Robustness of Fourier Estimator of Integrated Volatility in the Presence of Microstructure Noise

Maria Elvira Mancino

Dept. Mathematics for Decisions

University of Firenze, Italy

mariaelvira.mancino@dmd.unifi.it

Simona Sanfelici

Dept. Economics

University of Parma, Italy

simona.sanfelici@unipr.it

Abstract

We study the finite sample properties of the Fourier estimator of integrated volatility under market microstructure noise. We derive an analytic expression for the bias and the mean squared error of the contaminated estimator. These estimates can be practically used to design optimal MSE-based estimators, which are very robust and efficient in the presence of noise. Moreover an empirical analysis based on a simulation study and on high-frequency logarithmic prices of the Italian stock index futures (FIB30) validates the theoretical results.

1 Introduction

In the recent years the literature devoted to the model free measurement of volatility has received strong impulse by the availability of high frequency financial data (see the review paper by [Andersen, Bollerslev and Diebold, 2002]). Nevertheless the efficiency of all the methodologies proposed in estimating accurately the volatility builds on the observability of the true price process, while observed asset prices are contaminated by market microstructure effects, such as price discreteness, separate trading prices for buyers and sellers and other contaminations; as a consequence observed asset prices diverge from their efficient values (see, for instance, [Roll, 1984, Glosten and Milgrom, 1985, Harris, 1991], [O'Hara, 1995]).

The study of the implications of market microstructure noise for integrated volatility estimators has largely focused on the *realized volatility*, an estimator of the integrated (e.g. daily) volatility as the sum of intradaily squared returns, among others see [Zhang and al., 2005, Aït-Sahalia, Mykland and Zhang, 2005a, Bandi and Russell, 2005a,

Bandi and Russell, 2006a]. The theoretical justification for the use of this estimator is a classical result in semi-martingale process theory, i.e. the *quadratic variation* theorem, which essentially goes back to Wiener. In fact in the semi-martingale paradigm, when the asset prices are observed without errors, the realized volatility is a consistent estimator of the integrated volatility of the underlying price process and higher sampling frequencies over a fixed period of time result in more precise estimates of the integrated volatility of the price process. If this was the case, then the use of the highest possible frequency (tick by tick data) should give a precise identification of the volatility. Nevertheless, in the presence of microstructure noise the realized volatility estimator fails to converge to the integrated volatility of the underlying price process. This diverging behavior has been empirically observed by [Andersen and al., 1999] and theoretically analyzed in [Bandi and Russell, 2005a, Zhang and al., 2005]. Therefore, some methods have been proposed to correct the realized volatility estimator for the effect of market microstructure noise, in order to obtain unbiased estimators of the true integrated volatility [Zhou, 1996, Andersen, Bollerslev, Diebold and Ebens, 2001, Zhang and al., 2005, Barndorff-Nielsen and al., 2006a, Hansen and Lunde, 2006].

In this paper we will focus on the *Fourier estimator* of integrated volatility, which is obtained as a particular case of the estimator proposed in [Malliavin and Mancino, 2002]. The Fourier methodology allows to reconstruct the *instantaneous volatility* as a series expansion with coefficients gathered from the Fourier coefficients of the price variation. The consistency in probability uniformly in time and the asymptotic properties of the Fourier estimator of instantaneous volatility have been proved in [Malliavin and Mancino, 2005] in the absence of microstructure noise. Moreover the efficiency of the Fourier method to compute the *integrated volatility* has been analyzed in comparison with other estimators in [Barucci and Renò, 2001, Barucci and Renò, 2002, Hansen and Lunde, 2005, Kanatani, 2004], when the microstructure noise is ignored.

The aim of this paper is to study the *robustness* of the Fourier estimator of integrated volatility towards microstructure noise. We begin our analysis with a simple price formation mechanism which is typical of bid-ask bounce effects [Roll, 1984]. The logarithm of the observed price process is the sum of the log-price process in equilibrium and a component of microstructure noise which is represented by independent identically distributed random variables (*MA(1) model* in the sequel). In this setting we study the finite sample properties for the Fourier volatility estimator both through theoretical analysis and by simulation results. We recall that in this regard [Nielsen and Frederiksen, 2006] make an empirical comparison, through Monte Carlo simulations and using high frequency market data, between some estimators of integrated volatility: the realized volatility, the Fourier estimator, the wavelet estimator and some bias correction methods, namely the realized bipower variation by [Barndorff-Nielsen and Shephard, 2004], the kernel-based estimator by [Zhou, 1996] and the related unbiased estimator proposed by [Hansen and Lunde, 2006]. Their empirical analysis shows that the Fourier estimator is superior to the realized volatility and the wavelet estimator and, even compared to the bias correcting methods for microstructure noise, the Fourier estimator provides smaller root mean squared error, while having only slightly higher bias. Nevertheless their analysis is purely empirical and a precise treatment of market microstructure noise effects on

2

the Fourier estimator of the efficient price's volatility is needed. Therefore in this paper we take a first step towards the understanding of the Fourier estimator's properties when uncorrelated microstructure noises are included. More realistic microstructure noise dependence has been studied in [Hansen and Lunde, 2006] for the realized volatility estimator. However, the empirical work of [Hansen and Lunde, 2006] suggests this independence assumption is not too damaging statistically when we analyze data in tickly traded stocks recorded approximately every minute.

We derive the explicit analytical expression of the bias of the Fourier estimator for a given sample size n and a given number of Fourier coefficients N included in the estimation and we prove that the bias of the Fourier estimator converges to zero, for n, N increasing to $+\infty$, under the condition that $\frac{N^2}{n}$ goes to 0. Therefore, even if we do not proceed to any bias correction of the estimator, a suitable cutting of the highest frequencies makes the finite sample bias negligible. Even more strikingly, we prove this result holds both in the case of independent microstructure noise and in the case where the noise is correlated with the efficient returns. We note that usually a bias correction is accompanied by a larger asymptotic variance. In the present case, we obtain the analytical expression of the conditional (on the underlying volatility path) mean squared error (MSE) of the contaminated volatility estimator as a function of the sampling frequency and the number of Fourier coefficients. This expression shows that the MSE does not diverge for n, N increasing to $+\infty$, under the condition that $\frac{N^2}{n}$ goes to 0. These two results enlighten a peculiar feature of the Fourier estimator: while the MSE of the realized volatility estimator under microstructure noise diverges as the number n of observations increases, the MSE of the Fourier estimator is substantially unaffected by the presence of microstructure noise by choosing in a suitable way the number of Fourier coefficients to be included in the estimation, as indicated explicitly by the MSE computation.

We emphasize that the Fourier estimator needs no correction in order to be statistically efficient and robust to some kind of market frictions at the same time. This result is due to the following properties of the Fourier estimator: on one side it uses all available data by integration, therefore incorporating not only the squared increments of the prices but also the auto-covariances of all orders along the time window; on the other side the high-frequency noise or short-run noise is ignored by the Fourier estimator by cutting the highest frequencies in the construction of the estimator. In other words, when efficiently implemented, the Fourier estimator uses as much as possible of the available sample path without being excessively biased due to the impact of market frictions.

The contribution of various order auto-covariances has early been considered by [Zhou, 1996, Corsi et al., 2001] and very recently used to correct the bias of the realized variance type estimators in the presence of microstructure noise. In particular we refer to the subsampled estimator by [Zhang and al., 2005, Aït-Sahalia, Mykland and Zhang, 2005b] and the realized (subsampled) kernels by [Barndorff-Nielsen and al., 2006a, Barndorff-Nielsen and al., 20 Nevertheless the Fourier estimator works differently, since the Dirichlet (or the Fejer) kernel appearing in the Fourier estimator depends on both the number of frequencies N and the number of observations: this fact has a great number of important implications in terms of the efficiency of the estimator. Theorems 4.1 and 5.1 provide the tools to optimize the finite sample performance of the Fourier estimator in terms of the selection of

the number of frequencies by minimizing the mean squared error, for a given number of intra-daily observations (see [Bandi and Russell, 2006b] for a similar approach). We note that in the case of the realized kernels the optimal selection procedure is implemented with respect to the bandwidth, given the sample size.

Our theoretical results are confirmed by a simulation study. We simulate discrete data from a continuous time stochastic volatility model with microstructure contaminations as in [Nielsen and Frederiksen, 2006]. Consistently with the theoretical results, we show that when $N = n/2$ the Fourier estimator behaves like the realized volatility estimator and it does indeed explode as the sampling interval goes to zero. Nevertheless, both bias and MSE can be strongly reduced by choosing N conveniently and as $\frac{N^2}{n} \to 0$ the Fourier estimator turns out to be unbiased and the MSE converges to a small positive constant. Our analysis suggests to use quote-to-quote returns and try to minimize MSE as a function of the cutting frequency N_{cut}. In this way, we find optimal sampling at higher frequencies than those obtained with realized volatility. This is due to the fact that the Fourier estimator can utilize more information in the data, without being affected by a severe bias, simply choosing the lower frequencies of the Fourier estimator. By letting N free to vary, we show that it may be convenient to choose $N < \frac{n}{2}$, where $\frac{n}{2}$ is the Nyquist frequency, especially for large n. We provide an easily implementable procedure to select numerically a value for (n, N), which optimizes the Fourier estimator's finite sample performance in terms of its bias and MSE. The optimal MSE-based estimator turns out to be very attractive, even in comparison with methods specifically designed to handle market microstructure contaminations. More specifically, the Fourier estimator is competitive in terms of MSE for high sampling frequencies up to 30 sec, while having only a slightly higher bias. This procedure is then applied to high-frequency intraday returns on the Italian stock index futures (FIB30), where the choice of a suitable N_{cut} allows to render the Fourier estimator invariant to short-run noise introduced by market microstructure effects, with consequent efficiency gains. Moreover, our simulations indicate that the RV estimator is more biased than the Fourier estimator in the presence of market microstructure noise and, therefore, that the actual volatility might be higher on average than predicted by the much used realized volatility, as already noticed in [Nielsen and Frederiksen, 2006].

The paper is organized as follows. In section 2 we describe the assumptions for the underlying price formation mechanism. In section 3 we recall the definition of the Fourier estimator of the (integrated) volatility. Sections 4 and 5 contain the computation of the bias and the MSE as functions of the sampling frequency and the number of Fourier coefficients included in the construction of the estimator. Section 6 extends the analysis of the bias of the Fourier estimator to more general price formation rules. In section 7, we test our theoretical findings by means of a Monte Carlo simulation, while in Section 8 we apply our results to the optimal sampling of quote-to-quote FIB30 logarithmic prices. Section 9 concludes. The proofs are contained in the Appendix.

2 The model of prices with microstructure effects

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

$$\widetilde{p}(t) = p(t) + \eta(t) \tag{1}$$

where $p(t)$ is the efficient log-price process and $\eta(t)$ is the microstructure noise. We can think of $p(t)$ as the log-price in equilibrium, that is the price that would prevail in the absence of market microstructure frictions. The econometrician does not observe the returns of the true return series, but the returns contaminated by market microstructure effects. Therefore an estimator of the integrated volatility should be constructed using the contaminated returns.

We consider a fixed time period (e.g. a trading day), say $[0, T]$. Suppose that the process is observed at a discrete unevenly spaced grid $\{0 = t_{0,n} \leq t_{1,n} \leq \cdots \leq t_{k_n,n} \leq T\}$ for any $n \geq 1$. We make the following assumptions:

A.I $p(t)$ is a continuous semi-martingale satisfying the stochastic differential equation

$$dp(t) = \sigma(t)\,dW(t) + b(t)\,dt$$

where W is a Brownian motion on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t \in [0,T]}, P)$, σ and b are adapted stochastic processes such that

$$E[\int_0^T \sigma^4(t)dt] < \infty \quad , \quad E[\int_0^T b^2(t)dt] < \infty.$$

A.II The random shocks $\eta(t_{j,n})$, for $0 \leq j \leq k_n$ and for all n, are independent and identically distributed with mean zero and bounded fourth moment.

A.III The true return process $\delta_{j,n}(p) := p(t_{j,n}) - p(t_{j-1,n})$ is independent of $\eta(t_{j,n})$ for any j, n.

To simplify the notation, in the sequel we will write $\delta_j(p)$ and η_j instead of $\delta_{j,n}(p)$ and $\eta(t_{j,n})$.

Remark 2.1 *A structural model like (1), where the efficient price is considered a continuous process and the noise is observed on a fixed grid of prices, has been proposed in [Aït-Sahalia, Mykland and Zhang, 2005a] and [Bandi and Russell, 2005a]. It is coherent with the model free volatility estimation method introduced in [Malliavin and Mancino, 2002], but here we introduce explicitly microstructure effects. The instantaneous volatility process is allowed to display jumps, diurnal effects, high persistence, non-stationarities and leverage effects.*

Remark 2.2 *In general the noise return moments may depend on the sampling frequency (see [Hansen and Lunde, 2004, Bandi and Russell, 2005b]). Here we consider the simplified 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 [Roll, 1984]. It is known to be a realistic approximation in decentralized markets where traders arrive in a random fashion with idiosyncratic price setting behavior, the foreign exchange market being a valid example [Bai and al., 2005]. See [Zhang and al., 2005, Bandi and Russell, 2006a, Hansen and Lunde, 2006] for additional discussions.*

Remark 2.3 *The hypothesis that the η_j's are independent of the increments $\delta_j(p)$ is discussed in [Hansen and Lunde, 2006]. Their empirical work suggests that the independence assumption is not too damaging statistically when we analyze data in tickly traded stocks recorded every minute. Therefore our analysis is mainly developed in this setting, but we examine the robustness of the Fourier estimator under more general microstructure noise dependence in section 6.*

3   Fourier method for volatility estimation

The Fourier estimator of spot volatility was introduced in [Malliavin and Mancino, 2002]. The methodology has been generalized in [Malliavin and Mancino, 2005] where the following result is proved: the volatility function is computed by establishing a connection between the Fourier transform of the price process and the Fourier transform of the volatility process. This result is rigorously stated in the following theorem. We note that by a change of the origin of time and rescaling the unit of time we can always reduce ourselves to the case where the time window is $[0, 2\pi]$.

Theorem 3.1 *Consider a semimartingale p satisfying assumption* **A.I**. *Denote the Fourier transform of dp by*

$$\mathcal{F}(dp)(k) := \frac{1}{2\pi} \int_{]0,2\pi[} \exp(-ikt)\, dp(t), \tag{2}$$

and the Bohr convolution product between two functions Φ, Ψ defined on the integers by

$$(\Phi *_B \Psi)(k) := \lim_{N \to \infty} \frac{1}{2N+1} \sum_{s=-N}^{N} \Phi(s)\Psi(k-s). \tag{3}$$

Then we have

$$\frac{1}{2\pi} \mathcal{F}(\sigma^2)(k) = (\mathcal{F}(dp) *_B \mathcal{F}(dp))(k), \quad for\ all\ k \in \mathbf{Z}. \tag{4}$$

The equality (4) is true in probability, which means that the limit appearing in the r.h.s of (3) exists in probability.

For simplicity, we assume $b = 0$ because [Malliavin and Mancino, 2005] prove that the drift term gives zero contribution to the volatility estimation. In particular, by considering the case $k = 0$ in the formula (4), we obtain that the integrated volatility over the time interval $[0, 2\pi]$ can be computed as

$$\int_0^{2\pi} \sigma^2(t)dt = (2\pi)^2 \, (\mathcal{F}(dp) *_B \mathcal{F}(dp))(0). \tag{5}$$

In the sequel, we assume that the price process p is observed at a discrete unevenly spaced grid $\{0 = t_{0,n} \le t_{1,n} \le \cdots \le t_{k_n,n} \le 2\pi\}$ for any $n \ge 1$, where for simplicity we can take $k_n = n$, with the only condition that $\rho(n) := \max_{0 \le h \le n} |t_{h+1,n} - t_{h,n}|$ goes to zero as

6

$n \to \infty$. From (5) the finite sample Fourier estimator of the integrated volatility over $[0, 2\pi]$ is defined as follows

$$\frac{(2\pi)^2}{2N+1} \sum_{s=-N}^{N} \mathcal{F}(dp)_n(s)\mathcal{F}(dp)_n(-s) \tag{6}$$

$$\mathcal{F}(dp)_n(s) := \frac{1}{2\pi} \sum_{j=1}^{n} \exp(-ist_j)\delta_j(p). \tag{7}$$

The consistency in probability uniformly in time of the Fourier estimator of the instantaneous volatility is proved in the absence of microstructure noise under assumption **A.I** in [Malliavin and Mancino, 2005] (Theorem 3.2). In particular the following convergence in probability holds

$$\lim_{n,N \to \infty} \frac{1}{2N+1} \sum_{|s| \leq N} \sum_{j=1}^{n} \sum_{j'=1}^{n} e^{is(t_j-t'_j)}\delta_j(p)\delta_{j'}(p) = \int_0^{2\pi} \sigma^2(t)dt.$$

We emphasize two different features of the Fourier estimation method: the first one is that the Fourier estimator uses all available data by integration, thus incorporating not only the squared increments of the prices but also the products of disjoint increments along the time window (i.e. the auto-covariances of all orders), the second one is the convolution product in (6) which weights the cross-products at any given frequency.

In [Barucci and Renò, 2001, Barucci and Renò, 2002, Hansen and Lunde, 2005] the efficiency of the Fourier estimator of integrated volatility has been analyzed in comparison with other estimators only in the absence of microstructure.

In this paper we are interested in the estimation of the integrated volatility, given the observations of the *contaminated process* $\tilde{p}$ defined in (1). In the sequel we will denote

$$\hat{\sigma}_{n,N}^2 := \frac{(2\pi)^2}{2N+1} \sum_{s=-N}^{N} \mathcal{F}(d\tilde{p})_n(s)\mathcal{F}(d\tilde{p})_n(-s), \tag{8}$$

which we call the *Fourier estimator*. We note that, taking (7) into account, the Fourier estimator (8) can be re-written as

$$\hat{\sigma}_{n,N}^2 := \sum_{j=1}^{n} \sum_{j'=1}^{n} D_N(t_j - t'_j)\delta_j(\tilde{p})\delta_{j'}(\tilde{p}), \tag{9}$$

where $D_N(t)$ denotes the rescaled Dirichlet kernel defined by (see e.g. [Malliavin, 1995])

$$D_N(t) := \frac{1}{2N+1} \sum_{|s| \leq N} e^{ist} = \frac{1}{2N+1} \frac{\sin[(N+\frac{1}{2})t]}{\sin \frac{t}{2}}. \tag{10}$$

7

Remark 3.2 *In the numerical applications we will also consider the finite sample properties of the following version of the Fourier estimator, simply obtained by weighting the convolution product with the La Vallee Poussin kernel:*

$$\hat{\Sigma}^2_{n,N} := \frac{(2\pi)^2}{N+1} \sum_{s=-N}^{N} \left(1 - \frac{|s|}{N}\right) \mathcal{F}(d\tilde{p})_n(s)\mathcal{F}(d\tilde{p})_n(-s). \tag{11}$$

The mathematical results which will be proved with respect to the Dirichlet kernel in the next two sections, can be obtained in the same way in this context, as the Fourier estimator (11) can be re-written as

$$\hat{\Sigma}^2_{n,N} := \sum_{j=1}^{n} \sum_{j'=1}^{n} F_N(t_j - t'_j)\delta_j(\tilde{p})\delta_{j'}(\tilde{p}),$$

where $F_N(t) = \frac{\sin^2(Nt)}{(Nt)^2}$ is the Fejer kernel. Empirically we will see to what extent this modification improves the behavior of the Fourier estimator for very high observation frequencies.

The contribution of various order auto-covariances has early been considered by [Zhou, 1996] and recently used to correct the bias of the realized variance type estimators in the presence of microstructure noise. In particular we refer to the subsampled estimator by [Zhang and al., 2005] and the realized (subsampled) kernels by [Barndorff-Nielsen and al., 2006a, Barndorff-Nielsen and al., 2006b]. Nevertheless the Fourier estimator works differently. In order to illustrate this point we set in the hypothesis of [Barndorff-Nielsen and al., 2006b], where the time gap τ between two observations is constant and H is the bandwidth, then the realized kernel correction to realized variance estimator is

$$\sum_{h=1}^{H} k(\frac{h-1}{H}) \left\{\sum_{j=1}^{n} \delta_j(\tilde{p})\delta_{j-h}(\tilde{p}) + \sum_{j=1}^{n} \delta_j(\tilde{p})\delta_{j+h}(\tilde{p})\right\}.$$

The weight $k(x)$ is a function of the bandwidth only. On the other hand, by exploiting the same data set as for the above realized kernel estimator, the Fourier estimator correction to realized variance is

$$\sum_{h=1}^{H} D_N(\tau h) \left\{\sum_{j=1}^{n} \delta_j(\tilde{p})\delta_{j-h}(\tilde{p}) + \sum_{j=1}^{n} \delta_j(\tilde{p})\delta_{j+h}(\tilde{p})\right\}.$$

Thus the Dirichlet kernel in Fourier estimator depends on the number of frequencies N, besides the delay between two observations. Even though this seems to leave many degrees of freedom, we will see in the next sections that a single optimization over the number of frequencies N renders the Fourier estimator very efficient even in presence of microstructure noise. Theorems 4.1 and 5.1 provide a way to optimize the finite sample performance of the Fourier estimator as a function of the number of frequencies by the minimization of the mean squared error (MSE), for a given number of intra-daily observations, with a similar approach as in [Bandi and Russell, 2006b]. We note that for the

realized kernels estimator in [Bandi and Russell, 2006b, Barndorff-Nielsen and al., 2006a] the optimization is performed with respect to the bandwidth H, given the sample size.

In the following sections we derive analytical formulae of the bias and the MSE of the Fourier estimator under bid-ask microstructure noise. This computation will serve as a basis for the optimal choice of the cutting frequency for a given data sampling interval when considering financial return series data.

4   Bias Computation

Denote $\delta_j(\tilde{p}) := \tilde{p}(t_j) - \tilde{p}(t_{j-1})$ where $\tilde{p}$ is defined in (1) and $\varepsilon_j := \eta_j - \eta_{j-1}$ where η_j's are defined in **A.II**. Let

$$\hat{V}_n = \sum_{j=1}^{n} (\delta_j(\tilde{p}))^2,$$

where n is the number of observations in the trading interval $[0, T]$. Then $\hat{V}_n$ is taken as the *realized volatility* estimator of the integrated volatility $\int_0^T \sigma^2(t)dt$, henceforth denoted by V. The realized volatility is a consistent estimator of integrated volatility in the hypothesis that the prices are observed without measurement errors, but in practice, due to market microstructure noise, sampling at the highest frequency leads to a bias problem (see [Zhou, 1996]). Under the hypothesis that $\frac{T}{n}$ is the time distance between adjacent logarithmic prices, it is easy to prove that the realized volatility estimator $\hat{V}_n$ diverges as the number n of observations increases and the bias is the following

$$E[\hat{V}_n - V] = 2nE[\eta^2]. \tag{12}$$

The result (12) was established in [Bandi and Russell, 2005a, Zhang and al., 2005].

We consider the Fourier estimator defined in (8). The definition of the Fourier estimator does not require evenly spaced data. Anyway for simplicity of computation, we will suppose that the observations are equidistant in time and $\frac{2\pi}{n}$ is the distance between two observations, where $[0, 2\pi]$ is the trading period. Then the bias is computed as follows.

Theorem 4.1 *For any fixed integers n, N the following identity holds*

$$E[\hat{\sigma}_{n,N}^2 - V] = 2n\ E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(2N+1)\frac{\pi}{n}]}{\sin(\frac{\pi}{n})}\right). \tag{13}$$

We analyze now the r.h.s of (13). Observe that if we impose the condition $\frac{N^2}{n} \to 0$, we obtain

$$\lim_{n,N\to\infty} 2n\ E[\eta^2] \left(1 - \frac{1}{2N+1} \frac{\sin[(2N+1)\frac{\pi}{n}]}{\sin(\frac{\pi}{n})}\right) = 0.$$

Therefore, if we consider a fixed number of observations n and we choose N "small" with respect to n, the bias of the Fourier estimator is smaller than the bias of the realized volatility; furthermore it goes to zero for n, N increasing at the proper rate.

We can derive the following conclusion: the Fourier estimator is asymptotically unbiased under the condition $\frac{N^2}{n}$ goes to 0. Moreover the result (13) shows that for fixed n, that is for a finite sample, with a suitable choice of N there is lower bias with respect to the realized volatility estimator.

5   MSE Computation

In this section we compute the mean squared error (MSE) of the Fourier estimator conditional on the volatility path. For simplicity, we will suppose that the volatility process is independent of W, therefore we assume that the no leverage hypothesis holds (see [Andersen, Bollerslev and Diebold, 2002] and [Meddahi, 2002] for a justification of the no-leverage assumption in the literature).

In [Bandi and Russell, 2005a, Hansen and Lunde, 2006], under the hypothesis that $\frac{T}{n}$ is the time distance between adjacent logarithmic prices, it is proved that the MSE of the realized volatility estimator defined by $\hat{V}_n = \sum_{j=1}^{n}(\delta_j(\tilde{p}))^2$, is the following

$$E[(\hat{V}_n - V)^2] = 2\frac{T}{n}(Q + o(1)) + \Lambda_n, \tag{14}$$

where Q is the so-called *integrated quarticity* $\int_0^T \sigma^4(s)ds$, $o(1)$ is a term which goes to zero as n goes to infinity, and

$$\Lambda_n := n^2\alpha + n\beta + \gamma,$$

with

$$\alpha = (E[\varepsilon^2])^2, \quad \beta = E[\varepsilon^4] + 2E[\varepsilon^2\varepsilon_{-1}^2] - 3(E[\varepsilon^2])^2 \tag{15}$$
$$\gamma = 4E[\varepsilon^2]V - 2E[\varepsilon^2\varepsilon_{-1}^2] + 2(E[\varepsilon^2])^2.$$

We use the notation ε for $\eta_j - \eta_{j-1}$ for a generic j and ε_{-1} for $\eta_{j-1} - \eta_{j-2}$ for the same j. Easy computations show that under the assumption **A.II** the following identities hold

$$\alpha = 4E[\eta^2]^2, \quad \beta = 4E[\eta^4], \quad \gamma = 8E[\eta^2]V + 2E[\eta^2]^2 - 2E[\eta^4].$$

The following result contains the computation of the MSE of the Fourier volatility estimator.

Theorem 5.1 *For any fixed n, N the following relation holds*

$$E[(\hat{\sigma}_{n,N}^2 - V)^2] = 2\frac{2\pi}{n}(Q + o(1)) + n^2\hat{\alpha} + n\hat{\beta} + \hat{\gamma}, \tag{16}$$

where

$$\hat{\alpha} = \alpha \ (1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n}));$$

$$\hat{\beta} = \beta \ (1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n}));$$

$$\hat{\gamma} = \gamma + 4Q\frac{2\pi}{2N+1} + 4(E[\eta^2]^2 + E[\eta^4])(2D_N(\frac{2\pi}{n}) - D_N^2(\frac{2\pi}{n})) \tag{17}$$

with α, β, γ as in (15) and $D_N(t)$ is the rescaled Dirichlet kernel defined in (10).

Remark 5.2 *It is worth noting that the terms $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ depend only on the time gap between two observations of the price, which here is assumed to be equal $\frac{2\pi}{n}$, and on the number of Fourier coefficients. There is no dependence on the time period $[0, 2\pi]$ under consideration.*

The above result needs some comments. Denote by MSE_{RV} and MSE_{RVm} the mean squared error of the realized volatility estimator in the absence of microstructure noise and in the presence of microstructure noise, respectively. We have

$$MSE_{RV} = 2\frac{2\pi}{n}(Q + o(1))$$

and

$$MSE_{RVm} = MSE_{RV} + n^2\alpha + n\beta + \gamma.$$

Therefore it is clear that, in the absence of microstructure effects, the mean squared error of the realized volatility estimator goes to zero as $n \to \infty$, while in the presence of microstructure effects the mean squared error of the realized volatility estimator diverges as $n \to \infty$, due to the presence of the terms of order n^2 and n.

Analogously we make now the comparison of the mean squared error of the Fourier estimator without microstructure noise, denoted by MSE_F and with microstructure noise, denoted by MSE_{Fm}. We have

$$MSE_F = MSE_{RV} + c(n, N), \tag{18}$$

where $c(n, N)$ is a term which goes to zero as N, n go to infinity. In fact from the proof of Theorem 5.1 it is easily seen that $c(n, N)$ is equal to (39), therefore it is less or equal than $4Q\frac{2\pi}{2N+1}$. Moreover

$$MSE_{Fm} = MSE_F + n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) + \tilde{\gamma}(n, N)$$

where

$$\hat{\alpha}(n, N) = 4[\eta^2]^2\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right),$$

$$\hat{\beta}(n, N) = 4E[\eta^4]\left(1 + D_N^2(\frac{2\pi}{n}) - 2D_N(\frac{2\pi}{n})\right)$$

and

$$\tilde{\gamma}(n, N) = \gamma + 4(E[\eta^2]^2 + E[\eta^4])(2D_N(\frac{2\pi}{n}) - D_N^2(\frac{2\pi}{n})).$$

We note that if $\frac{N^2}{n} \to 0$ then

$$\lim_{n,N\to\infty} n^2\hat{\alpha}(n, N) + n\hat{\beta}(n, N) = 0$$

and

$$\lim_{n,N\to\infty} \tilde{\gamma}(n, N) = 8E[\eta^2]V + 2E[\eta^4] + 6E[\eta^2]^2. \tag{19}$$

11

It follows that the mean squared error of the Fourier estimator does not diverge and it is not significantly affected by microstructure noise, in fact choosing N conveniently we obtain that MSE_F and MSE_{Fm} differ about the positive constant term (19).

Finally we have proved that the Fourier estimator needs no correction in order to be asymptotically unbiased and robust to market frictions of $MA(1)$-type. In the next section we will see that the result extends to different types of microstructure noise.

6   The case with dependent noise

In this section we remove the Assumption **A.III** by considering the case when the noise η is correlated with efficient returns $\delta(p)$. We consider a particular form of dependent noise, more precisely we follow an example in [Hansen and Lunde, 2006] with market microstructure noise that is time-dependent in tick time and correlated with efficient returns.

Let $t_0 < t_1 < \ldots < t_n$ be the times at which prices are observed and consider the case where we sample intraday returns at the highest possible frequency in tick time. Suppose that the noise is given by

$$\widetilde{\eta}_j := \alpha \delta_j(p) + \eta_j$$

where α is a real constant and $\widetilde{\eta}_j$ and η_j are the shorten notation for $\widetilde{\eta}_{t_j}$ and η_{t_j}. The case $\alpha = 0$ corresponds to the case with independent noise assumption. Let $\varepsilon_j := \eta_j - \eta_{j-1}$ and denote

$$\widetilde{\varepsilon}_j := \alpha(\delta_j(p) - \delta_{j-1}(p)) + \varepsilon_j.$$

Then for any j we easily have

$$E[\widetilde{\varepsilon}_j^2] = \alpha^2 E[\int_{t_{j-1}}^{t_j} \sigma^2(s)ds] + \alpha^2 E[\int_{t_{j-2}}^{t_{j-1}} \sigma^2(s)ds] + 2E[\eta^2] \tag{20}$$

and

$$E[\widetilde{\varepsilon}_j \delta_j(p)] = \alpha E[\int_{t_{j-1}}^{t_j} \sigma^2(s)ds]. \tag{21}$$

In [Hansen and Lunde, 2006] the bias of the realized volatility estimator in this setting is computed as follows

$$E[\hat{V}_n - V] = 2\alpha^2 V + 2\alpha V + 2n \, E[\eta^2]. \tag{22}$$

This bias can be negative if $\alpha < 0$, which is actually in the case where $\widetilde{\eta}_j$ and $\delta_j(p)$ are negatively correlated. We consider now the Fourier estimator defined in (8) and we compute the bias of the Fourier estimator in this case.

Proposition 6.1 *For any integers n, N the following identity holds*

$$E[\hat{\sigma}_{n,N}^2 - V] = (2\alpha^2 V + 2\alpha V + 2n \, E[\eta^2])(1 - D_N(\frac{2\pi}{n})), \tag{23}$$

where D_N is defined in (10).

Again we observe that the bias can be negative if $\alpha < 0$, but we also note that the bias of the Fourier estimator goes to zero if $n, N \to \infty$ and $\frac{N^2}{n} \to 0$ even in the presence of dependent microstructure noise. Therefore the Fourier estimator turns out to be asymptotically unbiased under this kind of dependent microstructure noise. This result can be extended to the case where $\widetilde{\eta}_j$ and $\delta_j(p)$ are not only contemporaneously correlated, but $\widetilde{\eta}_j$ is correlated with lagged values of $\delta_j(p)$.

7   Monte Carlo simulations

In this section we simulate discrete data from a continuous time stochastic volatility model with microstructure contaminations as in [Nielsen and Frederiksen, 2006]. From the simulated data, Fourier estimates of the volatility can be compared to the value of the true integrated variance. Consistently with the theoretical results of previous sections, we show that when $N = n/2$ the Fourier estimator behaves like the realized volatility estimator and it does indeed explode as the sampling interval goes to zero. Nevertheless, both bias and MSE can be strongly reduced by choosing N conveniently and as $\frac{N^2}{n} \to 0$ the Fourier estimator turns out to be unbiased and the MSE converges to a small positive constant.

The infinitesimal variation of the true log-price process and spot volatility is given by the CIR square-root model [Cox et al., 1985]

$$
\begin{aligned}
dp(t) &= \sigma(t) \ dW_1(t) \\
d\sigma^2(t) &= \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t) \ dW_2(t),
\end{aligned}
\tag{24}
$$

where W_1, W_2 are independent Brownian motions. Moreover, we assume that the logarithmic noises η are i.i.d. Gaussian 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. Alternative, possibly discrete, distributions can be used to describe microstructure noise. In [Nielsen and Frederiksen, 2006], for instance, a bid-ask bounce effect is described by an order-driven indicator discrete variable. In our case, the contaminated process becomes

$$
\tilde{p}(t_j) = p(t_j) + \eta(t_j), \qquad \eta(t_j) \sim N(0, \xi^2)
$$

so that

$$
\delta_j(\tilde{p}) = \delta_j(p) + \varepsilon_j,
$$

where the ε_j follow a MA(1) process with negative first order autocorrelation. Since $\delta_j(p)$ and ε_j are independent, the variance and covariance of contaminated returns can be easily computed

$$
var(\delta_j(\tilde{p})) = \int_{t_{j-1}}^{t_j} \sigma^2(s)ds + 2\xi^2, \quad cov(\delta_j(\tilde{p}), \delta_{j-1}(\tilde{p})) = -\xi^2.
$$

Hence, we notice that $\delta_j(\tilde{p})$ exhibits spurious volatility and negative serial correlation as a consequence of noise contamination.
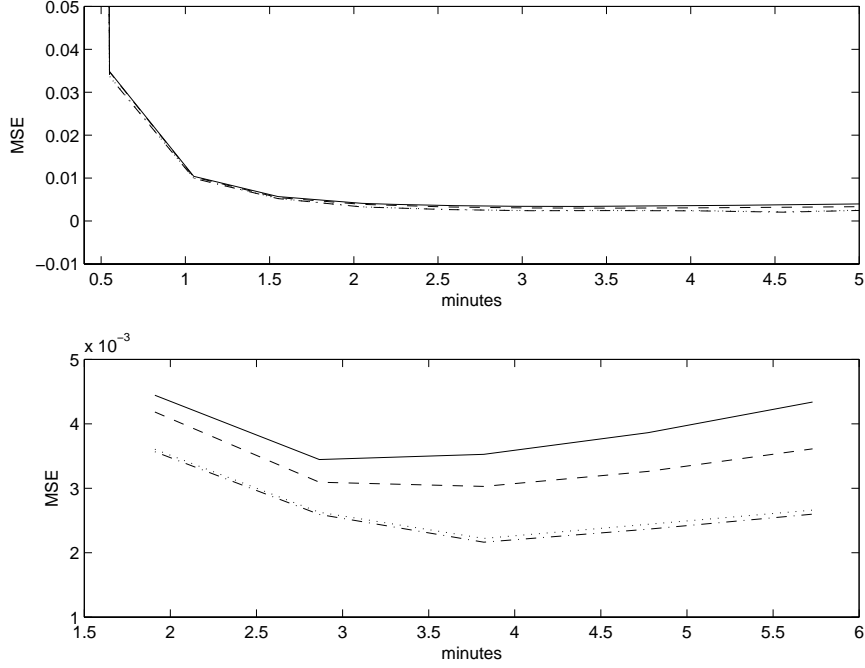
Figure 1: MSE plots as a function of the sampling period. '-' estimated MSE for $\hat{\sigma}^2_{n,N}$; ':' true MSE for $\hat{\sigma}^2_{n,N}$; '- -' estimated MSE for $\hat{V}_n$; '-.' true MSE for $\hat{V}_n$. Panel B is the same graph as Panel A, plotted on a different scale. The true MSE plots overlap because for $N = n/2$ the Fourier estimator behaves like the the realized volatility, i.e. they are both inconsistent at high sampling frequencies.

In order to avoid other data manipulations such as interpolation or imputation which might affect the numerical results, we generate (through simple Euler Monte Carlo discretization) high frequency evenly sampled true and observed returns by simulating second-by-second return and variance paths over a daily trading period of $T = 6$ hours, for a total of 21600 observations per day. Then we sample the observations for different choices of the uniform sampling interval $\rho(n) = T/n$ so that we obtain different data sets $(t_j, \ \tilde{p}(t_j), \ j = 0, 1 \ldots n)$ with σ recorded at every t. For instance, the choice $n = 360$ corresponds to a sampling period of $\rho(360) = 1$ minute. In implementing the Fourier estimator $\hat{\sigma}^2_{n,N}$, the smallest wavelength that can be evaluated in order to avoid aliasing effects is twice the smallest distance between two consecutive prices, which yields $N \leq n/2$ (*Nyquist frequency*). For 1 minute returns, it corresponds to $N \leq 180$.

Figure 1 shows the performance of the Fourier estimator $\hat{\sigma}^2_{n,N}$ and of the realized volatility $\hat{V}_n$ in terms of MSE as a function of the sampling period. For each estimator, the true and the estimated MSEs are plotted. In the true MSE, the value V is obtained from σ by numerical integration. The estimated MSE is computed by (16) for $\hat{\sigma}^2_{n,N}$ and by (14) for $\hat{V}_n$. In particular, for each sampling frequency, N is taken equal to $n/2$. The practical calculation hinges on the estimation of the relevant noise moments as well as on the preliminary identification of V and Q. Since the noise moments do not vary across frequencies under the MA(1) model, in computing the MSE estimates we use sample

14

moments constructed using quote-to-quote return data in order to estimate the relevant population moments of the noise components according to [Bandi and Russell, 2005a]

$$E[\varepsilon^2] = E\left[\frac{1}{n}\sum_{j=1}^{n}(\delta_j(\tilde{p}))^2\right] - \frac{V}{n},$$

$$E[\varepsilon^4] = E\left[\frac{1}{n}\sum_{j=1}^{n}(\delta_j(\tilde{p}))^4\right] - \frac{6E[\varepsilon^2]V}{n} + O\left(\frac{1}{n^2}\right),$$

$$E[\varepsilon^2\varepsilon_{-1}^2] = E\left[\frac{1}{n-1}\sum_{j=2}^{n}(\delta_j(\tilde{p}))^2(\delta_{j-1}(\tilde{p}))^2\right] - \frac{2E[\varepsilon^2]V}{n-1} + O\left(\frac{1}{n(n-1)}\right) + O\left(\frac{1}{n^2}\right).$$

These relations, together with the estimates of Sections 4 and 5 allow to measure the bias and MSE of the volatility estimators also in the case of empirical market quote data, where the efficient price and volatility and the noise contaminations are not available. Other possible estimators of these quantities are discussed in [Barndorff-Nielsen and al., 2006a], although the statistical gains are minor. Preliminary estimates of V and Q are obtained by computing $\hat{\sigma}_{n,N}^2$, $\hat{V}_n$ and the estimator $\hat{Q} = \frac{n}{6\pi}\sum_{j=1}^{n}(\delta_j(\tilde{p}))^4$ for the integrated quarticity using 2-minute returns. In the Fourier transform context, a suitable estimate of the integrated quarticity can be derived; nevertheless, the use of such an estimate instead of realized integrated quarticity does not affect much the numerical results. The parameter values used in the simulations are taken from the unpublished Appendix to [Bandi and Russell, 2005a] and reflect the features of IBM time series: $\alpha = 0.01$, $\beta = 1.0$, $\nu = 0.05$, $\xi = 0.000142$. The initial value of σ^2 is set equal to one, while $p(0) = \log 100$. The simulations are run for 500 daily replications, using the computer language Matlab. While in the absence of microstructure noise the MSE decreases as the sampling frequency increases, this is no longer true when microstructure effects are introduced (Fig. 1, Panel A). In agreement with the theoretical results, the sharp spike as ρ approaches zero shows that both the realized variance and the Fourier estimator with $N = n/2$ cannot be consistent estimates of the integrated variance of the underlying log price process in the presence of microstructure noise. Panel B is the same graph as Panel A, plotted on a different scale. The minimum of the true MSE for both estimators is 0.0022 attained for 3.82 minute returns.

As a matter of fact, when analyzing high frequency time series the diffusion model (24) does not hold for small time steps and microstructure effects can affect the computation of the Fourier coefficients. This is shown in Figure 2, where the average $\hat{\sigma}_{n,N}^2$ is plotted as a function of the highest frequency N employed in the Fourier expansion, when all the observations are used ($n = 21600$). As remarked in [Barucci and Renò, 2002], if $d\tilde{p}(t)$ was normally distributed then, as N increases, the plot should tend to the fixed (and known) average integrated variance value (i.e. 0.2499). This is not the case: for a frequency larger than a certain value (denoted by N_{cut}) the Fourier coefficients tend to increase inconsistently, as a consequence of the negative serial correlation of the contaminated returns. In our setting, this happens approximately for $N_{cut} = 90$, which corresponds roughly to a time step of $6 \cdot 60/(2N_{cut}) = 2$ min.
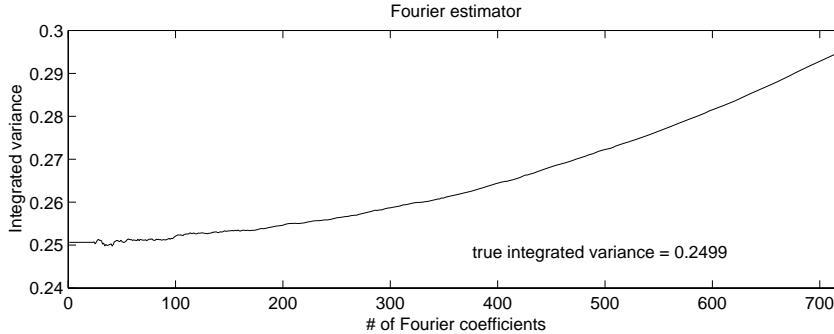
Figure 2: Average $\hat{\sigma}_{n,N}^2$ as a function of the highest frequency N employed in the Fourier expansion and $n = 21600$ (quote-to-quote returns).

This suggests to cut the highest frequencies in the computation of the integrated volatility, i.e. we compute the Fourier expansion for $N = min(n/2, N_{cut})$ when n grows too high, i.e. for high frequency data. Moreover, from the theoretical results of sections 4 and 5, this cutting procedure should result in a smaller bias and MSE of the Fourier estimator and, ultimately, it should provide "near" consistency of $\hat{\sigma}_{n,N}^2$ as $\frac{N^2}{n} \to 0$. Figure 3, Panel A, shows the true (dotted line) and estimated (solid line) mean integrated volatility across the 500 daily replications, obtained with the truncated ($N_{cut} = 90$) Fourier estimator for various sampling intervals ranging from 1 second to 5 minutes. Panels B and C show the true (dotted line) and estimated (solid line) bias and MSE respectively. We can see that now, when the sampling interval becomes smaller than 2 minutes, both the bias and the MSE start to decrease monotonically as a result of the cutting procedure. The minimum bias and MSE are then attained for quote-to-quote returns. As noticed in [Nielsen and Frederiksen, 2006], this behavior of the Fourier estimator can be attributed to the decomposition of the integrated variance into components of varying frequencies. That is, cutting the highest frequencies in the Fourier expansion implies that high-frequency noise or short-run noise is ignored by the estimator. Hence, by choosing a smaller number of low frequency ordinates to be used for estimation, i.e. by choosing N_{cut} small, it is in principle possible to render the Fourier estimator invariant to short-run noise introduced by market microstructure effects.

Figure 4, Panel A, shows the true (dotted line) and estimated bias as a function of the sampling interval for different values of N_{cut}. All the plots coincide for large sampling interval and start to separate from each other for the time step corresponding to the cutting frequency, i.e. for $\Delta t = 6 \cdot 60/(2N_{cut})$. Moreover, as the cutting frequency N_{cut} is reduced then the Fourier estimator is characterized by smaller bias for every choice of the sampling interval, with the minimum attained for quote-to-quote returns where the estimator is almost unbiased. On the contrary, the MSE (Panel B) shows a more complicated behavior: as the cutting frequency N_{cut} is reduced from 360 to 90, the MSE is reduced as well for every choice of the sampling interval; however, further reduction of N_{cut} results in a larger MSE, especially for very high frequency data. This is due to the $O((2N + 1)^{-1})$ term in (17).

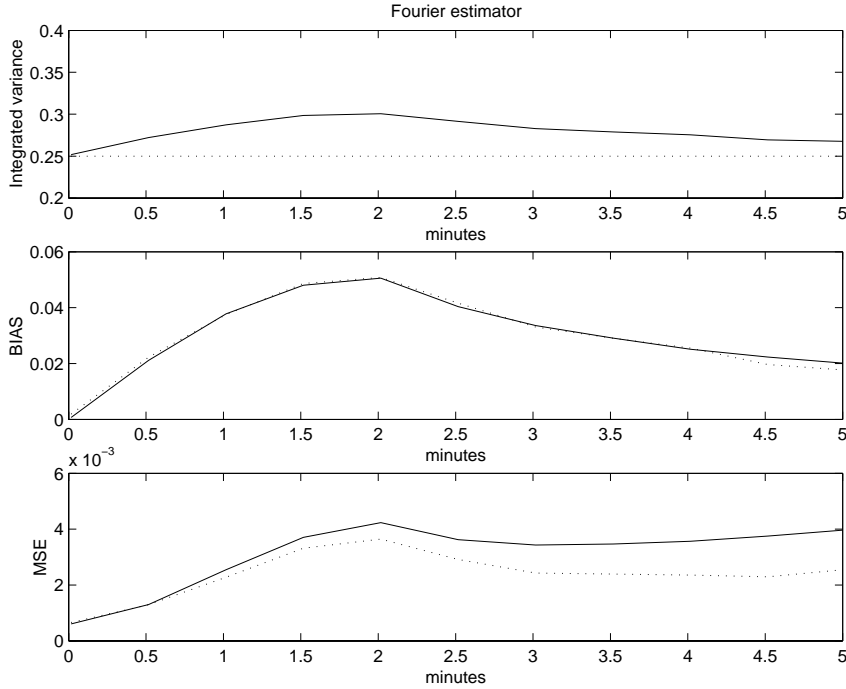The cutting procedure has good effects on the performance of the Fourier estimator

Figure 3: Panel A: true (:) and estimated (-) mean integrated volatility for $N_{cut} = 90$ and various sampling intervals. Panels B and C: true (:) and estimated (-) bias and MSE respectively.

also in terms of the following relative error statistics

$$\mu = E\left[\frac{\hat{\sigma}_{n,N}^2 - V}{V}\right], \quad RMSE = \left\{E\left[\left(\frac{\hat{\sigma}_{n,N}^2 - V}{V}\right)^2\right]\right\}^{1/2},$$

which can be interpreted as relative bias and root mean squared error of the estimator. These are precisely the statistics considered in [Barucci and Renò, 2002] and [Nielsen and Frederiksen, 20 Figure 5 shows the distribution of the relative error $(\hat{\sigma}_{n,N}^2 - V)/V$ in the case of a sampling frequency equal to 3.82 min and in the case of quote-to-quote returns. The latter is characterized by a smaller mean and standard deviation. Here N_{cut} is taken equal to 90.

The analysis above suggests to use quote-to-quote returns and try to minimize MSE as a function of the cutting frequency N_{cut}. This minimization over the integer variable N_{cut} can be performed easily by comparison of the computed MSE values. This is done in Fig 6, where the true and estimated bias and MSE of the Fourier estimator are plotted as a function of the number of the Fourier coefficients. The minimum of the true MSE is 2.88e-4 and is attained for $N_{cut} = 264$ which, at least theoretically, corresponds to a sampling frequency of $6 \cdot 60/(2 \cdot 264) = 0.68$ min, much higher than the value in Fig. 1. Nevertheless, we note this does not correspond either to the minimum bias or to the minimum mean and std values of the relative error. In particular, from Table 1 we can see that as the cutting frequency is decreased (yet kept over a suitable level) the mean percentage error of the Fourier estimator decreases, while its standard deviation increases.
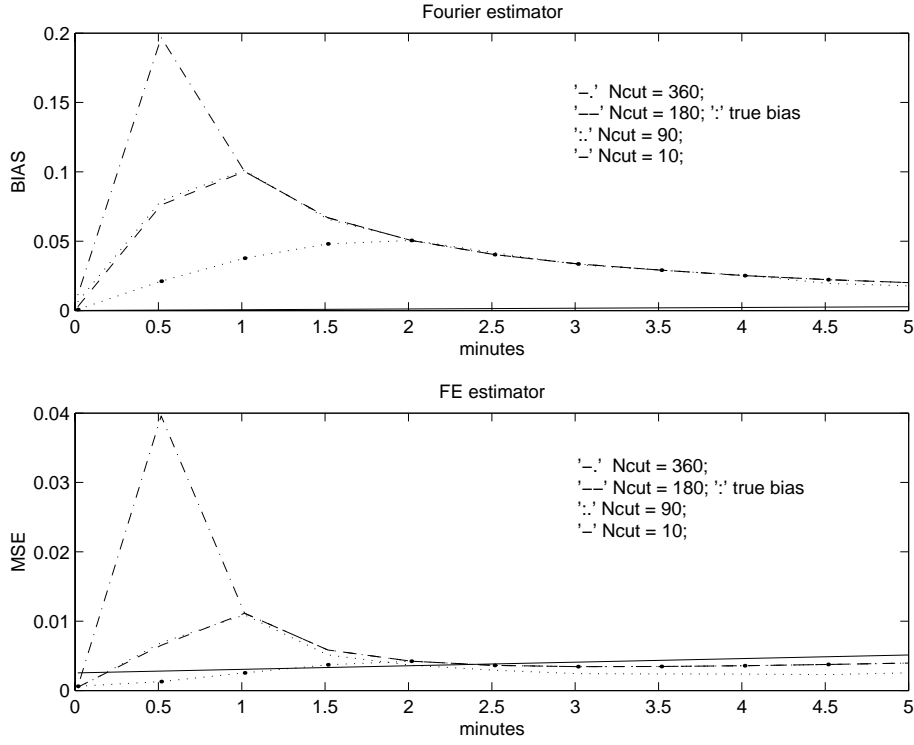
17

Figure 4: Panel A: true and estimated bias as a function of the sampling interval for different values of N_{cut}. Panel B: same format for the MSE.
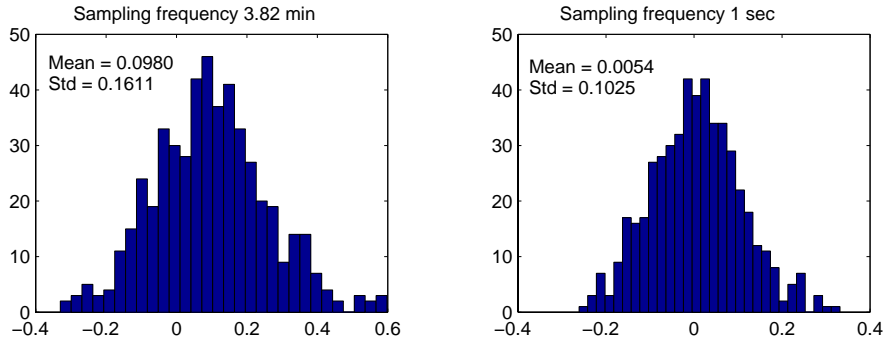

Figure 5: Distribution of $\hat{\sigma}_{n,N}^2 - V/V$ in the case of a sampling frequency equal to 3.82 min and in the case of quote-to-quote returns. $N_{cut} = 90$.
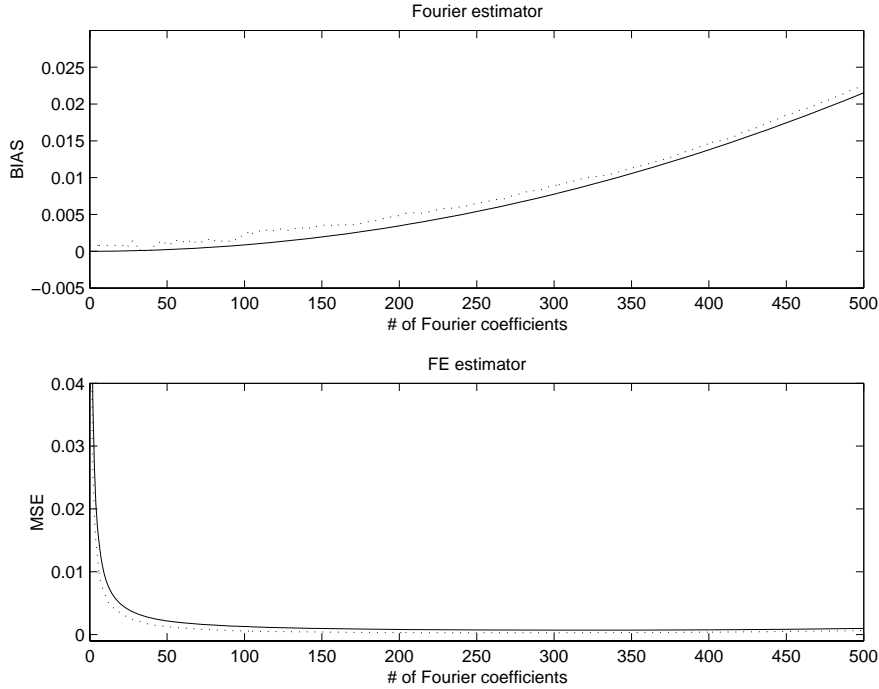
Figure 6: True (:) and estimated (-) bias and MSE of the Fourier estimator as a function of the number of the Fourier coefficients. Quote-to-quote returns.

N_{cut}	60	120	180	240	300	360	420	480	540	600	660
$\hat{\sigma}_{n,N}^2$	0.2512	0.2528	0.2539	0.2559	0.2587	0.2617	0.2658	0.2707	0.2757	0.2816	0.2881
μ	0.0054	0.0119	0.0160	0.0241	0.0355	0.0474	0.0639	0.0833	0.1034	0.1271	0.1532
std	0.1290	0.0881	0.0720	0.0640	0.0595	0.0550	0.0524	0.0494	0.0462	0.0456	0.0435
$RMSE$	0.1290	0.0888	0.0737	0.0684	0.0693	0.0726	0.0826	0.0969	0.1132	0.1350	0.1592

Table 1: Fourier integrated variance and other relative error statistics for different values of N_{cut}. True integrated variance $V = 0.2499$.

From these considerations, we can conclude that the optimal sampling frequency must account for both the bias and the variance of the sampling error through the minimization of the MSE, which is a combination of the two.

In Figure 7 we show the estimated bias and MSE as a function of both the sampling interval ρ ranging from 1 sec to 3 min and the cutting frequency N_{cut} ranging from 0 to 330. Again, the minimum of the estimated MSE is 4.69e-004 and is attained for quote-to-quote data ($\rho = 1$ sec) and $N_{cut} = 240$ which, at least theoretically, corresponds to a sampling frequency of $6 \cdot 60/(2 \cdot 240) = 0.75$ min. At this frequency, the true MSE is 2.92e-004.

Finally, in order to understand more deeply from an empirical point of view how the Fourier estimator relates to the other estimators that have been specifically proposed to handle the microstructure noise, we consider the *flat-top realized kernels* by [Barndorff-Nielsen and al., 2006a], [Barndorff-Nielsen and al., 2006b] with kernels of *Bartlett*, *Cubic* and TH_2 type, the *two-scale estimator* by [Zhang and al., 2005] and the bias cor-
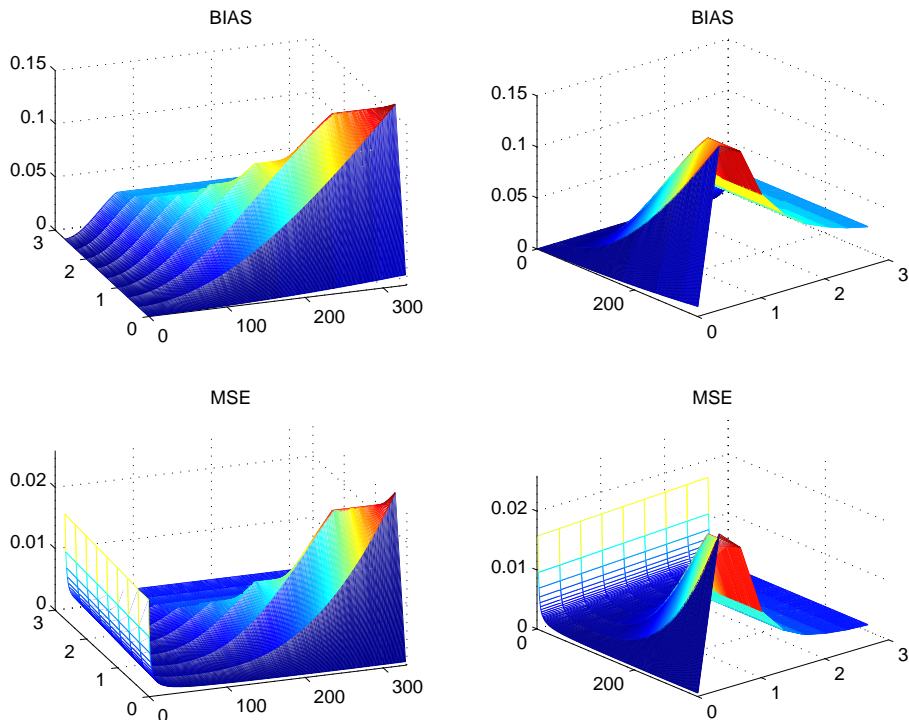
Figure 7: Estimated bias and MSE of the Fourier estimator as a function of the sampling frequency and of the number of the Fourier coefficients.

rected estimator by [Hansen and Lunde, 2006]. The realized kernels may be considered as unbiased corrections of the realized volatility by means of the first H autocovariances of the returns, while the two-scale (subsampling) estimator is a bias-adjusted average of lower frequency realized volatilities computed on S non-overlapping observation subgrids. In particular, when H is selected to be zero the realized kernels become the realized volatility. Finite sample MSE optimal rules for these estimators are considered in [Bandi and Russell, 2006b]. In our analysis, we differentiate from their study in that the optimal MSE-based estimators are designed relying on the true MSE. The comparative analysis of these methods is shown in Tables 2 and 3. The Fourier estimators with Dirichlet (DIR) or Fejer (FEJ) kernels are optimally designed in order to minimize the true MSE with respect to the number of Fourier coefficients N for a given sampling interval $\rho = 1$ sec, 30 sec, 1 min, 5 min. The realized kernels are optimized according to the same criterion with respect to the number of autocovariances H and the two-scale ZMA estimator with respect to the number of subgrids S.

We notice that, at a sampling frequency of 5 min the effects of microstructure noise are not evident. Therefore, the optimal MSE-based value of the parameter H is automatically selected to be zero and the realized kernels become the realized volatility. The optimal MSE-based two-scale ZMA estimator exhibits a larger MSE and negative bias, while the HL estimator shows a slightly larger MSE and a very small bias. Both the Fourier estimators perform comparably and, in particular, the FEJ kernel allows for smaller MSE and bias. At 1 min frequency, the noise induced autocorrelation of returns becomes effec-

20

	MSE				BIAS			
	1 sec	30 sec	1 min	5 min	1 sec	30 sec	1 min	5 min
Fourier DIR	2.88e-4	1.11e-3	1.51e-3	2.31e-3	7.04e-3	1.67e-2	1.49e-2	1.52e-2
Fourier FEJ	2.53e-4	9.64e-4	1.28e-3	2.05e-3	6.79e-3	1.29e-2	1.44e-2	7.41e-3
Bartlett Kernel	9.18e-5	7.50e-4	1.34e-3	2.32e-3	5.96e-4	-2.76e-4	1.56e-4	1.52e-2
Cubic Kernel	9.93e-5	7.50e-4	1.34e-3	2.32e-3	7.47e-4	-2.76e-4	1.56e-4	1.52e-2
TH_2 Kernel	8.99e-5	7.27e-4	1.26e-3	2.32e-3	6.26e-4	-2.56e-4	6.39e-5	1.52e-2
Two-scale ZMA	1.82e-4	2.00e-3	3.24e-3	9.31e-3	-6.45e-3	-2.75e-2	-3.64e-2	-6.47e-2
Real. Vol. HL	3.43e-3	9.24e-4	1.34e-3	5.51e-3	-3.55e-4	-4.49e-4	1.56e-4	-3.70e-3
Real. Vol.	3.76e+1	4.12e-2	1.13e-2	2.32e-3	6.13e+0	2.01e-1	1.03e-1	1.52e-2

Table 2: Comparison of optimized integrated volatility estimators.

	Optimal MSE-based parameter values for N, H, S								
	1 sec	30 sec	1 min	5 min		1 sec	30 sec	1 min	5 min
Fourier DIR	264	79	53	35	Bartlett Kernel	13	2	1	0
Fourier FEJ	386	107	84	50	Cubic Kernel	14	2	1	0
Two-scale ZMA	37	10	8	5	TH_2 Kernel	19	3	2	0

Table 3: Optimal MSE-based parameter values for N, H, S. When H is selected to be zero the realized kernels become the realized volatility.

tive and the realized volatility starts to strongly overestimates the underlying integrated volatility. In this setting, the optimal MSE-based values for H are 1 for the Bartlett and Cubic kernels and 2 for the TH_2 kernel. This correction results in a smaller MSE and negligible bias. Identical performance is obtained with the HL estimator, while the two-scale ZMA shows a larger MSE and negative bias. Both the optimal MSE-based Fourier estimators perform very well in terms of MSE, while having only a slightly higher bias. At higher sampling frequencies the impact from market microstructure effects becomes more evident and the realized volatility becomes progressively unstable. At the highest frequency, the realized kernels provide the best estimate both in terms of MSE and of bias. Moreover, as already observed in the literature, the finite sample performance of the cubic and Bartlett kernels is virtually identical and the Bartlett kernel is slightly preferable at 1 sec frequency. The smooth TH_2 kernel provides the best volatility estimate and tends to select more lags than the others. Very strikingly, for all the sampling frequencies the optimally designed Fourier estimators provide very good results and are practically unaffected by noise, having only a slightly higher MSE for quote-to-quote returns. Notice that the use of the FEJ kernel allows to slightly improve the behavior of the Fourier estimator for very high frequencies. Hence, the Fourier method remains a very attractive estimator even in comparison with methods specifically designed to handle market microstructure contaminations. More specifically, the Fourier estimator is competitive in terms of MSE for sampling frequencies up to 30 sec, while having only a slightly higher bias.
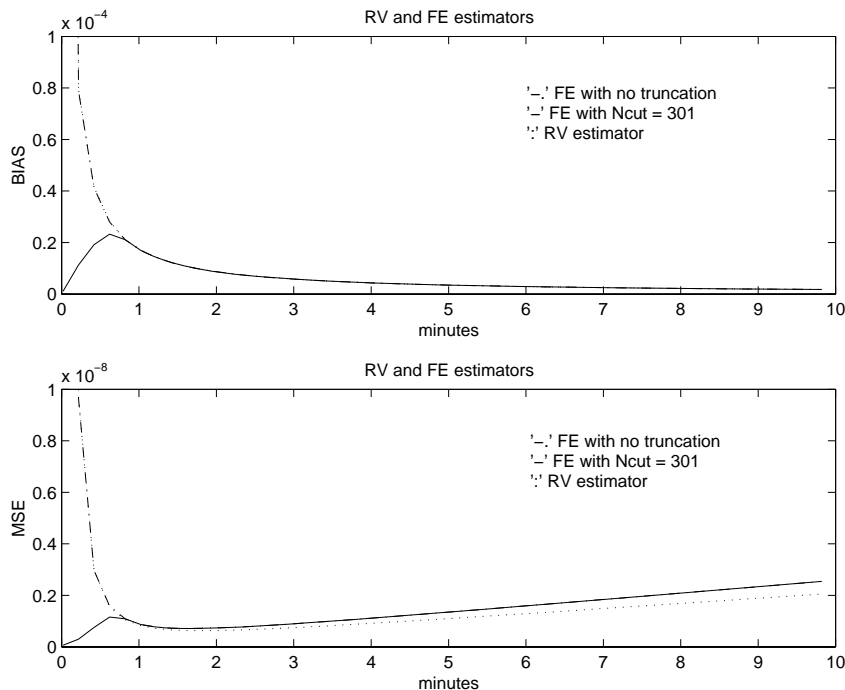
Figure 8: Estimated conditional bias and MSE of the RV estimator and of the Fourier estimator without truncation and with $N_{cut} = 301$, as a function of the sampling interval.

8 Empirical analysis

We analyze quote-to-quote logarithmic prices of the Italian stock index futures, named FIB30, for the period January 11, 2000 to January 31, 2001, for a total of 269 trading days. We use only the prices of the next-to-expiration contracts, which are the most liquid ones, with the FIB30 expiring quarterly. This time series is part of the data set used by [Bianco and Renò, 2006]. The advantage of using the futures is that it is a traded asset and, moreover, the stock index futures is always more liquid than the portfolio which constitutes the index.

Quotes prior to 10 a. m. are removed to eliminate opening quotes from our sample. We have a total of 1514523 quotes over the period and on average a new quote arrives every 5.67 seconds. We construct 10-seconds continuously-compounded log-returns. The smallest return is -0.29% and the largest is 0.30%. The first-order autocorrelation is significantly negative and equal to -0.1518; the second autocorrelation is 0.0144 and the third is 0.0095, with 95% confidence interval $[-0.0485, 0.0485]$. Thus, the MA(1) approximation seems to capture the main economic effects in the data.

Since the requirement of evenly spaced data is not essential for both estimators, we construct intraday returns using a sort of *tick time sampling* scheme [Hansen and Lunde, 2006], where the t_j's are chosen to be the time of the first transaction occurring a fixed period, say 2 minutes, after the previous one. Alternative sampling schemes, such as *calendar time sampling* combined with an interpolation or imputation procedure would give the same qualitative results, eventually introducing further sources of noise. In Fig. 8 we plot

the estimated conditional MSE and bias of the RV estimator and of the Fourier estimator without truncation and with $N_{cut} = 301$ as a function of the sampling interval. The chosen value for N_{cut} is the optimal cutting frequency for quote-to-quote returns, i.e. the one which minimizes the MSE. The MSE's are estimated using the methods discussed in Section 5, with the sample moments constructed using quote-to-quote returns to consistently estimate the moments of the noise. Preliminary estimates of V and Q are obtained using 15-minute returns. The minimum of the MSE for the RV estimator is 6.28e-010 attained at 1.8 minutes, while the minimum of the MSE for the Fourier estimator without truncation is 7.13e-010 at 1.66 minutes. Nevertheless, we see that the curve for the Fourier estimator with $N_{cut} = 301$ reaches the smallest MSE for quote-to-quote returns. The corresponding integrated volatility estimates are 0.1205e-3 for the optimal Fourier estimator without truncation and 0.12155e-3 for the optimal Fourier estimator with $N_{cut} = 301$ and quote-to-quote returns. Therefore, as suggested by our theory, by choosing a suitable N_{cut} it is possible to render the Fourier estimator invariant to short-run noise introduced by market microstructure effects, with consequent efficiency gains. Moreover, since our theoretical results and our simulations indicate that the RV estimator is more biased than the Fourier estimator in the presence of market microstructure noise, the fact that the optimal RV estimate is 0.12139e-3 and that RV estimates for commonly used sampling frequencies (e.g. 13 minutes) are 0.1205e-3 indicates that the actual volatility might be higher on average than predicted by the much used realized volatility, as already noticed in [Nielsen and Frederiksen, 2006].

9   Conclusions

In this paper we have studied the finite sample properties of the Fourier estimator of integrated volatility in the presence of market microstructure noise, both in the case of independent noise as in the case where the noise is correlated with the efficient returns. We find that the Fourier estimator is almost unbiased; in fact, even if we do not proceed to any bias correction of the estimator, we prove that the bias of a finite sample can be made negligible by a suitable cutting of the highest frequencies. Moreover we prove that the mean squared error of the Fourier estimator is substantially unaffected by the presence of microstructure noise by choosing in an appropriate way the number of Fourier coefficients to be included in the estimation, as indicated explicitly by the mean squared error computation. These properties prove the effectiveness of the Fourier estimator of volatility even in the presence of different type of microstructure noise effects. The theoretical results are confirmed by a careful empirical analysis, which suggests as a rule for building efficient Fourir estimators to use quote-to-quote returns and minimize the estimated MSE as a function of the cutting frequency N_{cut}. This minimization over the integer variable N_{cut} can be performed easily and with no computational effort by comparison of the computed MSE values.

10 Acknowledgments

References

[Aït-Sahalia, Mykland and Zhang, 2005a] Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2005) How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, **18**, 351–416.

[Aït-Sahalia, Mykland and Zhang, 2005b] Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2005) Ultra high-frequency volatility estimation with dependent microstructure noise. *Working Paper*.

[Andersen and Bollerslev, 1998] Andersen, T. and Bollerslev, T. (1998) Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, 39: 885-905.

[Andersen, Bollerslev and Diebold, 2002] Andersen, T., Bollerslev, T. and Diebold, F. (2002) Parametric and Nonparametric Volatility Measurement. Forthcoming *Handbook of Financial Econometrics*.

[Andersen, Bollerslev and Diebold, 2005] Andersen, T., Bollerslev, T. and Diebold, F. (2005) Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. *Working paper*.

[Andersen, Bollerslev, Diebold and Ebens, 2001] Andersen, T., Bollerslev, T., Diebold, F. and Ebens, H. (2001) The distribution of realized stock return volatility. *Journal of Financial Economics*, **61**, 43–76.

[Andersen and al., 1999] Andersen, T., Bollerslev, T., Diebold, F. and Labys, P. (1999) (Understanding, optimizing, using and forecasting) Realized volatility and correlation. *Working paper*.

[Bai and al., 2005] Bai X., Russell, J.R. and Tiao, G. (2005) Effects on non-normality and dependence on the precision of variance estimates using high frequency data. *Working paper*.

[Bandi and Russell, 2005a] Bandi, F.M. and Russell, J.R. (2005) Microstructure noise, realized variance and optimal sampling. *Working paper*.

[Bandi and Russell, 2005b] Bandi, F.M. and Russell, J.R. (2005) Volatility. *Working paper forthcoming in Handbook of Financial Engineering.* Editors J.R.Birge and V.Linetsky.

[Bandi and Russell, 2006a] Bandi, F.M. and Russell, J.R. (2006) Separating market microstructure noise from volatility. *Journal of Financial Economics*, Forthcoming.

[Bandi and Russell, 2006b] Bandi, F.M. and Russell, J.R. (2006) Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. *Working paper*.

[Barndorff-Nielsen and Shephard, 2002] Barndorff-Nielsen, O.E., and Shephard, N. (2002) Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society*, Series B, 64, 253–280.

[Barndorff-Nielsen and Shephard, 2004] Barndorff-Nielsen, O.E., and Shephard, N. (2004) Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics*, **2**, 1–48.

[Barndorff-Nielsen and al., 2006a] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2006) Designing realised kernels to measure ex-post variation of equity prices in the presence of noise. *Working paper*.

[Barndorff-Nielsen and al., 2006b] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N. (2006) Subsampling realised kernels. *Working paper*.

[Bartlett, 1946] Bartlett, M.S. (1946) On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Supplement*, 8, 27–41.

[Barucci, Malliavin, Mancino, Renò and Thalmaier, 2003] Barucci, E., Malliavin, P., Mancino, M.E., Renò, R. and Thalmaier A. (2003) The price-volatility feedback rate: an implementable mathematical indicator of market stability. *Mathematical Finance*, **13** p.17-35.

[Barucci and Renò, 2001] Barucci, E. and Renò, R. (2001) On measuring volatility of diffusion processes with high frequency data. *Economics Letters*, 74: 371-378.

[Barucci and Renò, 2002] Barucci, E. and Renò, R. (2002) On measuring volatility and the GARCH forecasting performance. *Journal of International Financial Markets, Institutions and Money*, 12: 183-200.

[Bianco and Renò, 2006] Bianco, S. and Renò, R. (2006) Dynamics of intraday serial correlation in the Italian futures market. *Journal of Futures Markets*, 26(1): 61-84.

[Campbell et al., 1997] Campbell, J., Lo, A., and MacKinlay, C. (1997) The econometrics of financial markets. Princeton University Press.

[Corsi et al., 2001] Corsi, F., Zumbach, G., Müller, U.A. and Dacorogna, M. (2001) Consistent high-precision volatility from high-frequency data. *Economic Notes*, **30**(2), 183–204.

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

[Glosten and Milgrom, 1985] Glosten, L. and Milgrom, P. (1985) Bid, ask, and transactions prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, **13**, 71–100.

[Hansen and Lunde, 2004] Hansen, P.R. and Lunde, A. (2004) An unbiased measure of realized variance. Working Paper, Brown University.

[Hansen and Lunde, 2005] Hansen, P.R. and Lunde, A. (2005) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics*, **20**, 873–889.

[Hansen and Lunde, 2006] Hansen, P.R. and Lunde, A. (2006) Realized variance and market microstructure noise (with discussions). *Journal of Business and Economic Statistics*, **24**, 127–218.

[Harris, 1991] Harris, L. (1991) Stock price clustering and discreteness. *Review of Financial Studies*, **4**(3), 389–415.

[Hayashi and Yoshida, 2005] Hayashi, T. and Yoshida, N. (2005) On covariance estimation of nonsynchronously observed diffusion processes. *Bernoulli*, **11**, n.2, 359–379.

[Hsieh, 1991] Hsieh, D. (1991) Chaos and nonlinear dynamics: application to financial markets. *Journal of Finance*, 46: 1839-1877.

[Kanatani, 2004] Kanatani, T. (2004) Integrated volatility measuring from unevenly sampled observations. *Economics Bullettin*, 3: 1-8.

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

[Malliavin, 1997] Malliavin, P. (1997) *Stochastic analysis.* A series of comprehensive studies in mathematics, vol.313. Springer-Verlag, 1997.

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

[Malliavin and Mancino, 2005] Malliavin, P. and Mancino, M.E. (2005). A Fourier transform method for nonparametric estimation of volatility. *Working paper.*

[Meddahi, 2002] Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics*, **17**, 475–508.

[Newey and West, 1987] Newey, W.K. and West, K.D. (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, **55**, 703–708.

[Nielsen and Frederiksen, 2006] Nielsen, M.O. and Frederiksen, P.H. (2006). Finite sample accuracy and choice of sampling frequency in integrated volatility estimation. *Journal of Empirical Finance,* forthcoming.

[O'Hara, 1995] O'Hara, M. (1995) *Market Microstructure Theory.* Blackwell.

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

[Zhang and al., 2005] Zhang, L., Mykland, P. and Aït-Sahalia, Y. (2005) A tale of two time scales: determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association,* 100, 1394–1411.

[Zhou, 1996] Zhou, B. (1996) High frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics,* **14**(1), 45–52.

11 Appendix: Proofs

Proof. (of Theorem 4.1) By the definitions of Fourier estimator and realized volatility estimator, using (12) and the following identities

$$E[\delta_j(p)\delta_{j'}(p)] = 0 \text{ if } j \neq j'$$

$$E[\varepsilon_j\varepsilon_{j'}] = -E[\eta^2] \text{ if } |j' - j| = 1$$

$$E[\varepsilon_j\varepsilon_{j'}] = 0 \text{ if } |j' - j| > 1$$

$$E[\delta_j(p)\varepsilon_{j'}] = 0,$$

we easily get

$$E[\hat{\sigma}_{n,N}^2 - V] = \tag{25}$$

$$= E[\hat{V}_n - V] + 2E[\sum_{j'=1}^{n}\sum_{j=1}^{j'-1}\frac{1}{2N+1}\sum_{|s|\leq N}e^{is(t_j-t_{j'})}(\delta_j(p)\delta_{j'}(p) + \varepsilon_j\varepsilon_{j'} + \delta_j(p)\varepsilon_{j'})] \tag{26}$$

$$= 2nE[\eta^2] + 2\frac{1}{2N+1}\sum_{j'=1}^{n}\sum_{|s|\leq N}e^{is\frac{2\pi}{n}}E[\varepsilon_{j'}\varepsilon_{j'-1}] \tag{27}$$

$$= 2nE[\eta^2] - 2n\frac{1}{2N+1}\sum_{|s|\leq N}e^{is\frac{2\pi}{n}}E[\eta^2] \tag{28}$$

$$= 2n\ E[\eta^2]\left(1 - \frac{1}{2N+1}\frac{\sin[(2N+1)\frac{\pi}{n}]}{\sin(\frac{\pi}{n})}\right). \qquad \square \tag{29}$$

Proof. (of Theorem 5.1) We introduce the following notation

$$MiX := 2\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}D_N(t_j - t_{j'})\left(\delta_j(p)\delta_{j'}(p) + \varepsilon_j\varepsilon_{j'} + 2\delta_j(p)\varepsilon_{j'}\right) \tag{30}$$

27

and

$$D_N(t) := \frac{1}{2N+1} \sum_{|s| \leq N} e^{ist}.$$

With these notations we expand the MSE of $\hat{\sigma}_{n,N}^2$ and we obtain:

$$E[(\hat{\sigma}_{n,N}^2 - V)^2] = E[(\hat{V}_n - V)^2] + E[(MiX)^2 + 2MiX(\hat{V}_n - V)]. \qquad (31)$$

In virtue of (14) the first addendum in (31) is equal to

$$2\frac{2\pi}{n}(Q + o(1)) + n^2\alpha + n\beta + \gamma \qquad (32)$$

where

$$\alpha := 4(E[\eta^2])^2, \qquad \beta := 4E[\eta^4]$$
$$\gamma := 8E[\eta^2]V + 2E[\eta^2]^2 - 2E[\eta^4].$$

Therefore we have to compute

$$E[(MiX)^2] + 2E[MiX(\hat{V}_n - V)].$$

Let us consider now the first term

$$E[(MiX)^2] = E[(2\sum_{j'=2}^{n}\sum_{j=1}^{j'-1} D_N(t_j - t_{j'})\delta_j(p)\delta_{j'}(p))^2] \qquad (33)$$

$$+E[(2\sum_{j'=2}^{n}\sum_{j=1}^{j'-1} D_N(t_j - t_{j'})\varepsilon_j\varepsilon_{j'})^2] \qquad (34)$$

$$+E[2(\sum_{j'=2}^{n}\sum_{j=1}^{j'-1} D_N(t_j - t_{j'})2\delta_j(p)\varepsilon_{j'})^2] \qquad (35)$$

$$+4E[\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}\sum_{i'\neq j'}\sum_{i<i'} D_N(t_j - t_{j'})D_N(t_i - t_{i'})\delta_j(p)\delta_{j'}(p)\varepsilon_i\varepsilon_{i'}] \qquad (36)$$

$$+4E[\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}\sum_{i'\neq j'}\sum_{i<i'} D_N(t_j - t_{j'})D_N(t_i - t_{i'})\delta_j(p)\delta_{j'}(p)\delta_i(p)\varepsilon_{i'}] \qquad (37)$$

$$+4E[\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}\sum_{i'\neq j'}\sum_{i<i'} D_N(t_j - t_{j'})D_N(t_i - t_{i'})\varepsilon_j\varepsilon_{j'}\delta_i(p)\varepsilon_{i'}]. \qquad (38)$$

We note that the terms (36), (37) and (38) are zero in virtue of the Assumption **A.III**. We start with the term (33). We have

$$4E[(\sum_{j'=2}^{n}\sum_{j=1}^{j'-1} D_N(t_j - t_{j'})\delta_j(p)\delta_{j'}(p))^2] \qquad (39)$$

$$= 4 \sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N^2(t_j - t_{j'}) \int_{t_{j-1}}^{t_j} \sigma^2(s) ds \int_{t_{j'-1}}^{t_{j'}} \sigma^2(s) ds \leq \frac{2\pi}{2N+1} 4Q, \tag{40}$$

being Q the integrated quarticity. We consider now the term (34):

$$E[(2 \sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N(t_j - t_{j'}) \varepsilon_j \varepsilon_{j'})^2] = 4E[\sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N^2(t_j - t_{j'}) \varepsilon_j^2 \varepsilon_{j'}^2] \tag{41}$$

$$+4E[\sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N(t_j - t_{j'}) \varepsilon_{j'} \varepsilon_j \sum_{\substack{i'=2 \\ i' \neq j'}}^{n} \sum_{i=1}^{i'-1} D_N(t_i - t_{i'}) \varepsilon_i \varepsilon_{i'}] \tag{42}$$

$$= 4 \sum_{j'=3}^{n} \sum_{j=1}^{j'-2} D_N^2(t_j - t_{j'}) E[\varepsilon_{j'}^2 \varepsilon_j^2] + 4 \sum_{j'=2}^{n} D_N^2(\frac{2\pi}{n}) E[\varepsilon_{j'}^2 \varepsilon_{j'-1}^2] \tag{43}$$

$$+8 \sum_{j'=3}^{n} \sum_{j=1}^{j'-2} D_N(t_j - t_{j'}) D_N(t_j - t_{j'-1}) E[\varepsilon_j^2 \varepsilon_{j'} \varepsilon_{j'-1}] \tag{44}$$

$$+4 \sum_{j=2}^{n-1} D_N^2(\frac{2\pi}{n}) E[\varepsilon_j^2 \varepsilon_{j-1} \varepsilon_{j+1}] + 8 \sum_{j=4}^{n} \sum_{i=2}^{j-2} D_N^2(\frac{2\pi}{n}) E[\varepsilon_j \varepsilon_{j-1} \varepsilon_i \varepsilon_{i-1}] \tag{45}$$

$$= 4D_N^2(\frac{2\pi}{n})(n-1)(3E[\eta^2]^2 + E[\eta^4]) + 4D_N^2(\frac{2\pi}{n})(n^2 - 5n + 6)E[\eta^2]^2 \tag{46}$$

$$+8D_N^2(\frac{2\pi}{n})(n-2)E[\eta^2]^2 + o(1) \tag{47}$$

$$= 4E[\eta^2]^2 D_N^2(\frac{2\pi}{n})(n^2 - 1) + 4E[\eta^4] D_N^2(\frac{2\pi}{n})(n-1) + o(1) \tag{48}$$

We consider now the term (35):

$$E[(2 \sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N(t_j - t_{j'}) 2\delta_j(p) \varepsilon_{j'})^2] = \tag{49}$$

$$= 16 \sum_{j'=2}^{n} \sum_{j=1}^{j'-1} D_N^2(t_j - t_{j'}) E[\delta_j(p)^2 \varepsilon_{j'}^2] + 16 \sum_{j'=3}^{n} \sum_{j=1}^{j'-2} D_N(t_j - t_{j'}) D_N(t_j - t_{j'-1}) E[\delta_j(p)^2 \varepsilon_{j'} \varepsilon_{j'-1}]$$

$$= 16 \sum_{j'=3}^{n} \sum_{j=1}^{j'-2} D_N^2(t_j - t_{j'}) E[\delta_j(p)^2](2E[\eta^2] - 2E[\eta^2]) + o(1).$$

This concludes the computation of $E[(MiX)^2]$. Now we turn to

$$2E[MiX \ (\hat{V}_n - V)]. \tag{50}$$

The term (50) is equal to

$$2E[MiX \left(\sum_{j=1}^{n} \delta_j(p)^2 - V\right)] + 2E[MiX \ (\sum_{j=1}^{n} \varepsilon_j^2 + 2 \sum_{j'} \sum_{j<j'} \delta_j(p) \varepsilon_{j'})]. \tag{51}$$

29

Notice that the first addendum in (51), which is equal to

$$2E[(2\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}D_N(t_j-t_{j'})\left(\delta_j(p)\delta_{j'}(p)+\varepsilon_j\varepsilon_{j'}+2\delta_j(p)\varepsilon_{j'}\right)\left(V-\sum_{i=1}^{n}\delta_i(p)^2\right)]$$

is zero in virtue of the assumption **A.III** and the fact that

$$E[V-\sum_{j=1}^{n}\delta_j(p)^2]=0.$$

Consider now the second addendum in (51). In virtue of the Assumption **A.III**, with a computation similar to (49), this term reduces to

$$2E[2\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}D_N(t_j-t_{j'})\varepsilon_j\varepsilon_{j'}\ (\sum_{i=1}^{n}\varepsilon_i^2)] \tag{52}$$

$$+2E[4\sum_{j'=2}^{n}\sum_{j=1}^{j'-1}D_N(t_j-t_{j'})\delta_j(p)\varepsilon_{j'}\ (2\sum_{j'}\sum_{j<j'}\delta_j(p)\varepsilon_{j'})] \tag{53}$$

$$=4D_N(\frac{2\pi}{n})(2\sum_{j=2}^{n-1}E[\varepsilon_j\varepsilon_{j+1}\sum_{i=1}^{j-1}\varepsilon_i^2]+4D_N(\frac{2\pi}{n})2\sum_{j=1}^{n-1}E[\varepsilon_{j-1}^3\varepsilon_j] \tag{54}$$

$$+8\sum_{j'=3}^{n}\sum_{j=1}^{j'-2}D_N(t_j-t_{j'})E[\delta_j(p)^2]E[\varepsilon_{j'}^2+2\varepsilon_{j'}\varepsilon_{j'-1}]+o(1) \tag{55}$$

$$=-8D_N(\frac{2\pi}{n})(n^2-3n+2)E[\eta^2]^2-8D_N(\frac{2\pi}{n})(n-1)(3E[\eta^2]^2+E[\eta^4])+o(1).\tag{56}$$

Finally we have

$$RMSE=2\frac{2\pi}{n}(Q+o_p(1))+n^2\hat{\alpha}+n\hat{\beta}+\hat{\gamma} \tag{57}$$

where

$$\hat{\alpha}=4[\eta^2]^2\left(1+D_N^2(\frac{2\pi}{n})-2D_N(\frac{2\pi}{n})\right);$$

$$\hat{\beta}=4E[\eta^4]\left(1+D_N^2(\frac{2\pi}{n})-2D_N(\frac{2\pi}{n})\right);$$

$$\hat{\gamma}=\gamma+4Q\frac{2\pi}{2N+1}+4(E[\eta^2]^2+E[\eta^4])(2D_N(\frac{2\pi}{n})-D_N^2(\frac{2\pi}{n})).$$

The proof is completed. □

Proof. (of Proposition 6.1) We split the computation as follows

$$E[\hat{\sigma}_{n,N}^2-V]=E[\hat{V}_n-V]+2E[\sum_{j'=1}^{n}\sum_{j<j'}D_N(t_j-t_{j'})\delta_j(\tilde{p})\delta_{j'}(\tilde{p})].$$

As showed in [Hansen and Lunde, 2006], firstly we have

$$E[\hat{V}_n - V] = E[\sum_{j=1}^{n}(\delta_j(p))^2 + \tilde{\varepsilon}_j^2 + 2\delta_j(p)\tilde{\varepsilon}_j - V]$$

$$= V + \sum_{j=1}^{n}\alpha^2 E[(\delta_j(p))^2] + \sum_{j=1}^{n}\alpha^2 E[(\delta_{j-1}(p))^2] + 2nE[\eta^2] + 2\alpha\sum_{j=1}^{n}E[(\delta_j(p))^2] - V$$

$$= (2\alpha^2 + 2\alpha)V + 2nE[\eta^2].$$

Secondly we have

$$2E[\sum_{j'=1}^{n}\sum_{j<j'}D_N(t_j - t_{j'})\delta_j(\tilde{p})\delta_{j'}(\tilde{p})] = D_N(\frac{2\pi}{n})\sum_{j=1}^{n}E[\tilde{\varepsilon}_j\delta_{j-1}(p)] + E[\tilde{\varepsilon}_j\tilde{\varepsilon}_{j-1}]. \qquad (58)$$

It is easily seen that

$$E[\tilde{\varepsilon}_j\tilde{\varepsilon}_{j-1}] = -\alpha^2 E[(\delta_{j-1}(p))^2] - E[\eta^2]$$

and

$$E[\tilde{\varepsilon}_j\delta_{j-1}(p)] = -\alpha E[(\delta_{j-1}(p))^2]$$

Therefore (58) is equal to

$$2D_N(\frac{2\pi}{n})\sum_{j=1}^{n}(-\alpha E[\int_{t_{j-2}}^{t_{j-1}}\sigma^2(s)ds] - \alpha^2 E[\int_{t_{j-2}}^{t_{j-1}}\sigma^2(s)ds] - E[\eta^2])$$

$$= -2\alpha D_N(\frac{2\pi}{n})V - 2\alpha^2 D_N(\frac{2\pi}{n})V - 2nD_N(\frac{2\pi}{n})E[\eta^2].$$

Finally the bias of the Fourier estimator becomes

$$2\alpha^2 V(1 - D_N(\frac{2\pi}{n})) + 2\alpha V(1 - D_N(\frac{2\pi}{n})) - 2nE[\eta^2](1 - D_N(\frac{2\pi}{n})). \quad \square$$