Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: www.tandfonline.com/journals/rquf20

High-frequency volatility of volatility estimation free from spot volatility estimates

Simona Sanfelici, Imma Valentina Curato & Maria Elvira Mancino

To cite this article: Simona Sanfelici, Imma Valentina Curato & Maria Elvira Mancino (2015) High-frequency volatility of volatility estimation free from spot volatility estimates, Quantitative Finance, 15:8, 1331-1345, DOI: 10.1080/14697688.2015.1032542

To link to this article: https://doi.org/10.1080/14697688.2015.1032542

Quantitative Finance, 2015
Vol. 15, No. 8, 1331–1345, http://dx.doi.org/10.1080/14697688.2015.1032542

Routledge
Taylor & Francis Group

High-frequency volatility of volatility estimation free from spot volatility estimates

SIMONA SANFELICI*†, IMMA VALENTINA CURATO‡ and MARIA ELVIRA MANCINO§

†Department of Economics, University of Parma, Via J.F. Kennedy 6, 43125 Parma, Italy
‡Institute of Mathematical Finance, Ulm University, Helmholtzstr. 18, 89069 Ulm, Germany
§Department of Economics and Management, University of Firenze, Via delle Pandette 32, 50127 Firenze, Italy

(*Received 27 December 2013; accepted 5 June 2014*)

We define a new consistent estimator of the integrated volatility of volatility based only on a pre-estimation of the Fourier coefficients of the volatility process. We investigate the finite sample properties of the estimator in the presence of noise contaminations by computing the bias of the estimator due to noise and showing that it vanishes as the number of observations increases, under suitable assumptions. In both simulated and empirical studies, the performance of the Fourier estimator with high-frequency data is investigated and it is shown that the proposed estimator of volatility of volatility is easily implementable, computationally stable and even robust to market microstructure noise.

Keywords: Stochastic volatility; Volatility of volatility; High-frequency data; Microstructure; Fourier analysis

JEL Classifications: C13, C14, C58, C22

1. Introduction

Motivated by empirical studies showing the patterns of volatilities in financial time series, in the last decades many stochastic volatility models have been proposed: such models are able to reproduce stylized facts such as variance heteroscedasticity, predictability, volatility smile and negative correlation between asset returns and volatility. Very recently Barndorff-Nielsen and Veraart (2013) proposed a new class of stochastic volatility of volatility models, introducing an extra source of randomness. The estimation of all these models is rather complicated, and the main difficulties are due to the fact that some factors are unobservable (e.g. the volatility in a standard stochastic volatility model or, even worse, the stochastic volatility of volatility in stochastic volatility of volatility models), thus we have to handle them as latent variables.

In this paper, we focus on the estimation of integrated stochastic volatility of volatility using high-frequency data and we define a consistent non-parametric estimator based on the Fourier series methodology introduced in Malliavin and Mancino (2002a), Malliavin and Mancino (2009), which works both in the case of classical stochastic volatility models and in the context of stochastic volatility of volatility models. The proposed estimator needs only to pre-estimate the Fourier coefficients of the volatility process from the observations of a price process and does not require a preliminary estimation of the instantaneous volatility.

An early application of the Fourier methodology to identify the parameters (volatility of volatility and leverage, i.e. the covariance between the stochastic variance process and the asset price process) of stochastic volatility models, including classical models such as Heston (1993), Hull and White (1987), Stein and Stein (1991), has been developed in Barucci and Mancino (2010). However, the problem of robustness with respect to microstructure noise is not addressed by these authors; hence, the numerical simulations assessing the performance of the method employ low-frequency observations.

The issue of estimating the volatility of volatility in the presence of jumps is studied in Cuchiero and Teichmann (2015): firstly, the authors combine jump robust estimators of integrated realized variance and the Fourier-Fejer inversion formula to get an estimator of the instantaneous volatility path; secondly, they use again jump robust estimators for integrated volatility in which they plug the estimated path of the volatility process in order to obtain an estimator of the volatility of volatility. Barndorff-Nielsen and Veraart (2013) define a class of stochastic volatility of volatility models and show that it 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 (2012) proposes an estimator for the

*Corresponding author. Email: simona.sanfelici@unipr.it

integrated volatility of volatility, which is also based on increments of the pre-estimated spot volatility process and attains the optimal convergence rate. The common feature of these estimators is that they first estimate the volatility path using some consistent estimate of the instantaneous volatility; secondly, they estimate the volatility of volatility using the estimated volatility process as a proxy of the unknown paths. However, these estimators do not take into account the microstructure noise effects, which would seriously affect the accuracy of the estimation as the spot volatility estimators are quite sensitive to noise.

In the present work, we define the *Fourier estimator of volatility of volatility*, we prove its consistency and we claim efficiency for our method when applied to compute the volatility of the volatility in the presence of microstructure noise. To this end, we compute the bias due to noise of the proposed estimator of volatility of volatility and we show that it converges to zero as the number of observations increases, by suitably deleting the highest frequencies in the Fourier expansions. This result is due to the intrinsic robustness of the Fourier estimator of volatility; in fact, the finite sample properties of the Fourier estimator of integrated volatility in the presence of market microstructure noise have been studied in Mancino and Sanfelici (2008), where the authors find that, even without any bias correction of the estimator, the bias of a finite sample can be made negligible by suitably deleting the highest frequencies in the Fourier expansion. Our procedure can be extended without any conceptual difficulties to the multidimensional setting.

We stress the point that the Fourier estimator of the volatility of volatility is notably different from the other proposed volatility of volatility estimators: in fact, the others all use some estimated instantaneous volatility path in order to define the volatility of volatility estimators by means of some numerical differentiation (more or less in spirit they are quadratic or power variations of the estimated spot volatilities). To the contrary, our approach relies only on *integrated* quantities, i.e. the Fourier coefficients of the volatility. As was early observed in Malliavin and Mancino (2002a), this is a peculiarity of the Fourier estimator that renders the proposed method easily implementable, computationally stable and even robust to market microstructure noise.

The finite sample performance of the Fourier estimator of volatility of volatility is tested in extensive numerical simulations, using both classical stochastic volatility models, where the spot variance follows a mean-reverting square-root process, and models with stochastic volatility of volatility, namely where the volatility of the variance process is driven by a second source of randomness. Our analysis is threefold. We first show the sensitivity of the Fourier estimator to the choice of cut frequencies, to which the consistency of the estimator is related, and we test the robustness of the estimator with respect to several noise settings. Then, we test the performance of the Fourier estimator using as a benchmark the pre-estimated spot variance-based realized variance of Barndorff-Nielsen and Veraart (2013) and the bias corrected realized variance estimator of Vetter (2012). Finally, we address the issue of parameter identification of stochastic volatility models and we consider an empirical application to S&P 500 index futures.

The paper is organized as follows. Section 2 reviews the Fourier methodology for estimating volatilities. In section 3, we define the Fourier estimator of volatility of volatility and prove its consistency. The asymptotically unbiasedness of the estimator with respect to (some kind of) microstructure noise is proved in section 4. In section 5, we test its performance in several scenarios. Section 6 concludes. The technical proofs are contained in the appendix 1.

2. The Fourier method for computing volatilities

We consider a fairly general class of stochastic volatility dW models. Suppose that the log price-variance processes satisfy

$$\textbf{(A.I)} \qquad \begin{cases} \mathrm{d}p(t) = \sigma(t)\mathrm{d}W(t) + a(t)\mathrm{d}t \\ \mathrm{d}v(t) = \gamma(t)\mathrm{d}Z(t) + b(t)\mathrm{d}t \end{cases}$$

where $p(t)$ is the logarithm of the asset price and $v(t) := \sigma^2(t)$ is the variance process. Let W and Z be correlated Brownian motions on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$, satisfying the *usual conditions*. Assume that $\sigma(t), \gamma(t)$ are non-negative adapted processes and $a(t), b(t)$ are adapted processes such that

$$\textbf{(A.II)} \qquad \begin{aligned} &E\left[\int_0^T a^2(t)\mathrm{d}t\right] < \infty, \quad E\left[\int_0^T b^2(t)\mathrm{d}t\right] < \infty \\ &E\left[\int_0^T \sigma^4(t)\mathrm{d}t\right] < \infty, \quad E\left[\int_0^T \gamma^4(t)\mathrm{d}t\right] < \infty. \end{aligned}$$

Therefore, our stochastic volatility models assume the variance process to be a continuous Brownian semimartingale, but the volatility of the variance process might have jumps. Further, we will show in section 5 that the proposed estimator of volatility of volatility works well also in the stochastic volatility of volatility models by Barndorff-Nielsen and Veraart (2013). In the sequel, we will often refer to the process $v(t)$ as the volatility, as is usually done in the econometric literature.

We briefly recall the Fourier volatility estimation method by Malliavin and Mancino (2009). By rescaling the unit of time, we can always reduce ourselves to the case where the time window $[0, T]$ becomes $[0, 2\pi]$. Then, define the kth Fourier coefficient of the price process

$$c_k(\mathrm{d}p) := \frac{1}{2\pi}\int_0^{2\pi} \exp(-\mathrm{i}kt)\,\mathrm{d}p(t),$$

and consider for all integers k, the *Bohr convolution product*

$$\lim_{N\to\infty} \frac{2\pi}{2N+1} \sum_{|s|\le N} c_s(\mathrm{d}p)c_{k-s}(\mathrm{d}p). \tag{1}$$

In Malliavin and Mancino (2009), it is proved that the limit (1) exists in probability and it is equal to the kth Fourier coefficient of the volatility process v, which we denote by $c_k(v)$.

The knowledge of the Fourier coefficients $c_k(v)$ of the unobservable instantaneous volatility process $v(t)$ allows us to handle this process as an observable variable and we can iterate the procedure in order to compute the volatility of the volatility process: given the price-variance model in (A.I), the kth Fourier coefficient of the volatility $\gamma^2(t)$ of the volatility process can be computed as the following limit in probability

$$c_k(\gamma^2) = \lim_{M\to\infty} \frac{2\pi}{2M+1} \sum_{|s|\le M} c_s(\mathrm{d}v)c_{k-s}(\mathrm{d}v), \tag{2}$$

where we can use the integration by parts formula to write the Fourier coefficients of $\mathrm{d}v$, that is, for any integer k, $k \neq 0$,

$$c_k(\mathrm{d}v) = \mathrm{i}kc_k(v) + \frac{1}{2\pi}(v(2\pi) - v(0)).$$

We start from this key property of Fourier estimation method, namely the possibility of iterating the Bohr convolution procedure, and we propose an estimator of the *integrated volatility of volatility*, indeed the zero Fourier coefficient of the process $\gamma^2(t)$, which is easily implementable with high-frequency market data.

The idea of using the estimated Fourier coefficients of the volatility as building blocks to obtain results for other related quantities has been applied in Malliavin and Mancino (2002b) to compute the price-volatility feedback-rate, in Mancino and Sanfelici (2012) to estimate the quarticity and in Curato and Sanfelici (2015) for the estimation of the leverage, i.e. the covariance between the stochastic variance process and the asset price process.

In this paper, we claim the effectiveness of Fourier estimation method when applied to compute the volatility of the volatility in the presence of microstructure noise, a result that is due to the intrinsic robustness of the Fourier estimator of volatility. In fact, Mancino and Sanfelici (2008) analyse the finite sample properties of the Fourier estimator of integrated volatility in the presence of market microstructure noise and find out that, even without any bias correction of the estimator, the bias on a finite sample can be made negligible by suitably deleting the highest frequencies in the Fourier expansion. In this paper, we analytically compute the bias of the Fourier estimator of the volatility of volatility due to the presence of noise and we show that this bias is asymptotically vanishing, under a suitable choice of the number of Fourier frequencies.

3. The Fourier estimator of volatility of volatility

In this section, we define the Fourier estimator of the volatility of volatility which relies on the convolution formulae (1) and (2); then we prove that it is consistent in probability.

For any positive integer n, let $\mathcal{S}_n := \{0 = t_0 \leq \cdots \leq t_n = 2\pi\}$ be the set of (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 $\rho(n) \to 0$ as $n \to \infty$. Moreover, let $\delta_i(p) := p(t_{i+1}) - p(t_i)$.

For any integer k, $|k| \leq 2N$, let

$$c_k(\mathrm{d}p_n) := \frac{1}{2\pi}\sum_{i=0}^{n-1} \exp(-\mathrm{i}kt_i)\delta_i(p), \tag{3}$$

then for any integer j, $|j| \leq N$, let

$$c_j(v_{n,N}) := \frac{2\pi}{2N+1}\sum_{|k|\leq N} c_k(\mathrm{d}p_n)c_{j-k}(\mathrm{d}p_n). \tag{4}$$

The following result states the consistency of the estimator (4) of the Fourier coefficients of the volatility process. The proof can be found in Malliavin and Mancino (2009).

THEOREM 3.1 *Under the assumptions* (A.I) *and* (A.II) *and the condition* $\rho(n)N \to 0$*, then, for any integer* j*, the following convergence in probability holds*

$$\lim_{n,N\to\infty} c_j(v_{n,N}) = c_j(v).$$

Given the estimated Fourier coefficients of the volatility process (4), we construct an estimator of the second-order quantity (i.e. the volatility of volatility) starting from (2). More precisely, we define the Fourier estimator of the (integrated) volatility of volatility $\int_0^{2\pi} \gamma^2(t)\mathrm{d}t$ as

$$\widehat{\gamma}^2_{n,N,M} := \frac{(2\pi)^2}{M+1}\sum_{|j|\leq M}\left(1 - \frac{|j|}{M}\right) j^2\, c_j(v_{n,N})c_{-j}(v_{n,N}). \tag{5}$$

In (5), we have chosen to add a Barlett kernel, which improves the behaviour of the estimator for very high observation frequencies.

We emphasize the fact that the estimator (5) does not require the preliminary estimation of the instantaneous volatility, but only the estimated Fourier coefficients of the volatility.

As far as we know, all the recently proposed estimators of volatility of volatility need the estimated volatility path in order to estimate the volatility of volatility, the ratio being that the reconstructed (estimated) path of the volatility is plugged into an estimator of integrated volatility, e.g. the realized volatility (see, for instance, Vetter 2012, Barndorff-Nielsen and Veraart 2013, Cuchiero and Teichmann 2015). Therefore, a large number of observations for the price process is necessary, as it is statistically clear that the integrated variance of the volatility process can be estimated only on a larger time scale than the one used for estimating the volatility path from the observed prices. This yields a huge loss of information contained in the original data-set. On the other side, it is well known that spot volatility estimation is quite unstable, especially in the presence of microstructure effects as it happens with high-frequency data. On the contrary, the Fourier estimator can reconstruct the integrated volatility of volatility using as input the Fourier coefficients of the observable log-returns, in other words using only integrated quantities from the whole data-set.

In order to prove the consistency of the proposed estimator, we add a further assumption on the values of the volatility at the end points (see Barndorff-Nielsen *et al*. 2008 for a similar idea):

(A.III) we redefine the two end values $v(0)$ and $v(2\pi)$ to be, respectively, equal to $\frac{v(0^+)+v(0^-)}{2}$ and $\frac{v(2\pi^+)+v(2\pi^-)}{2}$. Equivalently, we can use an average of m distinct observations in the intervals $(-\varepsilon, \varepsilon)$ and $(2\pi - \varepsilon, 2\pi + \varepsilon)$. This jittering is used to eliminate end-effects that would otherwise appear.

The following result proves that (5) is a consistent estimator of the integrated volatility of volatility and gives the growth rates between the highest Fourier frequencies N and M, which are needed for the construction of the estimators $c_j(v_{n,N})$ and $\hat{\gamma}^2_{n,N,M}$, respectively, and the initial mesh width $\rho(n)$ of the price process observations.

THEOREM 3.2 *Under the assumptions* (A.I)–(A.III) *and the conditions* $N\rho(n) \to 0$ *and* $\frac{M^4}{N} \to 0$*, then the following convergence in probability holds*

$$\lim_{n,N,M\to\infty} \widehat{\gamma}^2_{n,N,M} = \int_0^{2\pi} \gamma^2(t)\mathrm{d}t.$$

Remark 3.3 The multivariate extension of our results to obtain a high-frequency estimator of the covariance of the covariance matrix is essentially contained in the proposed theory. In fact,

the Fourier method was originally introduced by Malliavin and Mancino (2002a) for the estimation of multivariate volatility in order to overcome the difficulties intrinsic in the use of the quadratic covariation formula on true return data, due to the non-synchronicity of observed prices on different assets. We do not intend to develop this theory in the present paper, but we claim that the availability of a multivariate extension is an added important advantage of our estimator of second-order quantities.

4. Robustness to microstructure noise

In this section, we derive the analytical expression of the bias of the Fourier estimator of volatility of volatility due to the presence of microstructure noise, for a given sample size n and a given number of Fourier coefficients N and M included in the estimation, and we prove that the bias of the Fourier estimator converges to zero, for n, N, M increasing at suitable rates. Therefore, even if we do not proceed to any bias correction of the estimator, a suitable cutting of the highest frequencies can make the finite sample bias negligible.

We suppose that the logarithm of the observed price process is given by

$$\widetilde{p}(t) = p(t) + \eta(t) \tag{6}$$

where $p(t)$ is the efficient log price in equilibrium and $\eta(t)$ is the microstructure noise.

The following assumptions hold:

(M.I) the random shocks $\{\eta(t_i)\}_{0\le i\le n}$, for all n, are independent and identically distributed with mean zero and bounded fourth moment;

(M.II) the shocks $\{\eta(t_i)\}_{0\le i\le n}$ are independent of the price process p, for all n.

Remark 4.1 We consider here the simple case where the microstructure noise displays an MA(1) structure with a negative first-order autocorrelation. The MA(1) model is typically justified by bid-ask bounce effects (see Roll 1984). The hypothesis that the random noises are independent of the returns (see the discussion in Hansen and Lunde (2006)) is assumed here with the aim to obtain simple analytic expressions for the bias. Nevertheless, we expect that similar results would be observed under more general microstructure noise dependence, as a consequence of the robustness of the Fourier volatility estimator proved in Mancino and Sanfelici (2008) under general dependent noise structure. A specific simulation study confirming this intuition is developed in section 5.

To simplify the notation, in the sequel we will write η_i instead of $\eta(t_i)$. Denote $\delta_i(\widetilde{p}) := \widetilde{p}(t_{i+1}) - \widetilde{p}(t_i)$, where $\widetilde{p}$ is defined in (6). Then $\delta_i(\widetilde{p}) = \delta_i(p) + \varepsilon_i$, where $\varepsilon_i := \eta_{i+1} - \eta_i$.

We focus on the estimator of integrated volatility of volatility in the presence of microstructure noise defined by

$$\widetilde{\gamma}^2_{n,M,N} = \frac{(2\pi)^2}{M+1} \sum_{|j|\le M} \left(1 - \frac{|j|}{M}\right) j^2\, c_j(\widetilde{v}_{n,N}) c_{-j}(\widetilde{v}_{n,N}) \tag{7}$$

where

$$c_j(\widetilde{v}_{n,N}) = \frac{2\pi}{2N+1} \sum_{|k|\le N} c_k(d\widetilde{p}_n) c_{j-k}(d\widetilde{p}_n),$$

is the estimated jth Fourier coefficient of the volatility, given price observations contaminated by microstructure noise.

The following result contains the computation of the bias induced by the noise. For simplicity, we assume equally spaced data in the following theorem.

THEOREM 4.2 *Under the assumptions* (A.I), (A.II) *and* (M.I), (M.II), *let* $\widehat{\gamma}^2_{n,M,N}$ *and* $\widetilde{\gamma}^2_{n,M,N}$ *be defined respectively by* (5) *and* (7). *Then, it holds*

$$\begin{aligned} &E\left[\widetilde{\gamma}^2_{n,M,N} - \widehat{\gamma}^2_{n,M,N}\right] \\ &\quad = 2E\left[\eta^2\right] E\left[\int_0^{2\pi} \sigma^2(t)\mathrm{d}t\right] \Lambda(n,N,M) \\ &\qquad + 2\left(E\left[\eta^4\right] + 3E\left[\eta^2\right]\right) \Gamma(n,N,M) \\ &\qquad + 2E\left[\eta^2\right] \Psi(n,N,M), \end{aligned}$$

where $\Lambda(n,N,M)$, $\Gamma(n,N,M)$ *and* $\Psi(n,N,M)$ *are deterministic functions that go to* 0 *as* $n, N, M \to \infty$, *under the conditions* $\frac{M^2N^2}{n} \to 0$ *and* $\frac{M^2}{N} \to 0$.

Remark 4.3 From Theorems 3.2 and 4.2, the growth conditions ensuring both the consistency of the Fourier estimator of volatility of volatility (7) and its asymptotically unbiasedness in the presence of microstructure noise are that $N = O(n^\alpha)$ and $M = O(n^\beta)$ with $0 < \alpha < \frac{1}{2}$ and $0 < \beta < \frac{\alpha}{4}$.

5. Numerical results

In this section, we simulate discrete data from a continuous time stochastic volatility model with and without microstructure contaminations. From the simulated data, Fourier estimates of the integrated volatility of volatility can be compared to the value of the true quantity and to estimates obtained with other methods proposed in the literature. However, to the best of our knowledge, only very recently the literature has been focused specifically on the analysis of estimators for integrated volatility of volatility. We refer to the works of Barndorff-Nielsen and Veraart (2013), Vetter (2012), Cuchiero and Teichmann (2015). None of these contributions, however, consider the issue of microstructure effects which may be problematic in empirical applications and therefore they do not apply to a real high-frequency setting.

Another aspect that is worth mentioning is that, by their nature, all existing estimators of volatility of volatility rely on a preliminary estimation of the spot volatility path. It is well known that spot volatility estimation is particularly difficult and quite unstable, especially in the presence of microstructure effects. On the contrary, the Fourier estimator can reconstruct the Fourier coefficients of the volatility of the variance process starting from the observable log-prices. Therefore, our estimate is obtained by iterated convolutions of the Fourier coefficients of the log-returns, without resorting explicitly to any proxy of the latent spot variance of returns. We think that this can represent a strength of our approach, as it will be highlighted by the following numerical simulations.

As a benchmark for our estimator, we use the pre-estimated spot variance based realized variance of Barndorff-Nielsen and Veraart (2013), that we call *realized variance* in the following.

This estimator is consistent in the absence of microstructure frictions. To obtain roughly unbiased and valid estimates of the integrated volatility of volatility when microstructure effects play a role, we can resort to low-frequency sampling. However, the well-known bias-variance trade off comes up as sparse sampling eliminates information contained in the available data. For the reader's convenience, we recall the construction of the realized variance estimator.

Hypothetically, let us assume that we observe the volatility process σ^2 at equally spaced times $i\Delta_n$, $i = 0, 1, 2, \ldots, \lfloor T/\Delta_n \rfloor$, for some $\Delta_n > 0$ such that $\Delta_n \to 0$, as $n \to \infty$. The realized variance at time t is then defined as the sum of squared increments over the time interval $[0, t]$, for $0 \leq t \leq T$, i.e.

$$RV_t^n(\sigma^2) = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left(\Delta_i^n \sigma^2\right)^2,$$

where $\Delta_i^n \sigma^2 = \sigma^2(i\Delta_n) - \sigma^2((i-1)\Delta_n)$. Standard arguments assure that $RV_t^n(\sigma^2)$ converges in probability, uniformly on compacts, to the integrated volatility. However, since volatility is unobservable, we have to replace the squared volatility process by a consistent spot variance estimator. Barndorff-Nielsen and Veraart (2013) propose to use the locally averaged realized variance

$$\hat{\sigma}_s^2 = \frac{1}{K_n \delta_n} \sum_{i=\lfloor s/\delta_n \rfloor - K_n/2}^{\lfloor s/\delta_n \rfloor + K_n/2} \delta_i^n(p)^2,$$

where now $\delta_i^n(p) = p(i\delta_n) - p((i-1)\delta_n)$ is the ith log-return computed on a different time scale at which we observe the logarithmic asset price p, with mesh size $\delta_n > 0$. This estimator is constructed over a local window of size $K_n\delta_n$, where we require $K_n \to \infty$ such that $K_n\delta_n \to 0$. However, this only works when we estimate spot volatility on a finer time scale than the one used for computing the realized variance. Then we must assume $\delta_n < \Delta_n$. In particular, we can take

$$\Delta_n = O\left(\delta_n^C\right), \quad \text{for } 0 < C < 1,$$

and

$$K_n = O\left(\delta_n^B\right), \quad \text{for } -1 < B < 0.$$

In the presence of microstructure effects in the price process, besides sparse sampling, we can choose locally pre-averaged variance estimator to reduce the noise-induced bias as in Jacod *et al*. (2009). However, we limit our analysis to the realized variance estimator.

Vetter (2012) proposes a similar spot variance-based estimator and shows that it is possible to take $\Delta_n = K_n\delta_n$ preserving convergence at the optimal rate, provided that a bias correction is introduced. We will consider this estimator for integrated volatility of volatility as well in our analysis and we will call it *Corrected realized variance*.

In both cases, the necessary condition imposed on the choice of the time scales δ_n and Δ_n represents a limit for the efficiency of such procedures. On the one side, it requires using huge datasets of high-frequency returns, where market microstructure effects likely become manifest. On the other side, the choice of the second-level time scale Δ_n implies a loss of the information contained in the original time series.

Our simulation exercise is conducted using mainly two different stochastic volatility models. The first one is a classical stochastic volatility model, where the spot variance follows a mean-reverting square-root process. The second one is a model with stochastic volatility of volatility, namely the volatility of the variance process is driven by a second source of randomness. Our analysis is threefold. In section 5.1, we show the sensitivity of the Fourier estimator to the choice of the parameters M and N, to which the consistency of the estimator is related and we test the robustness of the estimator with respect to several noise settings. In section 5.2, we test the performance of the Fourier estimator with respect to the realized variance and the bias corrected realized variance estimators both on a standard stochastic volatility model and on a model with stochastic volatility of volatility. Finally, in sections 5.3 and 5.4, we address the issue of parameter identification of stochastic volatility models and we consider an empirical application to S&P 500 index futures.

5.1. *Parameter sensitivity and robustness to microstructure effects*

The definition of the Fourier estimator of volatility of volatility depends on the choice of two parameters characterizing the highest frequency Fourier coefficients of returns and of volatility, respectively, that enter in our estimator. We call these parameters the *cutting frequencies* at which the sums in (4) and (5) are truncated. Therefore, it is important to analyse the sensitivity of the estimator to the choice of the parameters M and N.

Let us consider a stochastic volatility model where the spot variance follows a mean-reverting square-root process. We simulate second-by-second return and variance paths over a daily trading period of $T = 6\,\text{h}$, for a total of 250 trading days and $n = 21\,600$ observations per day. The infinitesimal variation of the true log-price process and spot volatility is given by the CIR square-root model (see Cox *et al*. 1985),

$$\begin{cases} \mathrm{d}p(t) = \sigma(t)\,\mathrm{d}W(t) \\ \mathrm{d}\sigma^2(t) = \alpha(\beta - \sigma^2(t))\mathrm{d}t + \nu\sigma(t)\,\mathrm{d}Z(t), \end{cases} \tag{8}$$

where W, Z are two possibly correlated Brownian motions, with constant instantaneous correlation ρ. The parameter values used in the simulations are taken from the unpublished appendix to Bandi and Russell (2005) and reflect the features of IBM time series: $\alpha = 0.01$, $\beta = 1.0$ and $\nu = 0.05$. We take $\rho = -0.5$. The initial value of σ^2 is set equal to one, while $p(0) = \log 100$. Moreover, when microstructure effects are considered, we assume that the logarithmic noises η are Gaussian i.i.d. and independent from p; this is typical of bid-ask bounce effects in the case of exchange rates and, to a lesser extent, in the case of equities. We consider noise-to-signal ratios $\zeta = std(\eta)/std(r)$ equal to 0 in the no-noise case and to 2.5 for noisy data, where r are the 1-s returns.

In figure 1, we plot the real MSE of the Fourier estimator averaged over 250 days as a function of M and N, respectively, and of any combination (M, N) in the absence of microstructure effects. We notice that the Fourier estimator turns out to be on average quite robust to the choice of M in the interval $[0, 12]$. For larger values of M, both the MSE and bias rapidly increase. As regards to N, except for the lowest values up to about $N = 250$ and depending on M, the MSE exhibits small variability as well.

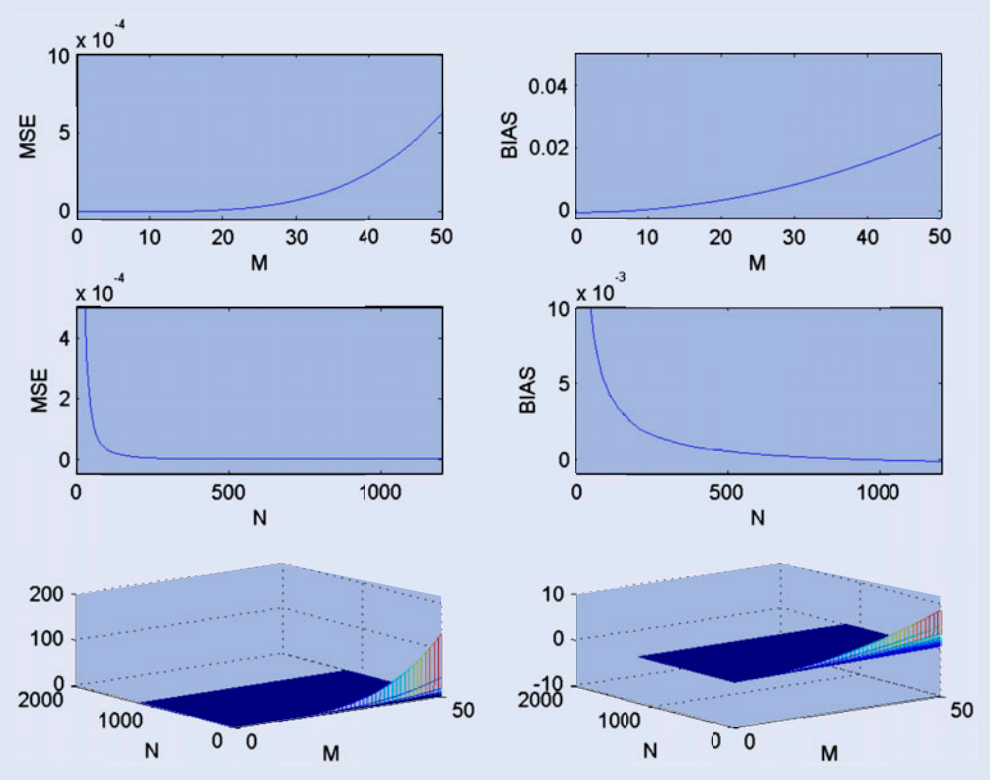

Figure 1. Real MSE and bias of the Fourier estimator of volatility of volatility averaged over the whole data-set (250 days) as a function of M and N, for the purely diffusive price process (8). True Integrated volatility of volatility 6.24e − 4.

Figure 2 shows the average MSE in the presence of i.i.d. noise, with $\zeta = 2.5$. The plots are qualitatively the same as in figure 1. We notice that the addition of noise does not seem to affect much the variability of the MSE as a function of N and the quality of estimation. However, the estimator seems to be more sensitive to the choice of M in the presence of noise than in the pure diffusive case. This is reflected by the MSE and bias, which show higher values for $M \geq 10$.

Usually, the minimum MSE is achieved for values of the cutting frequency N which turn out to be much smaller than the Nyquist frequency (i.e. $N \ll n/2$) both in the absence and in the presence of noise. Moreover, in complete agreement with the theory developed in section 3, the optimal value of M is very small. In these two simulations, we get that the optimal values of the cutting frequencies are $N = 995$, $M = 8$ and $N = 1230$, $M = 7$, respectively, and the minimum attained MSE is 5.75e − 8 and 6.63e − 8, respectively. As the noise-to-signal ratio increases, the choice of the parameter M has a more critical impact on the MSE and smaller values of M should be considered. We remark that the Fourier estimator makes use of all the n observed prices, because it reconstructs the signal in the frequency domain and therefore it can filter out microstructure effects by a suitable choice of M and N instead of reducing the sampling frequency.

Finally, we test the robustness of the Fourier estimator with respect to more general microstructure settings. Therefore, we relax both the assumptions (M.I) and (M.II) and analyse the behaviour of the Fourier estimator as a function of the sampling frequency. We consider again the model (8), with data featuring the IBM time series. Besides the case of pure diffusion, we consider three different microstructure models: the first one, denoted by UNC is the basic i.i.d. Gaussian model satisfying (M.I) and (M.II); in the second one, denoted by COR, we relax assumption (M.I) and allow first-order autocorrelation of the random shocks; in the third one, denoted by DEP, we relax assumption (M.II) and allow the random shocks $\eta(t_i)$ to be linearly dependent on the return $\delta_{i-1}(p)$, i.e. $\eta_i = \alpha_\eta \delta_{i-1}(p) + \hat{\eta}_i$ with $\hat{\eta}_i$ Gaussian i.i.d. random variables.

Table 1 lists the MSE of the Fourier estimates of the volatility of volatility as a function of the sampling frequency ranging from 1 s to 4 min. The parameters N and M of the Fourier estimator must be chosen conveniently. One possible criterion is the minimization of the true MSE. This procedures is unfeasible when applied to empirical data, where the actual volatility path is not observed. However, to evaluate the robustness of the estimator to different noise settings, we select the optimal parameters N and M by minimizing the true average MSE over 250 days.

We notice that in all the settings the optimal choice of the cutting frequencies M and N keeps the MSE low. In particular, the lowest MSE is always achieved at the highest frequency (1 s). This is due to the robustness of the Fourier estimator to microstructure effects which allows the method to use high-frequency data without resorting to sparse sampling.

Figure 3 shows the optimal cutting frequencies M and N as a function of the number of observations n and of the sampling interval $\rho(n)$. The presence of microstructure noise of any kind yields optimal values of both N and M that are lower than for the pure diffusive model. However, it has a larger effect on the choice of M rather than N. By inspecting both the MSE

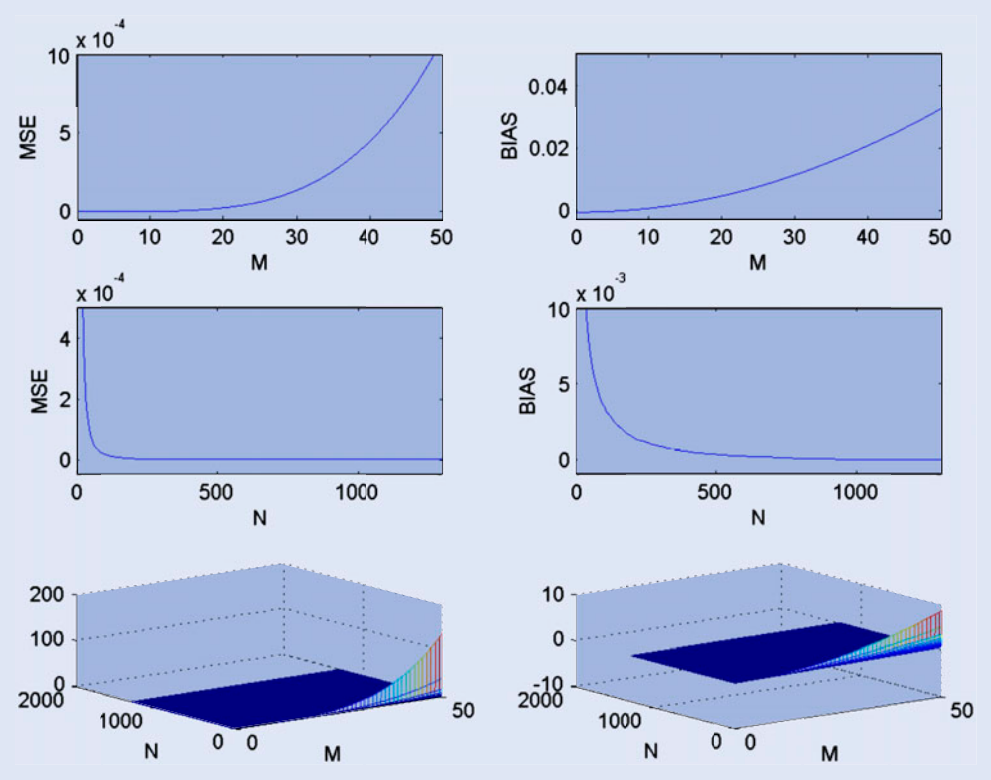

Figure 2. Real MSE and bias of the Fourier estimator of volatility of volatility averaged over the whole data-set (250 days) as a function of M and N, in the presence of microstructure effects, with $\zeta = 2.5$. True Integrated volatility of volatility 6.24e − 4.

Table 1. MSE of the Fourier volatility of volatility estimates under general noise settings.

	Fourier estimator MSE $\times 1.0e-6$						
Sampling freq.	1 s	15 s	30 s	1 m	2 m	3 m	4 m
NO NOISE	0.0547	0.0814	0.1010	0.1485	0.2001	0.1869	0.2680
UNC	0.0675	0.1060	0.1241	0.1609	0.2044	0.1792	0.2836
COR	0.0806	0.1224	0.1403	0.1675	0.1778	0.1931	0.2782
DEP	0.0679	0.1073	0.1237	0.1611	0.2043	0.1794	0.2851

Notes: Parameter values: $\alpha = 0.01$, $\beta = 1.0$, $\nu = 0.05$, $\rho = -0.5$, $\sigma^2(0) = 1$, $p(0) = \log 100$. When microstructure effects are considered, we consider a noise-to-signal ratio $\zeta = 2.5$. Moreover, in the case of autocorrelated noise, we assume a first-order autocorrelation coefficient $\rho_\eta = 0.5$, while in the case of dependent noise we assume $\alpha_\eta = 0.1$. True integrated volatility of volatility 6.240255e − 4.

in table 1 and the optimal choice of M and N in figure 3, we notice that the most problematic setting is provided by the case of correlated noise (COR), which entails smaller values of M and N in order to filter the microstructure effects.

5.2. *Fourier method efficiency*

Let us now consider the classical Heston model (see Heston 1993),

$$\begin{cases} \mathrm{d}p(t) = (\mu - \sigma^2(t)/2)\mathrm{d}t + \sigma(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}\sigma^2(t) = \alpha(\beta - \sigma^2(t))\mathrm{d}t + \nu\sigma(t)\,\mathrm{d}W_2(t), \end{cases} \tag{9}$$

where we assume the same data as in Vetter (2012), i.e. $\alpha = 5$, $\beta = 0.2$, $\nu = 0.5$, $\mu = 0.3$ and $\rho = -0.2$, which corresponds to a moderate leverage effects. Furthermore, we set $p(0) = 0$ and $\sigma_0^2 = \beta$. The trading period is set to $T = 1$ day. We generate $n = 10\,000$ daily observations, corresponding to a trading frequency of 8.64 s.

The sampling frequency δ_n and the other parameters M, N and K_n contained in the definition of the estimators considered in our analysis must be chosen conveniently, especially in the presence of noise. One possible criterion is the minimization of the true MSE. Another possible choice is the minimization of the expected asymptotic error variance. Both these procedures are unfeasible when applied to empirical data, where the actual volatility path is not observed. However, to evaluate the highest efficiency level that can be achieved by the analysed estimators, we select optimal parameters by minimizing the average MSE over 250 days. Table 2 displays the results of our analysis.

First, let us consider the case with no microstructure effects, i.e. $\zeta = 0.0$. The Fourier estimator is optimized with respect to M and N by minimizing the true MSE over a grid of discrete values of these parameters. Similarly, the optimal MSE-based realized variance estimator is obtained by choosing $\delta_n = 1/n$, $\Delta_n = \delta_n K_n/2$ and letting K_n vary in a suitable range of integer values around $2\sqrt{n}$. More precisely, the spot volatility

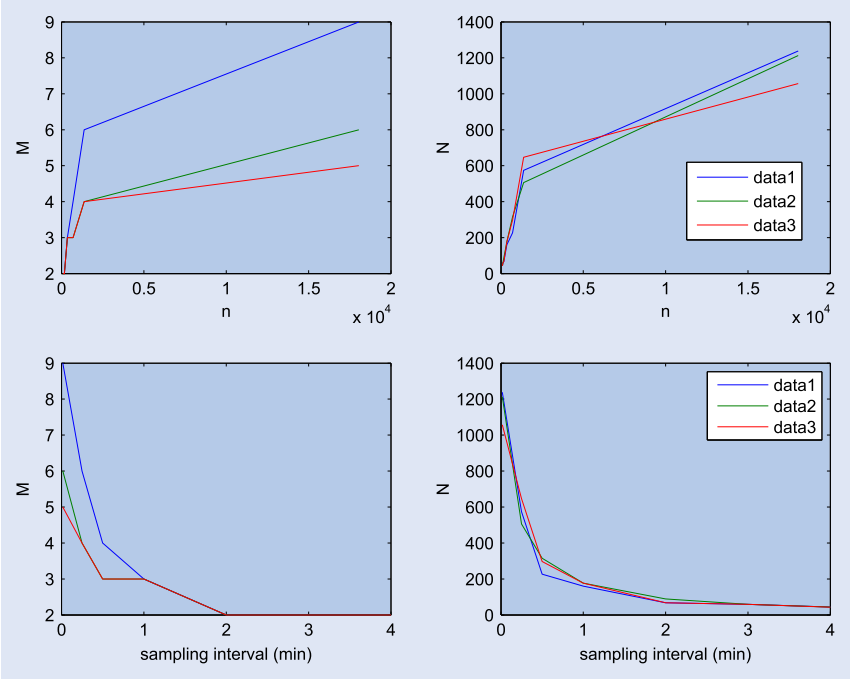

Figure 3. Optimal cutting frequencies M and N as a function of the number of observations n and of the sampling interval $\rho(n)$. 'data1' corresponds to pure diffusion; 'data2' corresponds to UNC and DEP noise settings; 'data3' corresponds to COR noise setting.

Table 2. Optimization procedures based on the minimization of the true average MSE (250 days, $n = 10\,000$ observations per day). True integrated vol of Vol 4.091552e − 2.

	Fourier–Fejer		Realized variance		C-Realized variance	
Noise-to-signal ratio	MSE	BIAS	MSE	BIAS	MSE	BIAS
$\zeta = 0.0$	1.39e − 4	−4.21e − 3	1.00e − 4	−3.58e − 3	9.68e − 4	−1.33e − 3
$\zeta = 0.5$	1.37e − 4	−5.93e − 3	1.46e − 4	−3.77e − 3	4.60e − 3	−1.05e − 2
$\zeta = 1.5$	1.26e − 4	−5.43e − 3	1.69e − 4	−2.17e − 3	5.55e − 3	−1.32e − 2
$\zeta = 2.5$	5.88e − 5	−1.61e − 3	9.79e − 5	−3.08e − 3	7.82e − 3	−1.53e − 2
$\zeta = 3.5$	7.32e − 5	−8.66e − 4	1.15e − 4	−2.14e − 3	1.26e − 2	−1.37e − 2

trajectory is estimated using tick-by-tick observations, while the realized variance of volatility is estimated at the frequency Δ_n corresponding to $K_n/2$ ticks, where the parameter K_n is chosen in order to minimize the daily MSE. The bias-corrected realized variance is constructed by choosing $K_n = \sqrt{n}$, as in section 4 of Vetter (2012). The Fourier and the realized variance estimators both provide low MSE and bias. The corrected realized variance performance is slightly worse in terms of MSE, although still acceptable, and provides the smallest bias. The optimal cutting frequency for the Fourier estimator are $N = 322$ e $M = 48$, while the optimal value for the window size in the realized variance estimator is $K_n = 240$ which entails $\Delta_n = 120\delta_n \cong 17$ m.

The case with microstructure effects is reported in table 2 as well. We consider four different levels of noise-to-signal ratio $\zeta = 0.5, 1.5, 2.5, 3.5$ and $\rho = -0.2$. The Fourier estimator is again optimized with respect to M and N by minimizing the true MSE over a grid of discrete values of these parameters. The realized variance estimator is not robust to microstructure noise; therefore, we have to resort to sparse sampling to keep the bias due to market microstructure low. The daily MSE is optimized with respect to both K_n and the sampling frequency δ_n at which the spot volatility path is estimated. However, we remark that sparse sampling may produce a loss of the rich information contained in the original high-frequency data-set. On the contrary, the Fourier estimator uses all the available data and seems to be invariant to the presence of increasing levels of noise.

As the noise-to-signal ratio increases, the optimal sampling frequency δ_n for the realized variance estimator δ_n passes from 276 to 492 s. The corresponding values for the second-level sampling interval Δ_n range from approximately 74 to 98 min. This keeps the bias of the realized variance estimator quite small, at the expenses of a slightly larger MSE. However, for $\zeta = 2.5$ and $\zeta = 3.5$ its performance gets worse than with the Fourier estimator.

The bias-corrected realized variance estimator shows a very poor performance. Using all the available data at the highest frequency would produce completely unreliable estimates, due to microstructure effects; for instance, in the case $\zeta = 1.5$, we would get an average MSE equal to 2.63 with average bias equal to 1.43. When the estimator is optimized in terms

of the MSE, then the log-prices are optimally sampled at the frequency of 467 s and K_n is chosen accordingly as the square root of the number of data in the sample; however, both bias and MSE of the bias-corrected realized variance estimator are still very large compared to the Fourier and realized variance estimators.

Finally, we consider a model with stochastic volatility of volatility, namely the volatility of the variance process is driven by a second source of randomness. This model expresses the possibility or fact that there is greater variability in the data structure that cannot be described by classical stochastic volatility models. We consider the following data-generating process

$$\begin{cases} \mathrm{d}p(t) = \sigma(t)\mathrm{d}W_1(t) \\ \mathrm{d}v(t) = \alpha_v(\beta_v - v(t))\mathrm{d}t + \gamma(t)\,\mathrm{d}W_2(t), \\ \mathrm{d}\gamma^2(t) = \alpha_\gamma(\beta_\gamma - \gamma^2(t))\mathrm{d}t + \nu_\gamma\gamma(t)\,\mathrm{d}W_3(t), \end{cases} \tag{10}$$

where W_1, W_2 are Brownian motions with correlation ρ, i.e. $\mathrm{d}\langle W_1, W_2\rangle_t = \rho\mathrm{d}t$ and W_3 is a third independent Wiener process. The process $(\gamma_t^2)_{t\geq 0}$ is driven by a CIR process and can be interpreted as the stochastic variability of variance. The Feller condition guarantees that both processes $v = \sigma^2$ and γ^2 stay positive. The parameter values used in the simulations are: $\alpha_v = 1.0$, $\beta_v = 1.0$, $\alpha_\gamma = 0.01$, $\beta_\gamma = 0.01$ and $\nu_\gamma = 0.0005$. We take $\rho = -0.5$. The initial value of σ^2 is set equal to one, while $p(0) = 0$ and $\gamma^2(0) = \beta_\gamma$. The trading period is set to $T = 1$ day and we generate $n = 10\,000$ daily observations, corresponding to a trading frequency of 8.64 s. We assume no microstructure effects and therefore no sparse sampling is needed when estimating the spot volatility path at the first level (i.e. $\delta_n = 1/n$). Numerical results are shown in table 3.

We notice that the Fourier estimator provides good estimates both in terms of bias and MSE. The optimal selected value of N is larger than for the Heston model, while the optimal value of M is smaller. The realized variance estimator seems to provide rather good estimates as well, but less efficient than the Fourier estimator. However, this is achieved by choosing a huge value of the parameter $K_n = 1686$, namely the time scale at which the second-level realized variance is computed is around $\delta_n K_n/2$ s, i.e. two hours. This has strong effects on the efficiency of the estimator, as it can be seen from figure 4 showing the histograms of the relative error $(\hat{\gamma}^2_{n,N,M} - \int_0^{2\pi}\gamma^2(t)\mathrm{d}t)/\int_0^{2\pi}\gamma^2(t)\mathrm{d}t$. The mean and standard deviation of the relative error for the realized variance is much larger than for the Fourier estimator.

Finally, looking again at table 3, the non optimized bias-corrected realized variance estimator with $K_n = 100$ shows a very poor performance. The inefficiency of both the realized variance estimators can be ascribed to the necessary condition imposed on the choice of the time scales δ_n and Δ_n. As already observed, the choice of the second-level time scale Δ_n implies a loss of the information contained in the original time series.

5.3. *Parameter identification of SV models*

Let us consider now the issue of parameter identification of stochastic volatility models. Suppose that the data-generating process for the log-price dynamics is the Heston model

$$\begin{cases} \mathrm{d}p(t) = \mu\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t) \\ \mathrm{d}\sigma^2(t) = \alpha(\beta - \sigma^2(t))\mathrm{d}t + \nu\sigma(t)\,\mathrm{d}Z(t), \end{cases}$$

where W and Z are two possibly correlated Brownian motions. Then, we can use our estimates to identify parameters of the stochastic volatility model from a finite sample. Using simple tools of Itô calculus, we can derive the following identity

$$\nu^2\sigma^2(t) = \gamma^2(t).$$

Therefore,

$$\nu^2\int_0^T \sigma^2(t)\mathrm{d}t = \int_0^T \gamma^2(t)\mathrm{d}t.$$

Using the Fourier analysis methodology, we get the following estimate of the parameter ν

$$\hat{\nu} = \left(\frac{\hat{\gamma}^2_{n,N,M}}{2\pi c_0(v_{n,N})}\right)^{\frac{1}{2}},$$

where $\hat{\gamma}^2_{n,N,M}$ is defined by (5) and $c_0(v_{n,N})$ by (4).

Therefore, by using the provided Fourier estimates of the integrated volatility and of the volatility of volatility, we can obtain estimations of the parameter ν identifying the diffusion component. This parameter does not change under equivalent measure changes and can be used to specify a model for purposes of pricing, hedging and risk management.

The data used in our simulations are taken from Barucci and Mancino (2010): $\alpha = 0.03$, $\beta = 0.25$, $\nu = 0.1$, $\mu = 0$ and $\rho = -0.2$. We simulate second-by-second return and variance paths over a daily trading period of $T = 6$ h, for a total of 100 trading days and $n = 21600$ observations per day. Numerical results are shown in table 4.

The table lists the estimated value of $\hat{\nu}$, obtained with the Fourier and the realized variance estimator, together with the relative error of the estimate, denoted by rel. error. Moreover, the table shows the relative bias and relative RMSE of the estimate $\hat{\gamma}^2_{n,N,M}$, defined as

$$\text{Rel. Bias} = E\left[\frac{\hat{\gamma}^2_{n,N,M} - \int_0^{2\pi}\gamma^2(t)\mathrm{d}t}{\int_0^{2\pi}\gamma^2(t)\mathrm{d}t}\right],$$

$$\text{Rel. RMSE} = \left(E\left[\left(\frac{\hat{\gamma}^2_{n,N,M} - \int_0^{2\pi}\gamma^2(t)\mathrm{d}t}{\int_0^{2\pi}\gamma^2(t)\mathrm{d}t}\right)^2\right]\right)^{1/2}.$$

We consider two different simulations: the first one with no microstructure effects and the second one with a noise-to-signal ratio $\zeta = 1.5$. In both cases, the performance of the Fourier estimator is better than the one of the realized variance estimator. In particular, the relative bias achieved with the Fourier estimator is one order of magnitude less than with the realized variance and the relative error over the Fourier estimated $\hat{\nu}$ is half the value obtained by the realized variance. We notice that, as in the previous section, in the case of no microstructure effects the optimal value of the parameter K_n is equal to 1680, namely the time scale at which the realized variance is computed is around $K_n/2$ s, i.e. 14 min, while the Fourier estimator uses second-by-second returns.

5.4. *An empirical application: the S&P 500 index futures*

We consider now a case study based on tick-by-tick data of the S&P 500 index futures recorded at the Chicago Mercantile

Table 3. Stochastic volatility of volatility model. Optimization procedures based on the minimization of the true average MSE (250 days, $n = 10\,000$ observations per day). True integrated vol of Vol 1.000018e − 2.

	Fourier-Fejer		Realized Variance		C-Variance	
No microstructure	MSE	BIAS	MSE	BIAS	MSE	BIAS
$\zeta = 0.0$	1.51e − 5	−1.71e − 3	9.81e − 5	5.86e − 3	5.16e − 1	−1.55e − 1
Parameter values	$N = 1180$	$M = 8$	$K_n = 1686$		$K_n = 100$	

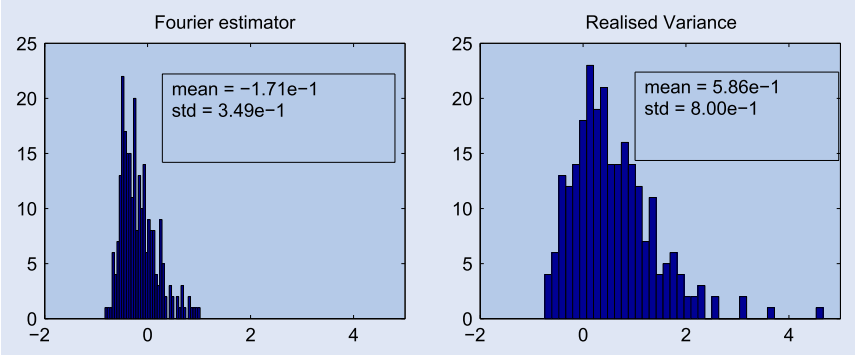

Figure 4. Stochastic volatility of volatility model. Histograms of the relative error $\left(\hat{\gamma}^2_{n,N,M} - \int_0^{2\pi} \gamma^2(t)\mathrm{d}t\right) / \int_0^{2\pi} \gamma^2(t)\mathrm{d}t$.

Table 4. Stochastic volatility model calibration. True value $\nu = 0.1$.

	Fourier-Fejer				Realized variance			
	$\hat{\nu}$	Rel. Error	Rel. Bias	Rel. RMSE	$\hat{\nu}$	Rel. Error	Rel. Bias	Rel. RMSE
$\zeta = 0.0$	0.0972	2.79e − 2	−4.35e − 2	2.16e − 1	0.1059	5.94e − 2	1.52e − 1	4.11e − 1
$\zeta = 1.5$	0.0964	3.60e − 2	−6.10e − 2	2.07e − 1	0.0921	7.92e − 2	−1.35e − 1	2.81e − 1

Table 5. Summary statistics for the sample of the traded CME S&P 500 index futures in the period 2 January 1990 to 29 December 2006 (11 611 297 trades). 'Std. Dev.' denotes the sample standard deviation of the variable.

Variable	Mean	Std. Dev.	Min	Max
S&P 500 index futures	893.97	366.24	295.60	1574.00
N. of ticks per minute	7.0433	3.5276	1	56

Exchange (CME). The sample covers the period from 2 January 1990 to 29 December 2006, a period of 4274 trading days, having 11 611 297 tick-by-tick observations. Table 5 describes the main features of our data set.

High-frequency returns are contaminated by transaction costs, bid-and-ask bounce effects, etc. leading to biases in the variance measures. Therefore, data filtering is necessary. Days with trading period shorter than 5 h have been removed. Jumps have been identified and measured using the Threshold Bipower Variation method (TBV) of Corsi *et al.* (2010), which is based on the joint use of bipower variation and threshold estimation of Mancini (2009). This method provides a powerful test for jump detection, which is employed at the significance level of 99.9%. We refer the reader to Mancino and Sanfelici (2012) for further details on the jump removal procedure. The number of days remaining after jump removal and filtering is 3078, for a total of 8 575 527 tick-by-tick data. The contribution coming from overnight returns is neglected.

Sparse sampling needed for the realized variance estimator can be performed either in *calendar time*, for instance with prices sampled every 5 or 15 min, or in *transaction time*, where prices are recorded every mth transaction. When we sample in calendar time, the x-min returns are constructed using the nearest neighbour to the x-minute tag. Figure 5 shows the average realized variance over the full sample period constructed for different sampling frequencies in calendar (Panel A) and transaction time (Panel B).

The volatility signature plots clearly indicate that the bias induced by market microstructure effects is relatively small for the highly liquid S&P 500 index futures, and dies out very quickly. Note that with a transaction taking place on average about every 8.57 s, the 1-min sampling interval corresponds to

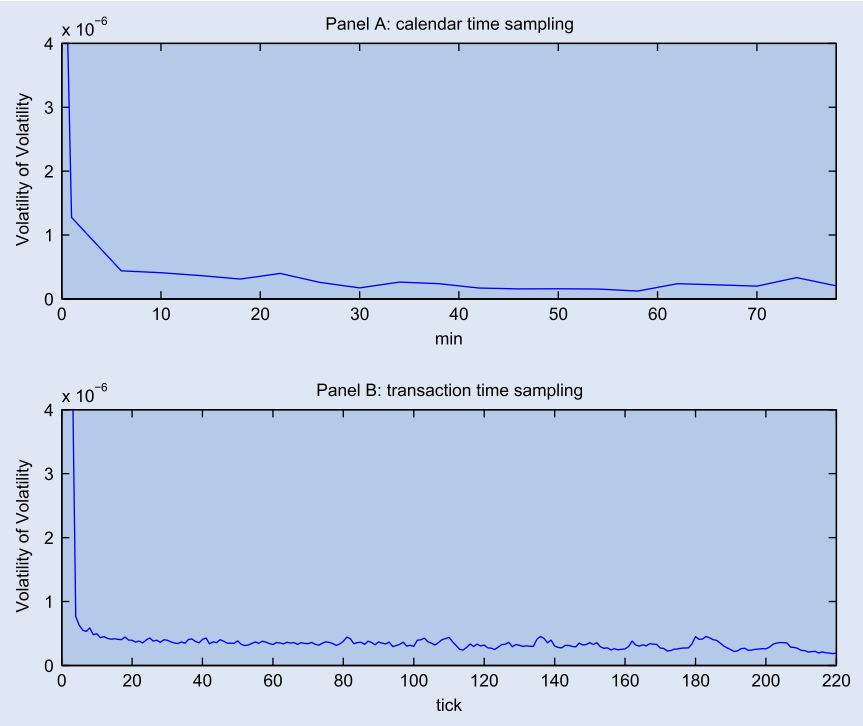

Figure 5. Realized variance: volatility signature plot of the S&P 500 index futures constructed over the full sample period. The graph shows average integrated volatility of volatility constructed for different frequencies measured in minutes (Panel A) and in number of ticks (Panel B). Note that there are about 8.57 s on average between trades, so that the average annualized 5-min based realized volatility corresponds to around the 35th tick.

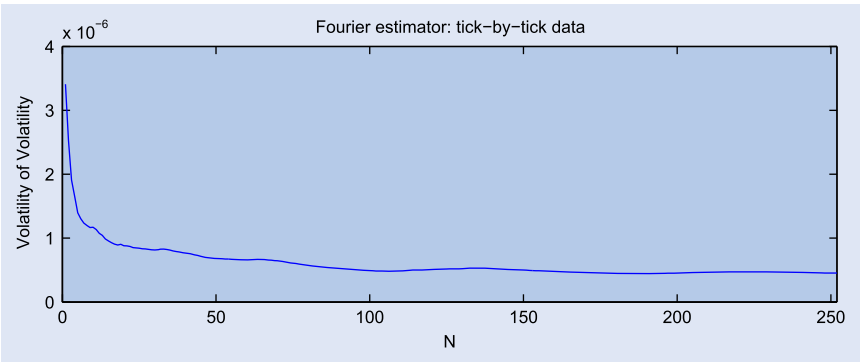

Figure 6. Fourier estimator: volatility signature plot of the S&P 500 index futures constructed over the full sample period. The graph shows average integrated volatility of volatility computed by means of the Fourier estimator, using tick-by-tick data, as a function of the parameter N. M is set equal to 3.

around the 7th tick presented in the figure, with large variability across the whole dataset. The impact of market microstructure effects on the five-min realized volatility measure for the S&P 500 index futures over the period from 1990 to 2006 can therefore be regarded as negligible. However, the estimates obtained by calendar time sampling are quite unstable and variable as the sampling frequency decreases. When sampling in transaction time, the most stable estimates are obtained for frequencies between 20 and 70 ticks that roughly correspond to 3–10 min. In both cases, for low frequencies the realized variance estimator becomes downwards biased because sparse sampling has a severe impact on the cardinality of the database. In particular, for any value of n we choose $K_n = 2\sqrt{n}$. This implies that most of the data are neglected when estimating the second-order quantities so that the volatility of volatility estimates are poor, especially when we start from sparse sampled data.

In figure 6, we plot the volatility signature plot as computed by means of $\tilde{\gamma}^2_{n,N,M}$ using tick-by-tick data, as a function of the parameter N. The value of the parameter M is set to 3. We can see that for N larger than 150 the estimates become much stable. Taking into account the mathematical properties of the Fourier estimator, when the trading period is $T = 6.5$ h, a parameter value of say $N = 200$ corresponds to sampling frequencies of $T/(2N) = 390/400 = 0.9750$ min, in the sense that the spectral decomposition in the frequency space allows to detect phenomena happening at the frequency of about 1 min, much higher than with the realized variance estimator.

6. Conclusions

We have introduced a new non-parametric estimator of the stochastic volatility of volatility which is particularly suited to work with high-frequency data. Our estimator is obtained in two steps: first we compute the Fourier coefficients of the volatility process using high-frequency observations of the log-returns, then we iterate the procedure with a convolution of the Fourier coefficients of the volatility process. An advantage of our method lies in the fact that it does not resort to the estimation of the path of the latent variance of returns, but it needs only integrated quantities. A theoretical and numerical study of the properties of our estimator highlights that cutting out the highest frequencies in the Fourier expansion makes this estimator robust to the presence of high-frequency noise components.

In conclusion, our discussion and numerical simulations show that the Fourier estimator of the volatility of volatility is robust to microstructure effects and efficient in finite samples.

Acknowledgements

We wish to thank Frederi Viens and an anonymous referee for their insightful comments and remarks. We thank the participants in the 5th Annual Modeling High Frequency Data in Finance conference at Stevens Institute, NJ, USA, in October 2013, for their interesting and stimulating discussion.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bandi, F.M. and Russell, J.R., Microstructure noise, realized variance and optimal sampling. Working Paper, University of Chicago, 2005. Available online at: http://faculty.chicagogsb.edu/federicobandi.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N., Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 2008, **76**(6), 1481–1536.

Barndorff-Nielsen, O.E. and Veraart, A.E.D., Stochastic volatility of volatility and variance risk premia. *J. Financ. Econ.*, 2013, **11**(1), 1–46.

Barucci, E. and Mancino, M.E., Computation of volatility in stochastic volatility models with high frequency data. *Int. J. Theor. Appl. Finance*, 2010, **13**(5), 1–21.

Corsi, F., Pirino, D. and Renó, R., Threshold bipower variation and the impact of jumps on volatility forecasting. *J. Econ.*, 2010, **159**(2), 276–288.

Cox, J.C., Ingersoll, J.E. and Ross, S.A., A theory of the term structure of interest rates. *Econometrica*, 1985, **53**, 385–408.

Cuchiero, C. and Teichmann, J., Fourier transform methods for pathwise covariance estimation in the presence of jumps. *Stoch. Proc. Appl.*, 2015, **125**(1), 116–160.

Curato, I. and Sanfelici, S., Measuring the leverage effect in a high-frequency trading framework. In *Handbook of High Frequency Trading*, edited by G.N. Gregoriou, pp. 425–446, 2015 (Elsevier: Plattsburgh, NY).

Hansen, P.R. and Lunde, A., Realized variance and market microstructure noise (with discussions). *J. Bus. Econ. Stat.*, 2006, **24**, 127–218.

Heston, S., A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud.*, 1993, **6**, 327–343.

Hull, J. and White, A., The pricing of options on assets with stcohastic volatilities. *J. Finance*, 1987, **42**, 281–300.

Jacod, J., Li, Y., Mykland, P.A., Podolskij, M. and Vetter, M., Microstructure noise in the continuous case: The pre-averaging approach. *Stoch. Proc. Appl.*, 2009, **119**, 2249–2276.

Malliavin, P. and Mancino, M.E., Fourier series method for measurement of multivariate volatilities. *Finance Stoch.*, 2002a, **4**, 49–61.

Malliavin, P. and Mancino, M.E., Instantaneous liquidity rate, its econometric measurement by volatility feedback. C.R. *Acad. Sci. Paris*, 2002b, **334**, 505–508.

Malliavin, P. and Mancino, M.E., A Fourier transform method for nonparametric estimation of multivariate volatility. *Ann. Stat.*, 2009, **37**(4), 1983–2010.

Mancini, C., Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.*, 2009, **36**, 270–296.

Mancino, M.E. and Sanfelici, S., Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. *Comput. Stat. Data Anal.*, 2008, **52**(6), 2966–2989.

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

Roll, R., A simple measure of the bid-ask spread in an efficient market. *J. Finance*, 1984, **39**, 1127–1139.

Stein, E. and Stein, J., Stock price distributions with stochastic volatility: Analytic approach. *Rev. Financ. Stud.*, 1991, **4**, 727–752.

Vetter, M., Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. Working Paper, 2012. Available from: http://arxiv.org/abs/1206.5761v1.

Appendix 1. Proofs

Along the proofs, C will denote a constant, not necessarily the same at the different occurrences.

Proof of Theorem 3.2. Under the model assumption (A.I)–(A.II), it is not restrictive to assume that the volatility process $v(t)$ is a.s. bounded.

We split

$$\widehat{\gamma}^2_{n,N,M} - \int_0^{2\pi} \gamma^2(t)\mathrm{d}t$$

as

$$\frac{(2\pi)^2}{M+1} \sum_{|j|\le M} \left(1 - \frac{|j|}{M}\right) j^2 \left[c_j(v_{n,N})c_{-j}(v_{n,N}) - c_j(v)c_{-j}(v)\right] \tag{A1}$$

$$+ \frac{(2\pi)^2}{M+1} \sum_{|j|\le M} \left(1 - \frac{|j|}{M}\right) j^2 \, c_j(v)c_{-j}(v) - \int_0^{2\pi} \gamma^2(t)\mathrm{d}t. \tag{A2}$$

Consider (A1). For any $|j| \le M$

$$\begin{aligned} &E[|c_j(v_{n,N})c_{-j}(v_{n,N}) - c_j(v)c_{-j}(v)|^2] \\ &\quad \le 2\Big(E[|\, c_j(v_{n,N})(c_{-j}(v_{n,N}) - c_{-j}(v))|^2] \\ &\qquad + E[|c_{-j}(v)(c_j(v_{n,N}) - c_j(v))|^2]\Big). \end{aligned}$$

By the definition (4), we write $c_j(v_{n,N}) - c_j(v)$ as the sum of the following two terms

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}j\phi_n(t)} v(t)\mathrm{d}t - \frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}jt} v(t)\mathrm{d}t \tag{A3}$$

$$+ \frac{1}{2\pi}\int_0^{2\pi}\int_0^t e^{-\mathrm{i}j\phi_n(u)} D_N(\phi_n(t) - \phi_n(u))\mathrm{d}p(u)\mathrm{d}p(t)$$

$$+ e^{-\mathrm{i}j\phi_n(t)} D_N(\phi_n(t) - \phi_n(u))\mathrm{d}p(u)\mathrm{d}p(t), \tag{A4}$$

where $\phi_n(t) = \sup\{t_k : t_k \le t\}$ and $D_N(x)$ is the rescaled Dirichlet kernel defined by $D_N(x) = \frac{1}{2N+1}\sum_{|k|\le N} e^{ikx}$.

Consider (A3)

$$E\left[\left|\frac{1}{2\pi}\int_0^{2\pi} e^{-\mathrm{i}jt}\left(1-e^{-\mathrm{i}j(\phi_n(t)-t)}\right)v(t)\mathrm{d}t\right|^2\right]$$
$$\le (ess\sup\|v\|_\infty)^2\frac{1}{2\pi}\int_0^{2\pi}|1-e^{-\mathrm{i}j(\phi_n(t)-t)}|^2\mathrm{d}t$$
$$\le (ess\sup\|v\|_\infty)^2 j^2\rho(n)^2.$$

Consider (A4)

$$E\left[\left|\frac{1}{2\pi}\int_0^{2\pi}\int_0^t e^{-\mathrm{i}j\phi_n(u)}D_N(\phi_n(t)-\phi_n(u))\mathrm{d}p(u)\mathrm{d}p(t)\right|^2\right]$$
$$\le (ess\sup\|v\|_\infty)^2\frac{1}{(2\pi)^2}\int_0^{2\pi}\int_0^t D_N^2(\phi_n(t)-\phi_n(u))\mathrm{d}u\mathrm{d}t$$
$$\le (ess\sup\|v\|_\infty)^2\frac{1}{2\pi}\frac{1}{2N+1}.$$

Therefore,

$$E\left[|c_j(v_{n,N})-c_j(v)|^2\right]$$
$$\le 2(ess\sup\|v\|_\infty)^2\left\{j^2\rho(n)^2+\frac{1}{\pi}\frac{1}{2N+1}\right\}.$$

Finally, for any $|j|\le M$

$$E\left[|c_{-j}(v)(c_j(v_{n,N})-c_j(v))|^2\right]$$
$$\le 2(ess\sup\|v\|_\infty)^4\left\{M^2\rho(n)^2+\frac{1}{\pi}\frac{1}{2N+1}\right\}. \tag{A5}$$

Consider now

$$E\left[|c_j(v_{n,N})(c_{-j}(v_{n,N})-c_{-j}(v))|^2\right]$$
$$\le 2\Big(E\left[|c_{-j}(v_{n,N})-c_{-j}(v)|^4\right]$$
$$+E\left[|c_j(v)(c_{-j}(v_{n,N})-c_{-j}(v))|^2\right]\Big). \tag{A6}$$

The second addend has been estimated in (A5). For the first addend consider the decomposition of $c_{-j}(v_{n,N})-c_{-j}(v)$ into (A3) and (A4): then a similar argument using Burkholder–Davis–Gundy inequality for the estimation of the fourth moment gives

$$E\left[|c_{-j}(v_{n,N})-c_{-j}(v)|^4\right]$$
$$\le C(ess\sup\|v\|_\infty)^4\left\{j^4\rho(n)^4+\frac{1}{\pi}\frac{1}{2N+1}\right\}.$$

Therefore, for any $|j|\le M$

$$E\left[|c_j(v_{n,N})(c_{-j}(v_{n,N})-c_{-j}(v))|^2\right]$$
$$\le C(ess\sup\|v\|_\infty)^4\{M^4\rho(n)^4+M^2\rho(n)^2+\frac{1}{2N+1}\}. \tag{A7}$$

By (A5) and (A7), for any $|j|\le M$

$$E[|c_j(v_{n,N})c_{-j}(v_{n,N})-c_j(v)c_{-j}(v)|^2]$$
$$\le C(ess\sup\|v\|_\infty)^4\left(M^2\rho(n)^2+M^4\rho(n)^4+\frac{1}{2N+1}\right).$$

Finally, the L^2 norm of (A1) is less or equal to

$$C\,M^4\,(ess\sup\|v\|_\infty)^4\left(M^2\rho(n)^2+M^4\rho(n)^4+\frac{1}{2N+1}\right),$$

which goes to 0 under the hypothesis $\rho(n)N\to 0$ and $\frac{M^4}{N}\to 0$.

Consider (A2). Using assumption (A.III), the periodic extension of $v(t)$ to $\mathbb{R}$ with period 2π (which we still denote by $v(t)$) satisfies $v(2\pi)-v(0)=0$ a.s. Therefore, applying Itô formula, we have

$$\frac{(2\pi)^2}{M+1}\sum_{|j|\le M}\left(1-\frac{|j|}{M}\right)j^2c_j(v)c_{-j}(v)-\int_0^{2\pi}\gamma^2(t)\mathrm{d}t$$
$$=2\int_0^{2\pi}\int_0^t F_M(s-t)\mathrm{d}v(s)\mathrm{d}v(t),$$

where $F_M(x)$ denotes the rescaled Fejer kernel $F_M(x)=\frac{1}{M+1}\times\sum_{|j|\le M}\left(1-\frac{|j|}{M}\right)e^{\mathrm{i}jx}$.

Then, we have

$$E\left[\left(\int_0^{2\pi}\int_0^t F_M(s-t)\mathrm{d}v(s)\mathrm{d}v(t)\right)^2\right]$$
$$=E\left[\int_0^{2\pi}\left(\int_0^t F_M(s-t)\mathrm{d}v(s)\right)^2\gamma^2(t)\mathrm{d}t\right]$$
$$\le E\left[\int_0^{2\pi}\gamma^4(t)\mathrm{d}t\right]^{\frac12}E\left[\int_0^{2\pi}\left(\int_0^t F_M(s-t)\mathrm{d}v(s)\right)^4\mathrm{d}t\right]^{\frac12}.$$

Applying Burkholder–Davis–Gundy inequality, we get

$$E\left[\int_0^{2\pi}\left(\int_0^t F_M(s-t)\mathrm{d}v(s)\right)^4\mathrm{d}t\right]$$
$$\le C\int_0^{2\pi}\int_0^t F_M^4(s-t)\mathrm{d}s\,\mathrm{d}t\;E\left[\int_0^{2\pi}\gamma^4(s)\mathrm{d}s\right]$$
$$\le C\frac{1}{M+1}E\left[\int_0^{2\pi}\gamma^4(s)\mathrm{d}s\right].$$

Finally, the L^2 norm of (A2) is less or equal to

$$C\frac{1}{\sqrt{M+1}}E\left[\int_0^{2\pi}\gamma^4(s)\mathrm{d}s\right].$$

□

Proof of Theorem 4.2. For any fixed j, $|j|\le M$, we have

$$E\left[c_j(\widetilde{v}_{n,N})c_{-j}(\widetilde{v}_{n,N})-c_j(v_{n,N})c_{-j}(v_{n,N})\right]$$
$$=E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\mathrm{d}p_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\mathrm{d}p_n)\right] \tag{A8}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\mathrm{d}p_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\mathrm{d}p_n)c_{-j-l}(\varepsilon_n)\right] \tag{A9}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\mathrm{d}p_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\varepsilon_n)\right] \tag{A10}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\mathrm{d}p_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\mathrm{d}p_n)\right] \tag{A11}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\mathrm{d}p_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\mathrm{d}p_n)c_{-j-l}(\varepsilon_n)\right] \tag{A12}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\mathrm{d}p_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\varepsilon_n)\right] \tag{A13}$$
$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\mathrm{d}p_n)\right] \tag{A14}$$

$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\mathrm{d}p_n)c_{-j-l}(\varepsilon_n)\right] \tag{A15}$$

$$+E\left[\frac{2\pi}{2N+1}\sum_{|h|\le N}c_h(\varepsilon_n)c_{j-h}(\varepsilon_n)\frac{2\pi}{2N+1}\sum_{|l|\le N}c_l(\varepsilon_n)c_{-j-l}(\varepsilon_n)\right] \tag{A16}$$

The terms (A8), (A9), (A11) and (A12) are similar. Consider (A8): it is equal to

$$E\Big[\frac{1}{2\pi}\sum_{u,u'}D_N(t_u-t_{u'})e^{-\mathrm{i}jt_{u'}}\varepsilon_u\delta_{u'}(p)\frac{1}{2\pi}\sum_{v,v'}D_N(t_v-t_{v'})\times e^{\mathrm{i}jt_{v'}}\varepsilon_v\delta_{v'}(p)\Big].$$

By using the independence between price and noise process, it can be written as

$$\frac{1}{2\pi}\sum_{u,u'}D_N(t_u-t_{u'})e^{-\mathrm{i}jt_{u'}}\frac{1}{2\pi}\sum_{v,v'}D_N(t_v-t_{v'})\times e^{\mathrm{i}jt_{v'}}E[\varepsilon_u\varepsilon_v]E[\delta_{u'}(p)\delta_{v'}(p)]. \tag{A17}$$

Observe that

$$E\left[\delta_{u'}(p)\delta_{v'}(p)\right]=0 \quad \text{if } u'\neq v'$$

and

$$E\left[\varepsilon_u^2\right]=2E\left[\eta^2\right],\quad E[\varepsilon_u\varepsilon_v]=\begin{cases}-E\left[\eta^2\right] & \text{if } |v-u|=1\\ 0 & \text{if } |v-u|>1.\end{cases}$$

Therefore, (A17) is equal to $\left(\text{every term is multiplied by } \left(\frac{1}{2\pi}\right)^2\right)$

$$\begin{aligned}&\sum_{u,u'}\sum_{v}D_N\left(t_u-t_{u'}\right)D_N\left(t_v-t_{u'}\right)E\left[\varepsilon_u\varepsilon_v\right]E\left[\left(\delta_{u'}(p)\right)^2\right]\\ &=\Bigg(\sum_{u,u'}D_N^2\left(t_u-t_{u'}\right)E\left[\varepsilon_u^2\right]+2\sum_{u,u'}D_N\left(t_u-t_{u'}\right)\\ &\qquad\times D_N\left(t_{u+1}-t_{u'}\right)E\left[\varepsilon_u\varepsilon_{u+1}\right]\Bigg)E\left[\left(\delta_{u'}(p)\right)^2\right]\\ &=2E\left[\eta^2\right]\sum_{u,u'}\Bigg(D_N^2\left(t_u-t_{u'}\right)-D_N\left(t_u-t_{u'}\right)\\ &\qquad\times D_N\left(t_{u+1}-t_{u'}\right)\Bigg)E\left[\left(\delta_{u'}(p)\right)^2\right].\end{aligned}$$

Using the inequality

$$|D_N\left(t_{u+1}-t_{u'}\right)-D_N\left(t_u-t_{u'}\right)|\le 1-D_N\left(\frac{2\pi}{n}\right)$$

and the following limit in probability

$$\lim_{N\to\infty}\int_0^{2\pi}\mathrm{d}u\int_0^{2\pi}D_N\left(u-u'\right)\sigma^2\left(u'\right)\mathrm{d}u'=C\int_0^{2\pi}\sigma^2(u)\mathrm{d}u,$$

we obtain for (A17) the asymptotic

$$2E\left[\eta^2\right]E\left[\int_0^{2\pi}\sigma^2(u)\mathrm{d}u\right]n\left(1-D_N\left(\frac{2\pi}{n}\right)\right).$$

Therefore, the sum over j in the definition of $\widetilde{\gamma}^2_{n,M,N}$ gives a term

$$2E\left[\eta^2\right]E\left[\int_0^{2\pi}\sigma^2(u)\mathrm{d}u\right]\Lambda(n,N,M),$$

where

$$\Lambda(n,N,M)=\frac{M(M+1)}{3}n\left(1-D_N\left(\frac{2\pi}{n}\right)\right),$$

which is $O\left(\frac{M^2N^2}{n}\right)$.

Now compute (A10). Note that the terms (A13)–(A15) are similar. For the independence between noise and price, it is equal to

$$\left(\frac{2\pi}{2N+1}\right)^2\sum_{|h|\le N}\sum_{|l|\le N}E\left[c_h(\varepsilon_n)c_l(\varepsilon_n)c_{-j-l}(\varepsilon_n)\right]\times E[c_{j-h}(\mathrm{d}p_n)]=0,$$

as $E[c_{j-h}(\mathrm{d}p_n)]=0$ for any j, h and n.

It remains to calculate (A16), which is equal to

$$\frac{1}{(2\pi)^2}\sum_{v,v'}\sum_{u,u'}e^{-\mathrm{i}j(t_{v'}-t_{u'})}D_N(t_v-t_{v'})D_N(t_u-t_{u'})E[\varepsilon_v\varepsilon_{v'}\varepsilon_u\varepsilon_{u'}]. \tag{A18}$$

We need the fourth moments for the noise process:

$$\begin{aligned}&E\left[\varepsilon^4\right]=2E\left[\eta^4\right]+6E\left[\eta^2\right]^2 \qquad\qquad (A19)\\ &E\left[\varepsilon_u^3\varepsilon_v\right]=\begin{cases}-E\left[\eta^4\right]-3E\left[\eta^2\right]^2 & \text{if } |u-v|=1\\ 0 & \text{if } |u-v|>1\end{cases}\\ &E\left[\varepsilon_u^2\varepsilon_v^2\right]=\begin{cases}E\left[\eta^4\right]+3E\left[\eta^2\right]^2 & \text{if } |u-v|=1\\ 4E\left[\eta^2\right]^2 & \text{if } |u-v|>1\end{cases}\\ &E\left[\varepsilon_u^2\varepsilon_v\varepsilon_{v+1}\right]=-2E\left[\eta^2\right]^2 \quad \text{if } v\ge u+1 \text{ or } v\le u-2\\ &E\left[\varepsilon_u\varepsilon_{u+1}^2\varepsilon_{u+2}\right]=2E\left[\eta^2\right]^2\\ &E\left[\varepsilon_u\varepsilon_{u+1}\varepsilon_v\varepsilon_{v+1}\right]=E\left[\eta^2\right]^2 \quad \text{if } |u-v|\ge 2.\end{aligned}$$

In order to compute (A18), we proceed as follows (each term has to be multiplied by $1/(2\pi)^2$):

(I) firstly, we add the terms with coefficient $E[\eta^4]+3E[\eta^2]^2$

$$\begin{aligned}&\sum_u E\left[\varepsilon_u^4\right]+8\sum_u\cos(j(t_{u+1}-t_u))D_N(t_u-t_{u+1})E\left[\varepsilon_u^3\varepsilon_{u+1}\right]\\ &\quad+6\sum_u\cos(j(t_{u+1}-t_u))E\left[\varepsilon_u^2\varepsilon_{u+1}^2\right]\\ &=2\left(E\left[\eta^4\right]+3E\left[\eta^2\right]^2\right)n\left[\left(1-\cos\left(j\frac{2\pi}{n}\right)\right)\right.\\ &\quad\left.+4\cos\left(j\frac{2\pi}{n}\right)\left(1-D_N\left(\frac{2\pi}{n}\right)\right)\right].\end{aligned}$$

Therefore, the sum over j in the definition of $\widetilde{\gamma}^2_{n,M,N}$ gives

$$\begin{aligned}&2\left(E\left[\eta^4\right]+3E\left[\eta^2\right]^2\right)\frac{1}{M+1}\sum_{|j|\le M}\left(1-\frac{|j|}{M}\right)j^2n\\ &\quad\times\left[\left(1-\cos\left(j\frac{2\pi}{n}\right)\right)+4\cos\left(j\frac{2\pi}{n}\right)\left(1-D_N\left(\frac{2\pi}{n}\right)\right)\right]\\ &=2\left(E\left[\eta^4\right]+3E\left[\eta^2\right]^2\right)\Gamma(n,N,M),\end{aligned}$$

where $\Gamma(n,N,M)=O\left(\frac{M^4}{n}\right)+O\left(\frac{M^2N^2}{n}\right)$ (we used: $\cos x=1-\frac{x^2}{2}+O(x^4)$).

(II) Secondly, we add the terms with coefficient $2E\left[\eta^2\right]^2$ and only one summation. We omit to consider a constant factor $c=30$ which multiply each addend, then we have:

$$
\begin{aligned}
&\sum_u \cos(j(t_{u+2}-t_u))D_N(t_{u+2}-t_{u+1})E\left[\varepsilon_u^2\varepsilon_{u+1}\varepsilon_{u+2}\right]\\
&\quad+\sum_u \cos(j(t_{u+2}-t_{u+1}))D_N(t_{u+2}-t_{u+1})D_N(t_u-t_{u+1})\\
&\qquad\times E\left[\varepsilon_u\varepsilon_{u+1}^2\varepsilon_{u+2}\right]\\
&\quad+\sum_u \cos(j(t_{u+2}-t_{u+1}))D_N(t_u-t_{u+1})E\left[\varepsilon_u\varepsilon_{u+1}\varepsilon_{u+2}^2\right]\\
&\quad+2\sum_u \cos(j(t_{u+3}-t_{u+1}))D_N(t_{u+3}-t_{u+2})D_N(t_u-t_{u+1})\\
&\qquad\times E\left[\varepsilon_u\varepsilon_{u+1}\varepsilon_{u+2}\varepsilon_{u+3}\right]\\
&=2E\left[\eta^2\right]^2 n\left\{-\cos\left(j\frac{4\pi}{n}\right)D_N\left(\frac{2\pi}{n}\right)\left(1-D_N\left(\frac{2\pi}{n}\right)\right)\right.\\
&\qquad\left.-\cos\left(j\frac{2\pi}{n}\right)D_N\left(\frac{2\pi}{n}\right)\left(1-D_N\left(\frac{2\pi}{n}\right)\right)\right\}\\
&=2E\left[\eta^2\right]^2 D_N\left(\frac{2\pi}{n}\right)\left(1-D_N\left(\frac{2\pi}{n}\right)\right)n\\
&\qquad\times\left\{-2\left(1-10\left(j\frac{\pi}{n}\right)^2+O\left(\frac{j}{n}\right)^4\right)\right\}.
\end{aligned}
$$

Finally, consider the sum over j

$$
\begin{aligned}
&2E\left[\eta^2\right]^2 n\left(1-D_N\left(\frac{2\pi}{n}\right)\right)\frac{1}{M+1}\\
&\qquad\times\sum_{|j|\le M}\left(1-\frac{|j|}{M}\right)j^2\left(1-10\left(j\frac{\pi}{n}\right)^2+O\left(\frac{j}{n}\right)^4\right)\\
&\quad=2E\left[\eta^2\right]^2\Psi_1(n,N,M),
\end{aligned}
$$

where $\Psi_1(n,N,M)=O\left(\frac{N^2M^2}{n}\right)$.

(III) Thirdly, we add the terms with coefficient $2E[\eta^2]^2$ and a double summation. We omit to write a constant $c=6$ which multiplies each term, then we have

$$
\begin{aligned}
&\sum_{u,v}\cos(j(t_v-t_u))E\left[\varepsilon_u^2\varepsilon_v^2\right]\\
&\quad+2\left(2\sum_{u,v}\cos(j(t_{v+1}-t_u))D_N(t_{v+1}-t_v)E\left[\varepsilon_u^2\varepsilon_v\varepsilon_{v+1}\right]\right)\\
&\quad+4\sum_{u,v}\cos(j(t_{v+1}-t_{u+1}))D_N(t_u-t_{u+1})D_M(t_v-t_{v+1})\\
&\qquad\times E\left[\varepsilon_u\varepsilon_{u+1}\varepsilon_v\varepsilon_{v+1}\right]\\
&=\sum_{u,v}\cos(j(t_v-t_u))4E\left[\eta^2\right]^2\\
&\quad+4\sum_{u,v}\cos(j(t_{v+1}-t_u))D_N\left(\frac{2\pi}{n}\right)\left(-2E\left[\eta^2\right]^2\right)\\
&\quad+4\sum_{u,v}\cos(j(t_{v+1}-t_{u+1}))D_N^2\left(\frac{2\pi}{n}\right)E\left[\eta^2\right]^2\\
&=4E\left[\eta^2\right]^2\sum_{u,v}\left\{\cos(j(t_v-t_u))-2\cos(j(t_{v+1}-t_u))\right.\\
&\qquad\left.\times D_N\left(\frac{2\pi}{n}\right)+\cos(j(t_{v+1}-t_{u+1}))D_N^2\left(\frac{2\pi}{n}\right)\right\}\\
&=4E\left[\eta^2\right]^2\left(1-D_N\left(\frac{2\pi}{n}\right)\right)^2 O\left(n^2\right).
\end{aligned}
$$

Then, considering the sum over j, we get a term

$$2E\left[\eta^2\right]^2\Psi_2(n,N,M),$$

where $\Psi_2(n,N,M)=O\left(\frac{M^2N^4}{n^2}\right)$. Finally, denote $\Psi:=\Psi_1+\Psi_2$. □