it is possible to derive an estimator of the integrated volatility of volatility by exploiting only the knowledge of the Fourier coefficients in (2), without the need of the preliminary estimation of the instantaneous volatility. This fact characterizes the Fourier method for estimating the volatility of volatility. In fact, as far as we know, all other existing methods rely on the pre-estimation of the spot volatility, see [Ait-Sahalia and Jacod, 2014, Cuchiero and Teichmann, 2015, Vetter, 2015, Li et al., 2021]. In general, these methods entail the pre-estimation of the spot volatility, in the presence or absence of noise contamination, as a first step; then, as a second step, a quadratic variation approach (e.g., the realised volatility formula) is applied to the pre-estimated spot volatility trajectory.

The estimator of the integrated volatility of volatility, defined in [Sanfelici et al., 2015], is given by 2π times

$$c_0(\gamma^2_{n,N,M}) := \frac{2\pi}{M+1} \sum_{|k|\leq M} \left(1 - \frac{|k|}{M+1}\right) k^2 c_k(v_{n,N}) c_{-k}(v_{n,N}), \tag{3}$$

as $c_0(\gamma^2_{n,N,M})$ is the estimator of $c_0(\gamma^2) = \frac{1}{2\pi}\int_0^{2\pi} \gamma^2(t)dt$. [Sanfelici et al., 2015] show that the estimator (3) is consistent under the assumptions (A.I)-(A.II)- (A.III) and the conditions $N/n \to 0$ and $M^4/N \to 0$ and asymptotically unbiased in the presence of microstructure noise[2]. However, the rate of convergence

[1] Hereinafter, we will follow the relevant econometric literature by using the term volatility as a synonym of variance, thus referring to $\sigma^2(t)$ as the volatility process. Similarly for the volatility of volatility.

[2] Note that these conditions on n, N and M are only sufficient for the consistency; the fact that they not sharp is due to the fact that the focus of the paper was not on the convergence rate, but on the finite-sample properties of the estimator in the presence of microstructure noise.

and the asymptotic normality are not established. We show in Theorem 3.1 that the convergence rate of the estimator (3) is not optimal. However, the estimator has a very good finite-sample performance, as shown in Section 4.

In order to obtain an estimator with the optimal rate of convergence in the absence of microstructure noise, a bias correction is needed and thus we consider the estimator

$$\widehat{\gamma}^2_{n,N,M} := \frac{2\pi}{M+1} \sum_{|k|\leq M} \left(1 - \frac{|k|}{M+1}\right) k^2 c_k(v_{n,N}) c_{-k}(v_{n,N}) - K\widehat{\sigma}^4_{n,N,M} \tag{4}$$

where the constant K is determined in (51) and $\widehat{\sigma}^4_{n,N,M}$ is the Fourier quarticity estimator

$$\widehat{\sigma}^4_{n,N,M} := 2\pi \sum_{|k|\leq M} c_k(v_{n,N}) c_{-k}(v_{n,N}). \tag{5}$$

The asymptotic normality of the estimator (5) is studied in [Livieri et al., 2019], while its properties in the presence of microstructure noise are studied in [Mancino and Sanfelici, 2012].

Note that the estimator (4) differs from (3) for the presence of the bias correction $K\,\widehat{\sigma}^4_{n,N,M}$. This bias correction, while ensuring a faster rate of convergence, destroys the positivity of the estimator, see also [Barndorff-Nielsen et al., 2011]. The estimator (3) is instead positive.

3 Central Limit Theorems

In this Section we study the asymptotic normality of the Fourier estimators of the integrated volatility of volatility defined by (3) and (4) and prove that the estimator (4) reaches the optimal rate of convergence $n^{1/4}$, at the cost of a de-biasing term, while the estimator (3) has a smaller asymptotic variance at the cost of a slower convergence rate.

Theorem 3.1 *Suppose that assumptions (A.I)-(A.II)-(A.III) hold. Let $N\rho(n) \sim c_N$ and $M\rho(n)^{1/2} \sim c_M$, where both c_M and c_N are positive constants*[3]*. Then, as $n, N, M \to \infty$, the following stable convergence in law holds:*

$$\rho(n)^{-1/4} \left(\widehat{\gamma}^2_{n,N,M} - \frac{1}{2\pi}\int_0^{2\pi} \gamma^2(t)dt\right)$$

$$\downarrow$$

$$\mathcal{N}\left(0, \frac{1}{2\pi}\int_0^{2\pi} K(c_M)\gamma^4(t) + K(c_M, c_N)\sigma^8(t) + \widetilde{K}(c_M, c_N)\,\gamma^2(t)\,\sigma^4(t)\,dt\right),$$

where $K(c_M) := \frac{4}{3}\frac{1}{c_M}$, $K(c_M, c_N) := \frac{16}{105}\,c_M^3(1 + 2\eta(c_N/\pi))^2$, $\widetilde{K}(c_M, c_N) := \frac{16}{15}\,c_M(1 + 2\eta(c_N/\pi))$ and $\eta(a) := \frac{1}{2a^2}r(a)(1 - r(a))$, with $r(a) = a - [a]$, being $[a]$ the integer part of a.

Note that if $c_N = \pi$ or, equivalently, $N = n/2$ (i.e., the cutting frequency N used for the estimation of the volatility coefficient given the log-prices is equal to the Nyquist frequency), then $\eta(c_N/\pi) = 0$ and the asymptotic variance in Theorem 3.1 becomes

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{4}{3}\frac{1}{c_M}\gamma^4(t) + \frac{16}{105}\,c_M^3\sigma^8(t) + \frac{16}{15}\,c_M\gamma^2(t)\,\sigma^4(t)\,dt. \tag{6}$$

[3] We stress the point that c_N and c_M are two positive constants, i.e., they do not depend on M, N. This notation is in line with the one used in [Mancino and Sanfelici, 2012].

Remark 3.2 *The realised volatility of volatility estimator (Th. 8.11 [Ait-Sahalia and Jacod, 2014]) is obtained as the quadratic variation of the estimated spot volatility, with a de-biasing term depending on the quarticity. The underlying model is a continuous semimartingale for the price, the volatility and the volatility of volatility. The convergence rate of the estimator is $n^{1/4}$ and the asymptotic variance is*

$$\int_0^T \frac{151}{70}\beta\gamma^4(t) + \frac{48}{\beta^3}\sigma^8(t) + \frac{12}{\beta}\sigma^4(t)\gamma^2(t)\,dt. \tag{7}$$

Letting $\beta = 1/c_M$, the correspondence between the asymptotic variances (6) and (7) is easily seen, with the second and third terms smaller in the case of the Fourier estimator. Note that the estimator in [Ait-Sahalia and Jacod, 2014] corresponds to (4) multiplied by 2π.
A similar approach as [Ait-Sahalia and Jacod, 2014] is considered in [Li et al., 2021], but it is extended to obtain a consistent estimator in the presence of noisy data. To this aim, the authors first build an estimator of the spot volatility by means of a pre-averaging method to get rid of the noise contamination, then they compute the realized variance from the spot volatility estimates to obtain an estimator of the integrated volatility of volatility. Finally, they also need to correct for the bias of the obtained estimator. The rate of convergence is $n^{1/8}$ in the presence of noise and $n^{1/4}$ without noise.

In order to obtain a feasible CLT from Theorem 3.1, a consistent estimator of the conditional variance is needed. We exploit again the Fourier methodology to build a consistent estimator of

$$\frac{1}{2\pi}\int_0^{2\pi} \Lambda(t)dt$$

with

$$\Lambda(t) := K(c_M)\gamma^4(t) + K(c_M, c_N)\sigma^8(t) + \widetilde{K}(c_M, c_N)\gamma^2(t)\,\sigma^4(t). \tag{8}$$

The result is obtained in Proposition 3.4, where the key ingredients are the following Remark 3.3 and the product formula for the Fourier coefficients, as studied in [Livieri et al., 2019].

Remark 3.3 *The estimation of the integrated volatility of volatility relies on the convolution product*

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1 - \frac{|k|}{M+1}\right)c_k(dv)c_{-k}(dv),$$

which allows computing the 0-th Fourier coefficient of the volatility of volatility process. The result is trivially extended to consider any continuous bounded function h as

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1 - \frac{|k|}{M+1}\right)c_k(dv)c_{-k}(h\,dv), \tag{9}$$

which leads to an estimator of

$$\frac{1}{2\pi}\int_0^{2\pi} h(t)\gamma^2(t)dt.$$

In particular, for $h(t) := e^{-\mathrm{i}kt}$, the convolution product (9) provides a formula for estimating the k-th Fourier coefficient of the volatility of volatility process $\gamma^2(t)$ (see also [Clement and Gloter, 2011] for the analogous result in the case of the multivariate Fourier volatility estimator).

Based on Remark 3.3, for any integer k, $|k| \le 2M$, we define

$$c_k(\sigma^4_{n,N,M}) := \sum_{|h|\le M} c_h(v_{n,N})c_{k-h}(v_{n,N}) \tag{10}$$

and

$$\bar{c}_k(\gamma^2_{n,N,M}) := \frac{2\pi}{M+1} \sum_{|h|\leq M} \left(1 - \frac{|h|}{M+1}\right) h(h-k)\, c_h(v_{n,N}) c_{k-h}(v_{n,N}) - K\, 2\pi\; c_k(\sigma^4_{n,N,M}), \tag{11}$$

where K is computed in (51). They are, resp., consistent estimators of $c_k(\sigma^4)$ and $c_k(\gamma^2)$, for any integer k. The following result holds.

Proposition 3.4 *Suppose that assumptions (A.I)-(A.II)-(A.III) hold. Let $N\rho(n) \sim c_N > 0$, $M\rho(n)^{1/2} \sim c_M > 0$ and $LM^{-1} \to 0$ as $n, N, M, L \to \infty$. Let $\Lambda(t)$ be defined as in (8) and, further, define*

$$\Lambda_{n,N,M,L} := K(c_M) V^{(1)}_{n,N,M,L} + \widetilde{K}(c_M, c_N) V^{(2)}_{n,N,M,L} + K(c_M, c_N) V^{(3)}_{n,N,M,L}, \tag{12}$$

where

$$\begin{aligned}
V^{(1)}_{n,N,M,L} &:= \sum_{|k|\leq L} \bar{c}_k(\gamma^2_{n,N,M}) \bar{c}_{-k}(\gamma^2_{n,N,M}), \\
V^{(2)}_{n,N,M,L} &:= \sum_{|k|\leq L} \bar{c}_k(\gamma^2_{n,N,M}) c_{-k}(\sigma^4_{n,N,M}), \\
V^{(3)}_{n,N,M,L} &=: \sum_{|k|\leq L} c_k(\sigma^4_{n,N,M}) c_{-k}(\sigma^4_{n,N,M}),
\end{aligned}$$

and $\bar{c}_k(\gamma^2_{n,N,M})$ is defined in (11), $c_k(\sigma^4_{n,N,M})$ in (10), and the constants $K(c_M)$, $\widetilde{K}(c_M, c_N)$ and $K(c_M, c_N)$ are specified in Theorem 3.1. As $n, N, M, L \to \infty$, the following convergence in probability holds:

$$\Lambda_{n,N,M,L} \to \frac{1}{2\pi} \int_0^{2\pi} \Lambda(t) dt.$$

Thus, we have the following feasible CLT.

Theorem 3.5 *Suppose that assumptions (A.I)-(A.II)-(A.III) hold. Let $N\rho(n) \sim c_N > 0$ and $M\rho(n)^{1/2} \sim c_M > 0$ and $LM^{-1} \to 0$ as $n, N, M, L \to \infty$. Let $\Lambda_{n,N,M,L}$ be defined as in (12). Then, as $n, N, M, L \to \infty$, the following stable convergence in law holds:*

$$\rho(n)^{-1/4} \frac{\widehat{\gamma}^2_{n,N,M} - \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} \gamma^2(t) dt}{\sqrt{\Lambda_{n,N,M,L}}} \to \mathcal{N}(0,1).$$

It is possible to obtain an estimator of the volatility of volatility without a bias-correction term and a smaller asymptotic variance, but the rate of convergence is slower, precisely $n^{\iota/2}$, with $\iota/2 \in (0, 1/5)$. The estimator is simply given by (3) multiplied by 2π and the following result holds.

Theorem 3.6 *Suppose that Assumptions (A.I)-(A.II)-(A.III) hold. Let $N\rho(n) \sim c_N > 0$ and $M\rho(n)^{\iota} \sim c_M > 0$, where $\iota \in (0, 2/5)$. Then, as $n, N, M \to \infty$, the following stable convergence in law holds:*

$$\rho(n)^{-\iota/2} \left(c_0(\gamma^2_{n,N,M}) - \frac{1}{2\pi} \int_0^{2\pi} \gamma^2(t) dt \right) \to \mathcal{N}\left(0, \frac{1}{2\pi} \int_0^{2\pi} \frac{4}{3} \frac{1}{c_M} \gamma^4(t)\, dt \right).$$

In order to build a feasible TLC it is enough to apply the same methodology as for Theorem 3.4. In particular, under the conditions $N\rho(n) \sim c_N$ and $M\rho(n)^\iota \sim c_M$, where $\iota \in (0, 2/5)$, a consistent estimator of the asymptotic variance is given by

$$\Gamma_{n,N,M,L} := \frac{4}{3}\frac{1}{c_M}\sum_{|k|\leq L} c_k(\gamma^2_{n,N,M})c_{-k}(\gamma^2_{n,N,M}), \tag{13}$$

where

$$c_k(\gamma^2_{n,N,M}) := \frac{2\pi}{M+1}\sum_{|h|\leq M}\left(1-\frac{|h|}{M+1}\right) h(h-k)\, c_h(v_{n,N})c_{k-h}(v_{n,N}). \tag{14}$$

Therefore, the following holds.

Theorem 3.7 *Suppose that Assumptions (A.I)-(A.II)-(A.III) hold. Let $N\rho(n) \sim c_N > 0$ and $M\rho(n)^\iota \sim c_M > 0$, where $\iota \in (0, 2/5)$. Then, as $n, N, M, L \to \infty$, the following stable convergence in law holds:*

$$\rho(n)^{-\iota/2}\frac{c_0(\gamma^2_{n,N,M}) - \frac{1}{2\pi}\int_0^{2\pi}\gamma^2(t)dt}{\sqrt{\Gamma_{n,N,M,L}}} \to \mathcal{N}(0,1).$$

Remark 3.8 *The result in Theorem 3.6 is in line with [Cuchiero and Teichmann, 2015], Th. 3.13, where an estimator without bias correction is considered. However, the proposed estimator relies on a smooth function of the plug-in spot volatility, where the spot volatility is estimated with the Fourier method. Therefore, it differs from our estimator, which is based on the convolution formula of the Fourier coefficients of the volatility process.*

4 Simulation Study

In this section we present a simulation study of the finite-sample performance of the rate-efficient estimator (4). The objective of the study is to provide support to the asymptotic result in Theorem 3.1, offer insight into the optimal selection of the frequency M, assess the robustness of the performance of the estimator to irregular sampling schemes and illustrate a comparison of its accuracy with that of the rate-efficient realized estimator by [Ait-Sahalia and Jacod, 2014].

4.1 Simulation design

We simulated discrete observations from two parametric models which satisfy Assumptions (A.I)-(A.II)-(A.III). The first model that we simulated is the Heston model (see [Heston, 1993]):

$$\begin{cases} dp(t) = \left(\mu - \frac{1}{2}v(t)\right)dt + \sigma(t)dW_t \\ dv(t) = \theta\left(\alpha - v(t)\right)dt + \sqrt{\gamma^2 v(t)}dZ_t, \end{cases} \tag{15}$$

where $v(t) := \sigma^2(t)$, $\mu \in \mathbb{R}$, $\theta, \alpha, \gamma > 0$, and ρ denotes the correlation between the Brownian motions W and Z. Under the Heston model, the volatility of volatility is given by $\gamma^2(t) = \gamma^2 v(t)$.

The second model that we simulated is the stochastic volatility of volatility model that appears in [Barndorff-Nielsen and Veraart, 2013] and [Sanfelici et al., 2015]. The model is as follows:

$$\begin{cases} dp(t) = \left(\mu - \frac{1}{2}v(t)\right) dt + \sigma(t) dW_t \\ dv(t) = \theta\left(\alpha - v(t)\right) dt + \gamma(t) dZ_t \\ d\gamma^2(t) = \chi\left(\eta - \gamma^2(t)\right) dt + \xi\gamma(t) dY_t, \end{cases} \tag{16}$$

where Y is a Brownian motion independent of W and Z, and $\mu \in \mathbb{R}$, $\theta, \alpha, \chi, \eta, \xi > 0$.

The parameter vectors used for the simulations of the models (15) and (16) are, resp.:

- $(\mu, \theta, \alpha, \gamma, \rho, x(0), v(0)) = (0.1, 5, 0.2, 0.5, -0.8, 1, 0.2)$;
- $(\mu, \theta, \alpha, \chi, \eta, \xi, \rho, x(0), v(0), \gamma^2(0)) = (0.1, 5, 0.2, 7, 0.1, 0.8, -0.8, 1, 0.2, 0.1)$.

Note that the selection of a negative ρ reproduces the presence of leverage effects. For each model we simulated 10^4 trajectories of length $T = 1/252$, i.e., one trading day. For each trajectory, observations were simulated on the equally-spaced grid with mesh equal to 1 second. For the simulations we have assumed that one trading day is 6.5-hour long.

4.2 Finite-sample performance

We assessed the finite-sample performance of the estimator (4) for increasing values of the sample size n, in order to provide numerical support to Theorem 3.1. Specifically, given $\rho(n) = T/n$, we chose a fixed estimation horizon T, corresponding to one day (as customary in econometric analyses), and let n vary. The resulting values of $\rho(n)$ considered range between 1 second and 5 minutes.

For what concerns the frequency N, which is needed for the convolution formula (2), we set $N = \left\lfloor c_N \rho(n)^{-1} \right\rfloor$ and selected $c_N = T/2$. This selection yields the value of N equal to the Nyquist frequency $[n/2]$ and allows obtaining the smallest variance of the asymptotic error, see (6) in section 3. As for the frequency M, we set $M = \left\lfloor c_M \rho(n)^{-1/2} \right\rfloor$ and optimized the value of the constant c_M based on the (unfeasible) numerical minimization of the mean squared error (MSE). In this regard, we found that the MSE-optimal value of c_M is equal to, resp., 0.05 and 0.07 for the models (15) and (16) (see subsection 4.3, where a feasible procedure to select M is also discussed).

Table 1 illustrates the finite-sample performance of the estimator (4) under the two different data-generating processes considered. Specifically, the Table illustrates the MSE and the bias for the different values of the sampling frequency $\rho(n)$. As expected, the bias and MSE improve as n is increased for both the data-generating processes considered, thereby providing numerical support to Theorem 3.1. In particular, note that the performance of the estimator is still satisfactory for $\rho(n)$ equal to 5 minutes, the sampling frequency typically used in the absence of noise with empirical data. Additionally, it is worth mentioning that the estimator never produced negative volatility of volatility estimates in this simulation study.

As additional support to the results in Thereom 3.1, the q-q plots in Figures 1 and 2 offer a comparison between the empirical quantiles of the unfeasible standardized estimation error from Theorem 3.1 and the theoretical quantiles of a standard normal distribution for different values of $\rho(n)$. Both in the

	Heston model		stochastic vol-of-vol model	
$\rho(n)$	MSE	Bias	MSE	Bias
5 minutes	$1.474 \cdot 10^{-8}$	$-5.674 \cdot 10^{-6}$	$3.886 \cdot 10^{-8}$	$1.455 \cdot 10^{-6}$
1 minute	$4.425 \cdot 10^{-9}$	$-4.142 \cdot 10^{-6}$	$1.939 \cdot 10^{-8}$	$1.304 \cdot 10^{-6}$
30 seconds	$2.913 \cdot 10^{-9}$	$-3.801 \cdot 10^{-6}$	$1.327 \cdot 10^{-8}$	$1.269 \cdot 10^{-6}$
5 seconds	$9.985 \cdot 10^{-10}$	$-2.932 \cdot 10^{-7}$	$8.113 \cdot 10^{-9}$	$4.161 \cdot 10^{-7}$
1 second	$4.229 \cdot 10^{-10}$	$-1.833 \cdot 10^{-7}$	$6.199 \cdot 10^{-9}$	$3.644 \cdot 10^{-7}$

Table 1: MSE and bias of the estimator (4) in correspondence of different values of $\rho(n)$. The true values of the daily integrated volatility of volatility are equal, on average, to $1.985 \cdot 10^{-4}$ and $3.957 \cdot 10^{-4}$ for, resp., the models (15) and (16).

case of the model (15) and the model (16), as $\rho(n)$ becomes smaller, the approximation to the the standard normal distribution improves. In particular, while the approximation of the body of the distribution is satisfactory also for the largest $\rho(n)$ considered, i.e., 5 minutes, the approximation in the tails becomes accurate for $\rho(n)$ smaller or equal than 5 seconds.

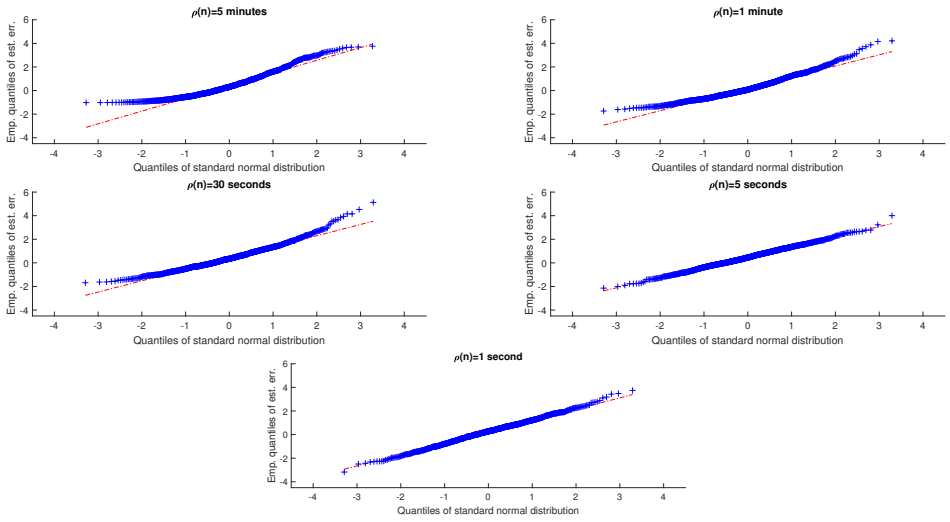

Figure 1: Comparison, for different values of $\rho(n)$, between the empirical quantiles of the standardized estimation error from Theorem 3.1, computed with simulated observations from the model (15), and the theoretical quantiles of a standard normal distribution.

4.3 Sensitivity to the frequency M

The careful selection of the frequency M, that is, the constant c_M, is key to efficiently implement the estimator (4) with finite samples. Given the selection $M = \left[c_M \rho(n)^{-1/2}\right]$, Figure 3 shows the sensitivity of the MSE of the estimator to different values of the constant c_M in the range $(0.01, 1)$, for different values of $\rho(n)$. Based on Figure 3, the optimal MSE is achieved when c_M is equal to, resp., 0.05 and 0.07 for the models (15) and (16), independently of $\rho(n)$. Further, it appears that the MSE is relatively flat

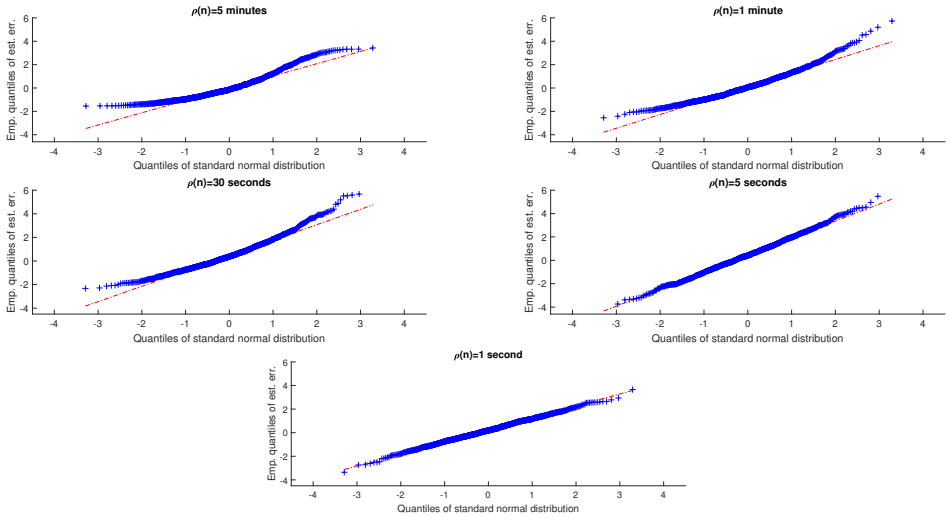

Figure 2: Comparison, for different values of $\rho(n)$, between the empirical quantiles of the standardized estimation error from Theorem 3.1, computed with simulated observations from the model (16), and the theoretical quantiles of a standard normal distribution.

for c_M in the intervals $(0.04, 0.06)$ and $(0.06, 0.07)$.

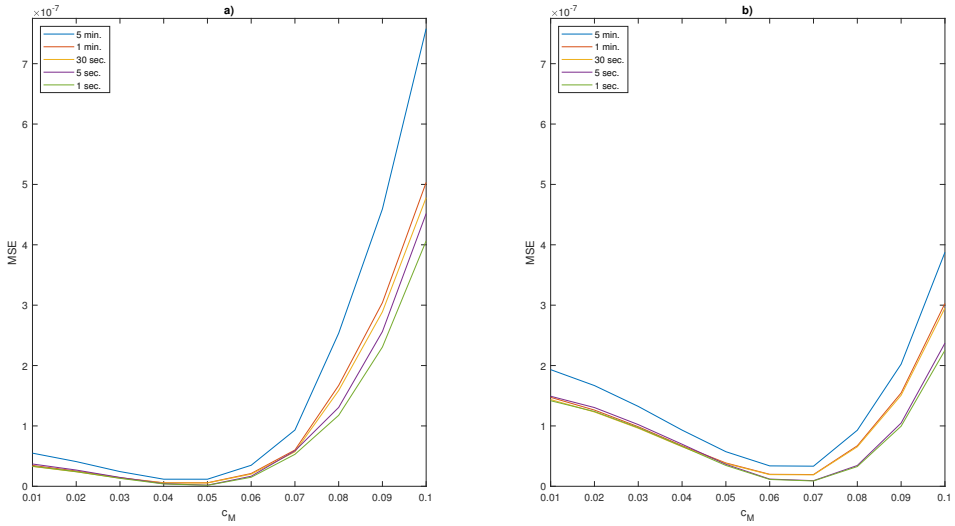

Figure 3: Sensitivity of the MSE of the estimator (4) to different values of the constant c_M in $M = [c_M \rho(n)^{-1/2}]$, for different values of $\rho(n)$. Panel a) refers to the model (15), while panel b) refers to the model (16).

As outlined in the following remark, it is possible to exploit the feasible Theorem 3.5 to optimize the selection of c_M with empirical data.

Remark 4.1 *The feasible selection of c_M can be performed based on the adaptive procedure described in Section 2.4 of [Li et al., 2021]. For the estimator (4), the procedure can be summarized as follows.*

First, given an initial value for c_M, denoted by $c_{M,0}$, one computes the standard error of the estimator as $\rho(n)^{1/4}\sqrt{\Lambda_{n,N,M,L}}$, exploiting the feasible estimator of the asymptotic variance in (12). Then, the computation of the standard error is iterated in correspondence of increasing values of c_M on a grid of step size s, that is, for $c_{M,j}$, $j \geq 1$, such that $c_{M,j} - c_{M,j-1} = s \ \forall j$. Iterations stop when the absolute value of the marginal decrease in the standard error (w.r.t. the standard error recorded in correspondence to the initial value $c_{M,0}$) becomes smaller than a given threshold ϑ. Unreported simulations show that, for $L = n^{1/4}$, the triple $(c_{M,0}, s, \vartheta) = (0.03, 0.01, 0.25)$ provides an estimation accuracy which is comparable to that demonstrated in subsection 4.2 with the unfeasible optimization of the MSE. Moreover, the same simulations suggest that the procedure is fairly robust to selections of the threshold ϑ in the interval $(0.2, 0.3)$.

4.4 Performance with irregular sampling

So far, in the simulation study we assumed that prices were observable on the equally-spaced grid with mesh size $\rho(n)$ equal to 1 second. However, the setup of Section 3 allows for an irregular sampling scheme. To assess the robustness of the performance of the estimator (4) to irregular sampling, we considered the case when observation times follow a Poisson process, that is, durations between observations are drawn from an exponential distribution with mean λ (see, e.g., [Mancino et al., 2017], Chapter 3.3).

Specifically, we considered three different values of λ, corresponding to an average duration δ of 1.25, 1.5 and 2 seconds, and compared the resulting MSE and bias values with the case of regular sampling on the 1-second grid. For the estimation with the Poisson scheme, we set $N = [n/2]$ and optimized M based on the minimization of the MSE. In this regard, we found that it is MSE-optimal to select a smaller M, compared to the regular-sampling case. Specifically, letting M^* denote the optimal selection with regular 1-second sampling, numerical results suggest that it is optimal to select $M = [M^*/2]$.

Table 2 reports the resulting bias and MSE, offering a comparison with the case where the a regular 1-second sampling scheme is adopted, corresponding to the last row of Table 1. The results in Table 2 suggest that the Fourier estimator (4) may still offer a satisfactory performance with irregular sampling schemes; in particular, it appears that the bias is relatively less affected than the MSE, compared to the regular-sampling case.

	Heston model		stochastic vol-of-vol model	
δ	MSE	Bias	MSE	Bias
2	$6.888 \cdot 10^{-10}$	$-1.910 \cdot 10^{-7}$	$9.672 \cdot 10^{-9}$	$3.840 \cdot 10^{-7}$
1.5	$5.250 \cdot 10^{-10}$	$-1.888 \cdot 10^{-7}$	$7.564 \cdot 10^{-9}$	$3.776 \cdot 10^{-7}$
1.25	$4.662 \cdot 10^{-10}$	$-1.865 \cdot 10^{-7}$	$6.696 \cdot 10^{-9}$	$3.730 \cdot 10^{-7}$
regular 1-sec. sampling	$4.229 \cdot 10^{-10}$	$-1.833 \cdot 10^{-7}$	$6.199 \cdot 10^{-9}$	$3.644 \cdot 10^{-7}$

Table 2: Comparison of MSE and bias of the estimator (4) when observation times follow a Poisson scheme (with different values of the average duration δ) with the case of regular 1-second sampling.

4.5 Comparison with the performance of the rate-efficient realized estimator

This subsection contains a comparative study of the finite-sample performance of the rate-efficient Fourier estimator (4) and the rate-efficient realized estimator by [Ait-Sahalia and Jacod, 2014] (see Remark 3.2). We recall the definition of the latter. Let $\kappa(n)$ denote a sequence of integers such that $\kappa(n) \sim \beta\rho(n)^{-1/2}$, $\beta > 0$, $\rho(n) := T/n$. The estimator reads

$$\hat{\gamma}_n^2 = \frac{3}{2\kappa(n)} \sum_{i=1}^{n-2\kappa(n)+1} \left(\left(\widehat{\sigma}_n^2(t_{i+k(n)}) - \widehat{\sigma}_n^2(t_i) \right)^2 - \frac{4}{\kappa(n)} (\widehat{\sigma}_n^2(t_i))^2 \right), \tag{17}$$

where

$$\widehat{\sigma}_n^2(t_i) = \frac{1}{\kappa(n)\rho(n)} \sum_{m=0}^{\kappa(n)-1} \left(p(t_{i+m}) - p(t_{i+m-1}) \right)^2$$

is the local estimator employed to pre-estimate the spot variance at time $t_i = iT/n$, $i = 0, ..., n$. The estimator (17) is also studied in [Vetter, 2015], where the author replaces $(\widehat{\sigma}_n^2(t_i))^2$ with

$$\widehat{\sigma}_n^4(t_i) = \frac{1}{3\kappa(n)\rho(n)^2} \sum_{m=0}^{\kappa(n)-1} \left(p(t_{i+m}) - p(t_{i+m-1}) \right)^4. \tag{18}$$

The rate-efficient realized estimator considered in [Vetter, 2015] thus reads

$$\hat{\tilde{\gamma}}_n^2 = \frac{3}{2\kappa(n)} \sum_{i=1}^{n-2\kappa(n)+1} \left(\left(\widehat{\sigma}_n^2(t_{i+k(n)}) - \widehat{\sigma}_n^2(t_i) \right)^2 - \frac{4}{\kappa(n)} \widehat{\sigma}_n^4(t_i) \right). \tag{19}$$

For the comparison, we replicated the simulation study carried out in subsection 4.2, this time using the realized estimators (17) and (19) to obtain estimates of the daily integrated volatility of volatility. The implementation of realized estimators requires the selection of the tuning parameter β. After setting $\kappa(n) = [\beta\rho(n)^{-1/2}]$, we selected $\beta = 0.04$ (resp., $\beta = 0.06$) in the case of the model (15) (resp., 16), based on the unfeasible optimization of the MSE with 1-second samples. Tables 3 and 4 summarize the results. By comparing the latter with Table 1 in subsection 4.2, it is immediate to see that the performance of the realized estimators (17) and (19) is not satisfactory, both in terms of bias and MSE, compared to the case of the Fourier estimator (4). See also [Sanfelici et al., 2015] and [Toscano and Recchioni, 2021] for similar considerations on the finite-sample performance of realized volatility of volatility estimators. Moreover, note that the comparison for $\rho(n)$ equal to 5 minutes is omitted, since the resulting bias and MSE of the realized estimators are larger than 1 in absolute value. Finally, simulations suggest that the use of the quarticity estimator in the de-biasing term in (19) does not improve the finite-sample performance.

	Heston model		stochastic vol-of-vol model	
$\rho(n)$	MSE	Bias	MSE	Bias
1 minute	$1.800 \cdot 10^{-3}$	$1.473 \cdot 10^{-2}$	$1.655 \cdot 10^{-3}$	$1.299 \cdot 10^{-2}$
30 seconds	$5.364 \cdot 10^{-4}$	$1.388 \cdot 10^{-2}$	$4.114 \cdot 10^{-4}$	$1.119 \cdot 10^{-2}$
5 seconds	$3.733 \cdot 10^{-4}$	$1.047 \cdot 10^{-2}$	$3.390 \cdot 10^{-4}$	$1.001 \cdot 10^{-2}$
1 second	$3.322 \cdot 10^{-4}$	$9.838 \cdot 10^{-3}$	$2.999 \cdot 10^{-4}$	$9.555 \cdot 10^{-3}$

Table 3: MSE and bias of the estimator (17) in correspondence of different values of $\rho(n)$. The true values of the daily integrated volatility of volatility are equal, on average, to $1.985 \cdot 10^{-4}$ and $3.957 \cdot 10^{-4}$ for, resp., the models (15) and (16).

	Heston model		stochastic vol-of-vol model	
$\rho(n)$	MSE	Bias	MSE	Bias
1 minute	$1.501 \cdot 10^{-3}$	$1.575 \cdot 10^{-2}$	$1.377 \cdot 10^{-3}$	$1.303 \cdot 10^{-2}$
30 seconds	$5.461 \cdot 10^{-4}$	$1.456 \cdot 10^{-2}$	$4.336 \cdot 10^{-4}$	$1.122 \cdot 10^{-2}$
5 seconds	$3.783 \cdot 10^{-4}$	$1.091 \cdot 10^{-2}$	$3.302 \cdot 10^{-4}$	$9.998 \cdot 10^{-3}$
1 second	$3.324 \cdot 10^{-4}$	$9.840 \cdot 10^{-3}$	$3.002 \cdot 10^{-4}$	$9.555 \cdot 10^{-3}$

Table 4: MSE and bias of the estimator (19) in correspondence of different values of $\rho(n)$. The true values of the daily integrated volatility of volatility are equal, on average, to $1.985 \cdot 10^{-4}$ and $3.957 \cdot 10^{-4}$ for, resp., the models (15) and (16).

Remark 4.2 *Unreported simulations show that the realized estimators (17) and (19) improve their finite-sample performance for larger values of the estimation horizon T. Specifically, realized estimators appear to achieve satisfactory accuracy, compared to the Fourier estimator (4), when T is equal to one year. This is in line with the selection of T equal to one year in the numerical and empirical high-frequency exercises by [Li et al., 2021], where the performance of the noise- and jump-robust version of the realized estimator volatility of volatility estimator is investigated.*

5 Empirical Analysis

To the best of our knowledge, the empirical properties of the volatility of volatility of financial assets have been scarcely explored in the literature. The aim of the empirical study presented in this Section is thus to provide insight into the existence of stylized facts pertaining to the daily dynamics of the volatility of volatility.

In fact, the numerical evidence presented in Section 4 suggests that the Fourier methodology allows reconstructing the integrated volatility of volatility with satisfactory accuracy on daily intervals by means of the rate-efficient estimator (4). Accordingly, in this section we use the estimator (4) to obtain the daily volatility of volatility series for two market indices: the S&P500 and the EUROSTOXX50. The periods considered for the analysis are, resp., May 1, 2007 - August 6, 2021 and June 29, 2005 - May 28, 2021.

5.1 Data description, estimation and sample statistics

For the empirical analysis we used the series of 5-minute trade prices, recorded during trading hours. Specifically, for the S&P500 index, we used the prices recorded between 9.30 a.m. and 4 p.m., while for the EUROSTOXX50 index we employed the prices recorded between 9 a.m. and 5.30 p.m. Days with early closure were discarded.

The estimation of the daily integrated volatility of volatility was performed via the rate-efficient Fourier estimator (4), without considering overnight returns. Before performing the estimation, we run the test by [Ait-Sahalia and Xiu 2014] on 5-minute series and found that the assumption of absence of noise could not be rejected at the 5% significance level for both indices. Moreover, following [Wang and Mykland, 2014], days with jumps were removed, based on the results of the test by [Lee and Mykland, 2008], which was applied at the 1% significance level. Overall, the number of days for which we estimated the volatility of volatility is 3343 and 3522 for, resp., the S&P500 and EUROSTOXX50.

For the estimation, given the choice of the sampling frequency $\rho(n) = 5$ minutes, we set $N = [c_N \rho(n)^{-1}]$ and $M = [c_M \rho(n)^{-1/2}]$. Then we selected c_N such that N equals the Nyquist frequency (see subsection 4.2) and c_M based on the feasible procedure illustrated in subsection 4.3. The resulting values of c_M are equal, on average, to 0.054 and 0.065 for, resp., the S&P500 and the EUROSTOXX50. Figures 4 and 5 display the reconstructed trajectories of the daily volatility of volatility of the two indices, while Table 5 compares sample statistics of volatility and volatility of volatility estimates. Daily volatility estimates were obtained via the Fourier estimator by [Malliavin and Mancino, 2002], applied to 5-minute returns[4]. We note that all volatility of volatility estimates obtained are strictly positive.

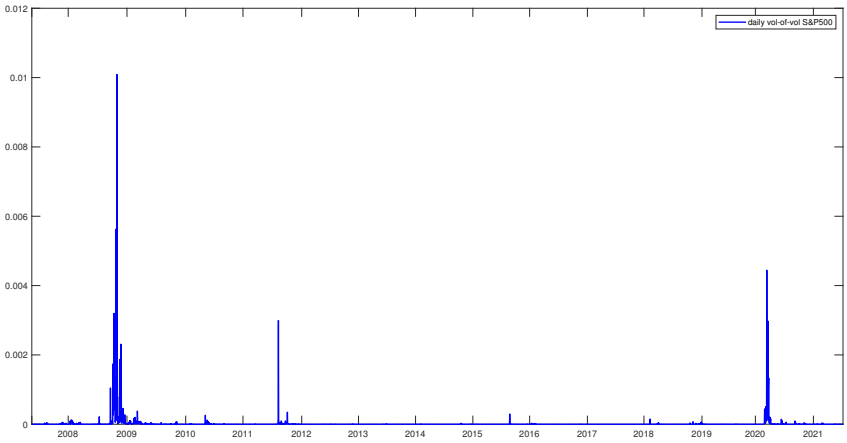

Figure 4: Daily volatility of volatility estimates for the S&P500 index over the period May 1, 2007- August 6, 2021.

Figures 4 and 5 show that the volatility of volatility spikes during financial crises, while remaining rather low and stable during tranquil periods. Indeed, both daily series reach their three highest peaks

[4]All the analyses appearing in this Section and involving the computation of the volatility were also performed using volatility estimates obtained via the 5-minute realized variance and the final outcome was pretty much the same.

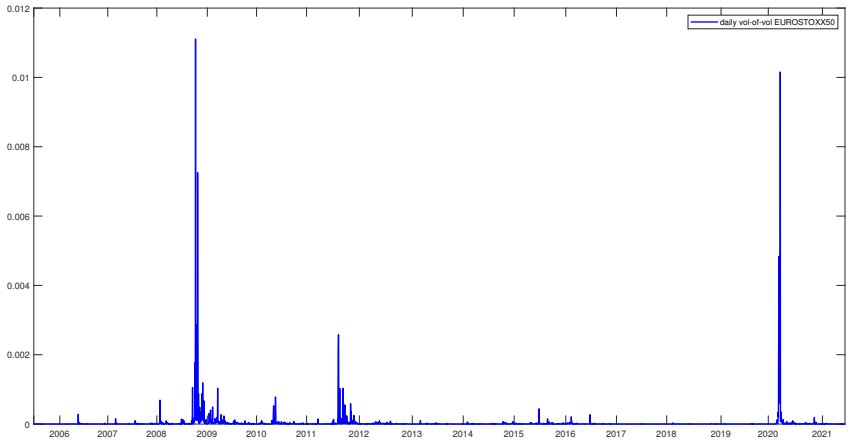

Figure 5: Daily volatility of volatility estimates for the EUROSTOXX50 index over the period June 29, 2005 - May 28, 2021

daily series	mean	median	st. dev.	min	max	skew.	kurt.
SPX vol.	$8.598 \cdot 10^{-5}$	$3.098 \cdot 10^{-5}$	$2.152 \cdot 10^{-4}$	$1.701 \cdot 10^{-6}$	$3.209 \cdot 10^{-3}$	7.951	83.902
SPX vol. of vol.	$2.351 \cdot 10^{-5}$	$3.867 \cdot 10^{-7}$	$2.608 \cdot 10^{-4}$	$9.545 \cdot 10^{-10}$	$1.009 \cdot 10^{-2}$	24.442	780.377
ESTX vol.	$1.093 \cdot 10^{-4}$	$5.848 \cdot 10^{-5}$	$1.944 \cdot 10^{-4}$	$2.845 \cdot 10^{-6}$	$3.916 \cdot 10^{-3}$	8.198	104.168
ESTX vol. of vol.	$3.233 \cdot 10^{-5}$	$1.841 \cdot 10^{-6}$	$3.267 \cdot 10^{-4}$	$5.054 \cdot 10^{-9}$	$1.630 \cdot 10^{-2}$	24.726	727.501

Table 5: Sample statistics (mean, median, standard deviation, minimum, maximum, skewness, kurtosis) of the estimated daily integrated volatility and volatility of volatility of the S&P500 (SPX) and the EUROSTOXX50 (ESTX) indices over, resp., the periods May 1, 2007- August 6, 2021 and June 29, 2005 - May 28, 2021.

in correspondence of, resp., the global financial crisis starting at the end of 2008, the instabilities of the Euro area in the second part of 2011 and the outbreak of the COVID pandemic in the first months of 2020.

Moreover, based on Table 5, we make the following remarks. First, the volatility of volatility is on average smaller than the volatility in the case of both indices. Secondly, the volatility of volatility appears to be more volatile than the volatility itself, as it displays larger sample standard deviations and maxima for both the estimated series. Finally, the volatility of volatility appears to be much more skewed and leptokurtic than the volatility for both the indices.

An analysis of the empirical regularities displayed by the reconstructed daily series of the volatility of volatility of the S&P500 and the EUROSTOXX50 is illustrated in the next subsection.

5.2 Insight into volatility-of-volatility stylized facts

The literature on the stylized facts related to the behavior of the volatility of financial assets is very rich (see, for instance, [Andersen et al., 2001], [Patton and Engle, 2001] and [Corsi, 2009], among many others). These include, e.g., clustering, long memory, mean-reversion, log-normality and leverage effects.

Nowadays, the volatility can be regarded, in some sense, as a traded asset itself. In fact, it is possible to "trade" the volatility of many financial asset classes via quoted and O.T.C. volatility derivatives (e.g., variance swaps, VIX futures and VIX options). Therefore, it may be of interest to evaluate which typical features of the daily volatility actually apply to the daily volatility of the volatility itself.

We have already observed in the previous subsection that the volatility of volatility shows clusters, being larger in correspondence of crises and smaller and less volatile during periods of economic stability. However, based on the observation of the sample auto-correlation function (see Figures 6 and 7), it appears to be less persistent than the volatility for both the indices considered. As for the mean-reversion property, the Augmented Dickey-Fuller test rejects the hypothesis of a unit root for both volatility of volatility series analysed, at the 0.01% significance level.

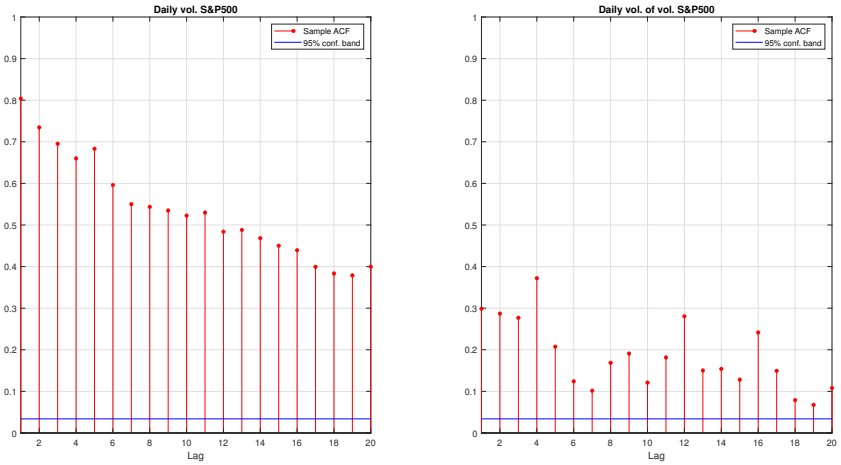

Figure 6: Sample autocorrelations (in red) of daily volatility (left panel) and volatility of volatility (right panel) for the S&P500 index over the period May 1, 2007- August 6, 2021. The 95% confidence band (in black) is computed under the null hypothesis of a Gaussian white noise process.

We also examine the year-by-year correlation of the daily volatility of volatility with, resp., the daily volatility and the daily log-return, computed as the difference between the closing and opening log-price. The dynamics of such correlations are summarized in Tables 6 and 7, where the values of return-variance correlations, a rough proxy of the leverage effect, are also displayed for comparison[5]. For both the indices, we observe that the yearly correlation between the log-return and the volatility of volatility tends to be negative and to follow the return-variance correlation closely, although being most often smaller in absolute value. This result may suggest the existence of a "second-order" leverage effect: what we observe is in fact that when the asset price decreases, not only the volatility increases, due the asset becoming riskier, but also the volatility of volatility - which can be seen as a proxy of the uncertainty about the amount of risk perceived by market operators, that is, the "volatility of risk" - becomes larger. The yearly correlation between the volatility of volatility and the volatility is instead positive and close

[5]Note that the correlations appearing in Tables 6 and 7 are typically pushed towards zero by the presence of a finite-sample bias, see [Ait-Sahalia et al., 2013]. For this reason, their true values are likely to be larger, in absolute value. However, obtaining unbiased and efficient estimates of these correlations goes beyond the scope of the exploratory analysis proposed.

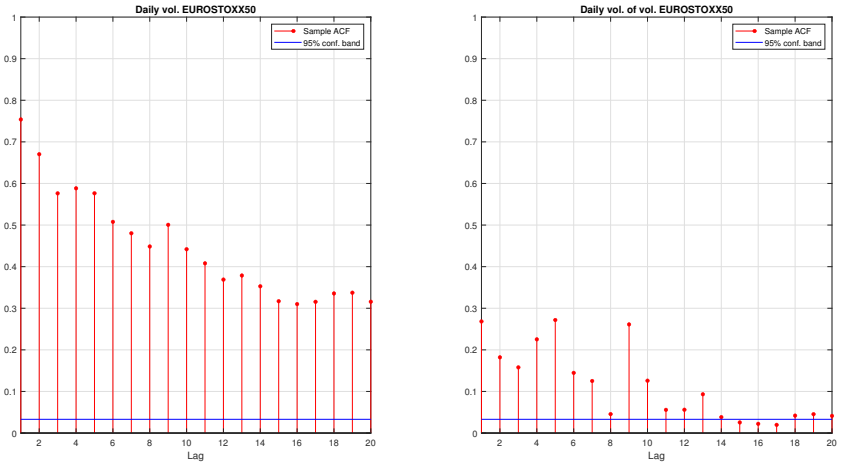

Figure 7: Sample autocorrelations (in red) of the daily volatility (left panel) and volatility of volatility (right panel) for the EUROSTOXX50 index over the period June 29, 2005 - May 28, 2021. The 95% confidence band (in black) is computed under the null hypothesis of a Gaussian white noise process.

to 0.7 on average for both the indices. This is consistent with the presence of volatility-of-volatility peaks in periods of higher market volatility.

year	vol. of vol. - vol.	vol.- ret.	vol. of vol. - ret.
2007	0.730	-0.126	-0.030
2008	0.689	0.049	0.054
2009	0.727	-0.063	0.067
2010	0.827	-0.199	-0.073
2011	0.728	0.019	0.213
2012	0.686	-0.231	0.022
2013	0.706	-0.354	-0.129
2014	0.842	-0.140	0.030
2015	0.582	-0.153	-0.281
2016	0.849	-0.122	-0.151
2017	0.638	-0.348	-0.160
2018	0.758	-0.218	-0.125
2019	0.858	-0.125	-0.026
2020	0.821	0.090	0.172
2021	0.913	-0.305	-0.219
average	0.757	-0.147	-0.043

Table 6: S&P500 index: sample yearly correlations of the daily volatility of volatility with the corresponding daily volatility and daily log-return.

Finally, we test for the Gaussianity of the logarithmic volatility and volatility of volatility estimates,

year	vol. of vol. - vol.	vol.- ret.	vol. of vol. - ret.
2005	0.6906	-0.1463	-0.1740
2006	0.593	-0.119	-0.101
2007	0.689	-0.237	-0.252
2008	0.833	-0.130	-0.053
2009	0.621	-0.144	-0.009
2010	0.804	-0.107	-0.040
2011	0.864	-0.072	-0.156
2012	0.718	-0.035	0.023
2013	0.685	-0.141	-0.052
2014	0.742	-0.128	-0.074
2015	0.584	-0.111	0.052
2016	0.732	-0.327	-0.275
2017	0.552	-0.172	-0.196
2018	0.791	-0.327	-0.303
2019	0.743	-0.180	-0.016
2020	0.705	-0.276	-0.084
2021	0.762	-0.185	-0.081
average	0.712	-0.167	-0.105

Table 7: EUROSTOXX50 index: sample yearly correlations of the daily volatility of volatility with the corresponding daily volatility and daily log-return.

using the Jarque-Bera and Anderson-Darling tests at the 5% significance level. The years in which both tests reject the null hypothesis of Gaussianity are 2008, 2011, 2016 and 2020, that is, the years in the sample that were the most characterized by market turmoil (in order: the global financial crisis, the Euro-area instability phase, Brexit and the outbreak of the COVID pandemic). This happens for both the quantities tested and both the indices analyzed, thus suggesting that the log-normal approximation for the distribution of the volatility and the volatility of volatility is more satisfactory in periods of market stability.

6 Conclusion

This paper fills a gap in the literature on financial econometrics by deriving the convergence rate of the Fourier estimator of the volatility of volatility. In this regard, we showed that the bias-corrected version of the estimator reaches the optimal rate $n^{1/4}$, while the estimator without bias-correction achieves a sub-optimal rate, but has a smaller asymptotic variance.

Further, we presented a numerical study that shows that the rate-optimal Fourier estimator of the volatility of volatility performs well in finite samples, even at the relatively small daily estimation horizon, where the competing rate-efficient realized estimator shows a poor performance.

Finally, we applied the Fourier estimator to multi-year samples of S&P500 and EUROSTOXX50 observations and gained some new knowledge about the empirical regularities that characterize the daily dynamics of the volatility of volatility, a topic which so far had been scarcely explored in the literature.

7 Appendix A: Proofs

The proofs of Theorems 3.1, 3.5, 3.6 and 3.7 are illustrated in the next subsections. Some preliminary remarks are useful.

Remark 7.1 *As every continuous process is locally bounded, thus all processes appearing are. Moreover, standard localization procedures (see, e.g., [Ait-Sahalia and Jacod, 2014]) allow to assume that any locally bounded process is actually bounded and almost-surely positive processes can be considered as bounded away from zero.*

Remark 7.2 *In [Malliavin and Mancino, 2009], Lemma 2.2, it is proved that the drift component of the semi-martingale model gives no contribution to the convolution formula (2). Therefore, we will refer to the drift-less model in Assumption (A.I). Moreover, as observed in [Malliavin and Mancino, 2002], we can assume that* $p(0) = p(2\pi)$ *and* $v(0) = v(2\pi)$*. In fact, if* $p(0) \neq p(2\pi)$ *(similarly for the process* v*), we introduce*

$$\widetilde{p}(t) = p(t) - \frac{t}{2\pi}[p(2\pi) - p(0)].$$

Then, while $\widetilde{p}$ *satisfies the required assumption, at the same time the volatility and co-volatilities estimations are not affected by a modification of the drift as above. From the point of view of the modeling, we may consider*

$$d\widetilde{p}(t) = \sqrt{v(t)}dW(t),$$

$$d\widetilde{v}(t) = \gamma(t)\, dZ(t),$$

being $v(t) = \widetilde{v}(t) - \frac{t}{2\pi}[\widetilde{v}(2\pi) - \widetilde{v}(0)]$*. In fact, for any* $k \neq 0$*, it holds* $c_k(dv) = c_k(d\widetilde{v})$*, while the* 0*-th Fourier coefficient,* $c_0(dv)$*, is not contributing to the definition of the Fourier volatility of volatility estimator. The situation would change in the case when one wishes to estimate the spot volatility. However, this is not an issue of the present study, as the estimation of the spot volatility is not required.*

7.1 Preliminary Decomposition

Given the discrete time observations $\{0 = t_0 \leq \ldots \leq t_i \leq \ldots \leq t_n = 2\pi\}$, we use the notation in continuous time by letting $\varphi_n(t) := \sup\{t_j : t_j \leq t\}$, for the sake of simplicity. From the Itô formula we have the following decomposition of the term (2)

$$c_k(v_{n,N}) := A_{k,n} + B_{k,n,N} + C_{k,n,N}, \tag{20}$$

where

$$A_{k,n} := \frac{1}{2\pi}\int_0^{2\pi} e^{-ik\varphi_n(s)}v(s)ds, \tag{21}$$

$$B_{k,n,N} := \frac{1}{2\pi}\int_0^{2\pi} e^{-ik\varphi_n(s)} \int_0^s D_N(\varphi_n(s)-\varphi_n(u))\sigma(u)dW_u\,\sigma(s)dW_s, \tag{22}$$

$$C_{k,n,N} := \frac{1}{2\pi}\int_0^{2\pi} \int_0^s e^{-ik\varphi_n(u)} D_N(\varphi_n(s)-\varphi_n(u))\sigma(u)dW_u\,\sigma(s)dW_s. \tag{23}$$

It follows that $\widehat{\gamma}^2_{n,N,M}$ is equal to the following sum:

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,A_{k,n}A_{-k,n} \tag{24}$$

$$+\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,2(A_{k,n}B_{-k,n,N}+A_{k,n}C_{-k,n,N}) \tag{25}$$

$$+\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,(B_{k,n,N}B_{-k,n,N}+2B_{k,n,N}C_{-k,n,N}+C_{k,n,N}C_{-k,n,N}). \tag{26}$$

In the following we denote:

$$Y_{n,N}(t,s) := \int_0^t D_N(\varphi_n(s)-\varphi_n(u))\sigma(u)dW_u, \tag{27}$$

and, for brevity, we will also use the following notation for the Dirichlet and the Fejér kernels (see also the Appendix B, Section 8):

$$D_{N,n}(s-u) := D_N(\varphi_n(s)-\varphi_n(u)), \qquad F_{M,n}(s-u) := F_M(\varphi_n(s)-\varphi_n(u)), \tag{28}$$

and similarly for the derivatives of the Fejér kernel[6] $F'_{M,n}$ and $F''_{M,n}$.

In order to identify the different contribution of all terms, we start with equation (24), which can be written as

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2 c_k(v)c_{-k}(v) \tag{29}$$

$$+\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,\left(A_{k,n}A_{-k,n}-c_k(v)c_{-k}(v)\right). \tag{30}$$

We consider now the term (25). An application of the Itô formula shows that it is equal to

$$2(AB^{(i)}_{M,n,N}+AB^{(ii)}_{M,n,N}+AC^{(i)}_{M,n,N}+AC^{(ii)}_{M,n,N}),$$

where

$$AB^{(i)}_{M,n,N} := \frac{1}{2\pi}\int_0^{2\pi}\int_0^s \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2 e^{\mathrm{i}k(\varphi_n(s)-\varphi_n(u))}v(u)du\,Y_{n,N}(s,s)\sigma(s)dW_s$$

$$= -\int_0^{2\pi}\frac{1}{2\pi}\int_0^s \frac{1}{M+1}F''_{M,n}(s-u)v(u)du\,Y_{n,N}(s,s)\sigma(s)dW_s, \tag{31}$$

$$AB^{(ii)}_{M,n,N} := \int_0^{2\pi}\frac{1}{2\pi}\int_0^s \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2 e^{-\mathrm{i}k(\varphi_n(s)-\varphi_n(u))}\,Y_{n,N}(u,s)\sigma(u)dW_u\,v(s)ds$$

$$= -\int_0^{2\pi}\frac{1}{2\pi}\int_0^s \frac{1}{M+1}F''_{M,n}(s-u)\,Y_{n,N}(u,s)\,\sigma(u)dW_u\,v(s)ds, \tag{32}$$

[6] We stress the point that the notation $D_{N,n}(s-u)$ refers to $D_N(\varphi_n(s)-\varphi_n(u))$, which is a function of two time variables, (s,u), as stated in (28). It is not to identify it with $D_N(\varphi_n(s-u))$. Notation (28) is maintained in order to highlight the role of the convolution product, which is the main characteristic of the Fourier estimation method in [Malliavin and Mancino, 2009].

and, letting

$$Y_{k,n,N}(t,s) := \int_0^t e^{-\mathrm{i}k\varphi_n(u)} D_N(\varphi_n(s) - \varphi_n(u))\sigma(u)dW_u,$$

then

$$AC^{(i)}_{M,n,N} := \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}\int_0^s e^{-\mathrm{i}k\varphi_n(u)}v(u)du\, Y_{-k,n,N}(s,s)\sigma(s)dW_s$$

$$= -\int_0^{2\pi}\frac{1}{2\pi}\int_0^s\int_0^u \frac{1}{M+1}F''_{M,n}(u-u')D_{N,n}(s-u')\sigma(u')dW_{u'}\, v(u)du\,\sigma(s)dW_s \tag{33}$$

$$-\int_0^{2\pi}\frac{1}{2\pi}\int_0^s\int_0^u \frac{1}{M+1}F''_{M,n}(u-u_1)v(u_1)du_1 D_{N,n}(s-u)\sigma(u)dW_u\,\sigma(s)dW_s, \tag{34}$$

and

$$AC^{(ii)}_{M,n,N} := \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}\int_0^s Y_{-k,n,N}(s',s)\sigma(s')dW_{s'}\, e^{-\mathrm{i}k\varphi_n(s)}v(s)ds$$

$$= -\int_0^{2\pi}\frac{1}{M+1}\frac{1}{2\pi}\int_0^s\int_0^{s'} F''_{M,n}(s-u)D_{N,n}(s-u)\,\sigma(u)dW_u\,\sigma(s')dW_{s'}\, v(s)ds. \tag{35}$$

Finally, we consider the term (26). Using the fact that the Fejér kernel is an even function, it can be re-written as

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\left(B_{k,n,N}B_{-k,n,N} + 2B_{k,n,N}C_{-k,n,N} + C_{k,n,N}C_{-k,n,N}\right)$$

$$= BB^{(i)}_{M,n,N} + 2BB^{(ii)}_{M,n,N} + 2(BC^{(i)}_{M,n,N} + BC^{(ii)}_{M,n,N} + BC^{(iii)}_{M,n,N}) + CC^{(i)}_{M,n,N} + 2CC^{(ii)}_{M,n,N},$$

where, using notation introduced in (27), each term is defined as follows:

$$BB^{(i)}_{M,n,N} := \frac{1}{2\pi}\frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\int_0^{2\pi} Y^2_{n,N}(s,s)v(s)ds, \tag{36}$$

$$BB^{(ii)}_{M,n,N} := -\int_0^{2\pi}\frac{1}{2\pi}\int_0^u \frac{1}{M+1}F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s Y_{n,N}(u,u)\sigma(u)dW_u, \tag{37}$$

$$BC^{(i)}_{M,n,N} := -\int_0^{2\pi}\frac{1}{2\pi}\int_0^s \frac{1}{M+1}F''_{M,n}(s-v)D_{N,n}(s-v)\sigma(v)dW_v\, Y_{n,N}(s,s)\sigma^2(s)ds, \tag{38}$$

$$BC^{(ii)}_{M,n,N} := \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}\int_0^u e^{-ik\phi_n(s)}Y_{n,N}(s,s)\sigma(s)dW_s\, Y_{k,n,N}(u,u)\,\sigma(u)dW_u, \tag{39}$$

$$BC^{(iii)}_{M,n,N} := -\int_0^{2\pi}\int_0^u\frac{1}{2\pi}\int_0^v \frac{1}{M+1}F''_{M,n}(v-s)D_{N,n}(v-s)\sigma(s)dW_s\,\sigma(v)dW_v\, Y_{N,n}(u,u)\sigma(u)dW_u, \tag{40}$$

$$CC^{(i)}_{M,n,N} := \frac{1}{2\pi}\frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\int_0^{2\pi} Y_{k,n,N}(s,s)Y_{-k,n,N}(s,s)v(s)ds, \tag{41}$$

$$CC^{(ii)}_{M,n,N} := \frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}\int_0^s Y_{k,n,N}(v,v)\sigma(v)dW_v\, Y_{-k,n,N}(s,s)\sigma(s)dW_s. \tag{42}$$

In summary, the estimation error

$$\widehat{\gamma}^2_{n,N,M} - \frac{1}{2\pi}\int_0^{2\pi}\gamma^2(t)dt \tag{43}$$

comprises the study of four main components:

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2c_k(v)c_{-k}(v)-\int_0^{2\pi}\gamma^2(t)dt \tag{44}$$

$$+\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,\left(A_{k,n}A_{-k,n}-c_k(v)c_{-k}(v)\right) \tag{45}$$

$$+2\left(AB^{(i)}_{M,n,N}+AB^{(ii)}_{M,n,N}+AC^{(i)}_{M,n,N}+AC^{(ii)}_{M,n,N}+BB^{(ii)}_{M,n,N}+BC^{(ii)}_{M,n,N}+BC^{(iii)}_{M,n,N}+CC^{(ii)}_{M,n,N}\right) \tag{46}$$

$$+BB^{(i)}_{M,n,N}+2BC^{(i)}_{M,n,N}+CC^{(i)}_{M,n,N}-K\widehat{\sigma}^4_{n,N,M}, \tag{47}$$

where $\widehat{\sigma}^4_{n,N,M}$ is defined in (5) and the constant K is determined in (51).

Accordingly, the proof of the theorem is divided into four steps. The first and second steps comprise the study of the bias correction term (47) and the asymptotic negligibility of the discretization error (45). The other two steps follow [Jacod, 1997] in order to identify the asymptotic variance and prove the stable convergence in law.

In the proof we consider the case of regular sampling, i.e., $\varphi_n(t)=\frac{2\pi}{n}j$, if $\frac{2\pi}{n}j\le t<\frac{2\pi}{n}(j+1)$, $j=0,\dots,n$. Further, C will always denote a positive constant, not necessarily the same.

7.2 Step I. The Bias Correction Term

Firstly, we show that the term (47) is $o_p(\rho(n)^{1/4})$, therefore proving that the error (43) is equal to

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2c_k(v)c_{-k}(v)-\int_0^{2\pi}\gamma^2(t)dt$$

$$+\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\,\left(A_{k,n}A_{-k,n}-c_k(v)c_{-k}(v)\right)$$

$$+2\left(AB^{(i)}_{M,n,N}+AB^{(ii)}_{M,n,N}+AC^{(i)}_{M,n,N}+AC^{(ii)}_{M,n,N}+BB^{(ii)}_{M,n,N}+BC^{(ii)}_{M,n,N}+BC^{(iii)}_{M,n,N}+CC^{(ii)}_{M,n,N}\right)$$

$$+o_p(\rho(n)^{1/4}).$$

We begin by studying the term $BB^{(i)}_{M,n,N}$ defined by (36). The term $BC^{(i)}_{M,n,N}$ defined by (38) and the term $CC^{(i)}_{M,n,N}$ defined by (41) are analogous to $BB^{(i)}_{M,n,N}$ and give the same contribution.

Using the Itô formula, the term

$$BB^{(i)}_{M,n,N}:=\frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}Y^2_{n,N}(s,s)v(s)ds$$

reads as

$$\frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}\int_0^{s}D^2_{N,n}(s-u)v(u)du\,v(s)ds \tag{48}$$

$$+\frac{1}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2\frac{1}{2\pi}\int_0^{2\pi}2\int_0^{s}Y_{n,N}(u,s)D_{N,n}(s-u)\sigma(u)dW_u\,v(s)ds. \tag{49}$$

The leading term is the first one, namely (48), which is easily seen to be equal to

$$\frac{1}{M+1}\frac{1}{6}M(M+1)(M+2)\,\frac{1}{2\pi}\int_0^{2\pi}\int_0^{s}D^2_{N,n}(s-u)v(u)du\,v(s)ds. \tag{50}$$

Now, using Lemma 8.2 and noting that $N/n \sim c_N/(2\pi)$, we have that (50) converges in probability to

$$\frac{c_M^2}{12}\,(1+2\eta(c_N/\pi))\frac{1}{2\pi}\int_0^{2\pi}\sigma^4(t)dt.$$

Consider now (49). Exploiting the boundedness of the process v, it is enough to observe that

$$E[(\int_0^s Y_{n,N}(u,s)D_{N,n}(s-u)\sigma(u)dW_u)^2] = E[\int_0^s Y_{n,N}^2(u,s)D_{N,n}^2(s-u)v(u)du]$$

$$\leq C\int_0^s\int_0^u D_{N,n}^2(s-s')ds'\,D_{N,n}^2(s-u)du = O(\rho(n)^3),$$

where we have used Lemma 8.2 and the property that $D_{N,n}^2(s-s') \leq CN^{-2}$ for $s' < s-\varepsilon$, $\varepsilon > 0$, for n large enough. Therefore, in probability the term (49) has order $M^2\rho(n)^{3/2} \sim c_M^2\rho(n)^{-1}\rho(n)^{3/2} = c_M^2\rho(n)^{1/2}$, and thus it converges to zero.

Finally, let

$$K := \left(\frac{1}{12}+\frac{2}{12}+\frac{1}{12}\right)\,\frac{c_M^2}{2\pi}(1+2\eta(c_N/\pi)) = \frac{1}{3}\,\frac{c_M^2}{2\pi}(1+2\eta(c_N/\pi)). \tag{51}$$

Then, following [Livieri et al., 2019], Theorem 3, the following convergence in probability holds:

$$\rho(n)^{-1/2}\left(K\hat{\sigma}_{n,N,M}^4 - (BB_{M,n,N}^{(i)} + 2BC_{M,n,N}^{(i)} + CC_{M,n,N}^{(i)})\right) \to K(X_{c_M} + Y_{c_M,c_N}),$$

where X_{c_M} are Y_{c_M,c_N} are defined as

$$X_{c_M} := -\frac{1}{c_M}\frac{1}{\pi}\int_0^{2\pi}\gamma^2(t)dt, \quad Y_{c_M,c_N} := 2c_M\frac{1}{\pi}(1+2\eta(c_N/\pi))\int_0^{2\pi}\sigma^4(t)dt.$$

Therefore, the proof that (47) is $o_p(\rho(n)^{1/4})$ is complete.

7.3 Step II. The Statistical Error

In this paragraph we consider the discretization error component (45), which reads

$$\frac{1}{M+1}\frac{1}{2\pi}\left(\int_0^{2\pi}2\int_0^s F''_{M,n}(s-u)v(u)du\,v(s)ds - \int_0^{2\pi}2\int_0^s F''_M(s-u)v(u)du\,v(s)ds\right)$$

$$= \frac{1}{M+1}\frac{1}{2\pi}\int_0^{2\pi}2\int_0^s \left(F''_{M,n}(s-u) - F''_M(s-u)\right)\,v(u)du\,v(s)ds \tag{52}$$

and prove that it is $o_p(\rho(n)^{1/4})$.

Up to a negligible multiplicative constant, we rewrite the latter as follows[7]:

$$E_{M,n}^{(1)} + E_{M,n}^{(2)},$$

where

$$E_{M,n}^{(1)} := \frac{1}{M+1}\sum_{|k|\leq M}\left(c_k(F''_{M,n}) - c_k(F''_M)\right)|c_k(v)|^2 \tag{53}$$

and

$$E_{M,n}^{(2)} := \frac{1}{M+1}\sum_{|k|>M}c_k(F''_{M,n})|c_k(v)|^2. \tag{54}$$

[7]With abuse of notation, in (53) we denoted by $c_k(F''_{M,n})$ the k-th Fourier coefficient of $\widetilde{F}''_{M,n}$, which is the quantity defined as $\widetilde{F}''_{M,n}(t) := F''_M(\varphi_n(t))$. A straightforward but lengthy proof, which is available from the authors, shows that the difference between the two kernels is negligible.

Note that (see, also, Remark 7.2), the term $|c_k(v)|^2$ reads

$$|c_k(v)|^2 = \frac{1}{4\pi^2 k^2}\left(\int_0^{2\pi}\gamma^2(s)\,ds + 2\int_0^{2\pi}\int_0^s \cos(k(s-u))\gamma(u)\,dZ_u\,\gamma(s)\,dZ_s\right). \tag{55}$$

Further, we compute the k-th Fourier coefficient of $F''_{M,n}$. First, note that, for any j, one has that

$$c_k\left(I_{[\frac{2\pi}{n}j,\frac{2\pi}{n}(j+1)[}\right) = \frac{1}{2\pi \mathrm{i} k}e^{-2\pi \mathrm{i}\frac{k}{n}j}\left(1-e^{-2\pi \mathrm{i}\frac{k}{n}}\right).$$

Therefore, it holds that

$$c_k(F''_{M,n}) = \sum_{j=0}^{n-1} F''_M\left(\frac{2\pi}{n}\,j\right)c_k\left(I_{[\frac{2\pi}{n}j,\frac{2\pi}{n}(j+1)[}\right) \tag{56}$$

$$= -\frac{1}{2\pi \mathrm{i} k}\left(1-e^{-2\pi \mathrm{i}\frac{k}{n}}\right)\sum_{|l|\le M}\left(1-\frac{|l|}{M+1}\right)l^2\sum_{j=0}^{n-1}e^{-2\pi \mathrm{i}\frac{k-l}{n}j}.$$

Now observe that, if n divides k, then $1-e^{-2\pi \mathrm{i}\frac{k}{n}}$ is equal to zero; accordingly, we assume that $k = nq + r$, $r \neq 0$ with either $q = 0$ if $|k| \le n$ or $q \neq 0$ otherwise. Moreover, note that the summation in (56) is either equal to n, if n divides $k-l$, or equal to zero, otherwise; hence, we set $l = r$, with $|r| \le M$. Thus, (56) reduces to

$$c_k(F''_{M,n}) = -\left(1-\frac{|r|}{M+1}\right)r^2\,\frac{1}{2\pi \mathrm{i} k}\,n\,(1-e^{-2\pi \mathrm{i}\frac{r}{n}}).$$

We now study the asymptotic behavior of the terms $E^{(1)}_{M,n}$ and $E^{(2)}_{M,n}$, separately. First, we prove that $E^{(2)}_{M,n}$ converges to zero in the L^1-norm. By taking into account the decomposition of $|c_k(v)|^2$ in (55), we have to prove that both terms resulting from such a decomposition are asymptotically negligible. However, here we explicitly compute the upper bound for the first term, which is the leading term. Using the bound $|1-e^{-2\pi \mathrm{i}\frac{r}{n}}| \le 2\pi\frac{|r|}{n}$, it holds

$$E\left[|E^{(2)}_{M,n}|\right] = \frac{1}{M+1}E\left[|2\sum_{k>M}c_k(F''_{M,n})\frac{1}{4\pi k^2}\int_0^{2\pi}\gamma^2(s)\,ds|\right] \tag{57}$$

$$\le C\frac{n}{M+1}E\left[\int_0^{2\pi}\gamma^2(s)\,ds\right]\sum_{k>M}\frac{r^2}{k^3}\left(1-\frac{r}{M+1}\right)|1-e^{-2\pi \mathrm{i}\frac{r}{n}}|$$

$$\le \frac{C}{M+1}\sum_{k>M}\frac{r^3}{k^3}\left(1-\frac{r}{M+1}\right) = \frac{C}{(M+1)^2}\sum_{q=1}^{\infty}\sum_{r=1}^{M}\frac{r^3(M+1-r)}{(nq)^3(1+\frac{r}{nq})^3}$$

$$\le \frac{CM^4}{n^3(M+1)^2}\sum_{q=1}^{\infty}\frac{1}{q^3}\sum_{r=1}^{M}\left(1+\frac{r}{nq}\right)^{-3}.$$

Now, note that:

$$\sum_{r=1}^{M}\left(1+\frac{r}{nq}\right)^{-3} \le \frac{1}{2}nq\left(1-(1+\frac{M}{nq})^{-2}\right) = (M+1) - \frac{3}{2}\frac{(M+1)^2}{nq} + o(1).$$

Therefore, (57) is $O(n^{-1})$.

For what concerns $E^{(1)}_{M,n}$, we also study the convergence in the L^1-norm. We set $k = r$. Again, we focus on the leading term from the decomposition in (55). For any fixed k, it holds:

$$c_k(F''_{M,n}) - c_k(F''_M) = -\left(1-\frac{|k|}{M+1}\right)k^2\left(\frac{1-e^{-2\pi \mathrm{i}\frac{k}{n}}}{2\,\pi\,\mathrm{i}\,\frac{k}{n}} - 1\right)$$

$$= -\left(1 - \frac{|k|}{M+1}\right) k^2 \left(\mathrm{i}\,\pi\,\frac{k}{n} - \frac{2}{3}\pi^2\frac{k^2}{n^2} + O\left(\frac{k^3}{n^3}\right)\right).$$

Therefore, we have:

$$E[|E^{(1)}_{M,n}|] = \frac{1}{M+1} E\left[\left|\sum_{|k|\le M}\left(c_k(F''_{M,n}) - c_k(F''_M)\right)\frac{1}{4\pi k^2}\int_0^{2\pi}\gamma^2(s)\,ds\right|\right]$$

$$\le \frac{1}{M+1}\sum_{|k|\le M}|c_k(F''_{M,n}) - c_k(F''_M)|\frac{1}{4\pi^2k^2}E\left[\int_0^{2\pi}\gamma^2(s)\,ds\right]$$

$$\le \frac{C}{n(M+1)}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)|k| = \frac{2C}{n(M+1)}\frac{M(M-2)}{6} = O(\rho(n)^{1/2}).$$

This completes the proof of the asymptotical negligibility of the statistical error with faster rate than $\rho(n)^{1/4}$.

7.4 Step III. Asymptotic Variance

This section follows [Jacod, 1997] in order to identify the asymptotic variance and prove the stable convergence in law.

First, consider the term (29), namely:

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)k^2c_k(v)c_{-k}(v) - \frac{1}{2\pi}\int_0^{2\pi}\gamma^2(t)dt. \tag{58}$$

Using the integration by parts formula and Remark 7.2, it holds $c_k(v) = \frac{1}{\mathrm{i}k}c_k(dv)$ and (58) is equal to

$$\frac{2\pi}{M+1}\sum_{|k|\le M}\left(1-\frac{|k|}{M+1}\right)c_k(dv)c_{-k}(dv) - \frac{1}{2\pi}\int_0^{2\pi}\gamma^2(t)dt.$$

By applying the Itô formula, the term (58) is equal to $2A_M(2\pi)$, where[8]

$$A_M(u) := \frac{1}{M+1}\,\frac{1}{2\pi}\int_0^{2\pi}\int_0^t F_M(t-s)dv(s)\,dv(t). \tag{59}$$

Then, according to [Jacod, 1997], we determine the variance of the asymptotic distribution by studying

$$\langle\rho(n)^{-1/4}2\,V_{M,n,N}, \rho(n)^{-1/4}2\,V_{M,n,N}\rangle_{2\pi}\,, \tag{60}$$

where

$$V_{M,n,N} := A_M + AB^{(i)}_{M,n,N} + AB^{(ii)}_{M,n,N} + AC^{(i)}_{M,n,N} + AC^{(ii)}_{M,n,N} \tag{61}$$
$$+BB^{(ii)}_{M,n,N} + BC^{(ii)}_{M,n,N} + BC^{(iii)}_{M,n,N} + CC^{(ii)}_{M,n,N}.$$

In the first step we study the bracket:

$$\rho(n)^{-1/2}\langle 2\,A_M, 2\,A_M\rangle_{2\pi}. \tag{62}$$

Noting that $M/c_M \simeq \rho(n)^{-1/2}$, we write (62) as

$$\frac{M}{c_M}4\,\frac{1}{(2\pi)^2}\int_0^{2\pi}\left(\frac{1}{M+1}\int_0^t F_M(t-s)dv(s)\right)^2\gamma^2(t)dt,$$

[8]To simplify the notation, in the following we will always omit the argument when it is equal to 2π, so we will write A_M instead of $A_M(2\pi)$. Similarly, for the processes in (61).

which, by using the Itô formula, is equal to

$$\frac{M}{c_M}\,\frac{1}{\pi^2}\int_0^{2\pi}\frac{1}{(M+1)^2}\int_0^t F_M^2(t-s)\gamma^2(s)ds\,\gamma^2(t)dt \tag{63}$$

$$+\frac{M}{c_M}\,\frac{1}{\pi^2}\int_0^{2\pi}\frac{1}{(M+1)^2}\int_0^t\int_0^s F_M(t-u)dv(u)\,F_M(t-s)dv(s)\,\gamma^2(t)dt. \tag{64}$$

Using Lemma 8.1, equation (121), it is seen that the term (63) converges in probability to

$$\frac{1}{2\pi}\int_0^{2\pi}\frac{4}{3}\,\frac{1}{c_M}\,\gamma^4(t)dt. \tag{65}$$

Further, in order to prove that the term (64) is $o_p(1)$, it is enough to compute

$$E[(\frac{1}{(M+1)^2}\int_0^t\int_0^s F_M(t-u)dv(u)\,F_M(t-s)dv(s))^2]^{1/2}$$

$$\leq C\,\frac{1}{M+1}\left(\frac{1}{M+1}\int_0^t E[\frac{1}{M+1}\int_0^s F_M^2(t-u)\gamma^2(u)du]\,F_M^2(t-s)ds\right)^{1/2}$$

and apply Lemma 8.1, which give that $(1/M)\int_0^s F_M^2(t-u)\gamma^2(u)du = o_p(1)$ and $(1/M)\int_0^t F_M^2(t-s)ds = O(1)$, because $(1/M)F_M^2(\cdot)$ is a *good kernel*.

The second step for identifying the asymptotic variance is to study

$$\rho(n)^{-1/2}\langle 2(AB^{(i)}_{M,n,N}+AB^{(ii)}_{M,n,N}+AC^{(i)}_{M,n,N}+AC^{(ii)}_{M,n,N}),2(AB^{(i)}_{M,n,N}+AB^{(ii)}_{M,n,N}+AC^{(i)}_{M,n,N}+AC^{(ii)}_{M,n,N})\rangle_{2\pi}. \tag{66}$$

The bracket contains 16 terms giving the same contribution for symmetry.

We consider the term $AB^{(i)}_{M,n,N}$ defined by (31). It holds

$$\langle\rho(n)^{-1/4}\,2AB^{(i)}_{M,n,N},\rho(n)^{-1/4}\,2AB^{(i)}_{M,n,N}\rangle_{2\pi} \tag{67}$$

$$=\rho(n)^{-1/2}\,\frac{4}{(2\pi)^2}\int_0^{2\pi}\left(\int_0^s\frac{1}{M+1}F''_{M,n}(s-u)v(u)du\right)^2 Y^2_{n,N}(s,s)v(s)ds$$

$$=\rho(n)^{-1/2}\,\frac{1}{\pi^2}\int_0^{2\pi}\left(\int_0^s\frac{1}{M+1}F''_{M,n}(s-u)v(u)du\right)^2\int_0^s D^2_{N,n}(s-u)v(u)du\,v(s)ds \tag{68}$$

$$+\rho(n)^{-1/2}\,\frac{2}{\pi^2}\int_0^{2\pi}\left(\int_0^s\frac{1}{M+1}F''_{M,n}(s-u)v(u)du\right)^2\int_0^s\int_0^u D_{N,n}(s-r)\sigma(r)dW_r\,D_{N,n}(s-u)\sigma(u)dW_u\,v(s)ds. \tag{69}$$

Consider (68). Applying Lemma 8.2, equation (130), and an analogous procedure to that applied in Step II for (52), by using the fact that $M/n\to 0$, it is equivalent to study

$$\rho(n)^{-1/2}\,\frac{1}{\pi^2}\,n^{-1}\pi(1+2\eta(c_N/\pi))\int_0^{2\pi}\left(\int_0^s\frac{1}{M+1}F''_M(s-u)v(u)du\right)^2 v^2(s)ds+o_p(1). \tag{70}$$

Applying the integration by parts and the boundedness of the volatility process v, it holds that

$$\int_0^s\frac{1}{M+1}F''_M(s-u)v(u)du=\int_0^s\frac{1}{M+1}F'_M(s-u)dv(u)+O_p(\rho(n)^{1/2}). \tag{71}$$

Therefore (70) gives

$$\rho(n)^{-1/2}\,\frac{1}{\pi^2}\pi\,n^{-1}(1+2\eta(c_N/\pi))\int_0^{2\pi}\left(\int_0^s\frac{1}{M+1}F'_M(s-u)dv(u)\right)^2 v^2(s)ds+o_p(1)$$

$$= \rho(n)^{-1/2}\,\frac{1}{\pi}\,n^{-1}(1+2\eta(c_N/\pi))\int_0^{2\pi}\int_0^s \frac{1}{(M+1)^2}|F'_M(s-u)|^2\gamma^2(u)du\,v^2(s)ds \tag{72}$$

$$+\rho(n)^{-1/2}\,\frac{2}{\pi}\,n^{-1}(1+2\eta(c_N/\pi))\frac{2}{(M+1)^2}\int_0^{2\pi}\int_0^s\int_0^u F'_M(s-r)dv(r)\,F'_M(s-u)dv(u)\,v^2(s)ds + o_p(1). \tag{73}$$

Consider (72). It is enough to observe that, by Lemma 8.1, it holds:

$$\frac{1}{M^3}\int_0^{2\pi}|F'_M(s-u)|^2\gamma^2(u)du\,v^2(s)ds \to \frac{2}{15}\pi\int_0^{2\pi}\gamma^2(u)du\,v^2(s)ds$$

and

$$\rho(n)^{-1/2}\,\frac{1}{\pi}\,n^{-1}(1+2\eta(c_N/\pi))\frac{M^3}{(M+1)^2} \to \frac{1}{\pi}\frac{c_M}{2}(1+2\eta(c_N/\pi)).$$

Then the term (72) converges to

$$\frac{1}{15}\,c_M\,(1+2\eta(c_N/\pi))\frac{1}{2\pi}\int_0^{2\pi}\gamma^2(s)\,v^2(s)ds. \tag{74}$$

It remains to prove that (73) is asymptotically negligible. To this aim, consider

$$E[(\int_0^s\int_0^u F'_M(s-r)dv(r)\,F'_M(s-u)dv(u))^2]^{1/2}$$

$$\leq C\left(\int_0^s|F'_M(s-u)|^2E[\int_0^u|F'_M(s-r)|^2\gamma^2(r)dr]\,du\right)^{1/2}.$$

Finally, using the fact that $\{K_M\}_M$ is a family of good kernels by Lemma 8.1, it holds that $\frac{1}{M^3}\int_0^u|F'_M(s-r)|^2\gamma^2(r)dr = o_p(1)$, for $r < s-\varepsilon$, $(\varepsilon>0)$, and $\frac{1}{M^3}\int_0^s|F'_M(s-u)|^2\gamma^2(u)du = O_p(1)$. Therefore, it follows that (73) has order $n^{-1/2}\,M^{-2}\,M^3\,o_p(1) = o_p(1)$.

We show now that (69) goes to zero. This result follows by noting that $D^2_{N,n}(s-r) < C\,N^{-2}$ for $r < s-\varepsilon$, for any $\varepsilon>0$ and n large enough, and the same remark used for (70). Thus the term (66) converges to

$$\frac{16}{15}\,c_M(1+2\eta(c_N/\pi))\frac{1}{2\pi}\int_0^{2\pi}\gamma^2(s)\,v^2(s)ds.$$

The last contribution to the variance of the asymptotic distribution is obtained by studying the bracket

$$\langle\rho(n)^{-1/4}2(BB^{(ii)}_{M,n,N}+BC^{(ii)}_{M,n,N}+BC^{(iii)}_{M,n,N}+CC^{(ii)}_{M,n,N}),\rho(n)^{-1/4}2(BB^{(ii)}_{M,n,N}+BC^{(ii)}_{M,n,N}+BC^{(iii)}_{M,n,N}+CC^{(ii)}_{M,n,N})\rangle_{2\pi}. \tag{75}$$

They are 16 terms giving the same contribution.

Consider

$$\langle\rho(n)^{-1/4}2BB^{(ii)}_{M,n,N},\rho(n)^{-1/4}2BB^{(ii)}_{M,n,N}\rangle_{2\pi} \tag{76}$$

$$= \rho(n)^{-1/2}4\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^u\frac{1}{M+1}F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s\right)^2Y^2_{n,N}(u,u)v(u)du.$$

By applying the Itô formula, we have that

$$\left(\int_0^u\frac{1}{M+1}F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s\right)^2 = \int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2Y^2_{n,N}(s,s)v(s)ds$$

$$+2\int_0^u\int_0^s\frac{1}{M+1}F''_{M,n}(u-s_1)Y_{n,N}(s_1,s)\sigma(s_1)dW_{s_1}\frac{1}{M+1}F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s.$$

Therefore, the bracket (76) splits as

$$\rho(n)^{-1/2}\int_0^{2\pi}\frac{4}{(2\pi)^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2Y^2_{n,N}(s,s)v(s)ds\,Y^2_{n,N}(u,u)v(u)du \tag{77}$$

$$+\rho(n)^{-1/2}\int_0^{2\pi}\frac{8}{(2\pi)^2}\int_0^u\int_0^s\frac{1}{(M+1)^2}F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\,F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s \tag{78}$$

$$\times Y^2_{n,N}(u,u)v(u)du.$$

Consider first (77), which gives the asymptotic variance term. As we show later, (78) goes instead to zero. By applying the Itô formula twice, it holds

$$\rho(n)^{-1/2}\int_0^{2\pi}\frac{1}{\pi^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2Y^2_{n,N}(s,s)v(s)ds\,Y^2_{n,N}(u,u)v(u)du$$

$$=\rho(n)^{-1/2}\int_0^{2\pi}\frac{1}{\pi^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2\int_0^s D^2_{N,n}(s-u')v(u')du'\,v(s)ds\int_0^u D^2_{N,n}(u-s')v(s')ds'\,v(u)du \tag{79}$$

$$+\rho(n)^{-1/2}\int_0^{2\pi}\frac{2}{\pi^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2\int_0^s\int_0^{s'}D_{N,n}(s-r)\sigma(r)dW_r\,D_{N,n}(s-s')\sigma(s')dW_{s'}\,v(s)ds \tag{80}$$

$$\times\int_0^u D^2_{N,n}(u-s')v(s')ds'\,v(u)du$$

$$+\rho(n)^{-1/2}\int_0^{2\pi}\frac{2}{\pi^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2\int_0^s D^2_{N,n}(s-u')v(u')du'\,v(s)ds \tag{81}$$

$$\times\int_0^u\int_0^{u'}D_{N,n}(u-r)\sigma(r)dW_r\,D_{N,n}(u-u')\sigma(u')dW_{u'}\,v(u)du$$

$$+\rho(n)^{-1/2}\int_0^{2\pi}\frac{4}{\pi^2}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2\int_0^s\int_0^{s'}D_{N,n}(s-r)\sigma(r)dW_r\,D_{N,n}(s-s')\sigma(s')dW_{s'}\,v(s)ds \tag{82}$$

$$\times\int_0^u\int_0^{u'}D_{N,n}(u-r)\sigma(r)dW_r\,D_{N,n}(u-u')\sigma(u')dW_{u'}\,v(u)du.$$

Using Lemma 8.2, equation (130), the term (79) gives

$$\rho(n)^{-1/2}\frac{1}{n^2}(1+2\eta(c_N/\pi))^2\int_0^{2\pi}\int_0^u\frac{1}{(M+1)^2}|F''_{M,n}(u-s)|^2\sigma^4(s)ds\,\sigma^4(u)\,du.$$

Finally, using Lemma 8.1, the term (79) converges to

$$\frac{1}{105}\,c_M^3\,(1+2\eta(c_N/\pi))^2\,\frac{1}{2\pi}\int_0^{2\pi}\sigma^8(t)dt.$$

A similar procedure as for (69) allows us to prove that (80), (81) and (82) go to zero in probability.

We verify now that the (78) goes to zero in probability. It is enough to show that

$$\int_0^u\int_0^s\frac{1}{(M+1)^2}F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\,F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s\,Y^2_{n,N}(u,u) \tag{83}$$

is $o_p(\rho(n)^{1/2})$. By the Itô formula and Lemma 8.2, it holds that

$$\int_0^u\int_0^s\frac{1}{(M+1)^2}F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\,F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s\,Y^2_{n,N}(u,u)$$

$$=\int_0^u\int_0^s\frac{1}{(M+1)^2}F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\,F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s\int_0^u D^2_{N,n}(u-r)v(r)dr \tag{84}$$

$$+o_p(\rho(n)^{1/2}).$$

Consider (84). By the Cauchy-Schwarz inequality,

$$E[|\int_0^u \int_0^s \frac{1}{(M+1)^2} F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\, F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s \int_0^u D^2_{N,n}(u-r)v(r)dr|]$$

$$\leq E[(\int_0^u \int_0^s \frac{1}{(M+1)^2} F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\, F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s)^2]^{1/2}$$

$$\times E[(\int_0^u D^2_{N,n}(u-r)v(r)dr)^2]^{1/2}.$$

Note that $\int_0^u D^2_{N,n}(u-r)v(r)dr = O_p(\rho(n))$, by Lemma 8.2. Consider then

$$E[(\int_0^u \int_0^s \frac{1}{(M+1)^2} F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\, F''_{M,n}(u-s)Y_{n,N}(s,s)\sigma(s)dW_s)^2]$$

$$= E[\int_0^u \left(\int_0^s \frac{1}{(M+1)^2} F''_{M,n}(u-r)Y_{n,N}(r,s)\sigma(r)dW_r\right)^2 |F''_{M,n}(u-s)|^2 Y^2_{n,N}(s,s)v(s)ds]$$

$$= E[\int_0^u \int_0^s \frac{1}{(M+1)^4} |F''_{M,n}(u-r)|^2 Y^2_{n,N}(r,s)v(r)dr\, |F''_{M,n}(u-s)|^2 Y^2_{n,N}(s,s)v(s)ds] + o(\rho(n))$$

$$= E[\int_0^u \int_0^s \frac{1}{(M+1)^4} |F''_{M,n}(u-r)|^2 \int_0^r D^2_{N,n}(s-u')v(u')du'\, v(r)dr\, |F''_{M,n}(u-s)|^2 Y^2_{n,N}(s,s)v(s)ds] + o(\rho(n)). \quad (85)$$

Again, using that $D^2_{N,n}(s-u') \leq C/N^2$ for $u' < s-\varepsilon$, for $\varepsilon > 0$ and n large enough, then (85) is smaller than

$$\frac{C}{N^2}\frac{1}{(M+1)^4} E[\int_0^u |F''_{M,n}(u-s)|^2 \int_0^s |F''_{M,n}(u-r)|^2 v(r)dr\, Y^2_{n,N}(s,s)ds]$$

$$= \frac{C}{N^2}\frac{1}{(M+1)^4} E[\int_0^u |F''_{M,n}(u-s)|^2 \int_0^s |F''_{M,n}(u-r)|^2 v(r)dr \left(\int_0^s D^2_{N,n}(s-u')v(u')du' + o_p(\rho(n)^2)\right) ds]$$

$$= \frac{C}{N^2}\frac{1}{(M+1)^4} E[\int_0^u |F''_{M,n}(u-s)|^2 \int_0^s |F''_{M,n}(u-r)|^2 v(r)dr\, v(s)ds + o_p(\rho(n))].$$

Then, using Lemma 8.1, and the fact that $(1/M^5)|F''_{M,n}(x)|^2$ is a good kernel, then this term has order $N^{-2}M^6 o(1) = M^2 o(1)$. Finally, we obtain the order of (83), which is $o_p(\rho(n)^{1/2})$.

We can conclude that the contribution of all the terms (75) is

$$\frac{16}{105}\, c_M^3\, (1+2\eta(c_N/\pi))^2 \frac{1}{2\pi}\int_0^{2\pi} \sigma^8(t)dt.$$

It remains to show that the other brackets in (60) give asymptotically negligible contributions. We study in detail the convergence in probability

$$\langle \rho(n)^{-1/4} A_M, \rho(n)^{-1/4} AB^{(i)}_{M,n,N}\rangle_{2\pi} \to 0, \quad (86)$$

and

$$\langle \rho(n)^{-1/4} AB^{(i)}_{M,n,N}, \rho(n)^{-1/4} BB^{(ii)}_{M,n,N}\rangle_{2\pi} \to 0. \quad (87)$$

The proof is analogous for the other terms. The bracket (86) is equal to

$$\rho(n)^{-1/2}\frac{1}{(2\pi)^2}\frac{1}{(M+1)^2}\int_0^{2\pi}\int_0^t F_M(t-s)dv(s)\int_0^t F''_{M,n}(t-s)v(s)ds\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt. \quad (88)$$

Omitting negligible constants, and using the result obtained in Step II for the term (52), by virtue of the fact that $M/n \to 0$, then (88) is equal to

$$\rho(n)^{-1/2}\left(\frac{1}{(M+1)^2}\int_0^{2\pi}\int_0^t F_M(t-s)dv(s)\int_0^t F_M''(t-s)v(s)ds\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt + o_p(\rho(n)^{1/2})\right).$$

Moreover, by (71), we are lead to study

$$\frac{1}{M+1}\int_0^{2\pi}\int_0^t F_M(t-s)dv(s)\int_0^t F_M'(t-s)dv(s)\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt. \tag{89}$$

Applying the Itô formula, (89) is equal to

$$\frac{1}{M+1}\int_0^{2\pi}\int_0^t F_M(t-s)F_M'(t-s)\gamma^2(s)ds\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt \tag{90}$$

$$+\frac{1}{M+1}\int_0^{2\pi}\int_0^t\int_0^s F_M(t-u)dv(u)\, F_M'(t-s)dv(s)\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt \tag{91}$$

$$+\frac{1}{M+1}\int_0^{2\pi}\int_0^t\int_0^s F_M'(t-u)dv(u)\, F_M(t-s)dv(s)\, Y_{n,N}(t,t)\sigma(t)\gamma(t)\rho\, dt. \tag{92}$$

Consider (90). It is equal to the sum

$$\frac{1}{M+1}\int_0^{2\pi}\int_0^t\int_0^s D_{N,n}(t-u)\sigma(u)dW_u\, F_M(t-s)F_M'(t-s)\gamma^2(s)ds\,\sigma(t)\gamma(t)\rho\, dt \tag{93}$$

$$+\frac{1}{M+1}\int_0^{2\pi}\int_0^t\int_0^s F_M(t-u)F_M'(t-u)\gamma^2(u)du\, D_{N,n}(t-s)\sigma(s)dW_s\,\sigma(t)\gamma(t)\rho\, dt. \tag{94}$$

We study (93). The term (94) is analogous. By the boundedness of the volatility and the volatility of volatility processes, it is enough to prove that

$$E[|\frac{1}{M+1}\int_0^t\int_0^s D_{N,n}(t-u)\sigma(u)dW_u\, F_M(t-s)F_M'(t-s)\gamma^2(s)ds|] \to 0.$$

The term (93) is smaller than

$$\frac{1}{M+1}\int_0^t E[(\int_0^s D_{N,n}(t-u)\sigma(u)dW_u)^2]^{1/2}\,|F_M(t-s)F_M'(t-s)|ds$$

$$\leq \frac{1}{M+1}\int_0^t E[\int_0^s D_{N,n}^2(t-u)v(u)du]^{1/2}\,|F_M(t-s)F_M'(t-s)|ds.$$

Noting that $D_{N,n}^2(t-u) < C/N^2$ for $u < t-\varepsilon$, for $\varepsilon > 0$, and n large enough, the previous term is smaller than

$$\frac{1}{M+1}\frac{C}{N}\left(\int_0^t |F_M(t-s)|^2ds\right)^{1/2}\left(\int_0^t |F_M'(t-s)|^2ds\right)^{1/2}.$$

Using Lemmas 8.1, this last term is $C\, M^{-1}n^{-1}M^{1/2}M^{3/2} = O(n^{-1/2})$.

Consider now (91). The term (92) is analogous. It is enough to show that

$$\frac{1}{M+1}E[|\int_0^t\int_0^s F_M(t-u)dv(u)\, F_M'(t-s)dv(s)\, Y_{n,N}(t,t)|] \to 0.$$

Therefore, using the Cauchy-Schwarz inequality, we consider

$$\frac{1}{M+1}E[(\int_0^t\int_0^s F_M(t-u)dv(u)\, F_M'(t-s)dv_s)^2]^{1/2}\, E[(Y_{n,N}(t,t))^2]^{1/2}.$$

Using the fact that $E[(Y_{n,N}(t,t))^2]^{1/2} = O(n^{-1/2})$, the Itô isometry and the boundedness of volatility of volatility, we have

$$E[(\int_0^t\int_0^s F_M(t-u)dv(u)\, F_M'(t-s)dv(s))^2] = E[\int_0^t(\int_0^s F_M(t-u)dv(u))^2\,|F_M'(t-s)|^2\gamma^2(s)ds]$$

$$\leq C\int_0^t |F'_M(t-s)|^2\, E[(\int_0^s F_M(t-u)dv(u))^2]\, ds = C\, M^4 \int_0^t \frac{1}{M^3}|F'_M(t-s)|^2\, E[\frac{1}{M}\int_0^s F_M^2(t-u)\gamma^2(u)du]\, ds.$$

By Lemma 8.1, it holds that for $u < t-\varepsilon$, $\frac{1}{M}\int_0^s F_M^2(t-u)\gamma^2(u)du = o_p(1)$, and $\frac{1}{M^3}\int_0^t |F'_M(t-s)|^2 ds = O(1)$. Then (91) has order $M^{-1}n^{-1/2}M^{3/2}M^{1/2}o_p(1) = o_p(1)$.

Consider now the bracket (87). Neglecting the irrelevant constants, this is equal to

$$\rho(n)^{-1/2}\int_0^{2\pi}\frac{1}{(M+1)^2}\int_0^s F''_{M,n}(s-u)v(u)du\int_0^s F''_{M,n}(s-u)Y_{n,N}(u,u)\sigma(u)dW_u\, Y^2_{n,N}(s,s)v(s)ds.$$

Applying the Itô formula, we have to study

$$\frac{1}{M}\int_0^{2\pi}\int_0^s\int_0^u F''_{M,n}(u-u_1)v(u_1)du_1\, F''_{M,n}(s-u)Y_{n,N}(u,u)\sigma(u)dW_u\, Y^2_{n,N}(s,s)v(s)ds \tag{95}$$

$$+\frac{1}{M}\int_0^{2\pi}\int_0^s\int_0^u F''_{M,n}(u-u_1)Y_{n,N}(u_1,u_1)\sigma(u_1)dW_{u_1}\, F''_{M,n}(s-u)v(u)du\, Y^2_{n,N}(s,s)v(s)ds.$$

We study (95). The other term is analogous. By the boundedness of the volatility process and applying the Cauchy-Schwarz inequality, it is enough to study

$$\frac{1}{M}E[(\int_0^s\int_0^u F''_{M,n}(u-u')v(u')du'\, F''_{M,n}(s-u)Y_{n,N}(u,u)\sigma(u)dW_u)^2]^{1/2}\, E[Y^4_{n,N}(s,s)]^{1/2}. \tag{96}$$

Using the Burholder-Davis-Gundy inequality and Lemma 8.2, it holds

$$E[Y^4_{n,N}(s,s)]^{1/2} \leq C\rho(n). \tag{97}$$

Then we compute:

$$E[(\int_0^s\int_0^u F''_{M,n}(u-u')v(u')du'\, F''_{M,n}(s-u)Y_{n,N}(u,u)\sigma(u)dW_u)^2]^{1/2}$$

$$= E[\int_0^s |\int_0^u F''_{M,n}(u-u')v(u')du'|^2 |F''_{M,n}(s-u)|^2 Y^2_{n,N}(u,u)v(u)du]^{1/2}$$

$$= E[\int_0^s |\int_0^u F''_{M,n}(u-u')v(u')du'|^2 |F''_{M,n}(s-u)|^2 \times$$

$$\times\left(\int_0^u D^2_{N,n}(u-r)v(r)dr + 2\int_0^u\int_0^r D_{N,n}(u-u_1)\sigma(u_1)dW_{u_1}\, D_{N,n}(u-r)\sigma(r)dW_r\right)v(u)du]^{1/2}.$$

By Lemma 8.2, it is enough to study:

$$E[\int_0^s |\int_0^u F''_{M,n}(u-u')v(u')du'|^2 |F''_{M,n}(s-u)|^2 \int_0^u D^2_{N,n}(u-r)v(r)drv(u)du]^{1/2}. \tag{98}$$

Using Lemma 8.2 and equation (71), the leading term of (98) gives

$$n^{-1/2}E[\int_0^s (\int_0^u F'_{M,n}(u-u')dv(u'))^2 |F''_{M,n}(s-u)|^2\, v^2(u)du]^{1/2}$$

$$\leq Cn^{-1/2}\left(\int_0^s |F''_{M,n}(s-u)|^2 E[|\int_0^u F'_{M,n}(u-u')dv(u')|^2]\, du\right)^{1/2}$$

$$\leq Cn^{-1/2}\left(\int_0^s |F''_{M,n}(s-u)|^2 \int_0^u |F'_{M,n}(u-u')|^2 du'\, du\right)^{1/2}$$

$$= Cn^{-1/2}\, M^{5/2}M^{3/2}\left(\int_0^s \frac{1}{M^5}|F''_{M,n}(s-u)|^2 \int_0^u \frac{1}{M^3}|F'_{M,n}(u-u')|^2 du'\, du\right)^{1/2}.$$

Finally, using Lemma 8.1, the term (95) is $O(M^{-1}n^{-1}n^{-1/2}\, M^{5/2}M^{3/2}) = O(n^{-1/2})$.

7.5 Step IV. Orthogonality

The final step requires to prove that in probability, as $n, N, M \to \infty$,

$$\langle \rho(n)^{-1/4} 2\, V_{M,n,N}, W \rangle_{2\pi} \to 0.$$

We provide a detailed proof for the convergence to 0 in probability of the bracket

$$\langle \rho(n)^{-1/4}\, BB^{(ii)}_{M,n,N}, W \rangle_{2\pi}. \tag{99}$$

The convergence of the other terms can be shown with an analogous procedure.

Consider (99), omitting the negligible constants. We show that

$$E[(\rho(n)^{-1/4} \int_0^{2\pi} \int_0^u \frac{1}{M+1} F''_{M,n}(u-s) Y_{n,N}(s,s)\sigma(s) dW_s\, Y_{n,N}(u,u)\sigma(u)du)^2] \to 0. \tag{100}$$

Let

$$Z_{N,M,n}(u,u) := \int_0^u \frac{1}{M+1} F''_{M,n}(u-s) Y_{n,N}(s,s)\sigma(s) dW_s.$$

Then (100) is equal to

$$\rho(n)^{-1/2} E[\int_0^{2\pi} \sigma(u) \int_0^{2\pi} \sigma(u') Z_{N,M,n}(u,u) Z_{N,M,n}(u',u')\, Y_{n,N}(u,u) Y_{n,N}(u',u')\, du'\, du]. \tag{101}$$

By symmetry, we can assume $u' \leq u$. By the Itô formula:

$$Z_{N,M,n}(u,u) Z_{N,M,n}(u',u')$$

$$= \int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) Y^2_{n,N}(s,s) v(s) ds \tag{102}$$

$$+2 \int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} \int_0^s F''_{M,n}(u'-r) Y_{N,n}(r,r)\sigma(r) dW_r\, F''_{M,n}(u-s) Y_{n,N}(s,s)\sigma(s) dW_s \tag{103}$$

and

$$Y_{n,N}(u,u) Y_{n,N}(u',u')$$

$$= \int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t) dt \tag{104}$$

$$+2 \int_0^u 1_{[0,u']}(t) \int_0^t D_{N,n}(u'-r)\sigma(r) dW_r\, D_{N,n}(u-t)\sigma(t) dW_t. \tag{105}$$

It is enough to consider the terms in (102) and (104). In fact, the terms (103) and (105) are of higher infinitesimal order. Substitute the terms (102) and (104) into (101) to obtain

$$\rho(n)^{-1/2} E[\int_0^{2\pi} \sigma(u) \int_0^{2\pi} \sigma(u') \int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) Y^2_{n,N}(s,s) v(s) ds \tag{106}$$

$$\times \int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t) dt\, du'\, du].$$

Consider

$$Y^2_{n,N}(s,s) = \int_0^s D^2_{N,n}(s-s_1) v(s_1) ds_1 + 2 \int_0^s \int_0^{s_1} D_{N,n}(s-s_2)\sigma(s_2) dW_{s_2}\, D_{N,n}(s-s_1)\sigma(s_1) dW_{s_1},$$

and substitute into equation (106). Finally, we have that

$$\rho(n)^{-1/2} E[\int_0^{2\pi} \sigma(u) \int_0^{2\pi} \sigma(u') \int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s)$$

$$\times \int_0^s D^2_{N,n}(s-s_1)v(s_1)ds_1\, v(s)ds \int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt\, du'\, du] + o(1)$$

$$\leq C\rho(n)^{1/2} \int_0^{2\pi}\int_0^{2\pi} E[|\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) v(s)ds \tag{107}$$

$$\times \int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt|]\, du'\, du + o(1),$$

where we have used the boundedness of volatility and Lemma 8.2. Then, using the Cauchy-Schwarz inequality, it holds that (107) is smaller than

$$C\rho(n)^{1/2} \int_0^{2\pi}\int_0^{2\pi} E[(\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) v(s)ds)^2]^{1/2} \tag{108}$$

$$\times E[(\int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt)^2]^{1/2}\, du'\, du.$$

First case, for any $\varepsilon > 0$, $u' < u - \varepsilon$:

$$E[(\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) v(s)ds)^2] \tag{109}$$

$$\leq E[\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} |F''_{M,n}(u-s)|^2 v(s)ds] E[\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} |F''_{M,n}(u'-s)|^2 v(s)ds]. \tag{110}$$

Using Lemma 8.1 and the fact that $u' < u - \varepsilon$, these terms have order $M^3\, M^3\, o(1)$. Then consider

$$E[(\int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt)^2] \tag{111}$$

$$\leq E[\int_0^u 1_{[0,u']}(t) D^2_{N,n}(u-t) v(t)dt] E[\int_0^u 1_{[0,u']}(t) D^2_{N,n}(u'-t) v(t)dt].$$

Using Lemma 8.2 and the fact that $D^2_{N,n}(u-t) \leq \frac{C}{N^2}$ for $t < u - \varepsilon$ and n large enough, we see that this term has order $N^{-2}\, \rho(n)$. Coming back to (108), it has order $\rho(n)^{1/2} M^{3/2}\, M^{3/2}\, N^{-1}\, \rho(n)^{1/2}\, o(1) = o(\rho(n)^{1/2}) \to 0$.

Now consider the case $u - \varepsilon \leq u' \leq u$, for any $\varepsilon > 0$. Starting from equation (108), we have

$$\rho(n)^{1/2} \int_0^{2\pi} du \int_0^{2\pi} 1_{[u-\varepsilon,u]}(u')\, du' E[(\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) v(s)ds)^2]^{1/2} \tag{112}$$

$$\times E[(\int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt)^2]^{1/2}.$$

In this case, it is easily seen that the order of

$$\rho(n)^{1/2}\, E[(\int_0^u 1_{[0,u']}(s) \frac{1}{(M+1)^2} F''_{M,n}(u-s) F''_{M,n}(u'-s) v(s)ds)^2]^{1/2}\, E[(\int_0^u 1_{[0,u']}(t) D_{N,n}(u-t) D_{N,n}(u'-t) v(t)dt)^2]^{1/2}$$

is $\rho(n)^{1/2} M^{3/2} M^{3/2} \rho(n)^{1/2} \rho(n)^{1/2} = O(1)$. Thus we have proved that (112) is smaller than $C\,\varepsilon$, for any $\varepsilon > 0$.

The proof of Theorem 3.1 is now completed.

7.6 Proof of Proposition 3.4

In the following, for any random function β, we denote as $c_0(\beta)$ the 0-th Fourier coefficient of β, that is, $c_0(\beta) = \frac{1}{2\pi} \int_0^{2\pi} \beta(t)dt$. Under the conditions on n, N, M, L, we prove the convergence in probability

$$V^{(1)}_{n,N,M,L} := \sum_{|k| \leq L} \bar{c}_k(\gamma^2_{n,N,M}) \bar{c}_{-k}(\gamma^2_{n,N,M}) \to c_0(\gamma^4). \tag{113}$$

We recall that

$$\bar{c}_k(\gamma^2_{n,N,M}) = c_k(\gamma^2_{n,N,M}) - K 2\pi c_k(\sigma^4_{n,N,M}), \tag{114}$$

according to definitions (11) and (14). Then, we plug (114) into (113) and, using the product formula for the Fourier coefficients (see [Livieri et al., 2019]), we observe the following convergences in probability:

$$\sum_{|k|\leq L} c_k(\gamma^2_{n,N,M}) c_{-k}(\gamma^2_{n,N,M}) \to c_0(\gamma^4) + \frac{1}{9} c_M^4 (1 + 2\eta(c_N/\pi))^2 c_0(\sigma^8) + \frac{2}{3} c_M^2 (1 + 2\eta(c_N/\pi)) c_0(\gamma^2\sigma^4),$$

$$2\pi K \sum_{|k|\leq L} c_k(\sigma^4_{n,N,M}) c_{-k}(\gamma^2_{n,N,M}) \to 2\pi K c_0(\gamma^2\sigma^4) + 2\pi K \frac{1}{3} c_M^2 (1 + 2\eta(c_N/\pi)) c_0(\sigma^8),$$

$$2\pi K \sum_{|k|\leq L} c_k(\sigma^4_{n,N,M}) c_{-k}(\sigma^4_{n,N,M}) \to c_0(\sigma^8).$$

Then, the convergence in (113) is ensured. We omit the proof for the convergence of $V^{(2)}_{n,N,M,L}$ and $V^{(3)}_{n,N,M,L}$, which does not contain any novel computation w.r.t. the previous one.

7.7 Proof of Theorem 3.5

The theorem follows straightforwardly from the stable convergence in Theorem 3.1 and Proposition 3.4.

7.8 Proof of Theorem 3.6

The proof relies on the basic decomposition as in Section 7.1. First of all, we prove that the bias correction is not needed because

$$BB^{(i)}_{M,n,N} + 2BC^{(i)}_{M,n,N} + CC^{(i)}_{M,n,N} = o_p(\rho(n)^{\iota/2}). \tag{115}$$

We study the term $BB^{(i)}_{M,n,N}$ defined by (36). The term $BC^{(i)}_{M,n,N}$ defined by (38) and the term $CC^{(i)}_{M,n,N}$ defined by (41) are analogous to $BB^{(i)}_{M,n,N}$.

Using the Itô formula, the term $BB^{(i)}_{M,n,N}$ is equal to

$$\frac{1}{M+1}\frac{1}{6} M(M+1)(M+2)\frac{1}{n} 2 \int_0^{2\pi} n \int_0^s D^2_{N,n}(s-u) v(u) du\, v(s) ds + o_p(\rho(n)^{1/2}), \tag{116}$$

where the order of the martingale part is obtained in Section 7.2.

Now, using Lemma 8.2 ii) and noting that $N/n \approx c_N$ and $M/n^\iota \approx c_M$, we have that (116) has order, in probability, equal to

$$n^{2\iota}\frac{1}{n}. \tag{117}$$

It is then enough to observe that $2\iota - 1 + \iota/2 < 0$, as soon as $\iota < 1/5$, as in the assumption. Thus (115) is proved.

The slower rate of M ensures that the discretization error is still negligible. As for the asymptotic variance, the only term which remains is the bracket

$$\langle \rho(n)^{-\iota/2} 2\, A_M, \rho(n)^{-\iota/2} 2\, A_M \rangle_{2\pi}.$$

Noting that $M/c_M \simeq \rho(n)^{-\iota}$, we obtain that its limit in probability is

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{4}{3}\frac{1}{c_M}\gamma^4(t) dt.$$

7.9 Proof of Theorem 3.7

The theorem follows straightforwardly from the stable convergence in Theorem 3.6 and the convergence in probability of $\Gamma_{n,N,M,L}$ to the asymptotic variance. The latter is immediately deduced from the following two remarks. First, under the conditions $N\rho(n) \sim c_N$ and $M\rho(n)^\iota \sim c_M$, where $\iota \in (0, 2/5)$, then $c_k(\gamma^2_{n,N,M})$, defined in (14), converges in probability to $c_k(\gamma^2)$, in virtue of the proof of Theorem 3.6 and Remark 3.3. Secondly, the product formula is applied.

8 Appendix B: Some properties of the Fejér and Dirichlet kernels

This section resumes some results about the rescaled Dirichlet kernel, defined as

$$D_N(x) := \frac{1}{2N+1} \sum_{|k|\leq N} e^{\mathrm{i}kx} = \frac{1}{2N+1} \frac{\sin((2N+1)x/2)}{\sin(x/2)} \tag{118}$$

and the Fejér kernel, defined as

$$F_M(x) := \sum_{|k|\leq M} \left(1 - \frac{|k|}{M+1}\right) e^{\mathrm{i}kx} = \frac{1}{M+1} \left(\frac{\sin((M+1)x/2)}{\sin(x/2)}\right)^2. \tag{119}$$

In the following we consider a regular partition of the time interval, still maintaining the continuous time notation used in the main proof of the TLCs.

Lemma 8.1 *Suppose that $M^2/n \to a$, as $n, M \to \infty$, for some constant $a > 0$. Then, it holds that:*

$$\lim_{n,M} \int_{-\pi}^{\pi} F_M(\varphi_n(x))\, dx = \lim_{M} \int_{-\pi}^{\pi} F_M(x)\, dx = 2\pi, \tag{120}$$

$$\lim_{M,n} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(\varphi_n(x)) dx = \lim_{M} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(x) dx = \frac{4\pi}{3}, \tag{121}$$

$$\lim_{n,M} \int_{-\pi}^{\pi} \frac{1}{M^3} |F'_M(\varphi_n(x))|^2\, dx = \lim_{M} \int_{-\pi}^{\pi} \frac{1}{M^3} |F'_M(x)|^2\, dx = \frac{2}{15}\pi, \tag{122}$$

$$\lim_{n,M} \int_{-\pi}^{\pi} \frac{1}{M^5} |F''_M(\varphi_n(x))|^2\, dx = \lim_{M} \int_{-\pi}^{\pi} \frac{1}{M^5} |F''_M(x)|^2\, dx = \frac{4}{105}\pi. \tag{123}$$

In addition, let

$$K_M(x) := \frac{15(M+1)}{M(4+6M+4M^2+M^3)} |F'_M(x)|^2, \tag{124}$$

$$L_M(x) := \frac{105(M+1)}{M(2M^5+12M^4+30M^3+40M^2+23M-2)} |F''_M(x)|^2. \tag{125}$$

Then, $\{K_M(x)\}_{M=1}^{\infty}$ and $\{L_M(x)\}_{M=1}^{\infty}$ are families of good kernels.[9]

Proof The results in (120) and (121) are proved in [Cuchiero and Teichmann, 2015], Lemma 5.1. Regarding the first equality in (122) and (123), it is sufficient to consider the Euler-MacLaurin formula applied to the squared first and second derivative of the Fejér kernel. For the sake of completeness, remind that for a function $f : [-\pi, \pi] \to \mathbb{R}$ of class C^{2p+1} it holds that

$$\int_{[-\pi,\pi]} f(x)\, dx - \frac{2\pi}{n} \left(\frac{f(\pi) + f(-\pi)}{2} + \sum_{j=1}^{n-1} f\left(a + j\frac{2\pi}{n}\right) \right) \tag{126}$$

[9] See [Stein and Shakarchi, 2011] for the definition.

$$= \sum_{k=1}^{p} \left(\frac{2\pi}{n}\right)^{2k} \frac{B_{2k}}{(2k)!} \left(f^{(2k-1)}(-\pi) - f^{(2k-1)}(\pi)\right) + R_{p,n,f},$$

where B_{2k} is the $(2k)$-th Bernoulli number and the rest $R_{p,n,f}$ satisfies

$$|R_{p,n,f}| \leq C_p \left(\frac{2\pi}{n}\right)^{2p+1} \int_{-\pi}^{\pi} |f^{(2p+1)}(x)|\, dx, \tag{127}$$

with C_p a constant depending only on p. In particular, let us consider positive integers k, h; then, we have that

$$(F_M^{(k)}(x))^h = \sum_{|j_1|,\ldots,|j_h| \leq M} \left(1 - \frac{|j_1|}{M+1}\right) \cdots \left(1 - \frac{|j_h|}{M+1}\right) (\mathrm{i}j_1)^k \ldots (\mathrm{i}j_h)^k e^{\mathrm{i}(j_1+\ldots+j_h)x}. \tag{128}$$

By observing that the number of terms in the summation is $(2M+1)^h$, that $|j_1 + \ldots + j_h| \leq hM$, using the bounds $\left(1 - \frac{|j_1|}{M+1}\right) \ldots \left(1 - \frac{|j_h|}{M+1}\right) \leq 1$ and $|e^{\mathrm{i}(j_1+\ldots+j_h)x}| \leq 1$, we have that $|f^{(2p+1)}| \leq 2^h h^{2p+1} M^{(k+1)h+p}$, where $f := (F_M^{(k)})^h$. It follows that $|R_{p,nf}| \leq C_p (2\pi)^{2p+2} 2^h h^{2p+1} M^{(k+1)h+p}/n^{2p+1}$. As a consequence, for k and h fixed and $M^2/n \to a$, we have $|R_{p,n,f}| = O(M^{(k+1)h-p-1})$. Therefore, for both $f = (1/M^3)|F_M^{(1)}|^2$ and $f = (1/M^5)|F_M^{(2)}|^2$, we have that $|R_{p,n,f}| = O(M^{-p})$.

For the result of interest, it is sufficient to consider $p = 1$. We show now that both the second term on the left hand side and the first on the right hand side of (126) converge to zero if $M^2/n \to a$ as $n, M \to \infty$. First, when $f := (1/M^3)|F_M^{(1)}|^2$, we have that the two terms are equal to zero since $F_M^{(1)}(\pi) = F_M^{(1)}(-\pi) = 0$. Instead, when $f := (1/M^5)|F_M^{(2)}|^2$ the term on the left hand side is an $O(M^{-3})$, whereas the term on the right hand side is equal to zero since $F_M^{(3)}(\pi) = F_M^{(3)}(-\pi) = 0$.

The assertion in (122) follows directly from the following calculation

$$\int_{-\pi}^{\pi} \frac{1}{M^3} |F_M'(x)|^2\, dx = 2\pi \frac{1}{M^3} \sum_{|k| \leq M} \left(1 - \frac{|k|}{M+1}\right)^2 k^2 = 2\pi \frac{M^3 + 4M^2 + 6M + 4}{15M^2(M+1)} \to \frac{2}{15}\pi. \tag{129}$$

Similarly, one obtains that

$$\int_{-\pi}^{\pi} \frac{1}{M^5} |F_M''(x)|^2\, dx = 2\pi \frac{1}{M^5} \sum_{|k| < M} \left(1 - \frac{|k|}{M+1}\right)^2 k^4$$

$$= 2\pi \frac{2M^5 + 12M^4 + 30M^3 + 40M^2 + 23M - 2}{105M^4(M+1)} \to \frac{4}{105}\pi.$$

It remains to prove that $\{K_M(x)\}_{M=1}^{\infty}$ and $\{L_M(x)\}_{M=1}^{\infty}$ are families of good kernels. First, consider K_M. We observe that $K_M(x) \geq 0$. Then, by using the previous computation, it is easy to show that

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} K_M(x)\, dx = 1.$$

Finally, by using the explicit expressions in terms of sine and cosine and the fact that $|\sin((M+1)x/2)| \leq C(M+1)|x|$ and $|\sin(x/2)| \geq c|x|$ for $|x| \leq \pi$, with $C, c > 0$ suitable constants, we have that

$$\int_{\delta \leq |x| \leq \pi} |K_M(x)|\, dx = \int_{\delta \leq |x| \leq \pi} \frac{15(M+1)}{M(4 + 6M + 4M^2 + M^3)} |F_M'(x)|^2\, dx$$

$$\leq \frac{15(M+1)}{M(4 + 6M + 4M^2 + M^3)} \int_{\delta \leq |x| \leq \pi} (Cx^{-2})^2 dx$$

$$\leq \frac{15(M+1)}{M(4 + 6M + 4M^2 + M^3)} C^2 \int_{\delta \leq |x| \leq \pi} x^{-4} dx \sim \frac{15(M+1)}{M(4 + 6M + 4M^2 + M^3)} C^2 \frac{1}{\delta^3},$$

which goes to 0 as $M \to \infty$.

Analogously, we prove that $\{L_M(x)\}_{M=1}^{\infty}$ is a family of good kernels. First, note that $L_M(x) \geq 0$ and $\frac{1}{2\pi}\int_{-\pi}^{\pi} L_M(x)\,dx = 1$. Finally, we have that

$$\int_{\delta\leq|x|\leq\pi} |L_M(x)|\,dx = \int_{\delta\leq|x|\leq\pi} \frac{105(M+1)}{M(-2+23M+40M^2+30M^3+12M^4+2M^5)}(F_M''(x))^2\,dx$$

$$\leq C^2 \frac{105(M+1)}{M(-2+23M+40M^2+30M^3+12M^4+2M^5)}(A(M+1)+B(M+1)^{-1})^2 \int_{\delta\leq|x|\leq\pi} x^{-4}\,dx$$

$$\sim \frac{105(M+1)^3}{M(-2+23M+40M^2+30M^3+12M^4+2M^5)}\frac{C^2}{\delta^3},$$

(for A, B suitable constants), which converges to zero as $M \to \infty$.

Lemma 8.2 *Under the condition $N/n \to a > 0$, the following results hold.*

i) For any $p > 1$, there exists a constant C_p such that

$$\lim_{n,N} n \sup_{x\in[0,2\pi]} \int_0^{2\pi} |D_N(\varphi_n(x)-\varphi_n(y))|^p dy \leq C_p.$$

ii) It holds that

$$\lim_{n,N} n \int_0^x D_N^2(\varphi_n(x)-\varphi_n(y))dy = \pi(1+2\eta(2a))$$

and, for any α-Hölder continuous function f, with $\alpha \in (0,1]$,

$$\lim_{N,n} n \int_0^x D_N^2(\varphi_n(x)-\varphi_n(y))f(y)dy = \pi(1+2\eta(2a))\,f(x), \tag{130}$$

where

$$\eta(a) := \frac{1}{2a^2} r(a)(1-r(a)), \tag{131}$$

being $r(a) = a - [a]$, with $[a]$ the integer part of a.

iii) for any $\varepsilon > 0$,

$$\lim_{N,n} n \int_0^{x-\varepsilon} |D_N(\varphi_n(x)-\varphi_n(y))|^2 dy = 0.$$

Proof. See [Clement and Gloter, 2011], Lemma 1 and Lemma 4.

References

[Ait-Sahalia et al., 2017] Aït-Sahalia, Y., Fan, J., Laeven, R.J.A., Wang, C.D. and Yang, X. (2017) Estimation of the continuous and discontinuous leverage effects *Journal of the American statistical association*, 112(520): 1744-1758.

[Ait-Sahalia et al., 2013] Aït-Sahalia, Y., Fan, J. and Li, Y.(2013) The leverage effect puzzle: disentangling sources of bias at high frequency. *Journal of Financial Economics*, 109:224-249.

[Ait-Sahalia and Jacod, 2014] Aït-Sahalia, Y. and Jacod, J. (2014) High-Frequency financial econometrics. Princeton University Press.

[Ait-Sahalia and Xiu 2014] Aït-Sahalia, Y. and Xiu, D. (2019) A Hausman test for the presence of market microstructure noise in high frequency data. *Journal of Econometrics*, 211: 176-205.

[Andersen et al., 2001] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001) The distribution of realized stock return volatility *Journal of Financial Economics*, 61(1): 43-76.

[Bandi et al., 2020] Bandi, F.M., Fusari, N. and Renò, R. (2020) Structural stochastic volatility *SSRN paper n. 3717015*.

[Barndorff-Nielsen and Veraart, 2009] Barndorff-Nielsen, O.E. and Veraart, A. (2009) Stochastic volatility of volatility in continuous time. CREATES Research Paper.

[Barndorff-Nielsen and Veraart, 2013] Barndorff-Nielsen, O.E. and Veraart, A. (2013) Stochastic volatility of volatility and variance risk premia. *Journal of Financial Econometrics*, 11(1): 1-46.

[Barndorff-Nielsen et al., 2011] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (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.

[Barucci and Mancino, 2010] Barucci, E. and Mancino, M.E. (2010) Computation of volatility in stochastic volatility models with high frequency data. *International Journal of Theoretical and Applied Finance*, 13(5): 1-21.

[Bollerslev et al., 2009] Bollerslev, T., Tauchen, G. and Zhou, H. (2009) Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11): 4463–4492.

[Corsi, 2009] Corsi, F. (2009) A simple approximate long-memory model of realized volatility *Journal of Financial Econometrics*, 7(2): 174-196.

[Clement and Gloter, 2011] Clement, E. and Gloter, A. (2011) Limit theorems in the Fourier transform method for the estimation of multivariate volatility. *Stochastic Processes and their Applications*, 121(5): 1097-1124.

[Cuchiero and Teichmann, 2015] Cuchiero, C. and Teichmann, J. (2015) Fourier transform methods for pathwise covariance estimation in the presence of jumps. *Stochastic Processes and Their Applications*, 125(1): 116-160.

[Heston, 1993] Heston, S. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2): 327-343.

[Jacod, 1997] Jacod, J. (1997) On continuous conditional Gaussian martingales and stable convergence in law. *Seminaire de Probabilites XXXI*, Springer, 232-246.

[Jing et al., 2014] Jing, B.Y., Liu, Z. and Kong, X.B. (2014) On the estimation of integrated volatility with jumps and microstructure noise. *Journal of Business and Economic Statistic*, 32(3): 457-467.

[Kalnina and Xiu, 2017] Kalnina, I. and Xiu, D. (2017) Nonparametric estimation of the leverage effect: a trade-off between robustness and efficiency. *Journal of the American Statistical Association*, 112(517): 384-399.

[Lee and Mykland, 2008] Lee, S.S. and Mykland, P.A. (2008) Jumps in financial markets: a new nonparametric test and jump dynamics *The Review of Financial Studies*, 21(6), 2535-2563.

[Li et al., 2021] Li, Y., Liu, G. and Zhang, Z. (2021) Volatility of volatility: estimation and tests based on noisy high frequency data with jumps. *Journal of Econometrics*, forthcoming.

[Livieri et al., 2019] Livieri, G., Mancino, M.E. and Marmi, S. (2019) Asymptotic results for the Fourier estimator of the integrated quarticity. *Decisions in Economics and Finance*, 42, 472-501.

[Malliavin, 1995] Malliavin, P. (1995). Integration and Probability. Springer Verlag.

[Malliavin and Mancino, 2002] Malliavin, P. and Mancino, M.E. (2002) Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics*, 4, 49-61.

[Malliavin and Mancino, 2009] Malliavin, P. and Mancino, M.E. (2009) A Fourier transform method for nonparametric estimation of multivariate volatility. *Annals of Statistics*, 37(4), 1983-2010.

[Mancino and Sanfelici, 2012] Mancino, M.E. and Sanfelici, S. (2012) Estimation of quarticity with high-frequency data. *Quantitative Finance,* 12(4), 607-622.

[Mancino et al., 2017] Mancino, M.E., Recchioni, M.C. and Sanfelici, S. (2017) Fourier-Malliavin volatility estimation: theory and practice. *Springer.*.

[Mancino and Toscano, 2021] Mancino, M.E. and Toscano, G. (2021) Rate-efficient asymptotyc normality for the Fourier estimator of the leverage process. *Statistics and its Interface (forthcoming).*

[Mykland and Zhang, 2009] Mykland, P.A. and Zhang, L. (2009) Inference for continuous semimartingales observed at high frequency. *Econometrica*, 77, 1403-1445.

[Patton and Engle, 2001] Patton, A.J. and Engle, R. (2001) What is a good volatility model? *Quantitative Finance*, 1(2): 237-245.

[Sanfelici et al., 2015] Sanfelici, S., Curato, I.V. and Mancino, M.E. (2015) High frequency volatility of volatility estimation free from spot volatility estimates. *Quantitative Finance*, 15(8):1-15

[Stein and Shakarchi, 2011] Stein, E.M. and R. Shakarchi, R. (2011) Fourier analysis: an introduction, Volume 1. Princeton University Press.

[Toscano and Recchioni, 2021] Toscano, G. and Recchioni, M.C. (2021) Bias-optimal vol-of-vol estimation: the role of window overlapping. *Decisions in Economics and Finance (forthcoming).*

[Vetter, 2015] Vetter, M. (2015) Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. *Bernoulli*, 21(4), 2393-2418.

[Wang and Mykland, 2014] Wang, C.D. and Mykland, P.A. (2014) The estimation of the leverage effect with high-frequency data *Journal of the American Statistical Association*, 109(505), 197-215.