

RESEARCH ARTICLE

# Fourier Spot Volatility Estimator: Asymptotic Normality and Efficiency with Liquid and Illiquid High-Frequency Data

Maria Elvira Mancino<sup>1</sup>\*, Maria Cristina Recchioni<sup>2</sup>\*

**1** Department of Economics and Management, University of Florence, Florence, Italy, **2** Department of Management, Polytechnical University of Marche, Ancona, Italy

\* These authors contributed equally to this work.

\* [m.c.recchioni@univpm.it](mailto:m.c.recchioni@univpm.it)



OPEN ACCESS

**Citation:** Mancino ME, Recchioni MC (2015) Fourier Spot Volatility Estimator: Asymptotic Normality and Efficiency with Liquid and Illiquid High-Frequency Data. PLoS ONE 10(9): e0139041. doi:10.1371/journal.pone.0139041

**Editor:** Guido Germano, University College London, UNITED KINGDOM

**Received:** June 9, 2015

**Accepted:** September 7, 2015

**Published:** September 30, 2015

**Copyright:** © 2015 Mancino, Recchioni. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data on S&P 500 index (illiquid data) are available from Bloomberg by subscription services at website: <http://www.bloomberg.com/professional/order-now?bbgsum=DG-WS-CORE-orderhpg>, <http://www.bloomberg.com/markets/world>. Data on FIB 30 index futures are available from Mancino, M.E. and Sanfelici, S. (2008). Robustness of the Fourier Estimator of Integrated Volatility in Presence of Microstructure Noise. Computational Statistics and Data Analysis, 52(6), 2966–2989 by contacting the corresponding author of that paper Simona Sanfelici (email: [simona.sanfelici@unipr.it](mailto:simona.sanfelici@unipr.it)).

## Abstract

The recent availability of high frequency data has permitted more efficient ways of computing volatility. However, estimation of volatility from asset price observations is challenging because observed high frequency data are generally affected by noise-microstructure effects. We address this issue by using the Fourier estimator of instantaneous volatility introduced in Malliavin and Mancino 2002. We prove a central limit theorem for this estimator with optimal rate and asymptotic variance. An extensive simulation study shows the accuracy of the spot volatility estimates obtained using the Fourier estimator and its robustness even in the presence of different microstructure noise specifications. An empirical analysis on high frequency data (U.S. S&P500 and FIB 30 indices) illustrates how the Fourier spot volatility estimates can be successfully used to study intraday variations of volatility and to predict intraday Value at Risk.

## 1 Introduction

The relevance of the estimation of time varying volatility in financial economics has been recognized for a long time but the recent availability of high-frequency financial data has given an enormous impulse to its investigation and application. As a matter of fact, estimates of spot volatility can be used to predict intraday Value at Risk (VaR) and to estimate stochastic volatility model parameters which are necessary for forecasting stock and futures prices, bond yields and so on.

Indeed, estimators capable of replicating satisfactorily intra-day volatility time variations may be useful tools for developing indicators of financial fragility and speculative behaviors. A first step towards this objective is the use of the spot volatility estimates to calibrate agent based models, such as the model proposed in Refs. [1–6]. In fact, knowledge of spot volatility allows us to reconstruct the agents' fractions giving insights into the main strategies used by the traders in the financial markets considered as illustrated in Refs. [1, 3, 4, 7]. Furthermore, the estimates of spot volatility can also be applied in medical research such as in cardiac and neuronal signal

**Funding:** The research leading to these results has been partially supported by the European Union, Seventh Framework Programme FP7, under grant agreement FinMaP no. 612955.

**Competing Interests:** The authors have declared that no competing interests exist.

processing. In fact, these estimates could be used to approximate heart rate volatility in order to design an early predictor of malign arrhythmias. In the recent scientific literature, we can find analysis of the short-term (5-minute) heart rate variability (HRV) based on the Fourier and fractal methodology for detecting anomalies in the variations of beat-to-beat interval series [8].

Volatility can be estimated through parametric or nonparametric methods as illustrated in the review [9]. Nonparametric methods address the computation of the historical volatility without assuming a specific functional form of the volatility while trying to reproduce well known stylized facts [7, 10]. Volatility is mainly computed over discrete time intervals relying on the *quadratic variation formula* and thus it is an *integrated volatility*. Instantaneous volatility estimation from high frequency data was first proposed in Ref. [11], by using rolling and block sampling filters. The benchmark for computing the volatility of a financial time series in a time interval with high frequency data (e.g. daily volatility) is provided by the sum of the squared intraday returns (i.e. *realized volatility*) [9, 12, 13]. In the limit, as the time interval between two consecutive observations converges to zero, the realized volatility converges to the quadratic variation of the process and its derivative provides the instantaneous volatility of the process. However, the approximation of the quadratic variation derivative required to get the instantaneous volatility generates appreciable numerical instabilities.

This paper deals with an alternative approach, based on the Fourier series and the Bohr convolution formula. This approach has been introduced in Refs. [14, 15] and it is mainly designed for measuring instantaneous multivariate volatility. The Fourier method reconstructs the instantaneous volatility as a series expansion with coefficients gathered from the Fourier coefficients of the price variation. For this reason it is based on the integration of the time series of returns rather than on its differentiation. Doing so, the Fourier estimator uses all the available observations and avoids any manipulation of the original data and any unstable numerical derivatives.

Several papers have studied the efficiency of the Fourier method in estimating the integrated volatility or co-volatilities even in the presence of microstructure noise, e.g. [16–22]. In fact, by considering the Fourier estimator of 0-th Fourier coefficient of the volatility function we obtain a consistent estimator of the integrated volatility. Moreover, the Fourier methodology has inspired some spot volatility estimators, such as the robust to jumps Fourier estimator in Ref. [23] and the spectral estimator in Ref. [24]. Conversely, the statistical properties and the empirical effectiveness of the Fourier method in estimating the whole path of the volatility process needs to be investigated further. This paper contributes to filling this gap.

Firstly, we prove the pointwise central limit theorem for the spot volatility estimator with a speed of convergence which is the optimal one for a spot volatility estimator. Furthermore, we show how to optimize the asymptotic variance through a suitable choice of ratio of the observation number to the number of the Fourier frequencies. These frequencies are those to be used in the Fejer series to reconstruct the volatility process.

Secondly, we study the efficiency of the Fourier estimator of spot volatility with high frequency data. In particular, we investigate the asymptotic normality of the estimator and the accuracy of the spot volatility estimates using 1-second returns. The accuracy of the estimates is tested comparing the empirical and theoretical distributions of the so called standardized returns [25]. We stress the point that the Fourier estimation method is a *global* method. In fact, it is designed to estimate the volatility path over the entire interval of interest. This fact may be relevant when the estimated volatility is used to calibrate stochastic volatility models. In contrast, most spot volatility estimators, especially the methods based on the quadratic variation formula, are defined as pointwise estimators; thus, their adjustment parameters are tuned to work well only at a specific point in time. Moreover, we show the robustness of the Fourier estimator to various microstructure noise specifications such as the additive noise and the rounding error. We prove that the estimator has a competitive edge even when compared to bias

adjusted estimators, such as the Two-Scales realized spot variance estimator [26, 27]. In fact, the Fourier estimator performs very well using high frequency data in all the different scenarios considered and without requiring any *ad hoc* adjustment.

Finally, we conduct an empirical analysis on high frequency liquid and illiquid data. As stressed in Ref. [25], high quality realized variance estimates can be constructed in a liquid market. Nevertheless, we show that the Fourier estimator provides accurate estimates even when illiquid assets are considered. This is probably due to its robustness to rounding errors. In fact, the price of an illiquid asset behaves like a price affected by rounding errors [28]. In addition, we illustrate some exercises to analyze the intraday volatility variations and to predict intraday Value at Risk.

The paper is organized as follows. Section 2 introduces the Fourier estimator. Section 3 contains the central limit theorem. Section 4 illustrates the finite sample properties of the Fourier estimator, including its robustness to microstructure noise. Section 5 presents an empirical analysis conducted on illiquid and liquid data. Section 6 concludes. Finally, Section 7 contains the proof of the central limit theorem and Section 8 summarizes some properties of Fejer and Dirichlet kernels.

## 2 Fourier Estimator of Spot Volatility

In this section we recall the definition of the Fourier estimator of spot volatility introduced in Ref. [14]. However, before to enter the technical definition of the estimator, we introduce the main argument underlying the Fourier methodology, that is the use of the *convolution formula* Eq (3).

Suppose that the asset log-price  $p(t)$  follows a semi-martingale satisfying the Itô stochastic differential equations

$$dp(t) = \sigma(t) dW(t) + b(t) dt, \tag{1}$$

where  $W$  is a Brownian motion on a filtered probability space satisfying the *usual conditions*, and  $\sigma$  and  $b$  are adapted random processes such that  $E[\int_0^T b^2(t)dt] < \infty$  and  $E[\int_0^T \sigma^4(t)dt] < \infty$ . Our model is very general, in particular it includes a fairly large class of stochastic volatility models which are widely used in finance, e.g. classical models such as [29–31]. In particular, leverage effects are allowed [32]. We stress the point that our approach is non-parametric, therefore we do not specify any functional form of the volatility process  $\sigma$ , we only assume the continuity of the paths (essentially, the Hölder-continuity of any Brownian path).

By change of the origin of time and rescaling the unit of time we can always reduce ourselves to the case where the time window  $[0, T]$  becomes  $[0, 2\pi]$ .

We now define the Fourier estimator of spot volatility introduced in Ref. [14]. For any positive integer  $n$ , let  $0 = t_0 \leq \dots \leq t_n = 2\pi$  be the (possibly unequally-spaced) trading dates of the asset, i.e., the observation times of the asset price. Denote  $\rho(n) := \max_{0 \leq i \leq n-1} |t_{i+1} - t_i|$  and suppose that  $\rho(n) \rightarrow 0$  as  $n \rightarrow \infty$ . Moreover, let  $\delta_i(p) := p(t_{i+1}) - p(t_i)$ .

For any integer  $k, |k| \leq 2N$ , define the discrete Fourier transform

$$c_k(dp_n) := \frac{1}{2\pi} \sum_{i=0}^{n-1} e^{-ikt_i} \delta_i(p), \tag{2}$$

then, for any integer  $k, |k| \leq N$ , consider the following *convolution* formula

$$c_k(\sigma_{n,N}^2) := \frac{2\pi}{2N+1} \sum_{|h| \leq N} c_h(dp_n) c_{k-h}(dp_n). \tag{3}$$

Formula (3) contains the identity relating the Fourier transform of the price process  $p(t)$  to the Fourier transform of the volatility  $\sigma^2(t)$ . By Eq (3) we gather all the Fourier coefficients of the volatility function by means of the Fourier transform of the log-returns. Then, the reconstruction of the volatility function  $\sigma^2(t)$  from its Fourier coefficients, can be obtained by the Fourier-Fejer summation. Finally, the *Fourier estimator of spot volatility* is defined: for any  $t \in (0, 2\pi)$

$$\hat{\sigma}_{n,N,M}^2(t) = \sum_{|k| \leq M} \left(1 - \frac{|k|}{M}\right) c_k(\sigma_{n,N}^2) e^{ik t}. \tag{4}$$

We note that the definition of the estimator  $\hat{\sigma}_{n,N,M}^2(t)$  depends on three parameters, the number of data  $n$  and the two *cutting frequencies*  $N, M$ . The choice of the relative growth of them will be discussed in the following paragraphs.

Note that we can write the estimated Fourier coefficients (3) as

$$c_k(\sigma_{n,N}^2) = \frac{1}{2\pi} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} D_N(t_j - t_i) e^{-ik t_j} \delta_i(p) \delta_j(p),$$

where  $D_N$  is the rescaled Dirichlet kernel defined as

$$D_N(x) = \frac{1}{2N + 1} \sum_{|h| \leq N} e^{ihx} = \frac{1}{2N + 1} \frac{\sin(2N + 1) \frac{x}{2}}{\sin \frac{x}{2}}. \tag{5}$$

Thus, the Fourier estimator of spot volatility (4) can be expressed as follows

$$\hat{\sigma}_{n,N,M}^2(t) = \frac{1}{2\pi} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} F_M(t - t_j) D_N(t_j - t_i) \delta_i(p) \delta_j(p), \tag{6}$$

where  $F_M$  is the Fejer kernel defined as

$$F_M(x) = \sum_{|k| \leq M} \left(1 - \frac{|k|}{M}\right) e^{ikx} = \frac{1}{M + 1} \left(\frac{\sin(M + 1) \frac{x}{2}}{\sin \frac{x}{2}}\right)^2. \tag{7}$$

We stress the point that the estimator (6) contains two terms: the quadratic part

$$\frac{1}{2\pi} \sum_{j=0}^{n-1} F_M(t - t_j) (\delta_j(p))^2 \tag{8}$$

and the cross terms

$$\frac{1}{2\pi} \sum_{i=0}^{n-1} \sum_{\substack{j=0 \\ j \neq i}}^{n-1} F_M(t - t_j) D_N(t_j - t_i) \delta_i(p) \delta_j(p). \tag{9}$$

The quadratic term (8) behaves like the Kernel-based spot volatility estimators seen in Refs. [33, 34]. Nevertheless, the second addend (9) is crucial in terms of robustness of the estimator in the presence of microstructure noise, through the choice of the frequency  $N$ , as it has also been pointed out for the Realised kernels estimator of integrated variance proposed in Ref. [35] and for the Laplace estimator proposed in Ref. [36]. A comparative analysis of the robustness of the Fourier spot volatility estimator with high frequency data is conducted in Section 4.

**Remark 2.1** *Although we have not considered the Fourier estimator of multivariate volatility in this paper, it is worth noting that the convolution formula (3) is directly applicable for*

obtaining the Fourier estimator of the covariance process between two asset price processes as explained in Ref. [15]. Moreover, the advantage of the convolution approach used by the Fourier estimator is that it is immune to the so called Epps effect [37]. In fact, when returns are recorded at the highest available observation frequency, they are asynchronous across different assets [38]. The realized covariance type estimators (e.g. [12, 25, 26, 39]) require choosing a “synchronization” method. Thus these estimators suffer from a downward bias, when the sampling interval is reduced. On the contrary, the Fourier estimator uses all the available observations and avoids any “synchronization” of the original data, because it is based on the integration of the time series of returns rather than on its differentiation.

### 3 Asymptotic Normality

In this section we study the pointwise asymptotic error distribution for the Fourier estimator of spot volatility defined in Eq (6). The central limit theorem assumes the ratio  $\frac{N}{n}$  between the cutting frequency and the number of data to be asymptotically constant (the Nyquist frequency  $N = \frac{n}{2}$  being a reference ratio), while the frequency  $M$  is slower increasing with respect to  $n$ . The speed of convergence is the optimal rate of convergence for a spot volatility estimator, at the cost of a (possible) bigger error variance. The discussion on how to optimize the asymptotic variance through a suitable choice of the ratio  $\frac{N}{n}$  is contained in Remark 3.3.

The limiting error distribution for the integrated multivariate volatility of the Fourier estimator with asynchronous trading in the absence of microstructure noise is studied in Ref. [21] and in the presence of microstructure noise in Ref. [22]. Note that the Fourier estimation of integrated quantities (integrated volatility or covariance) involves only the 0-th Fourier coefficient in the expansion.

We assume that the volatility process  $\sigma$  is a.s. continuous in  $[0, 2\pi]$  (e.g. driven by a second Brownian semimartingale), more precisely Hölder continuous with parameter  $\nu \in (0, \frac{1}{2})$ . For simplicity, we consider equally spaced observations, thus  $\rho(n) = \frac{2\pi}{n}$ .

**Theorem 3.1** Assume that the following conditions hold:  $\lim_{n, M \rightarrow \infty} \frac{M^2}{n} = a > 0$ , for some  $\gamma > 1$ , and  $\lim_{n, N \rightarrow \infty} \frac{N}{n} = c > 0$ . Then, for any fixed  $t \in (0, 2\pi)$ , as  $n, N, M \rightarrow \infty$ ,

$$\sqrt{\frac{n}{M}} \left( \hat{\sigma}_{n, N, M}^2(t) - \sigma^2(t) \right) \rightarrow \mathcal{N} \left( 0, \frac{4}{3} (1 + 2\eta(c)) \sigma^4(t) \right),$$

where the convergence is stable in law and the constant  $\eta(c)$  is defined in Eq (10).

**Remark 3.2** (Rate of convergence) The convergence rate in Theorem 3.1 is of order  $n^{\frac{\gamma-1}{2\gamma}}$ . It appears in the proof that  $1 < \gamma < 2\nu + 1$ , where  $\nu \in (0, \frac{1}{2})$ . For  $\gamma$  close to 2 the rate of convergence becomes  $\frac{1}{4}$ , which is the optimal rate of convergence for a non-parametric spot volatility estimator.

**Remark 3.3** (Optimal variance) The constant  $\eta(c)$  is equal to

$$\eta(c) := \frac{1}{2\tilde{c}^2} r(\tilde{c})(1 - r(\tilde{c})), \tag{10}$$

where  $\tilde{c} = 2c$  and  $r(x) = x - [x]$ , with  $[x]$  the integer part of  $x$ . The computation of the constant  $\eta(c)$  is presented in Lemma 1 formula (26) in Ref. [21]. Note that  $\eta(c)$  is nonnegative for any positive  $c$  and equal to zero when  $c = \frac{1}{2}k, k = 1, 2, \dots$

The case  $\eta(c) = 0$  is interesting since it provides the optimal asymptotic variance  $\frac{4}{3}\sigma^4(t)$  as discussed in Ref. [23]. The optimal asymptotic variance is obtained for  $c = \frac{1}{2}k, k = 1, 2, \dots$  and the choice  $k = 1$  (i.e.  $c = \frac{1}{2}$ ) corresponds to the natural choice of the Nyquist frequency for the Fourier estimator. Furthermore, for empirical purposes, the number of frequencies used in the Fourier

transform is chosen less than the number  $n$  of the available price observations, so that the values  $c = \frac{1}{2}k, k = 2, 3, \dots$  are not effective while the value  $c = \frac{1}{2}$  is appropriate. We stress that with this choice of  $c$  (in other words of  $N/n$ ) the Fourier estimator has the same rate of convergence and asymptotic variance of the Fejer kernel-based realized spot volatility considered in Refs. [34, 40]. Therefore, with an appropriate choice of  $N/n$ , the effect of adding the cross terms in Eq (6), which is essential in order to get an estimator robust to microstructure noise, is also not detrimental in view of the asymptotic efficiency. This particular feature of the Fourier estimator is analyzed in Section 4.

## 4 Simulation studies

In this section we study the efficiency of the Fourier estimator of spot volatility with high frequency data. Firstly, we illustrate some finite sample properties of this estimator. In particular, we investigate the asymptotic normality and the accuracy of the spot volatility estimates using 1-second returns. The accuracy of the estimates is tested by comparing the empirical and theoretical distributions of the so called standardized returns. Secondly, we study the robustness of the Fourier estimator to some microstructure noise specifications, namely, the additive (even dependent) noise and the rounding error. Our analysis shows that it has a competitive edge, even when compared with methods specifically designed to handle market microstructure contaminations.

### 4.1 Finite sample properties

In this section we investigate the finite sample properties of the Fourier spot volatility estimator. Consider the following one factor stochastic volatility model, which is also studied in Ref. [27]:

$$dp(t) = \mu dt + \sigma(t) dW(t), \tag{11}$$

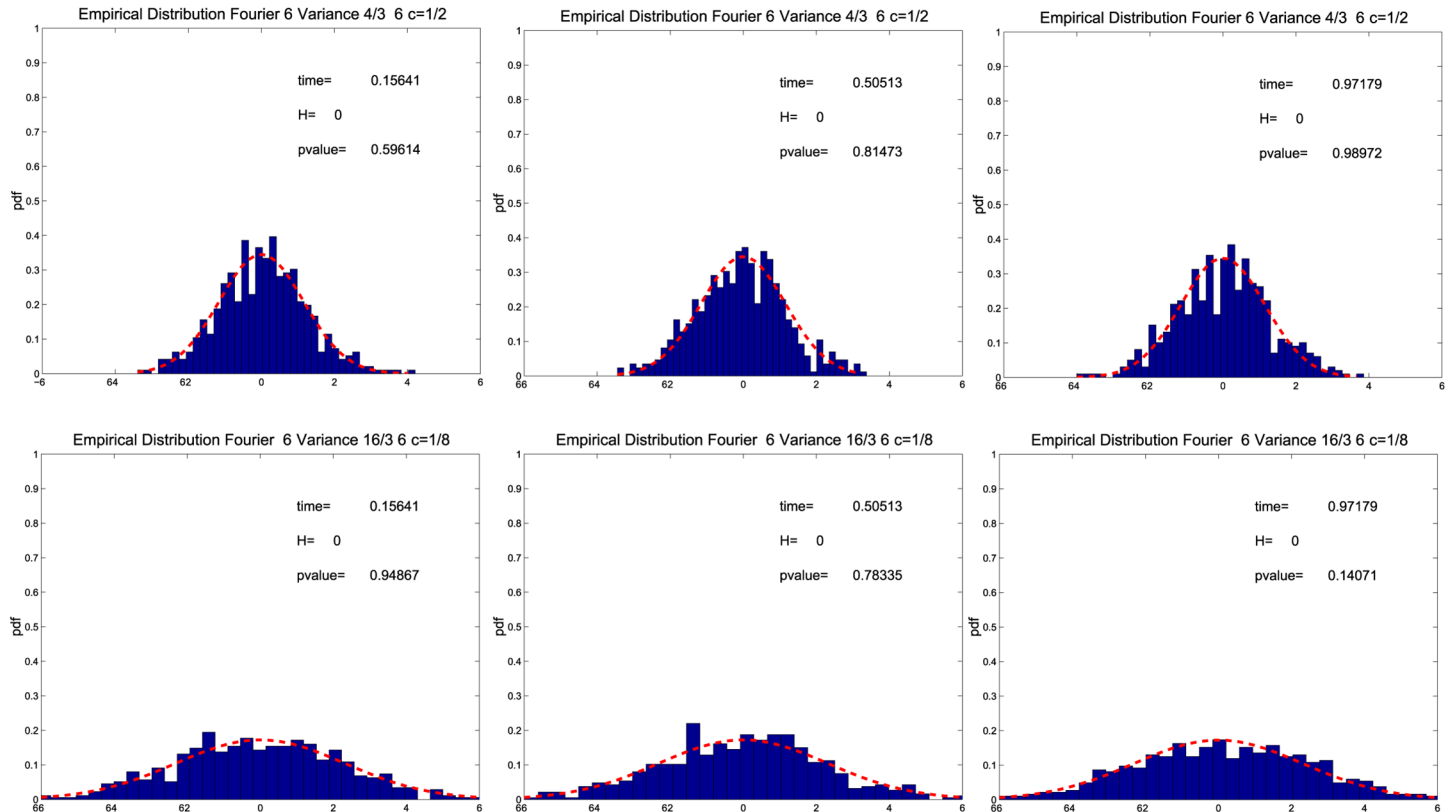
$$\sigma(t) = \exp(\beta_0 + \beta_1 \tau(t)), \tag{12}$$

$$d\tau(t) = \beta_2 \tau(t) dt + dZ(t), \tag{13}$$

where  $W$  and  $Z$  are correlated Brownian motions such that  $\langle dW, dZ \rangle_t = \lambda dt$ . The parameter  $\beta_0$  is chosen equal to  $\beta_1/(2\beta_2)$ . The initial random variable  $\tau_0$  is sampled from the distribution  $\mathcal{N}(0, -1/(2\beta_2))$  and the initial log-price is  $p_0 = \log(9)$ . The values of the parameters are as follows:  $\mu = 0.03, \beta_1 = 0.125, \beta_2 = -0.025, \lambda = -0.3$ . The second-by-second return and variance paths over a daily trading period of  $T = 6.5 \text{ hours} = 1 \text{ day}$  are computed using the explicit Euler discretization scheme with variable step-size. We simulate 504 trading days with  $n = 23400$  observations per day.

Firstly, we study the asymptotic error distribution of the Fourier estimator  $\hat{\sigma}^2(t)$ . Fig 1 shows the empirical distribution of  $\sqrt{n/M}(\hat{\sigma}^2(t) - \sigma^2(t))/\sigma^2(t)$ , when  $n = 23400, \rho(n) = T/n, T = 1 \text{ day}, M = \frac{1}{2\pi} \frac{1}{8} (\sqrt{n} \log n), N = c n$ . The bandwidth  $M$  is chosen in order to fulfill the requirement  $M^\gamma/n = O(1)$  with  $1 < \gamma < 2$ . Our choice is the one used in Ref. [33] to construct realised kernel spot volatility estimators. We consider  $c = 1/2$  (i.e.  $2c = 1$ ) (Fig 1  $t \approx 0.15$  (left upper panel),  $t \approx 0.5$  (middle upper panel),  $t \approx 0.9$  (right upper panel)) and  $c = 1/8$  (i.e.  $2c = 1/4$ ) (Fig 1  $t \approx 0.15$  (left lower panel),  $t \approx 0.5$  (middle lower panel),  $t \approx 0.9$  (right lower panel)). In each panel we also show the probability density function  $\mathcal{N}(0, \frac{4}{3}(1 + 2\eta(c)))$  prescribed by Theorem 3.1, where  $\eta(c)$  is defined in Eq (10). The empirical distributions shown in Fig 1 confirm the findings of Theorem 3.1, namely, the result that the asymptotic distribution of



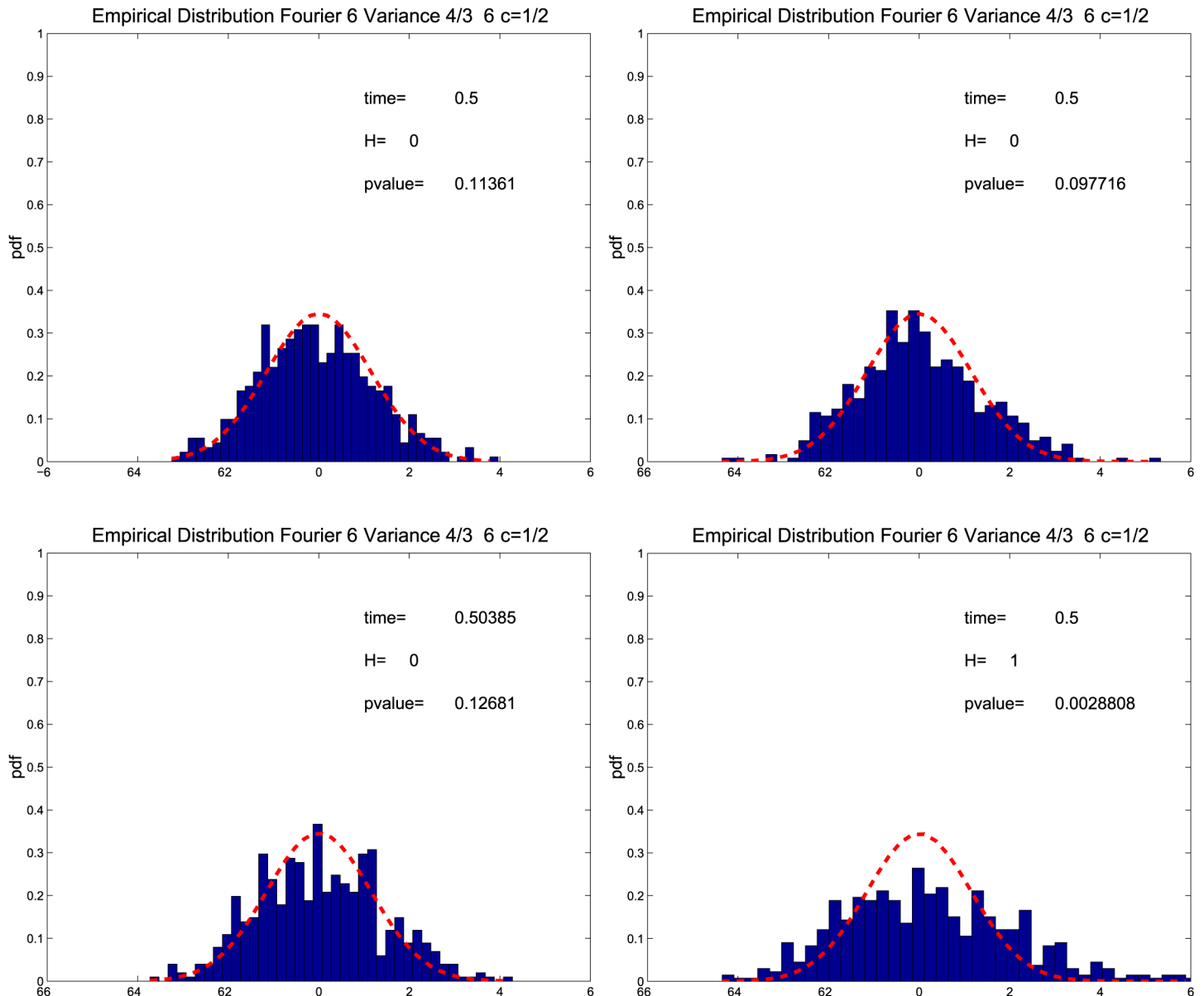


**Fig 1. Empirical distribution of the normalized asymptotic error.** The panels show the empirical distribution of  $\sqrt{n/M}(\hat{\sigma}^2(t) - \sigma^2(t))/\sigma^2(t)$  for  $t \approx 0.15$  left panels,  $t \approx 0.5$  middle panels,  $t \approx 0.9$  right panels and  $c = 1/2$  upper panels,  $c = 1/8$  lower panels with 1-second returns.

doi:10.1371/journal.pone.0139041.g001

$\sqrt{n/M}(\hat{\sigma}^2(t) - \sigma^2(t))/\sigma^2(t)$  is  $\mathcal{N}(0, \frac{4}{3}m)$ , when  $c = 1/(2m)$ ,  $m = 1, 2, \dots$ . In fact, when  $c = 1/2$  the variance attains its smallest value (i.e.  $4/3$ ), while the variance is larger when  $c = 1/8$ . The empirical distributions shown in each panel are tested for normality using the Bera-Jarque test at the significance level of 0.05. The test shows that the null hypothesis is not rejected and the p-values are shown in Fig 1.

We also try to determine the largest price frequency still capable of fitting satisfactorily the theoretical distribution prescribed by Theorem 3.1. We repeat the previous experiment choosing the Nyquist frequency  $N = n/2$  and  $M = \frac{1}{2\pi} \frac{1}{8} (\sqrt{n} \log n)$  which are shown to reduce the variance and a sampling interval of ten seconds (i.e.  $n = 23400/10$ ), thirty seconds (i.e.  $n = 23400/30$ ), one minute (i.e.  $n = 23400/60$ ) and five minutes (i.e.  $n = 23400/300$ ). We evaluate the empirical distribution at  $t \approx 0.5$  and we apply the Bera-Jarque test at the significance level of 0.05. The empirical and theoretical distributions are shown in Fig 2. We can observe that when the sampling interval increases from ten seconds to one minute (upper panels and left lower panel) the p-values remain substantially constant while the p-value of the five minute sample deteriorates and the null hypothesis is rejected. This suggests that the finite sample is able to reproduce theoretical properties of  $\sqrt{n/M}(\hat{\sigma}^2(t) - \sigma^2(t))/\sigma^2(t)$  when the price observations are more than one per minute.

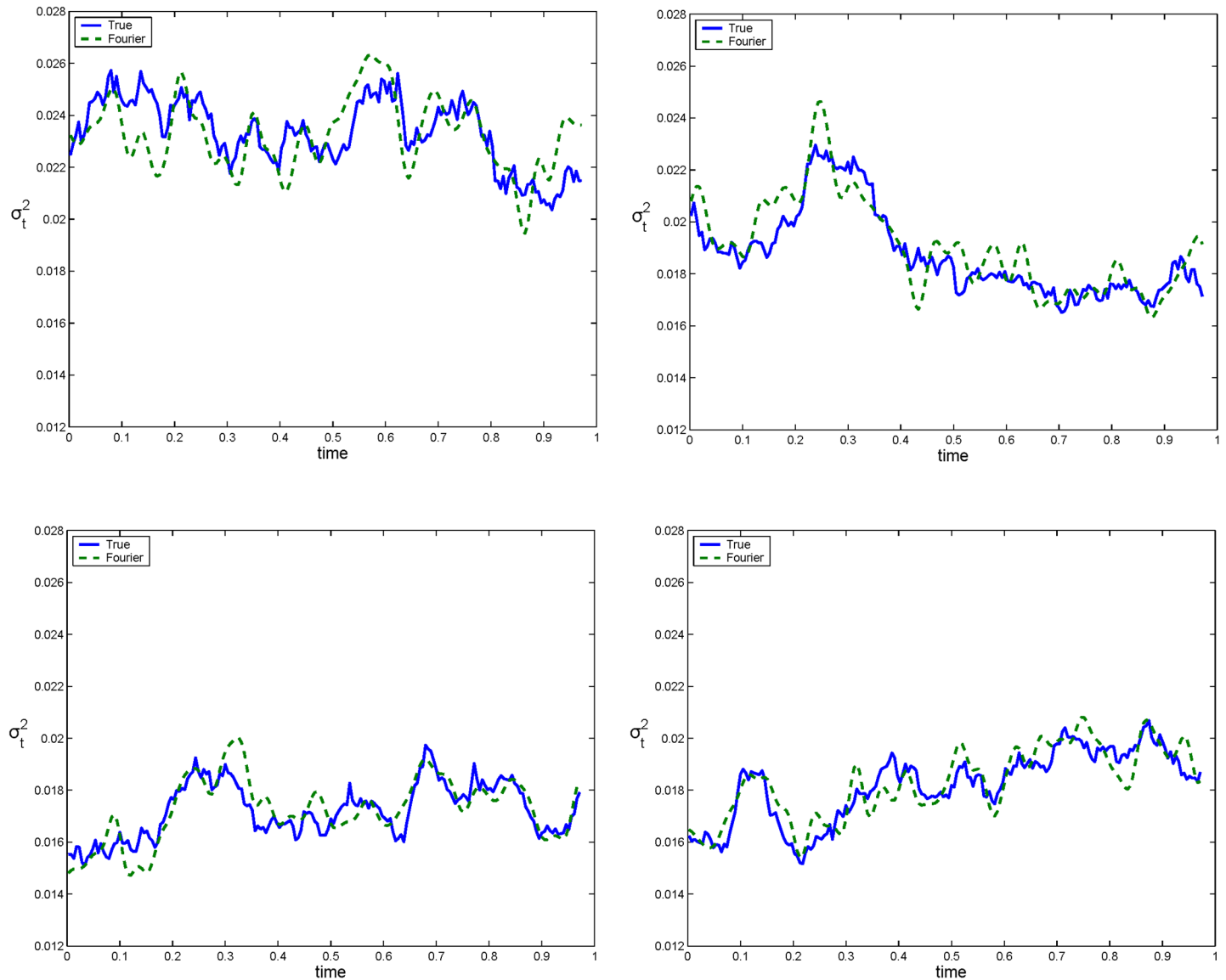


**Fig 2. Empirical distribution of the normalized asymptotic error for various sampling intervals.** The panels show the empirical distribution of  $\sqrt{n}/M(\hat{\sigma}^2(t) - \sigma^2(t))/\sigma^2(t)$  for  $t \approx 0.5$  and  $c = 1/2$  using 1-second returns (left upper panel), 30-second returns (right upper panel), 1-minute returns (left lower panel) 5-minute returns (right lower panel).

doi:10.1371/journal.pone.0139041.g002

Secondly, we prove that the Fourier estimator provides accurate spot volatility estimates. We use the simulated data  $p(t_i), t_i = i/n, i = 0, 1, \dots, n, n = 23400$ , to estimate the variance  $\sigma^2(t)$  on the time grid  $t_j^v = 0.5(2j - 1)\Delta t, j = 1, 2, \dots, 23400/120, \Delta t = 120/23400 = 1/195$ . That is, we estimate the spot variance using a sampling interval of two minutes. In this exercise the log-prices,  $p(t_i)$  are not affected by microstructure noise. Fig 3 shows four realizations of the true variance (solid line) and the corresponding estimates (dotted line) obtained with the Fourier estimator with  $N = n/2$  and  $M = \frac{1}{2\pi} \frac{1}{8} \sqrt{n} \log n$ . It is worth noting that Fig 3 shows that the Fourier estimator





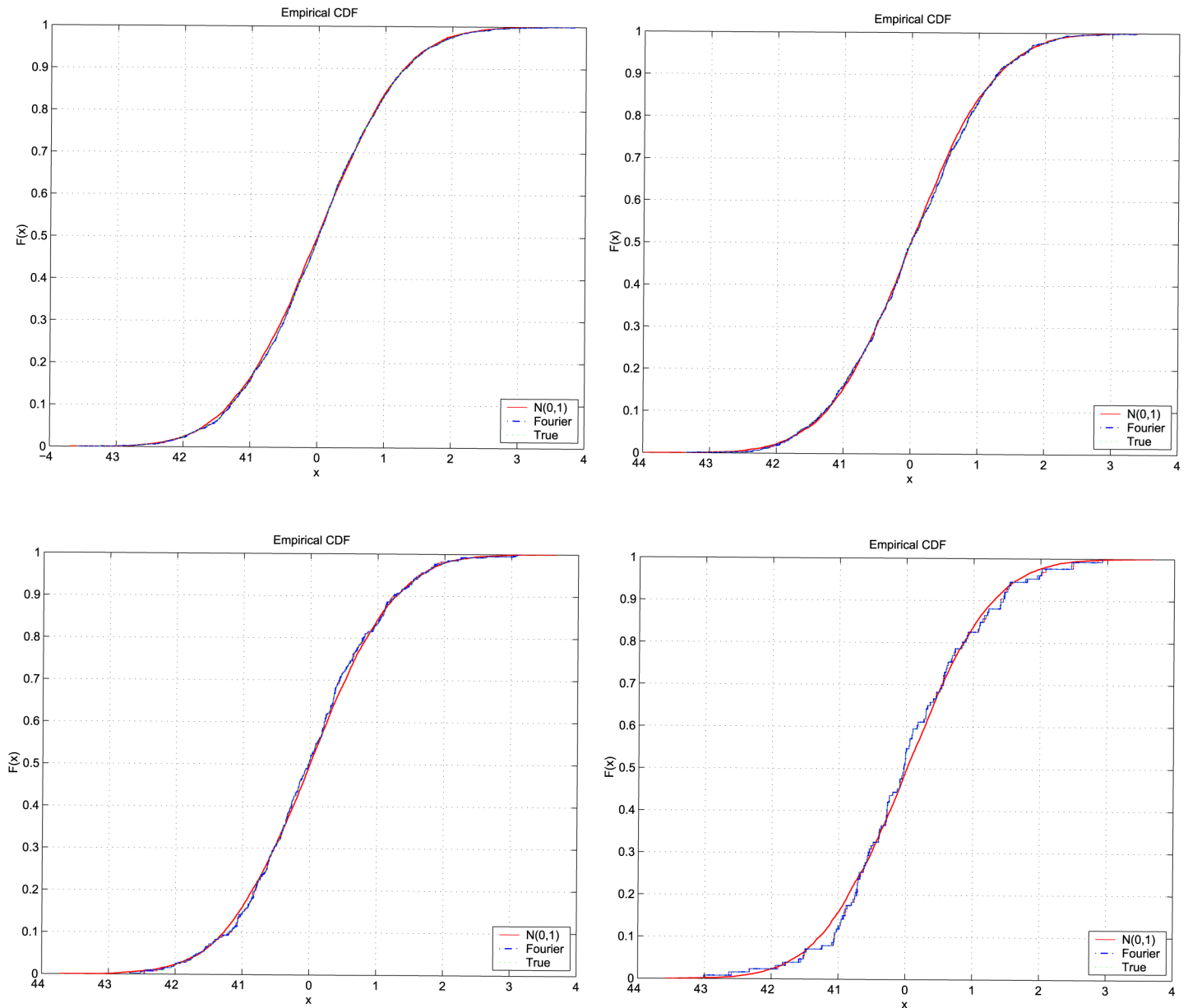
**Fig 3. True and estimated variance path.** The four graphs show the true variance,  $\sigma^2(t)$ , (solid line) and the Fourier estimated variance,  $\hat{\sigma}^2(t)$ , (dotted line) as a function of time for four realizations obtained with model (11)–(12) and a sampling interval of 2 minutes.

doi:10.1371/journal.pone.0139041.g003

approximates the true variance with a satisfactory accuracy over the entire interval. This property is a consequence of the fact that the Fourier method generates a *global* estimator. This accuracy is relevant for calibrating parametric models such as those of [29–31]. In fact, as shown in the empirical analysis, we can use the spot volatility estimates to reconstruct accurate standardized returns which could be used to efficiently estimate model parameters.

We further investigate the accuracy of the Fourier spot volatility estimates using the standardized returns defined by:

$$z_t = \frac{r_t}{\sigma_t \sqrt{\Delta t}}, \tag{14}$$



**Fig 4. Comparison of cumulative density functions.** Cumulative density functions of standard normal sample (red solid line), of the standardized returns obtained using the true volatility (green dotted line) and of the standardized returns obtained using the Fourier spot volatility estimates (blue dash-dot line) when  $\Delta t = 10$  seconds (left upper panel),  $\Delta t = 30$  seconds (right upper panel),  $\Delta t = 1$  minute (left lower panel) and  $\Delta t = 3$  minutes (right lower panel).

doi:10.1371/journal.pone.0139041.g004

where  $r_t = p_{t+\Delta t} - p_t$  is the log-return. These standardized returns are random variables normally distributed with zero mean and variance equal to one when the sample interval is sufficiently small. Specifically, we compute the standardized returns  $z_t^{\text{true}}$  and  $z_t^{\text{FR}}$  obtained using the true and the Fourier spot volatilities, respectively, and we compare their cumulative density functions with the theoretical one,  $\mathcal{N}(0,1)$ . Fig 4 shows the results of this comparison when  $\Delta t = 10$  seconds (Fig 4 left upper panel),  $\Delta t = 30$  seconds (Fig 4 right upper panel),  $\Delta t = 1$  minute (Fig 4 left lower panel) and  $\Delta t = 3$  minutes (Fig 4 right lower panel). The cumulative density functions are obtained elaborating the data of one realization of the log-price variable

**Table 1. Comparison of true and estimated standardized returns.** We compare the true standardized returns  $z_t^{\text{true}}$  and the estimated standardized returns  $z_t^{\text{FR}}$  using Kolmogorov-Smirnov and Jarque-Bera tests.

True standardized returns $z_t^{\text{true}}$				
$\Delta t$	KS rejections (%)	KS-pvalue	JB rejections (%)	JB-pvalue
10 secs	4%	0.47	6%	0.50
30 secs	3%	0.43	5%	0.48
1 min	8%	0.43	8%	0.48
3 mins	9%	0.46	6%	0.51
Fourier standardized returns $z_t^{\text{FR}}$				
$\Delta t$	KS rejections (%)	KS-pvalue	JB rejections (%)	JB-pvalue
10 secs	4%	0.47	5%	0.50
30 secs	4%	0.43	5%	0.48
1 min	8%	0.43	8%	0.48
3 mins	10%	0.46	6%	0.51

doi:10.1371/journal.pone.0139041.t001

observed in one day. In addition, we measure the performance of the Fourier estimator by generating 504 replications of the standardized return  $z_t^{\text{true}}$  and  $z_t^{\text{FR}}$  for various values of the sampling interval. Then we use the Kolmogorov-Smirnov (KS) and the Jarque-Bera (JB) test at the 5% significance level to determine whether the 504 random samples have the hypothesized standard normal cumulative density function.

Table 1 shows the results of the tests for the sampling intervals  $\Delta t = 10$  seconds, 30 seconds, 1 minute and 3 minutes. The upper panel in Table 1 shows the sampling interval  $\Delta t$ , the percentage of the KS test rejections, the average p-values, the percentage of the JB test rejections and the average p-values obtained using the true standardized returns  $z_t^{\text{true}}$ , while the lower panel shows the same quantities but for the estimated standardized returns  $z_t^{\text{FR}}$ . We can see that as we increase  $\Delta t$  the percentage of the test rejections increases and the fitting of the cumulative density functions deteriorates (see Fig 4). The poor fit observed when  $\Delta t = 3$  minutes is due to the fact that the assumption of standardized returns drawn by a standard normal distribution holds only for sufficiently small sampling intervals. When we correct the standardized returns by replacing  $r_t$  with  $r_t - 0.03 \Delta t$  in Eq (14), the percentage of the test rejections of the KS test and the average p-values are 4% and 0.47 respectively for both samples. For the JB test, the percentage of rejections and the average p-values are 3% and 0.51 respectively for both samples. That is, the percentage of rejections of the two tests is more than halved.

We highlight that the standardized returns obtained using the true and estimated volatility behave similarly for the four sampling rates considered. This finding indicates that the Fourier estimator is able to reproduce the statistical features of the true spot volatility.

#### 4.2 Robustness to microstructure noise effects

In this subsection we illustrate the robustness of the Fourier estimator to various microstructure noise specifications. The simulation study confirms the *global* character of the Fourier estimator. Indeed, it is possible to choose the frequencies  $N$  and  $M$  independently of time in order to get accurate spot volatility estimates over the time interval  $t \in (0, T)$ . Furthermore, even if we choose  $M$  and  $N$  independent of the noise specification we still get satisfactory results.

We compare the performance of the Fourier estimator with that of two alternative spot volatility estimators: the Fejer kernel-based realized estimator [33, 34, 40] and the Two-Scales realized spot variance estimator [26, 27]. We choose the Fejer kernel-based realized estimator because it coincides with the *quadratic part* of the Fourier estimator and is given by Eq (8). We

show that the Fourier estimator behaves substantially different from the Fejer kernel-based realized estimator, as the presence of the cross term (9) is essential to rendering the estimator efficient in the presence of microstructure noise. The numerical simulations show that the Fourier estimator is robust to microstructure noise effects while the Fejer kernel-based realized estimator is highly biased in the presence of noise. We choose the Two-Scales realized spot variance estimator because it is a suitable competitor to the Fourier estimator when noisy data are considered. In fact, the Two-Scales estimator is constructed by the localization of an integrated volatility estimator specifically designed to be robust to microstructure noise [26].

For the reader’s convenience we recall the definition of these two estimators. The Fejer kernel-based realized estimator is the quadratic part of the Fourier estimator given in Eq (8). We will denote it as  $\hat{\sigma}_{KF}^2(t)$ . The Two-Scales realized spot variance estimator (TS hereafter) has the following form:

$$\hat{\sigma}_{TS}^2(t) := \frac{1}{h} \sum_{t-h \leq t_i \leq t} \frac{(p(t_i) - p(t_{i-R}))^2}{R} - \left( \frac{nh - R + 1}{nRh} \right) \frac{1}{h} \sum_{t-h \leq t_i \leq t} (p(t_i) - p(t_{i-1}))^2. \tag{15}$$

The parameter  $R$  is the sub-sampling size parameter (*scale parameter*) and  $h$  is the interval length parameter (*bandwidth parameter*) [27]. The parameters  $R$  and  $h$  play a relevant role as the parameters  $N$  and  $1/M$ , which appear in the Fourier estimator’s definition. We rescale the Fourier and Fejer kernel-based realized estimators by the quantity  $\sum_{i=1}^n F_M(t - t_i)(t_i - t_{i-1})$ , where  $F_M$  is the Fejer kernel. This scaling improves the performance of both estimators without changing their asymptotic properties [34, 40].

Let us now describe the data-set considered in this subsection. The log-prices,  $p(t_i)$ ,  $i = 0, 1, \dots, n$ , are generated simulating the following stochastic volatility model:

$$dp(t) = (\mu - \sigma(t)^2/2)dt + \sigma(t) dW(t), \tag{16}$$

$$d\sigma^2(t) = \gamma(\theta - \sigma^2(t))dt + \nu\sigma(t) dZ(t), \tag{17}$$

where  $W(t)$  and  $Z(t)$  are standard Brownian motions with correlation  $\lambda$  (i.e.  $\langle dW, dZ \rangle_t = \lambda dt$ ). This model (with jumps) has been simulated in Ref. [28] to study the effect of rounding errors on integrated volatility estimators. Following [28], we choose  $T = 1$ ,  $\nu = 0.5/252$ ,  $\gamma = 5/252$ ,  $\theta = 0.1$ ,  $\mu = 0.05/252$ ,  $\lambda = -0.5$ ,  $\sigma_0^2 = 1$ ,  $p(0) = \log(9)$ .

We now introduce the four models of microstructure noise considered in the simulation study. The first three kinds consist of additive noises experienced in the financial prices and already extensively analyzed in the literature ([35, 41–43] and the reference therein). The last one is the rounding error [28, 44]. The rounding error is very critical when high frequency data are used. In fact, the financial prices are often rounded at 1% and this can result in a “piecewise constant” time series. In the first three specifications we suppose that the logarithm of the observed price  $\tilde{p}(t_i)$  is given by:

$$\tilde{p}(t_i) = p(t_i) + \eta(t_i), \quad i = 0, \dots, n, \tag{18}$$

where  $p$  is the efficient (latent) log-price process defined by Eq (16) and  $\eta$  describes the microstructure noise component.

The first noise specification, (M), is based on the following assumptions:

(M.I) the random shocks  $\eta(t_i)$  for any  $i = 0, 1, \dots, n$  are independent and identically distributed with Gaussian distribution  $\mathcal{N}(0, \tilde{\eta}^2)$

(M.II) the true return process  $\delta_i(p)$  is independent of  $\eta(t_i)$  for any  $i = 0, 1, \dots, n - 1$  and for any  $n$ .

The second specification, (M)', assumes (M.I)' and (M.II)', where (M.I)' the random shocks  $\eta(t_i)$  for any  $i = 0, 1, \dots, n$  are allowed for negative first order autocorrelation of the random shocks.

The third noise specification, (MD), is based on (M.I) and (M.II)', where (M.II)' the random shocks  $\eta(t_i)$  are linearly dependent on the return  $\delta_i(p)$ , namely:

$$\eta(t_i) = \zeta \delta_i(p) + \hat{\eta}_i,$$

where  $\hat{\eta}_i$  are Gaussian i.i.d. random variables with variance equal to  $\tilde{\eta}^2$  for any  $i$  and  $\zeta = 0.1$ . When we consider these three noise specifications, we choose  $\tilde{\eta} = \zeta \text{std}(r)$  where  $\text{std}(r)$  is the standard deviation of the 1-second returns. The quantity  $\zeta$  is the so called noise-to-signal ratio (see Ref. [42] for further details).

The last noise specification, (MR), takes into account the fact that prices involve rounding errors. The observed log-price are defined as follows:

$$\tilde{p}(t_i) = \log \left( \left[ \frac{\exp(p(t_i))}{l_\alpha} \right] l_\alpha \right), \quad i = 0, \dots, n, \tag{19}$$

where  $[\cdot]$  denotes rounding to the nearest integer and  $l_\alpha$  is the fixed rounding error level. As highlighted in Ref. [28], given that stock prices are often rounded to the cent, the choice  $l_\alpha = 0.01$  mimics the financial markets.

We apply the explicit Euler discretization scheme with variable step-size to compute second-by-second return and variance paths over a daily trading period of  $T = 1$  day. We simulate a total of 504 trading days (about two years) and  $n = 23400$  observations per day (approximately one observation per second). The volatility is estimated using a sampling interval of 1 minute since this sampling has shown a good accuracy of the standardized returns (see Fig 4 in Section 4.1).

We measure the performance of the spot volatility estimator,  $\hat{\sigma}^2(t)$ , over the entire interval  $[0, T]$  and near the right boundary by evaluating numerically the relative mean squared error  $\text{RMSE}(t) = E[(\hat{\sigma}^2(t) - \sigma^2(t))^2 / \sigma^4(t)]$  and the bias  $\text{BIAS}(t) = E[\hat{\sigma}^2(t) - \sigma^2(t)]$ . Specifically, the performance over the interval  $[0, T]$  is evaluated using the integrated relative mean squared error

$$\text{IRMSE} = \frac{1}{T} \int_0^T E[(\hat{\sigma}^2(t) - \sigma^2(t))^2 / \sigma^4(t)] dt$$

and the integrated bias

$$\text{IBIAS} = \frac{1}{T} \int_0^T E[\hat{\sigma}^2(t) - \sigma^2(t)] dt.$$

We investigate whether there exists an easily implementable formula for choosing the so called *cutting frequencies*  $N$  and  $M$  of the Fourier estimator and the bandwidths  $R$  and  $h$  of the TS estimator in order to minimize the integrated relative mean squared error and the relative error at  $t = T_b = 0.98$ . More specifically, we explore how the optimal choice of  $N, M, R$  and  $h$  depends on  $n$ . Consider the following values of  $N$  and  $M$ :

$$N = c_N n^\alpha, \quad \alpha = 1, \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \tag{20}$$

$$M = c_M n^\beta, \quad \beta = \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \tag{21}$$

where  $c_N$  and  $c_M$  are positive constants. We choose  $c_N = \frac{1}{2}$  and  $c_M = \frac{1}{2\pi} \frac{1}{8}$  for the Fourier estimator and  $c_M = \frac{1}{2\pi} \frac{1}{8}$  for the Fejer kernel-based realized estimator which depends only on the parameter  $M$ . We note that  $\beta := 1/\gamma$ , where  $\gamma$  is the parameter used in Theorem 3.1 to parameterize the cutting frequency  $M$ . Theorem 3.1 requires  $\frac{1}{2} < \beta < 1$ . In our exercise we also explore the behavior of the Fourier estimator for values of  $\beta$  not within the interval prescribed by the central limit theorem. This is theoretically justified by the fact that the highest frequencies should be cut in order to filter out microstructure noise effects arising from high frequency data as shown in Ref. [20]. We check whether 1-second observed data are able to match the asymptotic behavior.

The parameters  $R$  and  $h$  of the TS estimator are chosen in a slightly different way by taking into consideration the plug-in approach illustrated in Ref. [27]. More precisely, we choose the following values of  $R$  and  $h$ :

$$R = c_R n^\alpha, \quad \alpha = \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \quad c_R = 0.05(\xi + 1)/i, \quad i = 1, 2, \dots, 10, \quad (22)$$

$$h = c_h/n^\beta, \quad \beta = \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \quad c_h = 0.5 i/(\xi + 1), \quad i = 1, 2, \dots, 10, \quad (23)$$

where  $\xi$  is the noise to signal ratio. Note that, for a fixed value of  $n$ , we consider forty values of  $R$  and sixty values of  $h$ . This is done since we want to explore not only the dependence of  $R$  and  $h$  on  $n$  but also on the noise level  $\xi$ . For this reason we allow the constants  $c_R$  and  $c_h$  to vary with the noise level. Tuning the constants  $c_R$  and  $c_h$  on the noise level strongly improves the performance of the TS estimator, while this is not needed in the case of the Fourier estimator, which turns out to have a good performance even without such *ad hoc* choices of the constants  $c_N$  and  $c_M$ .

Note that we are looking for a bandwidth/frequency choice that can provide satisfactory volatility approximations over the entire time interval without being dependent on time and on specific properties of the volatility process. This fact may be relevant when the estimated volatility is used to calibrate stochastic volatility models. In fact, given the sample, if the value of  $M$  and  $N$  are independent of time (and, possibly, of noise) the estimator  $\hat{\sigma}^2(t)$  is a continuous function of  $t$  and this makes it suitable for the calibration of stochastic volatility models.

We examine the performance of the Fourier estimator, the Fejer-kernel based estimator and the TS estimator both in the absence ( $\xi = 0$ ) and in the presence of noise. In the simulations conducted by using the first three noise specifications, we consider two values of the noise-to-signal ratio,  $\xi = 0.8$  and  $\xi = 3.2$ . That is, we consider a total of six different noises for the additive case. Regarding the fourth specification (i.e. rounded prices) we consider only the rounding levels  $l_\alpha = 0.01$  (commonly observable in financial prices) and  $l_\alpha = 0.1$ . Tables 2, 3 and 4 show, from left to right, the noise to signal ratio, the pair  $(\alpha^i, \beta^i)$  which minimizes the integrated mean squared error, the integrated bias, the pair  $(\alpha^b, \beta^b)$  which minimizes the mean square error and the bias at  $t = T_b$ .

We can see that both the Fourier and the TS estimators are very robust to any noise specifications considered. The TS estimator performs slightly worse than the Fourier one in approximating the volatility over the entire interval, while the TS volatility estimates near the boundary are slightly more accurate than the Fourier ones. This is probably due to the fact that the bandwidths,  $R$  and  $h$ , are chosen independently of some specific properties of the volatility process (e.g., integrated quarticity, integrated volatility of volatility, as in Ref. [27]). In other words, the Fourier estimator fits the term *global* slightly better than the TS estimator.

**Table 2. Robustness of the Fourier estimator under microstructure noise effects.** The sampling interval is 1 second.

Fourier estimator						
No Noise						
Noise-to-signal ratio $\xi$	$(\alpha^i, \beta^i)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.0 (no noise)	$(1, \frac{2}{3})$	3.08e-4	-8.64e-4	$(1, \frac{2}{3})$	3.32e-4	-7.31e-3
Noise specification (M.I)-(M.II)						
Noise-to-signal ratio $\xi$	$(\alpha^i, \beta^i)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{2}{3})$	8.58e-3	4.38e-3	$(\frac{2}{3}, \frac{2}{3})$	8.64e-3	1.87e-2
3.2	$(\frac{2}{3}, \frac{2}{3})$	2.06e-2	4.86e-2	$(\frac{2}{3}, \frac{2}{3})$	1.84e-2	6.55e-2
Noise specification (M.I)'-(M.II)						
Noise-to-signal ratio $\xi$	$(\alpha^i, \beta^i)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{2}{3})$	8.48e-3	2.22e-3	$(\frac{2}{3}, \frac{2}{3})$	8.43e-3	1.64e-2
3.2	$(\frac{2}{3}, \frac{2}{3})$	1.83e-2	1.67e-2	$(\frac{2}{3}, \frac{2}{3})$	1.38e-2	3.38e-2
Noise specification (M.I)-(M.II)'						
Noise-to-signal ratio $\xi$	$(\alpha^i, \beta^i)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{2}{3})$	8.59e-3	5.01e-3	$(\frac{2}{3}, \frac{2}{3})$	8.61e-3	1.95e-2
3.2	$(\frac{2}{3}, \frac{2}{3})$	1.98e-2	4.98e-2	$(\frac{2}{3}, \frac{2}{3})$	1.73e-2	6.81e-2
Noise specification (MR)						
Rounding level $l_\alpha$	$(\alpha^i, \beta^i)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.01	$(\frac{2}{3}, \frac{2}{3})$	8.36e-3	3.15e-3	$(\frac{2}{3}, \frac{2}{3})$	8.51e-3	1.75e-2
0.1	$(\frac{1}{2}, \frac{1}{2})$	4.99e-2	3.36e-2	$(\frac{1}{2}, \frac{1}{2})$	1.55e-2	4.25e-2

doi:10.1371/journal.pone.0139041.t002

**Table 3. Robustness of the Fejer-kernel based realized estimator to microstructure noise effects.** The sampling interval is 1 second.

Fejer-kernel based realized estimator						
No Noise						
Noise-to-signal ratio $\xi$	$\beta^i$	IRMSE	IBIAS	$\beta^b$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.0 (no noise)	$\frac{1}{2}$	2.19e-4	-1.01e-3	$\frac{1}{2}$	2.64e-4	7.49e-3
Noise specification (M.I)-(M.II)						
Noise-to-signal ratio $\xi$	$\beta^i$	IRMSE	IBIAS	$\beta^b$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$\frac{1}{6}$	1.01e+1	3.16e+0	$\frac{1}{6}$	1.01e+1	3.16e+0
3.2	$\frac{1}{6}$	2.29e+3	4.73e+1	$\frac{2}{3}$	2.23e+3	4.72e+1
Noise specification (M.I)'-(M.II)						
Noise-to-signal ratio $\xi$	$\beta^i$	IRMSE	IBIAS	$\beta^b$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$\frac{1}{6}$	4.90e+0	2.21e+0	$\frac{1}{2}$	4.89e+0	2.12e+0
3.2	$\frac{1}{6}$	1.25e+3	3.54e+1	$\frac{1}{2}$	1.25e+3	3.53e+1
Noise specification (M.I)-(M.II)'						
Noise-to-signal ratio $\xi$	$\beta^i$	IRMSE	IBIAS	$\beta^b$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$\frac{1}{3}$	2.19e+0	1.48e+0	$\frac{1}{3}$	2.19e+0	1.48e+0
3.2	$\frac{1}{6}$	4.18e+3	2.04e+1	$\frac{1}{2}$	4.18e+3	2.04e+1
Noise specification (MR)						
Rounding level $l_\alpha$	$\beta^i$	IRMSE	IBIAS	$\beta^b$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.01	$\frac{1}{2}$	9.59e-2	2.58e-1	$\frac{1}{2}$	2.27e-1	3.15e-1
0.1	$\frac{1}{2}$	1.58e+1	2.17e+0	$\frac{1}{2}$	4.75e+1	3.57e+0

doi:10.1371/journal.pone.0139041.t003



**Table 4. Robustness of the Two-Scales realized spot variance estimator to microstructure noise effects.** The sampling interval is 1 second.

Two Scales Realized Spot variance estimator						
No Noise						
Noise-to-signal ratio $\xi$	$(\alpha^j, \beta^j)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.0 (no noise)	$(\frac{2}{3}, \frac{1}{6})$	4.73e-2	-9.60e-2	$(\frac{2}{3}, \frac{1}{6})$	3.92e-3	-4.50e-2
Noise specification (M.I)-(M.II)						
Noise-to-signal ratio $\xi$	$(\alpha^j, \beta^j)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{1}{6})$	3.87e-2	-9.77e-2	$(\frac{2}{3}, \frac{1}{6})$	1.31e-2	-6.48e-2
3.2	$(\frac{2}{3}, \frac{1}{6})$	5.46e-2	-6.54e-2	$(\frac{2}{3}, \frac{1}{6})$	1.41e-2	-1.71e-2
Noise specification (M.I)'-(M.II)						
Noise-to-signal ratio $\xi$	$(\alpha^j, \beta^j)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{1}{6})$	4.66e-2	-1.16e-1	$(\frac{2}{3}, \frac{1}{6})$	7.86e-3	-3.79e-2
3.2	$(\frac{2}{3}, \frac{1}{6})$	1.14e-1	-1.73e-1	$(\frac{2}{3}, \frac{1}{6})$	4.39e-2	-1.19e-1
Noise specification (M.I)-(M.II)'						
Noise-to-signal ratio $\xi$	$(\alpha^j, \beta^j)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.8	$(\frac{2}{3}, \frac{1}{6})$	3.89e-2	-9.85e-2	$(\frac{2}{3}, \frac{1}{6})$	5.35e-3	-3.48e-2
3.2	$(\frac{2}{3}, \frac{1}{6})$	5.47e-2	-6.53e-2	$(\frac{2}{3}, \frac{1}{6})$	1.41e-2	-1.70e-2
Noise specification (MR)						
Rounding level $l_\alpha$	$(\alpha^j, \beta^j)$	IRMSE	IBIAS	$(\alpha^b, \beta^b)$	RMSE( $T_b$ )	BIAS( $T_b$ )
0.01	$(\frac{2}{3}, \frac{1}{6})$	4.13e-2	-1.06e-1	$(\frac{2}{3}, \frac{1}{6})$	5.29e-3	-3.26e-2
0.1	$(\frac{2}{3}, \frac{1}{6})$	1.02e-1	-5.67e-2	$(\frac{2}{3}, \frac{1}{6})$	6.61e-2	6.95e-2

doi:10.1371/journal.pone.0139041.t004

Furthermore, the comparison between the results shown in Tables 2 and 3 confirms that the absence of the cross terms appearing in the Fourier estimator significantly affects its robustness.

The test used to select the optimal pair  $(\alpha, \beta)$  is infeasible since it is based on the minimization of the unknown mean squared error. However, in view of the empirical applications, we propose an elementary feasible test for the bandwidth (frequency) selection. For each  $t$  and for each pair  $(\alpha, \beta)$  we construct 504 realization of the standardized log-return  $z_t^{\alpha, \beta}$  and we apply the Bera-Jarque test at the significance level 0.05 with null hypothesis that  $z_t^{\alpha, \beta}$  is normal with unspecified mean and variance. We compute the rejection percentage of the null hypothesis for each pair  $(\alpha, \beta)$  and we select the optimal pair to be the one with the smallest rejection percentage. Proceeding with this simple test the selected optimal pairs of the Fourier estimator are the same as those shown in Table 2. Indeed, there is one exception at the rounding error at level 0.1 where the optimal pair selected by the new test is  $(2/3, 1/4)$ . The smallest percentage of rejection varies in the interval [5%, 85%] depending on the noise specifications.

### 5 Empirical Analysis

In this section we present an empirical analysis conducted on illiquid and liquid data to illustrate the ability of the Fourier spot volatility estimator to capture intraday variations of volatility and to predict high-frequency Value-at-Risk (VaR) [26–28].

The same empirical analysis is carried on two data-sets and we have  $T = 1$  day. The first data-set consists of 5-second observations of the U.S. S&P 500 index obtained by Bloomberg. We consider 5 trading days from March 4, 2013 to March 7, 2013 corresponding to about one week data. Each trading day starts at 15.30 (Rome local time) and ends at 22:30 (Rome local time). We use the observations from 15:30 to 22:00 since the others remain substantially

unchanged. This choice corresponds to 4680 (5-second) observations. The second data-set consists of the nearby Italian stock index futures, FIB30 in January 2001 observed every 5.67 seconds. Specifically, we consider four consecutive trading days, January 27, 28, 29 and 30, 2001. As shown in subsection 5.1 the first data-set is characterized by illiquid data while the second one by liquid data.

We will show that the Fourier estimator has a good performance when both illiquid and liquid data are used. Furthermore, the optimal frequencies  $M$  and  $N$  are chosen by using the simple test illustrated in Section 4.2.

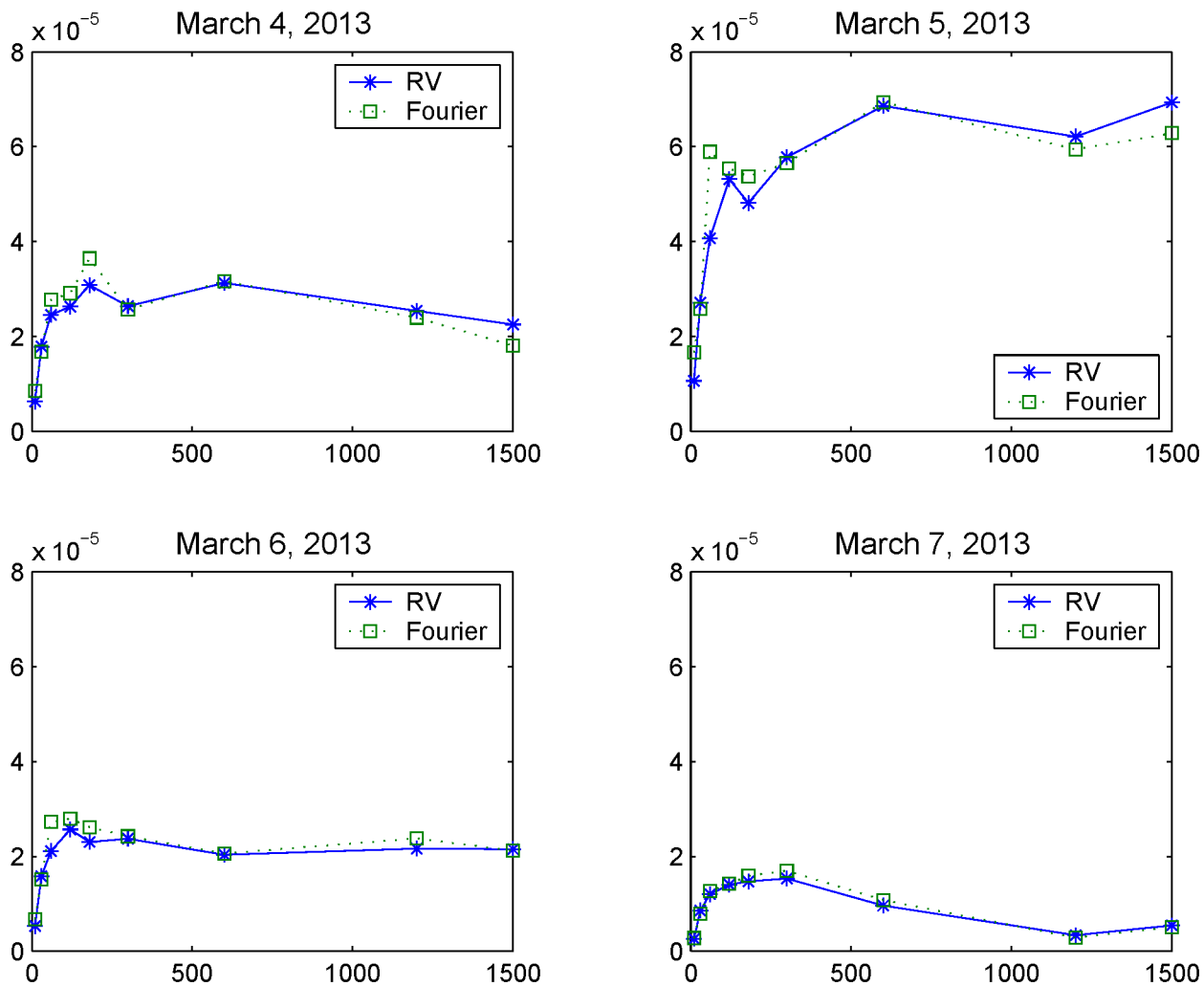
### 5.1 Intraday variations of volatility when data are illiquid

In this subsection we consider the first data-set of U.S. S&P 500 index 5-second returns observed on March 4-7, 2013. We first carry out a preliminary analysis to study the main features of the observed data. This analysis is conducted through the volatility signature plot and the autocorrelation functions. Fig 5 shows the volatility signature plots corresponding to March 4, 5, 6 and 7, 2013 evaluated using the Realized Volatility (RV) and the Fourier estimator for integrated volatility. Roughly speaking, the volatility signature plots of Fig 5 are plots of the realized variance against sampling intervals and, as explained in Ref. [25], these sampling intervals are chosen to be multiples of the smallest sampling interval. Ref. [25] highlights that highly liquid assets display the largest realized variance estimates at the highest sampling rates (i.e. 5-second returns) while illiquid assets display the largest realized variance estimates at the lowest sampling rates (i.e. 20-minute returns). In fact, liquid assets show a negative serial autocorrelation so that the oscillating swings in the returns reduce for larger sampling intervals by the effect of cancelation. The signature plots illustrated in Fig 5 show that we are dealing with an illiquid asset.

This finding is confirmed by Figs 6 and 7. Fig 6 shows the 5-second return autocorrelation function on March 4-7, 2013 while Fig 7 shows the autocorrelation function of the returns on March 4, 2013 for four different sampling frequencies (i.e. upper left panel 5-second returns, upper right panel 10-second returns, lower left panel 30-second returns and lower right panel 1-minute returns). These two figures show a positive serial autocorrelation at high frequencies and this implies smaller estimates of the realized variance for these frequencies. As stressed in Ref. [25], high quality realized variance estimates can be constructed in a liquid market. However, we investigate the performance of the Fourier estimator also against illiquid assets.

Fig 8 shows the autocorrelation function of a sampled Gaussian distribution  $\mathcal{N}(0, \Delta t)$  with sampling frequency of 1-minute (upper panel) and the autocorrelation of 1-minute observed returns on March 4, 2013 (lower panel). The results shown in the two panels and the preliminary analysis illustrated above suggest the use of 1-minute returns in order to reconstruct the distribution of the standardized return  $z_t$ , defined in Eq (14). As previously mentioned, the standardized return,  $z_t$ , is a standard Gaussian random variable when the prices are not affected by microstructure noise effects.

We compute an estimate,  $z_t^{\text{FR}}$ , of the standardized return  $z_t$  using the spot volatility estimated with the Fourier method as done in Subsection 4.1. The frequencies  $N$  and  $M$  for the Fourier estimator are chosen according to  $N = n/2$  and  $M = \frac{1}{2\pi} \frac{1}{8} \sqrt{n} \log n$ . Fig 9 shows the empirical cumulative distribution function for  $z_t^{\text{FR}}$  and the expected cumulative distribution function (i.e.  $\mathcal{N}(0,1)$ ), with  $\Delta t = T/n = (1/390)$  seconds. Moreover, the BJ and the KS tests applied to the random sample  $z_t^{\text{FR}}$  do not reject the null hypothesis at the significance level 0.05 and their p-values are 0.3, 0.6, respectively. The Fourier estimator shows a good performance in interpreting changes of volatilities despite the fact that the asset is illiquid. This finding is confirmed also by the second application, that is, the Value at Risk (VaR) prediction. This

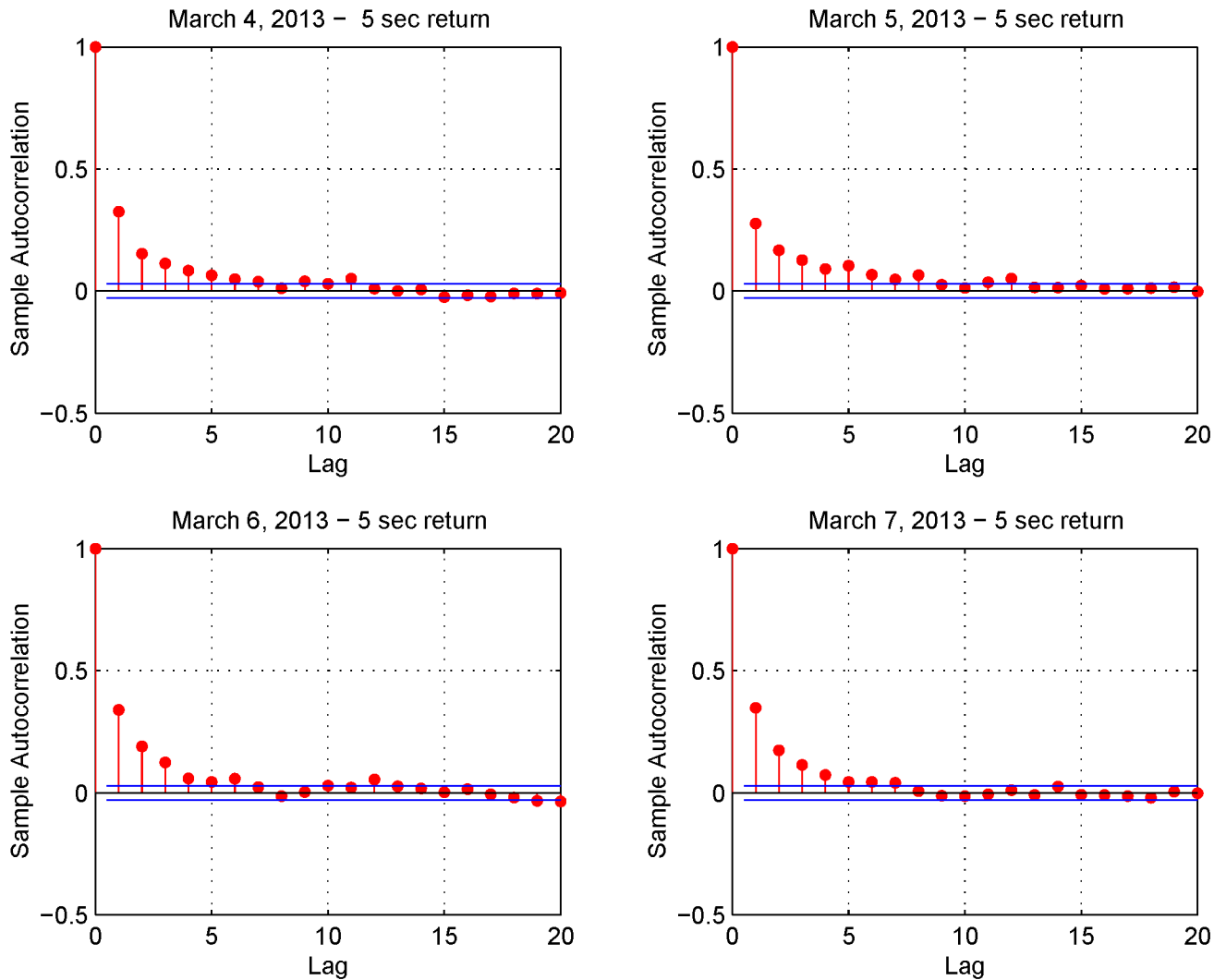


**Fig 5. Volatility signature plots versus sampling frequency in seconds.** The graphs show the signature plots of illiquid market data on March 4-7, 2013.

doi:10.1371/journal.pone.0139041.g005

application has already been illustrated in Refs. [26, 28] to measure the performance of integrated volatility estimators and in Ref. [27] for the spot volatility estimators.

We apply the Fourier and TS estimators to predict 1%, 5%, 95%, 99% Value at Risk of 1-minute returns. As explained in Ref. [28], the conditional coverage probability must be the same as the theoretical level  $\alpha$  of VaR in the left tail and  $1 - \alpha$  in the right tail. Hence, a measure of the performance is given by the difference between  $-Q(\alpha)$  and  $Q(1 - \alpha)$  where  $Q(\cdot)$  is the quantile function of the empirical distribution of the standardized returns  $z_t^{FR}$  and  $z_t^{TS}$ . Table 5 shows the conditional coverage probability (i.e. the empirical frequency that the failure of VaR prediction occurs) for the left and right tails of the 1-minute return distribution. Theoretically, we should expect a 1% and 5% rate of violation when we consider the 1% and 5% VaRs. We highlight the fact that the VaR predictions at 5% and 95% are satisfactory since we have two correct significant digits in the predictions.



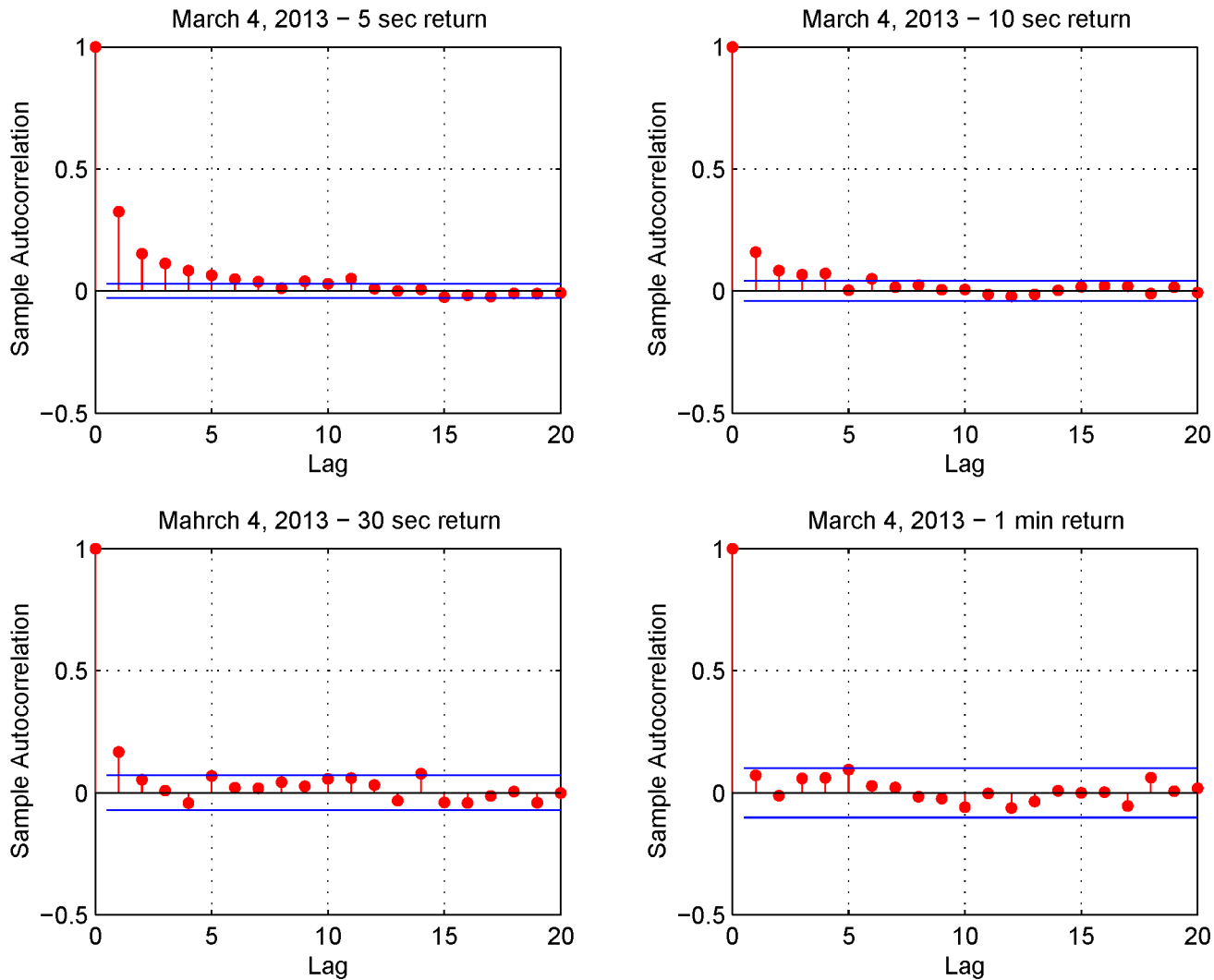
**Fig 6. Autocorrelation functions.** The four panels show the autocorrelation in 5-second returns (March 4-7, 2013).

doi:10.1371/journal.pone.0139041.g006

## 5.2 Intraday variations of volatility when data are liquid

In this subsection we repeat the exercises of Subsection 5.1 using the nearby Italian stock index futures, FIB30, recorded every 5.67 seconds. Specifically, we consider four consecutive trading days (January 27-30, 2001). Quotes prior to 10 a.m. are removed to eliminate opening quotes from our sample.

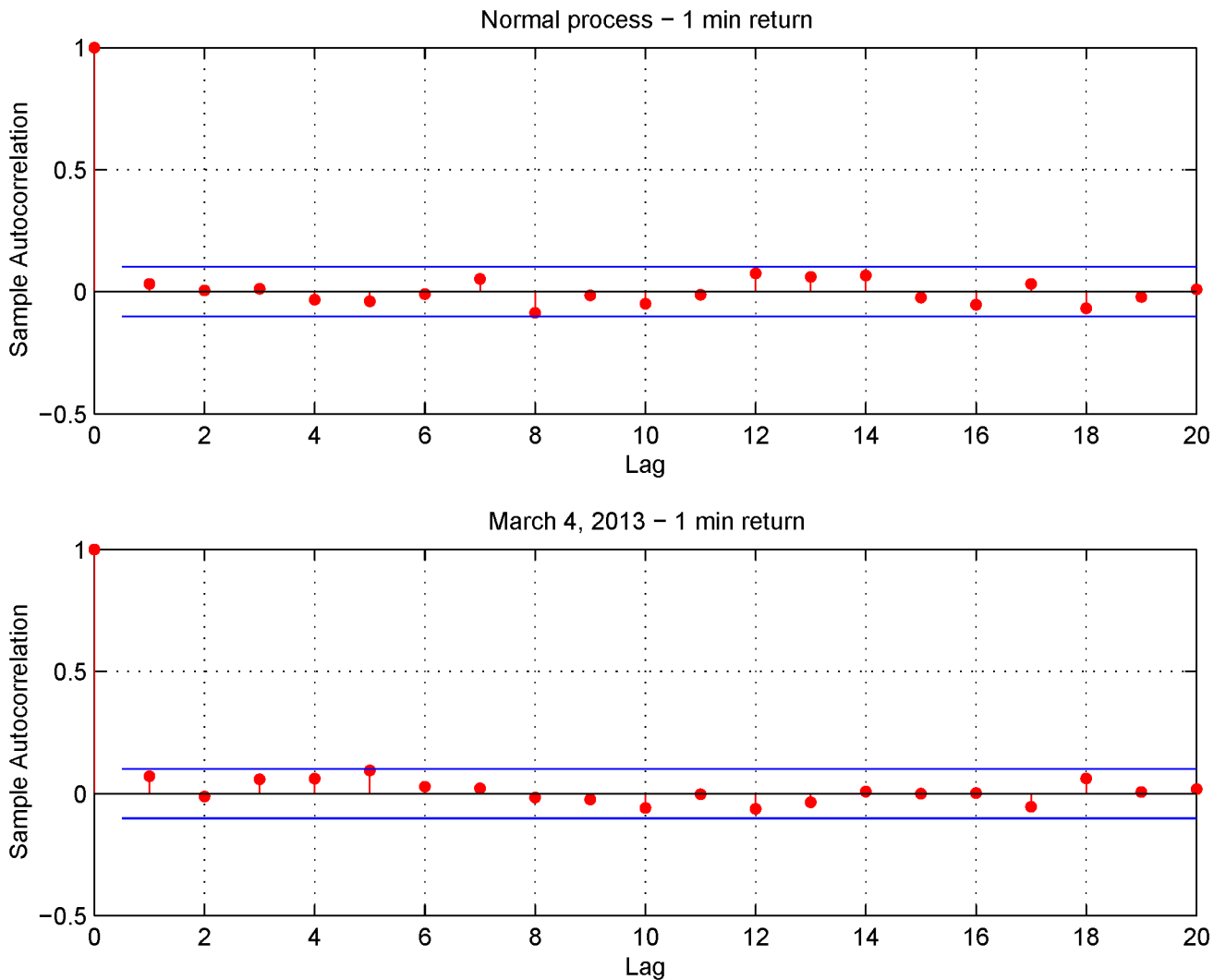
Fig 10 shows the volatility signature plots corresponding to January 27, 28, 29 and 30, 2001 evaluated using the Realized Variance (RV) and the Fourier estimator for the integrated volatility. The four plots in Fig 11 show the autocorrelation functions corresponding to the 5.67-second returns on January 27-30, 2001. We can see that the first-order autocorrelation is significantly negative while the second and third autocorrelations are slightly positive. The blue lines denote the 95% confidence interval. The shape of the signature plot of Fig 10 and the negative first order autocorrelation in Fig 11 show that FIB30 returns are liquid data. Fig 12 shows



**Fig 7. Autocorrelation for various sampling intervals on March 4, 2013.** Upper left panel (5-second returns), upper right panel (10-second returns), lower left panel (30-second returns) and lower right panel (1-minute returns).

doi:10.1371/journal.pone.0139041.g007

the autocorrelation function of the returns on January 27, 2001 for four different sampling frequencies (i.e. upper left panel 5.67-second returns, upper right panel 15-second returns, lower left panel 30-second returns and lower right panel 1-minute returns). As in Subsection 5.1, the preliminary analysis suggests the use of 1-minute returns to sample the standardized return,  $z_t^{FR}$ . Furthermore, the simple test illustrated in Section 4.2 selects the frequencies  $N = n/2$  and  $M = \frac{1}{2\pi} \frac{1}{8} \sqrt{n} \log n$ . Fig 13 shows the empirical distributions of the 1-minute standardized returns compared with the Gaussian density function obtained using illiquid data (Fig 13 left panel) and liquid data (Fig 13 right panel). The comparison of left and right panels highlights that the Fourier estimator provides accurate spot volatility estimates for both liquid and illiquid data. This is confirmed also by the VaR predictions shown in Table 5 (Liquid Data column), where the VaR predictions have at least two correct significant digits. Note that VaR predictions obtained using the TS estimators slightly outperform those obtained with Fourier estimator in the case of liquid data while the opposite happens in the case of illiquid data.

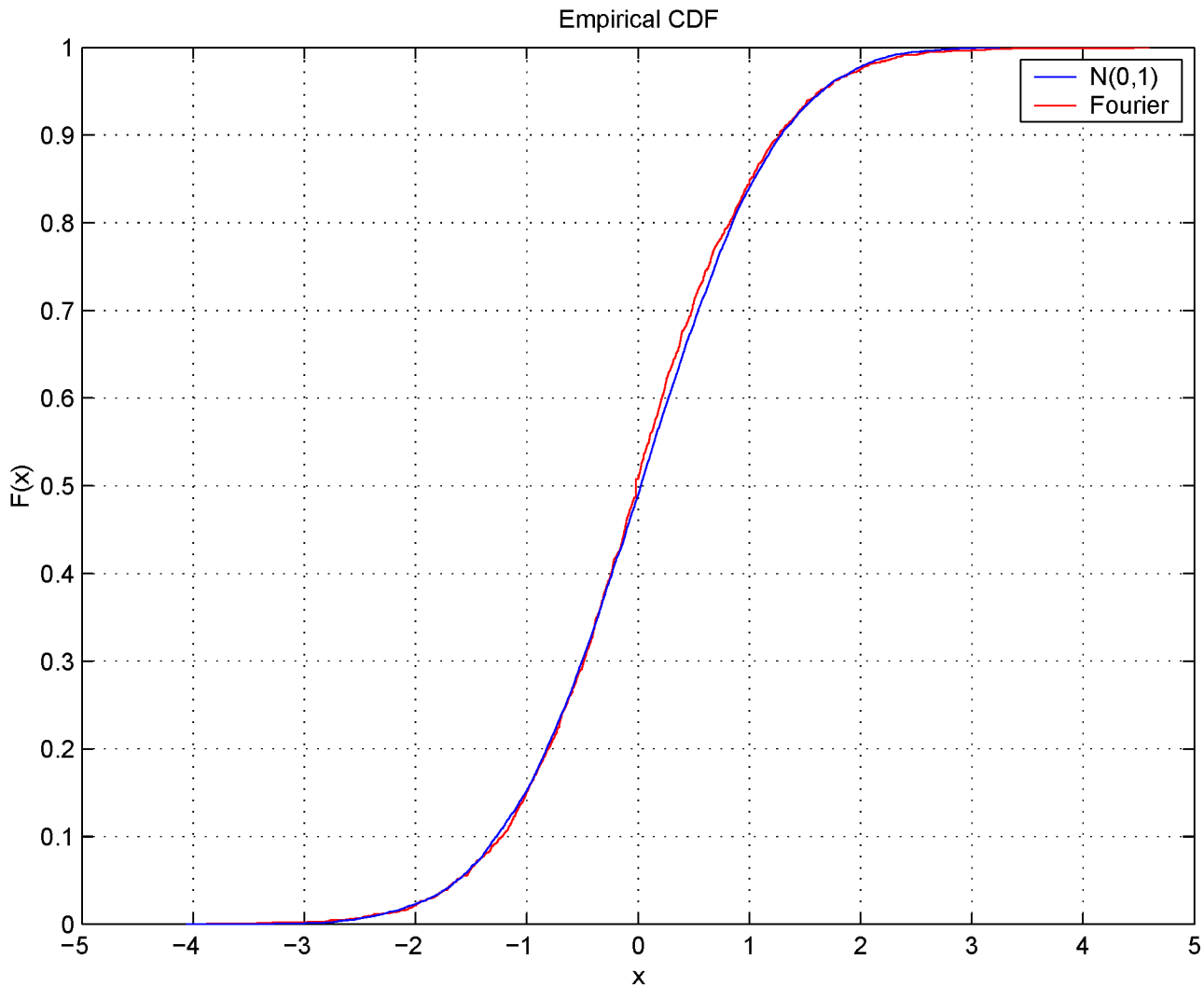


**Fig 8. Comparison of the autocorrelation functions.** Autocorrelation function of a sampled normal process  $\mathcal{N}(0, \Delta t)$  with sampling interval of 1 minute (upper panel) and of the 1-minute returns observed on March 4, 2013 (lower panel).

doi:10.1371/journal.pone.0139041.g008

## 6 Conclusions

We have obtained the asymptotic error distribution of the Fourier spot volatility estimator with optimal rate of convergence and variance. Furthermore, extensive simulation studies and two empirical analysis have been proposed to show the efficiency of the Fourier estimator with high frequency data. The results illustrated in Sections 4 and 5 show that it is robust to various noise specifications. In addition, it is a *globally robust* estimator in the sense that it allows us to choose values of the so called cutting frequencies independently of the specific instant of time in the observed time window, of the specific features of the volatility process and of different noise models and levels. This feature of the Fourier estimator makes it particularly suited for empirical applications, such as the calibration of stochastic volatility models for asset prices, futures prices and models where the price volatility can provide insights into the traders' strategies in the financial market. Finally, the high frequency intraday variations of volatility could



**Fig 9. Comparison of the empirical cumulative distribution function and the standard normal one.** Cumulative distribution functions of a sampled normal process  $N(0, 1)$  (blue line) and of the sampled process  $z_t^{FR}$  (red line) with a sampling interval of 1 minute (March 4-7, 2013).

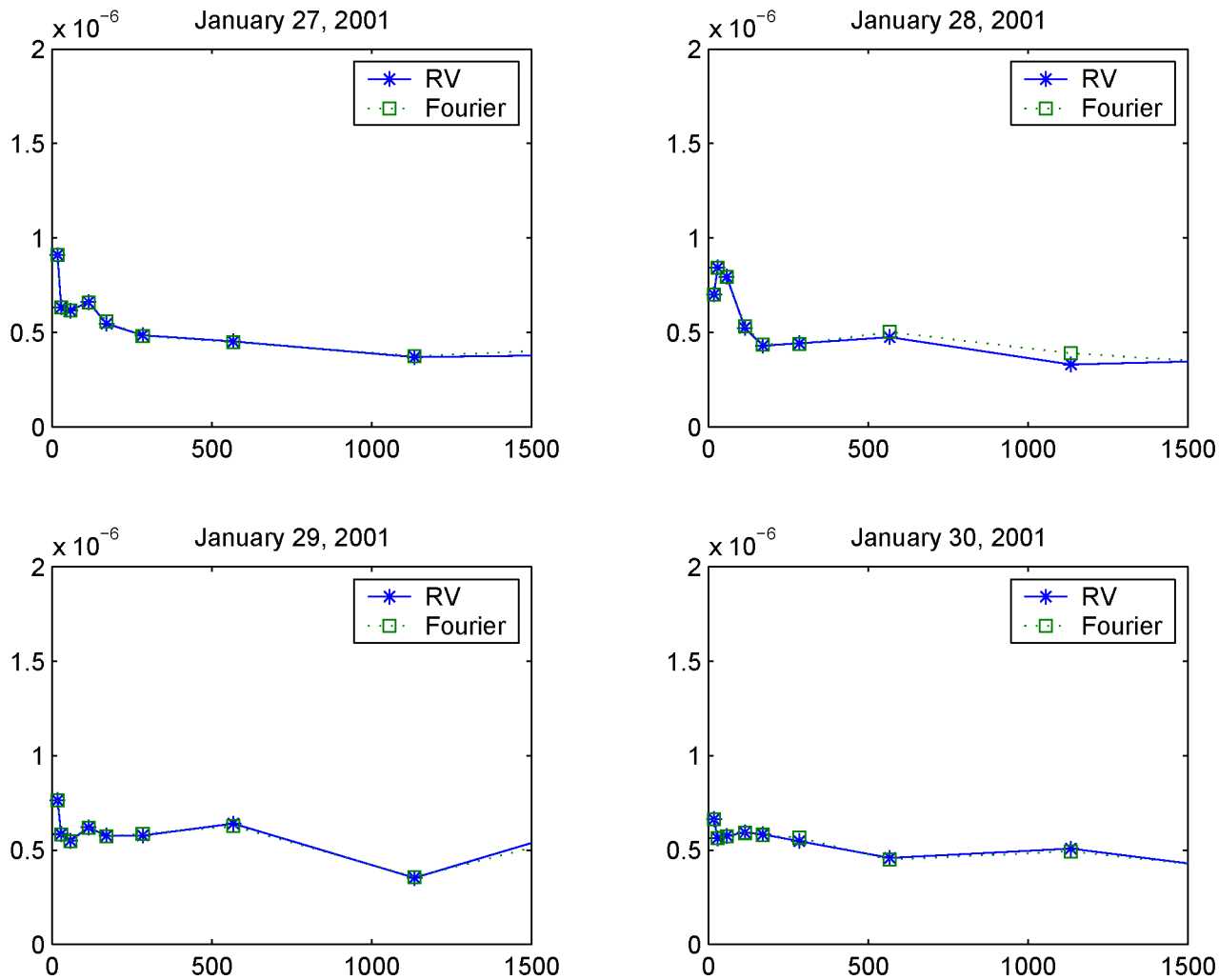
doi:10.1371/journal.pone.0139041.g009

**Table 5. VaR prediction.** Intraday VaR predictions obtained analyzing illiquid and liquid data with the Fourier and the TS estimators (sampling rate: one observation per minute).

	Illiquid Data		Liquid Data	
	Fourier	TS	Fourier	TS
1%	0.0119	0.0159	0.0087	0.0175
99%	0.0103	0.0144	0.0081	0.0148
5%	0.0484	0.0445	0.0508	0.0545
95%	0.0495	0.0520	0.0548	0.0572

doi:10.1371/journal.pone.0139041.t005





**Fig 10. Volatility signature plots versus sampling frequency in seconds.** The graphs show the signature plots of liquid market data on January 27-30, 2001

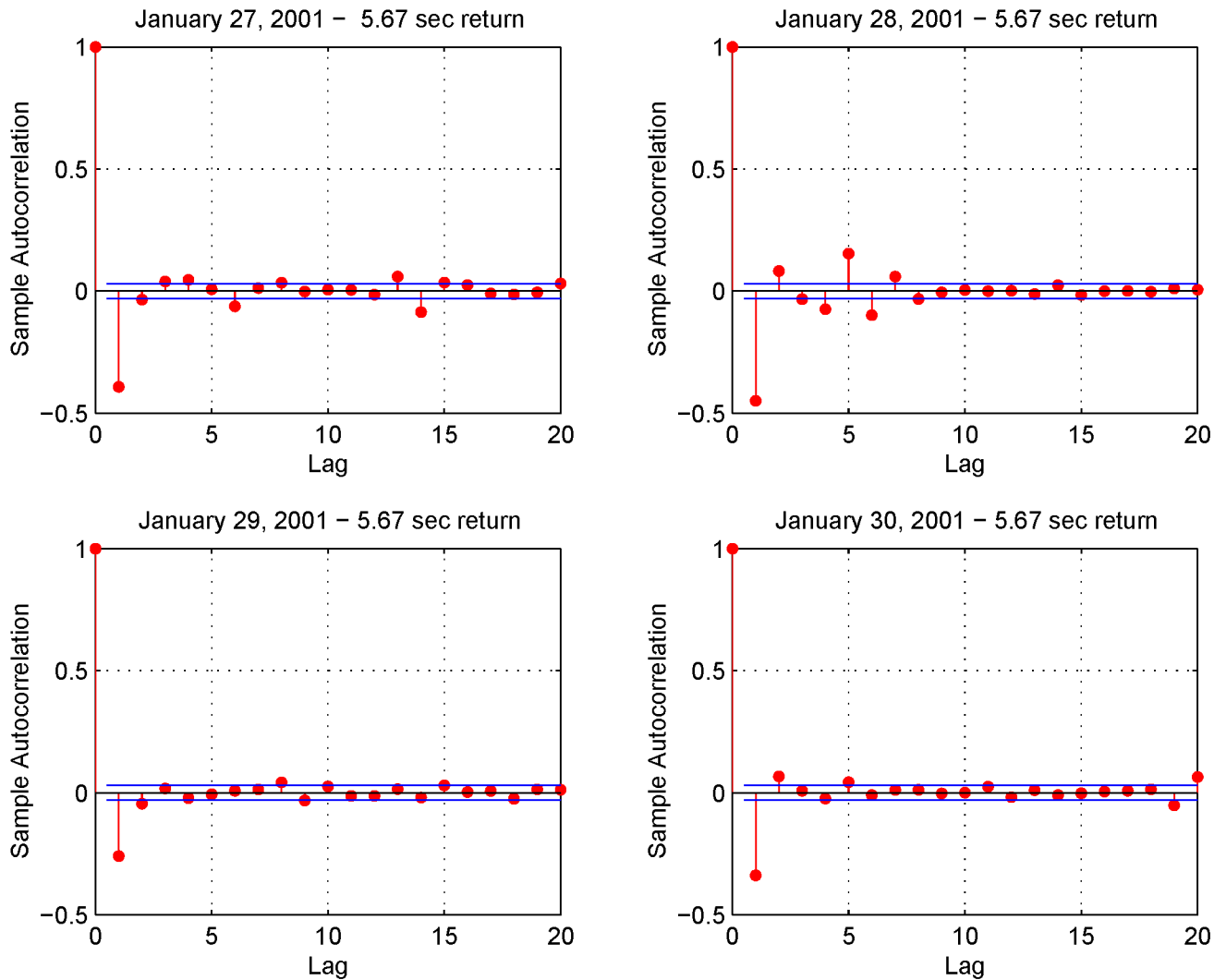
doi:10.1371/journal.pone.0139041.g010

be used to detect the onset of flash crashes (i.e. a deep fall in security prices occurring in a few seconds). The most relevant flash crash occurred on May 6th, 2010 when 4.1 billion dollars traded on the NYSE resulted in a drop of the Dow Jones Industrial Average of over 1000 points and then a rise to approximately the previous value. The mechanism which causes these events has been studied in depth and the high frequency spot volatility estimators could provide useful insights into this speculative trading.

### 7 Appendix A

**Remark 7.1** Given the discrete time observations  $\{t_j\}_{j=0, \dots, m}$  denote  $\phi_n(\tau) := \sup\{t_j \in [0, \tau]: t_j \leq \tau\}$ , then using Itô formula we can write Eq (6) as

$$\hat{\sigma}_{n,N,M}^2(t) = \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v)) \sigma^2(v) dv + F_1(t) + F_2(t), \tag{24}$$



**Fig 11. Autocorrelation functions.** The four panels show the autocorrelation in 5.67-second returns on January 28-31, 2001.

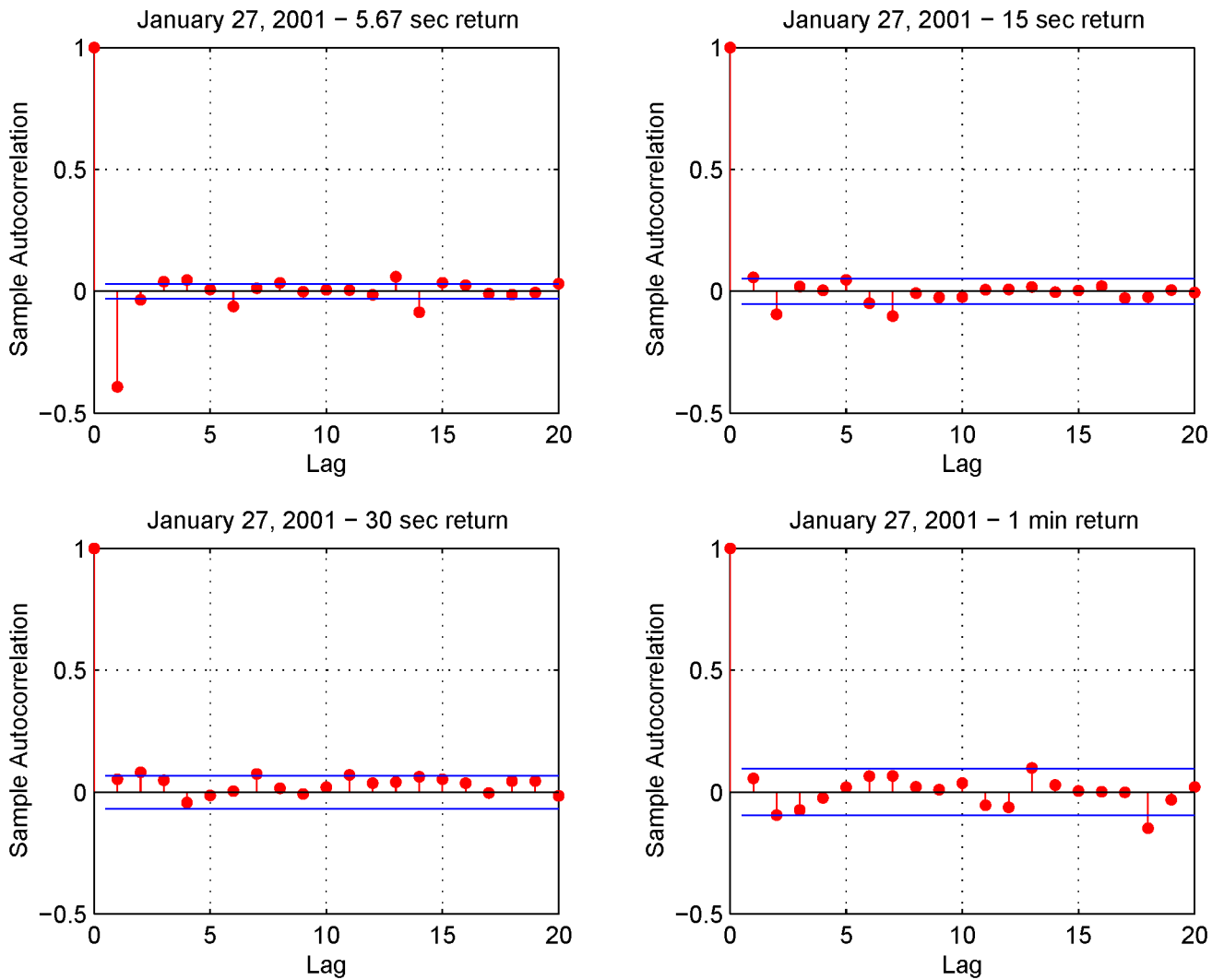
doi:10.1371/journal.pone.0139041.g011

where

$$F_1(t) = \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v)) \int_0^v D_N(\phi_n(v) - \phi_n(u)) \sigma(u) dW_u \sigma(v) dW_v \quad (25)$$

$$F_2(t) = \frac{1}{2\pi} \int_0^{2\pi} \int_0^v F_M(t - \phi_n(u)) D_N(\phi_n(v) - \phi_n(u)) \sigma(u) dW_u \sigma(v) dW_v. \quad (26)$$

*Proof.* (of Theorem 3.1) According to Lemma 2.2 in Ref. [15], we can assume that  $b = 0$ . Further, it is not restrictive to assume that the volatility process is a.s. bounded in  $[0, 2\pi]$ , then  $\text{ess sup} \|\sigma^2\|_\infty < \infty$ , where  $\|\sigma^2\|_\infty = \sup_t \sigma^2(t)$ .



**Fig 12. Autocorrelation for various sampling intervals on January 27, 2001.** Upper left panel (5.67-second returns), upper right panel (15-second returns), lower left panel (30-second returns) and lower right panel (1-minute returns).

doi:10.1371/journal.pone.0139041.g012

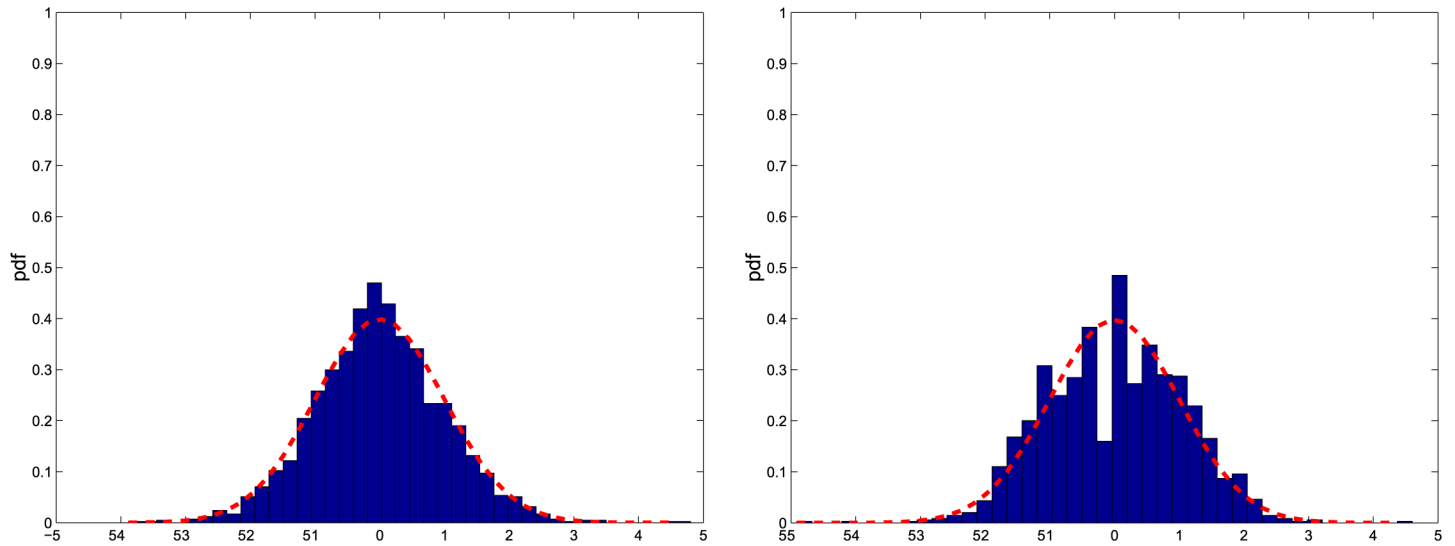
Using the decomposition by Eqs (24), (25), (26), we write

$$\begin{aligned} \sqrt{\frac{n}{M}}(\hat{\sigma}_{n,N,M}^2(t) - \sigma^2(t)) &= \sqrt{\frac{n}{M}} \left( \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v))\sigma^2(v)dv - \sigma^2(t) \right) \\ &\quad + \sqrt{\frac{n}{M}} (F_1(t) + F_2(t)). \end{aligned}$$

The proof is divided into three steps.

I) We prove that

$$\sqrt{\frac{n}{M}} \left( \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v))\sigma^2(v)dv - \sigma^2(t) \right) \tag{27}$$



**Fig 13. Empirical distribution functions.** Left panel: pdf of a random variable  $\mathcal{N}(0,1)$  (red dashed line) and of the sampled process  $z_t^{\text{FR}}$  with a sampling interval of 1 minute (illiquid data). Right panel: pdf of a random variable  $\mathcal{N}(0,1)$  (red dashed line) and of the sampled process  $z_t^{\text{FR}}$  with a sampling interval of 1 minute (liquid data).

doi:10.1371/journal.pone.0139041.g013

converges to 0 in probability. We write Eq (27) as

$$\sqrt{\frac{n}{M}} \left( \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v)) \sigma^2(v) dv - \frac{1}{2\pi} \int_0^{2\pi} F_M(t - v) \sigma^2(v) dv \right) \quad (28)$$

$$+ \sqrt{\frac{n}{M}} \left( \frac{1}{2\pi} \int_0^{2\pi} F_M(t - v) \sigma^2(v) dv - \sigma^2(t) \right). \quad (29)$$

For the term (28) it holds a.s.:

$$\begin{aligned} & \left| \sqrt{\frac{n}{M}} \left[ \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v)) \sigma^2(v) dv - \frac{1}{2\pi} \int_0^{2\pi} F_M(t - v) \sigma^2(v) dv \right] \right| \\ & \leq \text{ess sup } \|\sigma^2\|_\infty \sqrt{\frac{M}{n}} \rightarrow 0, \end{aligned}$$

by Lemma 8.1 i). Then, consider the  $L^1$ -norm of the term (29) and split it in two terms:

$$\begin{aligned} & \sqrt{\frac{n}{M}} E \left[ \left| \frac{1}{2\pi} \int_0^{2\pi} F_M(t - v) \sigma^2(v) dv - \sigma^2(t) \right| \right] \\ & \leq \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_{[0, 2\pi] \cap \{|t-v| \leq \frac{2\pi}{M+1}\}} F_M(t - v) E[|\sigma^2(v) - \sigma^2(t)|] dv \end{aligned} \quad (30)$$

$$+ \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_{[0, 2\pi] \cap \{|t-v| > \frac{2\pi}{M+1}\}} F_M(t - v) E[|\sigma^2(v) - \sigma^2(t)|] dv. \quad (31)$$

Consider Eq (30). It is smaller than

$$\begin{aligned} & \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_{[0,2\pi] \cap \{|t-v| \leq \frac{2\pi}{M+1}\}} F_M(t-v) dv E \left[ \sup_{[0,2\pi] \cap \{|t-v| \leq \frac{2\pi}{M+1}\}} |\sigma^2(v) - \sigma^2(t)| \right] \\ & \leq \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_{[0,2\pi] \cap \{|t-v| \leq \frac{2\pi}{M+1}\}} F_M(t-v) dv \frac{C}{(M+1)^\nu} \leq C \sqrt{\frac{n}{M}} \frac{1}{M^\nu} \end{aligned}$$

which goes to zero as soon as  $1 < \gamma < 1+2\nu$  (this holds under the given assumptions). Note that the first inequality follows because the volatility path is  $\nu$ -Hölder continuous. Consider Eq (31):

$$\begin{aligned} & \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_{[0,2\pi] \cap \{|t-v| > \frac{2\pi}{M+1}\}} F_M(t-v) |\sigma^2(v) - \sigma^2(t)| dv \\ & \leq \sqrt{\frac{n}{M}} 2 \operatorname{ess\,sup} \|\sigma^2\|_\infty \frac{1}{2\pi} \int_{[0,2\pi] \cap \{|t-v| > \frac{2\pi}{M+1}\}} F_M(t-v) dv \leq C \sqrt{\frac{n}{M}} \frac{1}{M} \end{aligned}$$

where we have used Lemma 8.1 ii). This converges to zero as soon as  $1 < \gamma < 3$ .

II) According to Ref. [45] we determine the variance of our asymptotic distribution, by studying

$$\left\langle \sqrt{\frac{n}{M}} (F_1(t) + F_2(t)), \sqrt{\frac{n}{M}} (F_1(t) + F_2(t)) \right\rangle$$

where  $\langle \cdot, \cdot \rangle$  denotes the quadratic covariation in  $[0, 2\pi]$ . This is composed of four terms, each of them leads to the same limit. We study the first one in detail, the remaining terms are similar. Consider

$$\begin{aligned} & \left\langle \sqrt{\frac{n}{M}} F_1(t), \sqrt{\frac{n}{M}} F_1(t) \right\rangle \\ & = \frac{n}{M} \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) \left( \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \right)^2 \sigma^2(v_2) dv_2. \end{aligned}$$

Applying Itô formula, we get

$$\begin{aligned} & \left( \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \right)^2 = \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 \\ & + 2 \int_0^{v_2} \left( \int_0^{v_1} D_N(\phi_n(v_2) - \phi_n(u)) \sigma(u) dW_u \right) D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1}. \end{aligned}$$

Let

$$\begin{aligned} F_{11}(t) & := \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 \sigma^2(v_2) dv_2 \\ F_{12}(t) & := \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) \int_0^{v_2} \left( \int_0^{v_1} D_N(\phi_n(v_2) - \phi_n(u)) \sigma(u) dW_u \right) \\ & \times D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \sigma^2(v_2) dv_2. \end{aligned}$$

We prove that, for any  $t$  fixed, in probability:

$$\frac{n}{M} F_{11}(t) \rightarrow \frac{1}{3} (1 + 2\eta(c)) \sigma^4(t) \tag{32}$$

where  $\eta(c)$  is equal to [Eq \(10\)](#), and

$$\frac{n}{M} F_{12}(t) \rightarrow 0. \tag{33}$$

We begin with [Eq \(32\)](#). Let  $V := 2\pi \frac{1}{2} (1 + 2\eta(c)) = \pi (1 + \eta(c))$ . We have:

$$\begin{aligned} & \left| \frac{n}{M} F_{11}(t) - V \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right| \\ &= \left| \frac{n}{M} \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 \sigma^2(v_2) dv_2 - V \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right| \tag{34} \\ &\leq \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \phi_n(v_2)) \left( n \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 - V \sigma^2(v_2) \right) \sigma^2(v_2) dv_2 \right| \\ &\quad + V \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \phi_n(v_2)) \sigma^4(v_2) dv_2 - \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right|. \tag{35} \end{aligned}$$

Consider [Eq \(34\)](#): it is less than

$$\frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \phi_n(v_2)) \left| n \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 - V \sigma^2(v_2) \right| \sigma^2(v_2) dv_2$$

which is  $o_p(1)$  in virtue of [Lemmas 8.1 iv\)](#) and [8.3](#).

Consider now [Eq \(35\)](#): by [Lemma 8.1 iv\)](#), it holds in probability

$$\lim_{n, M \rightarrow \infty} \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \phi_n(v_2)) \sigma^4(v_2) dv_2 - \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right| = 0.$$

Finally, observe that  $V \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) = \frac{1}{3} (1 + 2\eta(c)) \sigma^4(t)$ .

We prove now [Eq \(33\)](#). Consider

$$\begin{aligned} & \frac{n}{M} \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) \int_0^{v_2} \left( \int_0^{v_1} D_N(\phi_n(v_2) - \phi_n(u)) \sigma(u) dW_u \right) \\ & \quad \times D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \sigma^2(v_2) dv_2. \end{aligned}$$

Using Itô isometry, we have in  $L^2$ -norm:

$$\begin{aligned} & n^2 E \left[ \left( \int_0^{v_2} \left( \int_0^{v_1} D_N(\phi_n(v_2) - \phi_n(u)) \sigma(u) dW_u \right) D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \right)^2 \right] \\ &= n^2 E \left[ \int_0^{v_2} \left( \int_0^{v_1} D_N(\phi_n(v_2) - \phi_n(u)) \sigma(u) dW_u \right)^2 D_N^2(\phi_n(v_2) - \phi_n(v_1)) \sigma^2(v_1) dv_1 \right] \\ &\leq \| \sigma^2 \|_\infty^2 n^2 \int_0^{v_2} \int_0^{v_1} D_N^2(\phi_n(v_2) - \phi_n(u)) du D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1. \end{aligned}$$

Therefore, it is enough to observe that, as  $\frac{N}{n} \rightarrow c$ ,

$$n \int_0^{v_2} \left( n \int_0^{v_1} D_N^2(\phi_n(v_2) - \phi_n(u)) du \right) D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1 = o(1)$$

combining Lemmas 8.3 ii) and iii). Finally, remark that  $\frac{1}{M} \int_0^{2\pi} F_M^2(t - \phi_n(v_2)) dv_2 = O(1)$ , by Lemma 8.1 iii). This concludes the proof.

III) The last step of the proof requires to prove the convergence in probability

$$\left\langle \sqrt{\frac{n}{M}} (F_1(t) + F_2(t)), W \right\rangle \rightarrow 0.$$

Observe that:

$$\begin{aligned} \left\langle \sqrt{\frac{n}{M}} F_1(t), W \right\rangle &= \\ & \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v_2)) \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \sigma(v_2) dv_2. \end{aligned}$$

Consider the  $L^1$ -norm:

$$\begin{aligned} E \left[ \left| \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v_2)) \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \sigma(v_2) dv_2 \right| \right] \\ \leq \| \sigma \|_\infty \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v_2)) E \left[ \left| \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \right| \right] dv_2. \end{aligned} \tag{36}$$

Moreover, we have

$$E \left[ \left| \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \right| \right] \leq \| \sigma^2 \|_\infty^{\frac{1}{2}} \left( \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1 \right)^{\frac{1}{2}}.$$

Thus

$$\begin{aligned} \sqrt{\frac{n}{M}} E \left[ \left| \frac{1}{2\pi} \int_0^{2\pi} F_M(t - \phi_n(v_2)) \int_0^{v_2} D_N(\phi_n(v_2) - \phi_n(v_1)) \sigma(v_1) dW_{v_1} \sigma(v_2) dv_2 \right| \right] \\ \leq C \sqrt{\frac{1}{M}} \int_0^{2\pi} F_M(t - \phi_n(v_2)) \left( n \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1 \right)^{\frac{1}{2}}. \end{aligned} \tag{37}$$

Using Lemma 8.3 i)

$$\lim_{n, N \rightarrow \infty} n \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1 \leq C_2$$

therefore, Eq (37) is less than

$$C \sqrt{\frac{1}{M}} \int_0^{2\pi} F_M(t - v_2) dv_2 \leq C \sqrt{\frac{1}{M}} 2\pi \rightarrow 0 \text{ as } M \rightarrow \infty.$$

□

## 8 Appendix B

This Appendix contains some results about the Fejer and Dirichlet kernels: these results are known but we place them here for the reader's convenience.

**Lemma 8.1** (*Fejer kernel properties*)



i. For any  $M$

$$\int_{-\pi}^{\pi} F_M(x) dx = 2\pi.$$

Moreover, under the assumption  $\lim_{n,M \rightarrow \infty} \frac{M^\gamma}{n} = a > 0$  for some  $\gamma > 1$ , then

$$\lim_{n,M \rightarrow \infty} \int_{-\pi}^{\pi} F_M(\phi_n(x)) dx = 2\pi,$$

in particular,

$$\left| \int_{-\pi}^{\pi} F_M(\phi_n(x)) dx - \int_{-\pi}^{\pi} F_M(x) dx \right| \leq C \frac{M}{n}.$$

ii. For any  $M \geq 1$

$$\int_{\frac{2\pi}{M+1}}^{\pi} F_M(x) dx \leq \frac{C}{M}.$$

iii. If  $\lim_{n,M \rightarrow \infty} \frac{M^\gamma}{n} = a > 0$  for some  $\gamma > 1$

$$\lim_{M,n \rightarrow \infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(\phi_n(x)) dx = \lim_{M \rightarrow \infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(x) dx = \frac{4\pi}{3}. \tag{38}$$

iv. If  $\lim_{n,M \rightarrow \infty} \frac{M^\gamma}{n} = a > 0$  for some  $\gamma > 1$  and  $\sigma^2$  is Holder continuous with parameter  $\nu \in (0,1]$ , then

$$\lim_{n,M \rightarrow \infty} \frac{1}{M} \int_{-\pi}^{\pi} F_M^2(t - \phi_n(s)) \sigma^2(s) ds = \lim_{M \rightarrow \infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(t - s) \sigma^2(s) ds = \frac{4\pi}{3} \sigma^2(t). \tag{39}$$

*Proof.* The proofs can be found e.g. in Ref. [23], in order (5.9), Lemma 6.1, (5.10) or Remark 5.2, Lemma 5.1.  $\square$

**Remark 8.2** As in Ref. [23] Remark 5.2 all above results are true in  $[0, T]$  very similarly.

**Lemma 8.3**

i. If  $\lim_{n,N \rightarrow \infty} \frac{N}{n} = c \neq 0$ , then for any  $p > 1$  there exists a constant  $C_p$  such that

$$\lim_{n,N} n \sup_{v_2 \in [0,2\pi]} \int_0^{2\pi} D_N^p(\phi_n(v_2) - \phi_n(v_1)) dv_1 \leq C_p.$$

ii. If  $\lim_{n,N \rightarrow \infty} \frac{N}{n} = c \neq 0$ , then

$$\lim_{n,N \rightarrow \infty} n \int_0^{v_2} D_N^2(\phi_n(v_2) - \phi_n(v_1)) dv_1 = \pi(1 + \eta(c)).$$

iii. If  $\lim_{n,N \rightarrow \infty} \frac{N}{n} = c \neq 0$ , then for any  $v_1 < v_2$  we have

$$\lim_{N,n} n \int_0^{v_1} D_N^2(\phi_n(v_2) - \phi_n(u)) du = 0.$$

Moreover, if  $\sigma^2$  is Holder continuous with parameter  $\nu \in (0,1]$ , also holds for any  $v_1 < v_2$

$$\lim_{n,N \rightarrow \infty} n \int_0^{v_1} D_N^2(\phi_n(v_2) - \phi_n(u)) \sigma^2(u) du = 0.$$

*Proof.* The proofs can be found e.g. in Ref. [21] Lemma 3, Lemma 1 and Lemma 4(1), in order.  $\square$

## Acknowledgments

The research leading to these results has been partially supported by the European Union, Seventh Framework Programme FP7, under grant agreement FinMaP no. 612955.

## Author Contributions

Conceived and designed the experiments: MEM MCR. Performed the experiments: MEM MCR. Analyzed the data: MEM MCR. Contributed reagents/materials/analysis tools: MEM MCR. Wrote the paper: MEM MCR.

## References

1. Alfarano S, Lux T, Wagner F. Estimation of agent based models: the case of an asymmetric herding model. *Computational Economics* 2005; 26: 19–49. doi: [10.1007/s10614-005-6415-1](https://doi.org/10.1007/s10614-005-6415-1)
2. Tedeschi G, Mazlounian A, Gallegati M, Helbing D. Bankruptcy cascades in interbank markets. *PLoS One* 2012; 7: e52749. doi: [10.1371/journal.pone.0052749](https://doi.org/10.1371/journal.pone.0052749) PMID: [23300760](https://pubmed.ncbi.nlm.nih.gov/23300760/)
3. Brock WA, Hommes CH. A rational route to randomness. *Econometrica* 1997; 65: 1059–1095. doi: [10.2307/2171879](https://doi.org/10.2307/2171879)
4. Tedeschi G, Iori G, Gallegati M. Herding effects in order driven markets: the rise and fall of gurus. *Journal of Economic Behaviour and Organization* 2012; 81: 82–96. doi: [10.1016/j.jebo.2011.09.006](https://doi.org/10.1016/j.jebo.2011.09.006)
5. Kirman A. Ants, rationally, and recruitment. *The Quarterly Journal of Economics* 1993; 108: 137–156. doi: [10.2307/2118498](https://doi.org/10.2307/2118498)
6. Gontis V, Kononovicius A. Consentaneous agent-based and stochastic model of the financial markets. *PLoS One* 2014; 9: e102201. doi: [10.1371/journal.pone.0102201](https://doi.org/10.1371/journal.pone.0102201) PMID: [25029364](https://pubmed.ncbi.nlm.nih.gov/25029364/)
7. Tedeschi G, Iori G, Gallegati M. The role of communication and imitation in limit order markets. *The European Physical Journal B* 2009; 71: 489–497. doi: [10.1140/epjb/e2009-00337-6](https://doi.org/10.1140/epjb/e2009-00337-6)
8. Voss A, Schroeder R, Heitmann A, Peters A, Perz S. Short-term heart rate variability influence of gender and age in healthy subjects. *PLoS One*, 2015; 10: e0118308. doi: [10.1371/journal.pone.0118308](https://doi.org/10.1371/journal.pone.0118308) PMID: [25822720](https://pubmed.ncbi.nlm.nih.gov/25822720/)
9. Andersen T, Bollerslev T, Diebold F. Parametric and Nonparametric Volatility Measurement. In Ait-Sahalia Y, Hansen L editors. *Handbook of Financial Econometrics*. Oxford, North Holland; 2010. pp. 67–137.
10. Cont R. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* 2000; 1: 223–236. doi: [10.1080/713665670](https://doi.org/10.1080/713665670)
11. Foster DP, Nelson DB. Continuous record asymptotics for rolling sample variance estimators. *Econometrica* 1996; 64: 139–174. doi: [10.2307/2171927](https://doi.org/10.2307/2171927)
12. Barndorff-Nielsen OE, Shephard N. (2002) Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Society* 2002; Series B 64: 253–280.
13. Zheng Z, Qiao Z, Takaishi T, Stanley HE, Li B. Realized Volatility and Absolute Return Volatility: A Comparison Indicating Market Risk. *PLoS One* 2014; 9: e102940. doi: [10.1371/journal.pone.0102940](https://doi.org/10.1371/journal.pone.0102940) PMID: [25054439](https://pubmed.ncbi.nlm.nih.gov/25054439/)

14. Malliavin P, Mancino ME. Fourier series method for measurement of multivariate volatilities. *Finance and Stochastics* 2002; 4: 49–61. doi: [10.1007/s780-002-8400-6](https://doi.org/10.1007/s780-002-8400-6)
15. Malliavin P, Mancino ME. A Fourier transform method for nonparametric estimation of volatility. *The Annals of Statistics* 2009; 37: 1983–2010. doi: [10.1214/08-AOS633](https://doi.org/10.1214/08-AOS633)
16. Barucci E, Renò R. On measuring volatility and the GARCH forecasting performance. *Journal of International Financial Markets, Institutions and Money* 2002; 12: 183–200. doi: [10.1016/S1042-4431\(02\)00002-1](https://doi.org/10.1016/S1042-4431(02)00002-1)
17. Precup OV, Iori G. A comparison of high-frequency cross-correlation measures. *Physica A* 2004; 344: 252–256. doi: [10.1016/j.physa.2004.06.127](https://doi.org/10.1016/j.physa.2004.06.127)
18. Hansen PR, Lunde A. Realized variance and market microstructure noise (with discussions). *Journal of Business and Economic Statistics* 2006; 24: 127–218. doi: [10.1198/073500106000000071](https://doi.org/10.1198/073500106000000071)
19. Nielsen MO, Frederiksen PH. Finite sample accuracy and choice of sampling frequency in integrated volatility estimation. *Journal of Empirical Finance* 2006; 15: 265–286. doi: [10.1016/j.jempfin.2006.12.005](https://doi.org/10.1016/j.jempfin.2006.12.005)
20. Mancino ME, Sanfelici S. Robustness of the Fourier Estimator of Integrated Volatility in Presence of Microstructure Noise. *Computational Statistics and Data Analysis* 2008; 52: 2966–2989. doi: [10.1016/j.csda.2007.07.014](https://doi.org/10.1016/j.csda.2007.07.014)
21. Clement E, Gloter A. Limit theorems in the Fourier transform method for the estimation of multivariate volatility. *Stochastic Processes and their Applications* 2011; 121: 1097–1124. doi: [10.1016/j.spa.2010.11.016](https://doi.org/10.1016/j.spa.2010.11.016)
22. Park S, Linton O. Estimating the quadratic covariation matrix for an asynchronously observed continuous time signal masked by additive noise. Working Paper 2012. Available: <http://www.lse.ac.uk/fmg/workingPapers/discussionPapers/fmgdps/dp703.pdf>
23. Cuchiero C, Teichmann J. Fourier transform methods for pathwise covariance estimation in presence of jumps. *Stochastic Processes and their Applications* 2015; 125: 116–160. doi: [10.1016/j.spa.2014.07.023](https://doi.org/10.1016/j.spa.2014.07.023)
24. Bibinger M, Winkelmann L. Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics* 2015; 184: 361–378. doi: [10.1016/j.jeconom.2014.10.004](https://doi.org/10.1016/j.jeconom.2014.10.004)
25. Andersen T, Bollerslev T, Diebold F, Labys P. The distribution of exchange rate volatility. *Journal of the American Statistical Association* 2001; 96: 42–55. doi: [10.1198/016214501750332965](https://doi.org/10.1198/016214501750332965)
26. Zhang L, Mykland PA, Ait-Sahalia Y. A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 2005; 100: 1394–1411. doi: [10.1198/016214505000000169](https://doi.org/10.1198/016214505000000169)
27. Zu Y, Boswijk HP. Estimating spot volatility with high-frequency financial data. *Journal of Econometrics* 2014; 18: 117–135. doi: [10.1016/j.jeconom.2014.04.001](https://doi.org/10.1016/j.jeconom.2014.04.001)
28. Li Y, Mykland PA. Rounding Errors and Volatility Estimation. *Journal of Financial Econometrics* 2014; 13: 478–504. doi: [10.1093/jfinec/nbu005](https://doi.org/10.1093/jfinec/nbu005)
29. Heston S. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 1993; 6: 327–343. doi: [10.1093/rfs/6.2.327](https://doi.org/10.1093/rfs/6.2.327)
30. Hull J, White A. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 1987; 42: 281–300. doi: [10.1111/j.1540-6261.1987.tb02568.x](https://doi.org/10.1111/j.1540-6261.1987.tb02568.x)
31. Stein E, Stein J. Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies* 1991; 4: 727–752. doi: [10.1093/rfs/4.4.727](https://doi.org/10.1093/rfs/4.4.727)
32. Barucci E, Mancino ME. Computation of Volatility in Stochastic Volatility Models with High Frequency Data. *International Journal of Theoretical and Applied Finance* 2010; 13: 1–21. doi: [10.1142/S0219024910005991](https://doi.org/10.1142/S0219024910005991)
33. Fan J, Wang Y. Spot volatility estimation for high frequency data. *Statistics and its Interface* 2008; 1: 279–288. doi: [10.4310/SII.2008.v1.n2.a5](https://doi.org/10.4310/SII.2008.v1.n2.a5)
34. Kristensen D. Nonparametric filtering of the realized spot volatility: a kernel-based approach. *Econometric Theory* 2010; 26: 60–93. doi: [10.1017/S0266466609090616](https://doi.org/10.1017/S0266466609090616)
35. Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica* 2008; 76: 1481–1536. doi: [10.3982/ECTA6495](https://doi.org/10.3982/ECTA6495)
36. Curato IV, Mancino ME, Recchioni MC. Spot Volatility Estimation using the Laplace Transform. 2014 Preprint. Available <http://ssrn.com/abstract=2572340>
37. Epps T. Comovements in stock prices in the very short run. *Journal of the American Statistical Association* 1979; 74: 291–298. doi: [10.2307/2286325](https://doi.org/10.2307/2286325)

38. Tóth B, Kertész J. Accurate estimator of correlations between asynchronous signals. *Physica A* 2009; 388: 1696–1705. doi: [10.1016/j.physa.2008.12.062](https://doi.org/10.1016/j.physa.2008.12.062)
39. Jacod J. Limit of random measures associated with the increments of a Brownian semimartingale. 1994 Technical Report, Université de Paris VI.
40. Mancini C, Mattiussi V, Renó R. Spot volatility estimation using delta sequences. *Finance and Stochastics* 2015; 19: 261–293. doi: [10.1007/s00780-015-0255-1](https://doi.org/10.1007/s00780-015-0255-1)
41. Bandi FM, Russell JR. Separating microstructure noise from volatility. *Journal of Financial Economics* 2006; 79: 655–692. doi: [10.1016/j.jfineco.2005.01.005](https://doi.org/10.1016/j.jfineco.2005.01.005)
42. Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N. Realized kernels in practice: trades and quotes. *Econometrics Journal* 2009; 12: 1–32. doi: [10.1111/j.1368-423X.2008.00275.x](https://doi.org/10.1111/j.1368-423X.2008.00275.x)
43. Griffin JE, Oomen RCA. Covariance measurement in presence of non-synchronous trading and market microstructure noise. *Journal of Econometrics* 2011; 160: 58–68. doi: [10.1016/j.jeconom.2010.03.015](https://doi.org/10.1016/j.jeconom.2010.03.015)
44. Delattre S, Jacod J. A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli* 1997; 3: 1–28. doi: [10.2307/3318650](https://doi.org/10.2307/3318650)
45. Jacod J. On continuous conditional Gaussian martingales and stable convergence in law. *Séminaire de Probabilités, XXXI* in LNM, vol. 1655, Berlin: Springer; 1997. pp. 232–246.