

Chapter 3

Estimation of Integrated Volatility

The financial econometrics literature mainly focuses on the *integrated* volatility and cross-volatility on a fixed time horizon. Therefore, this chapter is devoted to the estimation of these quantities. In the context of the Fourier estimation method, the integrated volatilities are computed by simply taking the 0-th Fourier coefficient in formula (2.13). We begin with the study of the univariate estimator, for the ease of notation; nevertheless, the results holding for this case can be easily extended to the multivariate estimator that will be studied in Section 3.3, with special care to be paid for the asynchronous data case. Then, the issue of feasibility for these results is discussed by providing an estimator of the error asymptotic variance, called *quarticity*. Finally, the properties of the Fourier estimator versus different integrated volatility estimators proposed in the literature are outlined.

3.1 Univariate Estimator

The main results shown in Chapter 2 explain why the Fourier estimation method is defined to deal with multivariate problems by its own nature. However, for the sake of simplicity, we first define the Fourier estimator of volatility in the univariate case.

Let the asset price process follow the Itô stochastic differential equation

$$dp(t) = \sigma(t) dW(t) + b(t) dt, \quad (3.1)$$

where W is a Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let σ and b be random processes, adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$ which supports also the Brownian motion driving the asset price $p(t)$ (see Definitions A.2, A.3 and Remark A.1) and satisfying the following integrability conditions

$$E\left[\int_0^T b^2(t) dt\right] < \infty, \quad E\left[\int_0^T \sigma^4(t) dt\right] < \infty. \quad (3.2)$$

Remark 3.1. Model (3.1) is referred to as a Brownian semimartingale¹ and includes all the financial models considered in the following chapters, except for the jump-diffusion studied in Section 4.3 which belongs to a more general class of semimartingales. Non-semimartingale models for financial markets are also actively studied, such as those based on fractional Brownian motion introduced by Mandelbrot and Van Ness (1968) and, more recently, the Volatility Modulated Volterra processes. The latter have found applications not only in finance but also in the modeling of the dynamics of turbulence, where the volatility is also a key concept (see, e.g., Barndorff-Nielsen and Schmiegel (2008)).

For any positive integer n , let $\mathcal{S}_n := \{0 = t_{0,n} \leq \dots \leq t_{k_n,n} = 2\pi\}$ be the trading dates of the asset, i.e., the observation times of the asset price. For simplicity, in the following we take $k_n = n$ and we often omit the second subscript. Denote by $\rho(n)$ the mesh size of the partition \mathcal{S}_n , which is defined as $\rho(n) := \max_{0 \leq i \leq n-1} |t_{i+1} - t_i|$. Moreover, let $\delta_i(p) := p(t_{i+1}) - p(t_i)$.

For any integer s , $|s| \leq 2N$, consider the discretized Fourier coefficients of the asset returns

$$c_s(dp_n) := \frac{1}{2\pi} \sum_{i=0}^{n-1} e^{-ist_i} \delta_i(p). \quad (3.3)$$

Then, for any integer k , $|k| \leq N$, define

$$c_k(\sigma_{n,N}^2) := \frac{2\pi}{2N+1} \sum_{|s| \leq N} c_s(dp_n) c_{k-s}(dp_n). \quad (3.4)$$

Note that (3.4) coincides with (2.13) for $i = j$. We will see in (3.11) that it converges in probability to the k -th Fourier coefficient of $\sigma^2(t)$. In particular, for $k = 0$, it converges to

$$\mathcal{F}(\sigma^2)(0) := \frac{1}{2\pi} \int_0^{2\pi} \sigma^2(t) dt.$$

Therefore, according to (3.4), the *Fourier estimator of the integrated volatility* over $[0, 2\pi]$, namely the random variable $\int_0^{2\pi} \sigma^2(t) dt$, is defined as

$$\hat{\sigma}_{n,N}^2 := \frac{(2\pi)^2}{2N+1} \sum_{|s| \leq N} c_s(dp_n) c_{-s}(dp_n), \quad (3.5)$$

where $c_s(dp_n)$ is given in (3.3). By substituting formula (3.3) into (3.5), the Fourier estimator can be equivalently expressed as

$$\hat{\sigma}_{n,N}^2 = \sum_{j=0}^{n-1} \sum_{j'=0}^{n-1} D_N(t_j - t_{j'}) \delta_j(p) \delta_{j'}(p), \quad (3.6)$$

where $D_N(x)$ is the rescaled Dirichlet kernel defined by (2.9). The representation (3.6) helps us to compare the Fourier estimator with the volatility estimators based

¹ Interested readers can find a deep study of semimartingale theory in Protter (1992).

on the quadratic variation formula (2.15). In fact, we can rewrite (3.6) as

$$\widehat{\sigma}_{n,N}^2 = RV_n + \sum_{j=0}^{n-1} \sum_{\substack{j'=0 \\ j \neq j'}}^{n-1} D_N(t_j - t_{j'}) \delta_j(p) \delta_{j'}(p), \quad (3.7)$$

where RV_n denotes the *Realized Volatility* estimator, defined by

$$RV_n := \sum_{j=0}^{n-1} (\delta_j(p))^2. \quad (3.8)$$

Different features of the Fourier estimation method are highlighted by (3.7).

The Fourier estimator incorporates not only the squared log-returns but also the auto-covariances of any order along the time window.

The cross-terms (namely, the second addend in (3.7)) contribute to render the estimator robust to microstructure noise effect (this point will be discussed in Section 5). This feature has early been considered by Zhou (1996) and recently used to correct the bias of the realized volatility-type estimators in the presence of microstructure noise, as in particular for the realized (subsampled) kernels by Barndorff-Nielsen et al. (2008).

The convolution product leading to (3.6) weights the auto-covariances of any order, the weight being dependent on the number of frequencies N , in addition to the lag between observations.

The convolution product (3.5) can be weighted with different smoothing kernels in order to filter progressively high modes, for instance,

$$\frac{(2\pi)^2}{N+1} \sum_{|s| \leq N} \left(1 - \frac{|s|}{N}\right) c_s(dp_n) c_{-s}(dp_n).$$

This modified convolution formula, which will be considered in Section 5, leads to the following version of the Fourier estimator

$$\widetilde{\sigma}_{n,N}^2 := \frac{1}{N+1} \sum_{j=0}^{n-1} \sum_{j'=0}^{n-1} F_N(t_j - t_{j'}) \delta_j(p) \delta_{j'}(p), \quad (3.9)$$

where

$$F_N(x) := \sum_{|s| \leq N} \left(1 - \frac{|s|}{N}\right) e^{isx} = \frac{1}{N+1} \frac{\sin^2((N+1)x/2)}{\sin^2(x/2)} \quad (3.10)$$

is the Fejér kernel.

3.1.1 Asymptotic Results

Assume that the price process is described by model (3.1). Asymptotic conditions required for the irregular time grid are stated in Malliavin and Mancino (2009) Theorem 4.1, as well as the details of the proofs.

Consistency. Suppose that $\rho(n) \rightarrow 0$ as $n \rightarrow \infty$. The following asymptotic results hold in probability.

- (i) Let $c_k(\sigma_{n,N}^2)$ be defined in (3.4). For any k , it holds

$$\lim_{n,N \rightarrow \infty} c_k(\sigma_{n,N}^2) = \mathcal{F}(\sigma^2)(k). \quad (3.11)$$

The consistency in probability of the Fourier estimator of the integrated volatility (3.5) immediately follows from (3.11) for $k = 0$. Due to the relevance of the integrated volatility estimator for applied purposes, we separately state this result.

- (ii) Let $\widehat{\sigma}_{n,N}^2$ be defined in (3.5), then it holds

$$\lim_{n,N \rightarrow \infty} \widehat{\sigma}_{n,N}^2 = \int_0^{2\pi} \sigma^2(t) dt. \quad (3.12)$$

Central Limit Results. The asymptotic error distribution is Gaussian, with optimal rate² and variance under the assumption that the relative growth rate between the number of the Fourier frequencies N and the number of data n converges to $(1/2)k$, $k = 1, 2, \dots$. In this case, the following stable convergence in law³ holds

$$\rho(n)^{-1/2} \left(\widehat{\sigma}_{n,N}^2 - \int_0^{2\pi} \sigma^2(t) dt \right) \rightarrow \mathcal{N} \left(0, 2 \int_0^{2\pi} \sigma^4(t) dt \right). \quad (3.13)$$

Remark 3.2. The value of the asymptotic variance is linked to the ratio between n and N ; more precisely, if we assume that $N/n \rightarrow c > 0$ as $N, n \rightarrow \infty$, the asymptotic variance is $(1 + 2\eta(c)) 2 \int_0^{2\pi} \sigma^4(t) dt$ with

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

where $\widetilde{c} := 2c$ and $r(x) := x - [x]$, being $[x]$ the integer part of x (see Clement and Gloter (2011) Lemma 1 for details of this computation). Note that $\eta(c) = 0$ if we choose $c = (1/2)k$, k being a positive integer, and $\eta(c)$ is positive, otherwise. This remark justifies the choice $c = (1/2)k$ in the limit (3.13). Moreover, it is known

² The optimal rate of convergence for a non-parametric estimator of volatility is $O(n^{1/2})$.

³ For an introduction of the concept of stable convergence in law see, e.g., Aldous and Eagleson (1978) and Jacod and Shiryaev (2003).

that N must be chosen less or equal to the Nyquist frequency $n/2$ in order to avoid aliasing effects,⁴ which leads to the choice of $c = 1/2$ as the most suitable here.

In Section 5 it will be shown that the possibility of choosing the cutting frequency $N \ll n$ is an important feature of the Fourier estimator when dealing with high-frequency data. In fact, the market microstructure effects contained in high-frequency data are ruled out with the Fourier estimator by cutting the highest frequencies in the construction of the estimator. When $N/n \rightarrow 0$, the following limit theorem with slightly suboptimal rate holds:

$$\rho(n)^{-1/(2\gamma)} \left(\hat{\sigma}_{n,N}^2 - \int_0^{2\pi} \sigma^2(t) dt \right) \rightarrow \mathcal{N} \left(0, 2 \int_0^{2\pi} \sigma^4(t) dt \right) \quad (3.14)$$

where $\gamma > 1$ is such that $N^\gamma = O(n)$. The proof of (3.14) can be found in Clement and Gloter (2011).

Remark 3.3. The central limit results (3.13) and (3.14) are unfeasible, as the asymptotic variance $2 \int_0^{2\pi} \sigma^4(t) dt$ is not known. However, Section 3.2 will study the estimation of the integrated fourth power of the volatility process (named *quarticity*) by exploiting the estimated Fourier coefficients (3.4).

3.1.2 Finite Sample Properties

The efficiency of the Fourier method for computing the integrated volatility has been analyzed in several papers, see Barucci and Renò (2001, 2002), Hansen and Lunde (2006a,b), Nielsen and Frederiksen (2008), Griffin and Oomen (2011). It has been highlighted that one of the advantages of the Fourier method relies on the fact that it allows us to compute volatility measures from unevenly spaced data. On the contrary, most of the volatility estimators base their theoretical properties on data uniformity so that the values of the process must be imputed on a uniform grid. Nevertheless, such imputation has no negligible effects on the quality of volatility estimates, as the following numerical exercise shows.

The analysis is based on Monte Carlo simulations. Suppose that the asset log-price follows the continuous time GARCH(1,1) model

$$\begin{aligned} dp(t) &= \sigma(t) dW_1(t) \\ d\sigma^2(t) &= \theta(\omega - \sigma^2(t))dt + \sqrt{2\lambda\theta} \sigma^2(t) dW_2(t), \end{aligned} \quad (3.15)$$

where W_1, W_2 are independent standard Brownian motions. We set $\theta = 0.035$, $\omega = 0.6365$, $\lambda = 0.2962$ which are based on the daily Deutschemark-US dollar exchange rate from October 1, 1987 to September 30, 1992 (see Andersen et al. (1999b)). We consider 1000 Monte Carlo repetitions starting from the initial values $p(0) = \log 100$, $\sigma^2(0) = 0.6365$. High-frequency unevenly sampled observations have been

⁴ Section A.2.3 in the Appendix A contains a quick review of the Nyquist frequency.

generated as follows: a 6-hour trading period has been simulated by discretizing (3.15) with a time step of one second, for a total of 21600 observations per day. Then, observation times have been extracted in such a way that the duration between different trades is drawn from an exponential distribution with mean equal to $\tau = 5$ seconds, which corresponds to a value observed for many financial time series. As a result, we have a dataset $\{t_j, p(t_j), j = 0, \dots, n\}$ with t_j unevenly sampled.

Provided a uniform grid with $m + 1$ points $(0, \Delta t, 2\Delta t, \dots, T)$, with $\Delta t = T/m$, the daily integrated volatility can be computed by means of the Realized Volatility as follows⁵

$$RV_{\Delta t} := \sum_{i=0}^{m-1} (p((i+1)\Delta t) - p(i\Delta t))^2. \quad (3.16)$$

The convergence of (3.16) to the integrated volatility relies on a basic result for stochastic integrals (essentially, on Itô formula recalled in Theorem A.2) and holds under general assumptions.⁶ Then, theoretically, an arbitrary precision in the estimate of the integrated volatility can be reached by increasing the frequency of observations. Nevertheless, when a process is discretely observed, in order to compute the sum of squared returns one has to impute the values of the process on the uniform grid. This can be done, for instance, by linear interpolation of adjacent prices. However, interpolation introduces a bias that invalidates the consistency of the quadratic variation estimator. Note also that while the Fourier method uses all the observations, the sum of squared intraday returns may use only a fraction of them, i.e., for low m some observations are lost.

On the other side, when using high-frequency data, market microstructure effects make the asymptotic result useless because the microstructure effects swamp the integrated volatility contribution (see Bandi and Russell (2006, 2011) and the analysis in Chapter 5). Therefore, when estimating volatility with high-frequency data, (3.16) is usually computed with $m = 72$ corresponding to 5-minute returns over a 6-hour trading period, as indicated in Andersen and Bollerslev (1998). In fact, at the 5-minute frequency the effects of microstructure noise are negligible. In our simulation setting we also consider $m = 360$, corresponding to 1-minute returns, and $m = 720$ corresponding to 30-second returns.

The performance of the Fourier method is compared to that of (3.16) with $m = 72, 360, 720$ by the statistics

$$RBIAS = E \left[\frac{\int_0^T \sigma^2(s) ds - \widehat{\sigma}^2}{\int_0^T \sigma^2(s) ds} \right], \quad RRMSE = \sqrt{E \left[\left(\frac{\int_0^T \sigma^2(s) ds - \widehat{\sigma}^2}{\int_0^T \sigma^2(s) ds} \right)^2 \right]},$$

where $\widehat{\sigma}^2$ is $\widehat{\sigma}_{n,N}^2$ or $RV_{\Delta t}$, and $\int_0^T \sigma^2(s) ds$ is the integrated volatility generated in a simulation, whose value is known in the simulation setting. In each simulation,

⁵ Note that the definition of the Realized Volatility is the same as (3.8), but we prefer here to point out the time step size Δt instead of the number of observations.

⁶ A more detailed discussion of the convergence of the Realized Volatility-type estimators can be found in Aït-Sahalia and Jacod (2014) Section 6.

the Fourier estimator $\hat{\sigma}_{n,N}^2$ is built by taking $N = n/2$. The results are shown in Figure 3.1. First, let us consider the Realized Volatility. The 5-minute estimator RV_5

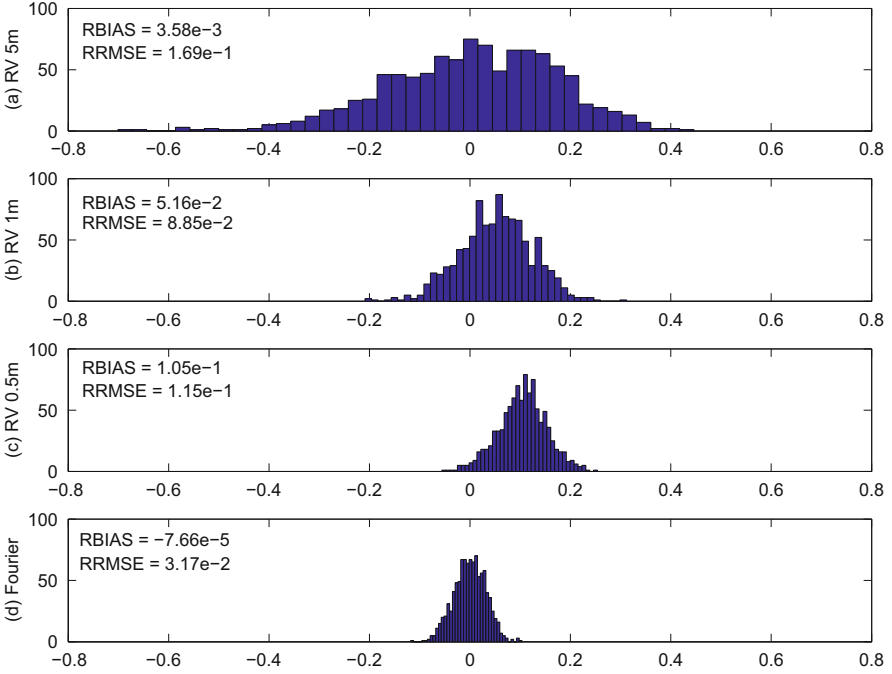


Fig. 3.1 Distribution of $\frac{\int_0^T \hat{\sigma}^2(t)dt - \hat{\sigma}^2}{\int_0^T \hat{\sigma}^2(t)dt}$, where $\hat{\sigma}^2$ are three different estimators of the integrated volatility: (a) estimator (3.16) with $\Delta t = 5$ -minute; (b) estimator (3.16) with $\Delta t = 1$ -minute; (c) estimator (3.16) with $\Delta t = 30$ -second; (d) Fourier estimator. Parameter values: $\theta = 0.035$, $\omega = 0.6365$, $\lambda = 0.2962$. The distribution is computed with 1000 “daily” replications.

provides a downward biased estimate of the integrated volatility ($RBIAS > 0$) with an RRMSE larger than the bias. The 1-minute RV_1 is also downward biased with an RRMSE of the same order of magnitude as the bias in mean. Increasing further the frequency, the estimator $RV_{0.5}$ is characterized by a small variance but a larger bias is observed. This effect can only be due to the interpolation scheme described above, since no other form of noise is present, therefore, it can be linked to non-uniform trading. The Fourier estimator has very small relative bias and relative root mean squared error due to its specific structure that allows for the use of the original non-uniform observations without preliminary manipulation.

3.2 Feasibility

In order to produce feasible central limit theorems for all the estimators of volatility, hence feasible confidence intervals, it is necessary to obtain efficient estimators of the so-called *quarticity*, which appears as conditional variance in the central limit results (3.13) and (3.14). Nevertheless, obtaining reasonably efficient estimators of integrated quarticity is a tougher problem than estimating the integrated volatility when high-frequency data are used, as the effect of microstructure noise is magnified, as remarked by Barndorff-Nielsen et al. (2008). A quite intuitive estimator of quarticity is the *Realized Quarticity* proposed by Barndorff-Nielsen and Shephard (2002) and defined as

$$RQ_n := \frac{n}{3T} \sum_{i=0}^{n-1} (\delta_i(p))^4. \quad (3.17)$$

However, it is consistent only in the absence of noise and sparse sampling is usually employed to face microstructure noise problems (see also Bandi and Russell (2011)). Therefore, it is not reliable with high-frequency data. Mykland (2012) proposed an improved estimator of quarticity, based on a local pre-averaging technique (which will be discussed in Section 5.2.3), which generalizes the estimator (3.17). Recently, Andersen et al. (2014) proposed a new family of neighborhood truncation estimators, that extends existing nearest neighbor estimators based on the minimum of two adjacent absolute returns or on the median of three adjacent absolute returns. Functionals of volatility are also studied by Jacod and Rosenbaum (2013). The next paragraph will show how to estimate quarticity by exploiting the knowledge of the Fourier coefficients of volatility and a basic product formula (see Section A.2.2).

3.2.1 Fourier Estimator of Quarticity

In the previous sections we have seen that all the Fourier coefficients of the variance function can be obtained by (3.4); therefore, these estimated coefficients can now be used as building blocks to estimate different (non-linear) functions of the volatility. In this section the estimated Fourier coefficients of the volatility will be employed to compute the integrated fourth power of the volatility function.

First step: Estimate the Fourier coefficients of the volatility function $\mathcal{F}(\sigma^2)(k)$ by means of (3.4).

Second step: Compute the k -th Fourier coefficient of the fourth power of the volatility, $\sigma^4(t)$, using the product rule of the Fourier series

$$\mathcal{F}(\sigma^4)(k) = \sum_{s+h=k} \mathcal{F}(\sigma^2)(s) \mathcal{F}(\sigma^2)(h). \quad (3.18)$$

Once again, the knowledge of all the Fourier coefficients of the function of interest, $\sigma^4(t)$, allows us to reconstruct the function itself. We focus here on the integrated fourth power. Considering the $k = 0$ Fourier coefficient is enough being interested in the integrated quantity.

Starting from (3.18), the *Fourier estimator of quarticity* is defined by

$$\hat{\sigma}_{n,N,Q}^4 := 2\pi \sum_{|s| < Q} c_s(\sigma_{n,N}^2) c_{-s}(\sigma_{n,N}^2), \quad (3.19)$$

where the $c_s(\sigma_{n,N}^2)$ are the estimated Fourier coefficients of the volatility, in their turn functions of the log-returns $\delta_i(p)$ ($i = 1, \dots, n$) according to (3.3)–(3.4).

Remark 3.4. In order to improve the behavior of the estimator for very high observation frequencies and in the presence of microstructure noise effects, the sum is weighted with a Barlett kernel, as follows:

$$\hat{\sigma}_{n,N,Q}^4 := 2\pi \sum_{|s| < Q} \left(1 - \frac{|s|}{Q}\right) c_s(\sigma_{n,N}^2) c_{-s}(\sigma_{n,N}^2). \quad (3.20)$$

Under the bandwidth conditions $\rho(n)NQ \rightarrow 0$ and $Q^2/N \rightarrow 0$ as $n, N, Q \rightarrow \infty$, the estimator (3.19) (equivalently, (3.20)) is consistent in probability, as proved by Mancino and Sanfelici (2012). The authors provide also a practical way to optimize the finite sample performance of the Fourier estimator as a function of the number of frequencies Q and N , by the minimization of the estimated MSE for a given number n of intra-day observations.

Remark 3.5. Notice that when $Q = 1$ the Fourier estimator of quarticity is simply the squared Fourier estimator of integrated volatility. Indeed, recognizing the considerable imprecision of quarticity estimators, other authors such as Jiang and Oomen (2008) opted for simply squaring integrated variance estimators. In this regard, higher order Fourier coefficients $c_s(\sigma_{n,N}^2)$ for $s \geq 1$ contribute to increase the precision of the quarticity estimator with respect to that naive approach.

By means of the Fourier quarticity estimator, it is possible to show evidence of a feasible version of the Central Limit theorem (3.13). We have repeated the Monte Carlo experiment of Section 3.1.2 for 5000 daily replications and the histograms and QQ plots of the normalized error

$$\rho(n)^{-1/2} \frac{\hat{\sigma}_{n,N}^2 - \int_0^{2\pi} \sigma^2(t) dt}{\left(2\hat{\sigma}_{n,N,Q}^4\right)^{1/2}} \quad (3.21)$$

are plotted in Figure 3.2. On each trading day (24 hours), 1-minute returns are recorded and volatility measures are computed according to the choice $N = n/2 = 720$. The value of the parameters N and Q for the quarticity estimate in the denominator of (3.21) are chosen according to the following criterion: N is set approximately equal to $n^{3/4}$ and, consequently, Q is determined by minimizing the estimate

of the MSE of the quarticity estimator provided by Corollary 3.3 in Mancino and Sanfelici (2012). This yields $N = 234$ and $Q = 2$. The plots reveal that the normalized error is approximately normally distributed with mean 0 and standard deviation 1. The kurtosis and skewness are equal to 3.0623 and -0.1164.

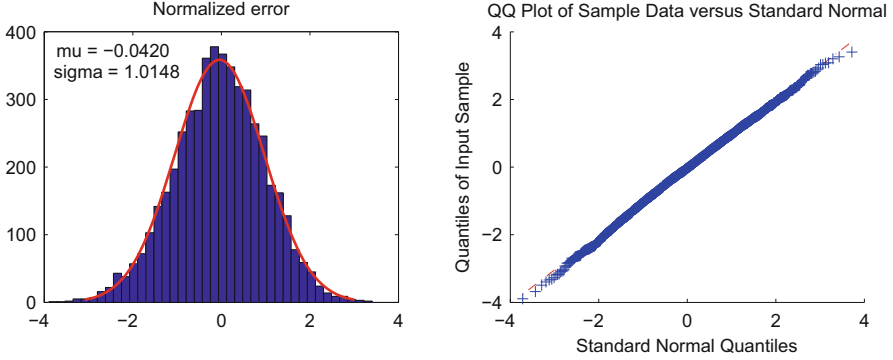


Fig. 3.2 Distribution of $\rho(n)^{-1/2} \frac{\hat{\sigma}_{n,N}^2 - \int_0^{2\pi} \sigma^2(t) dt}{(2\hat{\sigma}_{n,N,M}^4)^{1/2}}$. Parameter values: $\theta = 0.035$, $\omega = 0.6365$, $\lambda = 0.2962$. The distribution is computed with 5000 “daily” replications.

3.3 Multivariate Estimator

The computation of the covariance of financial-asset returns plays a central role for many issues in finance, both in terms of the theoretical understanding of market structure and for its relevant applications. In this respect, the potential of using high-frequency data for the computation of covariances has been shown, among others, by Andersen et al. (2003), Bollerslev and Zhang (2003), Fleming et al. (2003), Bouchaud and Potters (2003).

There are two crucial points pertaining to practical estimation of covariances. First, actual transaction data are recorded at random times. Thus, transaction prices of different assets are usually not observed (or recorded) at the same time. Second, due to such randomness of spacing, a significant portion of the original data sets would be missing at prespecified grid points. However, most of the covariance estimators available in the literature base their statistical properties on uniformity and synchronization of observation times. Consequently, we must choose the common sampling interval length first, and impute or interpolate the missing observations in some way. Then the cleaned data sets are used for the estimation as if they were regularly and concurrently observed, even if the two original processes may have very different observation frequencies. This preprocessing of data sets is called *synchronization*. The choice of the sampling interval and of the methods of imputation may be potential sources of bias, as already highlighted in Section 3.1.2. This may

provide a partial account for the *Epps effect*: the non-synchronicity of the arrival times of trades across markets leads to a bias towards zero in correlations among stocks as the sampling frequency increases.

Following the study in Martens (2004), the different approaches to estimate covariances can be split in two groups. The first group uses interpolation of data, in order to obtain synchronous returns, among them, Dimson (1979), Cohen et al. (1983), Scholes and Williams (1997). A different approach to data synchronization is given by the *refresh time* procedure proposed by Barndorff-Nielsen et al. (2011a) in order to construct the multivariate realized kernels; this synchronization procedure is employed also by Jacod et al. (2009), Christensen et al. (2010). The second group utilizes all transaction data and does not rely on any synchronization methods (see, e.g., Harris et al. (1995), De Jong and Nijman (1997), Hayashi and Yoshida (2005), Brandt and Diebold (2006)).

The Fourier covariance estimator belongs to the second class, because it uses all the available observations, being based on the integration of the time series of returns, as we highlighted in Section 2.2. Therefore, from the practitioner's point of view it is easy to implement as it does not rely on any choice of synchronization methods or sampling schemes.

Assume that the asset prices are described by model (2.1) and integrability conditions analogous to (3.2) hold. Let the trading times be

$$0 = t_{0,n_j}^j < t_{1,n_j}^j < \dots < t_{k_{n_j},n_j}^j = 2\pi, \quad j = 1, \dots, d.$$

For simplicity, we assume $k_{n_j} = n_j$, for any j , and omit the second subscript. For any $j = 1, \dots, d$, set $\rho(n_j) := \max_{0 \leq h \leq n_j-1} |t_{h+1}^j - t_h^j|$.

For any $|k| \leq N$ and $i, j = 1, \dots, d$, the estimator of the k -th Fourier coefficient of the covariance $\Sigma^{i,j}(t)$ is given by (2.13). Therefore, the Fourier estimator of integrated covariance between two assets, labeled by i and j , derives directly from (2.13) by taking the 0-th Fourier coefficient and it is defined as

$$\widehat{\Sigma}_{n_i, n_j, N}^{i,j} := \frac{(2\pi)^2}{2N+1} \sum_{|s| \leq N} c_s(dp_{n_i}^i) c_{-s}(dp_{n_j}^j). \quad (3.22)$$

By substituting (2.12) into (3.22), the estimator (3.22) can be rewritten as

$$\sum_{l=0}^{n_i-1} \sum_{r=0}^{n_j-1} D_N(t_l^i - t_r^j) \delta_{t_l^i}^i(p^i) \delta_{t_r^j}^j(p^j), \quad (3.23)$$

where $D_N(x)$ is the rescaled Dirichlet kernel (2.9).

3.3.1 Asymptotic Results

In order to simplify the notations, we consider two assets, labeled 1 and 2 and focus on the off-diagonal terms of the volatility matrix $\Sigma(t)$.

Suppose that the volatility process paths $\sigma_j^i(t)$ in (2.1) are continuous, e.g. the volatilities are driven by a second diffusion process. Let $\rho(n) := \rho(n_1) \vee \rho(n_2) \rightarrow 0$ as $n \rightarrow \infty$. The following results hold.⁷

Consistency. Assume that $N\rho(n) \rightarrow 0$ as $N, n \rightarrow \infty$. The following limiting results hold in probability.

(i) Let $c_k(\Sigma_{n_1, n_2, N}^{1,2})$ be defined by (2.13). Then, for any k , it holds

$$\lim_{n, N \rightarrow \infty} c_k(\Sigma_{n_1, n_2, N}^{1,2}) = \mathcal{F}(\Sigma^{1,2})(k). \quad (3.24)$$

(ii) In particular, for $k = 0$, (i) implies consistency for the Fourier estimator of integrated covariance given by (3.22). More precisely, it holds

$$\lim_{n, N \rightarrow \infty} \widehat{\Sigma}_{n_1, n_2, N}^{1,2} = \int_0^{2\pi} \Sigma^{1,2}(t) dt. \quad (3.25)$$

Central Limit Theorem. Assume that $N\rho(n) \rightarrow 0$ as $N, n \rightarrow \infty$, then the following stable convergence in law holds:

$$\rho(n)^{-\frac{1}{2\gamma}} \left(\widehat{\Sigma}_{n_1, n_2, N}^{1,2} - \int_0^{2\pi} \Sigma^{1,2}(t) dt \right) \rightarrow \mathcal{N} \left(0, \int_0^{2\pi} \Sigma^{1,1}(t) \Sigma^{2,2}(t) + (\Sigma^{1,2}(t))^2 dt \right) \quad (3.26)$$

where $\gamma > 1$ is such that $N^{-\gamma} = O(\rho(n))$.

Remark 3.6. The rate of convergence is slightly suboptimal, because $1/(2\gamma) < 1/2$. A bias-correction of the Fourier estimator permits to recover the optimal rate of convergence under the condition $N\rho(n) \rightarrow c > 0$; however, this correction can be explicitly computed only under very special sampling schemes, see Theorem 1 in Clement and Gloter (2011) for details. On the contrary, in Sections 3.3.2 and 5.3 it will be shown that it is advisable, when dealing with real data, to use the noncorrected (asymptotically unbiased) estimator with an appropriate cutting frequency to face two features of high-frequency data, namely, the asynchronicity of the observations and the presence of microstructure noise effects.

⁷ Asymptotic conditions required for the irregular/asynchronous time grids and detailed proof can be found in Malliavin and Mancino (2009) Theorem 4.4.

3.3.2 Asynchronicity Issues

This section deeply analyzes the effect of asynchronicity on the Fourier covariance estimator while showing that a suitable choice of the cutting frequency in the series expansion can make it negligible.

Some preliminary remarks are in order. When we want to estimate the covariance of two discretely observed processes using the *Realized Covariance* estimator

$$RC^{1,2} = \sum_{i=0}^{n-1} \delta_i(\bar{p}^1) \delta_i(\bar{p}^2), \quad (3.27)$$

data, if not equally spaced, must be preprocessed in order to make them synchronous. This can be obtained either by linear interpolation or by piecewise constant (previous-tick) interpolation over a (uniform) grid, giving \bar{p}^1, \bar{p}^2 as the interpolated processes. In particular, the second form of imputation of missing data is reasonable for it does not produce extraneous bias when estimating quadratic variations of univariate processes, i.e. when $p^1 = p^2$. However, the synchronization process as well as the choice of the spacing of the interpolation grid is a potential source of bias, especially when the (regular) interval size is small relative to the frequency of actual observations. The downward bias of the realized covariance estimator derives from the fact that each product $\delta_i(\bar{p}^1) \delta_i(\bar{p}^2)$ contributes to the sum if and only if a new observation occurs for both processes in the interval $[t_i, t_{i+1}[$. Otherwise, at least one increment is equal to zero and is ignored in the sum. Such occasions of zero increment will become dominant if the mesh becomes finer. On the other hand, large mesh spacing leads to inefficient use of data.

Realized Covariance (3.27) with linearly interpolated returns may be less biased, but this is because of the downward bias in the volatility measurement due to the linear interpolation illustrated in Section 3.1.2. The spurious positive serial correlation induced by the linear interpolation technique lowers the volatility estimates. Since variances are spuriously measured to be lower, correlations turn out to be spuriously higher, thus compensating in some way the bias due to asynchronicity.

The bias of the Fourier covariance estimator can be easily derived by (3.23) and takes the form

$$E[\widehat{\Sigma}_{n_1, n_2, N}^{1,2} - \int_0^{2\pi} \Sigma^{1,2}(t) dt] = \sum_{l=1}^{n_1-1} \sum_{r=1}^{n_2-1} (D_N(t_l^1 - t_r^2) - 1) E[\int_{I_l^1 \cap I_r^2} \Sigma^{1,2}(t) dt]. \quad (3.28)$$

Remark 3.7. For synchronous observations it holds $D_N(t_l^1 - t_r^2) = D_N(0) = 1$ if $l = r$, otherwise $I_l^1 \cap I_r^2 = \emptyset$, thus implying the right-hand side of (3.28) is equal to zero and the estimator is unbiased. In fact, when data are synchronous, the Fourier estimator of integrated covariance has the same statistical properties of the univariate volatility estimator.

In the general asynchronous case, the Fourier covariance estimator turns out to be *asymptotically unbiased* under the condition $\rho(n)N \rightarrow 0$ as $n, N \rightarrow \infty$, which implies that $(D_N(t_l^1 - t_r^2) - 1)$ in (3.28) converges to 0. Thus, the bias (3.28) can be reduced by tuning the cutting frequency N with the sampling interval $\rho(n)$. Otherwise, a bias may appear.

A suitable choice of small values of N allows one to design rate suboptimal estimators (in the spirit of (3.26)), that are optimal in MSE terms, thus controlling the combined effects of bias and variance in the finite sample.

The following Monte Carlo study clarifies this important point. Assume for simplicity that $p^1 = p^2 = W$, where W is a Brownian motion. The process p^1 is observed at time $t_k^1 = 2\pi k/n$, for $k = 0, 1, \dots, n$. The process p^2 is observed at time $t_k^2 = 2\pi k/n + \pi/n$, for $k = 0, 1, \dots, n-1$ and we observe $p^2(0), p^2(2\pi)$. We want to estimate the integrated volatility and co-volatility divided by 2π , namely the constant $(2\pi)^{-1} \int_0^{2\pi} \Sigma^{i,j}(t) dt = 1$. Assume that $n = 100$ and N ranges from 0 to n , although in practice the condition $N \leq n/2$ should be fulfilled in order to avoid aliasing effects. We consider 10000 replications. In Figure 3.3 the bias and MSE of the Fourier estimators $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,1}$ and $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,2}$ as a function of the number of the Fourier coefficients are shown. It is evident that the estimator $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,1}$ has no bias regardless of the choice of N . However, for small values of N the MSE increases due to a greater variance of the estimator. The situation is clearly different for $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,2}$. In this case, asynchronicity yields an increasing bias as N increases. Obviously, this has effects on the MSE as well. To keep the bias low, we are forced

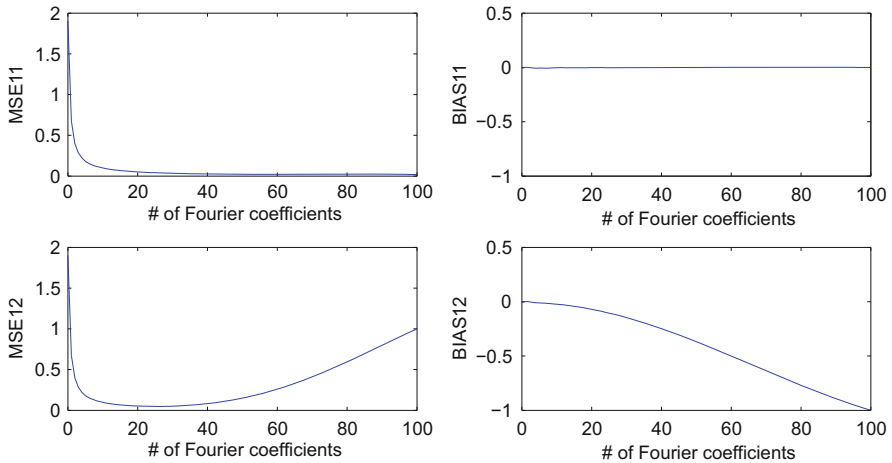


Fig. 3.3 MSE and BIAS of the Fourier estimators $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,1}$ and $(2\pi)^{-1} \hat{\Sigma}_{n,N}^{1,2}$ as functions of the number of the Fourier coefficients.

Table 3.1 Comparison of the MSE-optimal Fourier estimator and the Realized Covariance in terms of bias and variance for the asynchronously observed Brownian motion model.

	$n = 100$	$n = 500$	$n = 1000$	$n = 5000$
Optimal N	25	90	159	492
$c = N/n$	0.25	0.18	0.16	0.09
$\gamma = \log n / \log N$	1.43	1.38	1.36	1.37
$n \times \text{Variance}$	3.68	5.20	6.03	9.60
$N \times \text{Variance}$	0.92	0.94	0.96	0.94
	Variance BIAS		Variance BIAS	
Fourier	3.68e-2 -1.05e-1	1.04e-2 -5.40e-2	6.03e-3 -4.25e-2	1.92e-3 -1.63e-2
RC	1.26e-2 -5.05e-1	2.47e-3 -5.01e-1	1.23e-3 -5.01e-1	2.48e-4 -5.00e-1

to choose small values of N . The best criterion for this choice is the minimization of the MSE, since $\text{argmin} \text{MSE}(N)$ provides the best tradeoff between the different behavior of bias (that is increasing for large N) and variance (that is increasing for small N). Table 3.1 shows a numerical evidence of the Fourier estimator's behavior for an increasing number of data when the MSE-optimal cutting frequency is chosen. The table shows the variance and bias of the Fourier estimator of covariance for the asynchronously observed Brownian motions specified above. The last line lists the same quantities for the Realized Covariance, in order to emphasize the different behavior of the two estimators. Our Monte Carlo experiment consists of 10000 replications for increasing number of observations $n = 100, 500, 1000, 5000$. The Fourier estimator is optimized according to the MSE criterion, i.e., the cutting frequency N is chosen in order to minimize the MSE for any given n . The values of the cutting frequency N are listed in the table as well. It is evident that the ratio $c = N/n$ between the optimal cut-off frequency and the number of observations of each process is decreasing as n increases, in line with the condition $\rho(n)N \rightarrow 0$ prescribed for the asymptotic result (3.26). In this example, assuming the relation $N^\gamma = n$ holds (where γ appears in the asymptotic result (3.26)), we obtain $\gamma \simeq 1.37$. Thus, the relation between the MSE-based optimal N and n in Table 3.1 seems to be well represented by $N = Cn^{3/4}$, with $C \simeq 0.85$. The rate of convergence found in (3.26) is witnessed by the fact that the quantity $n \times \text{Variance} = \text{Var}(\sqrt{n}(\hat{\Sigma}_{n,N}^{1,2} - 1))$ is increasing, while $N \times \text{Variance} = \text{Var}(n^{1/(2\gamma)}(\hat{\Sigma}_{n,N}^{1,2} - 1))$ is stable as n increases. The results obtained with the Realized Covariance, after synchronizing observations by previous-tick interpolation over a uniform grid of n elements, are totally biased.

3.3.3 Comparison Study

A better understanding of the features of the Fourier covariance estimator in the presence of irregular and asynchronous data can be obtained through a comparison with other estimators. In Precup and Iori (2007) two interpolation based methods (the traditional Pearson coefficient and the Co-volatility weighted method proposed by Dacorogna et al. (2001)) are compared with the Fourier one. The authors show

that the Fourier method outperforms the two others in terms of generating more accurate estimates, not oversensitive to the choice of returns time scale in any narrow range. A different approach is proposed by Oya (2005), who applies the sub-sampling bias-correction method of Zhang et al. (2005) to the Fourier estimator of the univariate integrated volatility and obtains smaller MSEs than with other bias-corrected estimators.

Besides the Realized Covariance with different low-frequency sampling, we consider here the estimator proposed by Hayashi and Yoshida (2005) to circumvent the drawbacks caused by asynchronicity

$$AO^{1,2} = \sum_{l,r} \delta_{l^1}^1(p^1) \delta_{l^2}^2(p^2) 1_{\{l^1 \cap l^2 \neq \emptyset\}}, \quad (3.29)$$

where the product of the price increments contributes to the summation so long as the corresponding observation intervals are overlapping. We refer to this estimator as the *All-overlapping* (returns) estimator, as suggested by Corsi and Audrino (2010). In Hoshikawa et al. (2008) an empirical comparison between Realized Covariance, the All-overlapping and the Fourier estimator is performed under no market microstructure noise. Nevertheless, the analysis is conducted by letting the frequency N go to infinity without establishing any criterion for the optimal choice of N . The following study corrects this point.

We simulate discrete data from the continuous time bivariate GARCH model

$$\begin{bmatrix} dp^1(t) \\ dp^2(t) \end{bmatrix} = \begin{bmatrix} \beta_1 \sigma_1^2(t) \\ \beta_2 \sigma_4^2(t) \end{bmatrix} dt + \begin{bmatrix} \sigma_1(t) & \sigma_2(t) \\ \sigma_3(t) & \sigma_4(t) \end{bmatrix} \begin{bmatrix} dW_5(t) \\ dW_6(t) \end{bmatrix} \quad (3.30)$$

$$d\sigma_i^2(t) = (\omega_i - \theta_i \sigma_i^2(t))dt + \alpha_i \sigma_i^2(t) dW_i(t), \quad i = 1, \dots, 4,$$

where $\{W_i(t)\}_{i=1}^6$ are independent Wiener processes. The parameters of the model are: $\alpha_1 = 0.1$, $\alpha_2 = 0.1$, $\alpha_3 = 0.2$, $\alpha_4 = 0.2$, $\beta_1 = 0.02$, $\beta_2 = 0.01$, $\omega_1 = 0.1$, $\omega_2 = 0.1$, $\omega_3 = 0.2$, $\omega_4 = 0.2$, $\theta_1 = 0.1$, $\theta_2 = 0.1$, $\theta_3 = 0.1$, $\theta_4 = 0.1$. The initial values for prices and volatilities are $p^1(0) = \log 100$, $p^2(0) = \log 90$, $\sigma_1(0) = 0.9604$, $\sigma_2(0) = 0.5616$, $\sigma_3(0) = 1.2171$, $\sigma_4(0) = 1.3$.

High-frequency evenly sampled returns are generated (through simple Euler Monte Carlo discretization) by simulating second-by-second return and variance paths over a daily trading period of $h = 6$ hours, for a total of 21600 observations per day. Then the observations are sampled according to two different trading scenarios: *regular non-synchronous trading* (Reg-NS) with duration ρ_1 between trades for the first asset and $\rho_2 = 2\rho_1$ for the second and displacement $\delta \cdot \rho_1$ between the two, i.e. the second asset starts trading $\delta \cdot \rho_1$ seconds later but no trade of asset 1 occurs at the same time of a trade of asset 2; specifically, the link between the trading times of the two assets is the following: $t_j^2 = t_{2(j-1)+1}^1 + \delta \frac{\pi}{n_1-1}$ for $j = 1, \dots, n_2$. Moreover, we assume $t_1^1 = 0$, $t_{n_1}^1 = 2\pi$ and $n_2 = n_1/2$. The second trading scenario is *Poisson trading*, where durations between observations are drawn from an exponential distribution with means λ_1 and λ_2 for the two assets, respectively.

In implementing the Fourier estimator $\hat{\Sigma}_{n_1, n_2, N}^{1,2}$, the smallest wavelength that can be evaluated in order to avoid aliasing effects is twice the smallest distance between two consecutive prices (called *Nyquist frequency*),⁸ which under uniform sampling yields $N \leq \min((n_1 - 1)/2, (n_2 - 1)/2)$. Nevertheless, as already pointed out, smaller values of N may provide better variance/covariance measures. Specifically, we choose $N \simeq 0.85 \min(n_1^{3/4}, n_2^{3/4})$.

The Fourier covariance estimator is compared to the Realized covariance $RC_{0.5min}^{12}$ (resp. RC_{1min}^{12} and RC_{5min}^{12}) based on half a minute (resp. 1 minute and 5 minutes) returns and the All-overlapping estimator AO^{12} . The low frequency returns necessary for the Realized covariance-type estimators are obtained by imputation on a uniform grid. The Fourier and All-overlapping estimators use all tick-by-tick data.

The results are reported in Table 3.2. Within each row, entries are the values of the MSE and bias, using 500 Monte Carlo replications. When we consider covariance estimates, an important effect to deal with is the Epps effect. In fact, from Table 3.2 we see that in the Reg-NS setting the effects imputable to non-synchronicity are evident and spoil all the Realized covariance-type estimates based on synchronization.

Table 3.2 Comparison of integrated volatility estimators: $\rho_1 = 5$ sec, $\rho_2 = 10$ sec with a displacement of 3 seconds for Reg-NS trading ($\delta = 2/3$); $\lambda_1 = 5$ and $\lambda_2 = 10$ for Poisson trading scheme.

	Reg-NS		Poisson	
	MSE	BIAS	MSE	BIAS
$\hat{\Sigma}_{n_1, n_2, N}^{1,2}$	2.39e-3	-1.54e-2	3.65e-3	-3.88e-2
$RC_{0.5min}^{12}$	2.78e-2	-1.61e-1	3.13e-2	-1.71e-1
RC_{1min}^{12}	9.29e-3	-8.32e-2	1.01e-2	-8.87e-2
RC_{5min}^{12}	1.31e-2	-1.66e-2	1.25e-2	-2.33e-2
AO^{12}	5.91e-4	-2.74e-3	1.07e-3	1.34e-3

The best performance is given by the AO estimator, immediately followed by the Fourier estimator. Similar considerations hold for the Poisson trading scheme. The AO estimator still ranks first, immediately followed by the Fourier estimator. However, in the latter case the difference between the AO and Fourier estimator in terms of MSE is strongly reduced although the Fourier estimator is more biased.

In conclusion, the Fourier covariance estimator is rather efficient when considering a semimartingale model, in particular when realistic (i.e., non regular) asynchronous trading times are allowed. In Chapter 5 further gains given by a suitable implementation of the Fourier technique will be showed when the observed processes are affected by microstructure noise.

⁸ The notion of Nyquist frequency is discussed in Section A.2.3

3.3.4 Positive Definiteness

From a practical point of view, the choice of which estimators to use should not be only based on the rate of convergence to their asymptotic distributions, which is not necessarily a reliable guide to finite sample performance. In fact, this approach to the comparison of covariance estimators does not have an economic basis and treats overestimates and underestimates of volatility of the same magnitude as equally important. Recent works in the direction of focusing on comparisons which specifically use economic criteria, like forecasting properties, are Andersen et al. (2011), Ghysels and Sinko (2011), or in an asset-allocation context Fleming et al. (2001), Engle and Colacito (2006), Bandi et al. (2008), De Pooter et al. (2008), and Mancino and Sanfelici (2011a). The latter authors study the economic impact of volatility timing versus unconditional mean-variance efficient static asset allocation strategies and of selecting the appropriate sampling frequency or choosing between different bias and variance reduction techniques for the Realized Covariance matrices.

To this end the fact that the estimated covariance matrix preserves its positive semi-definiteness is a primary issue. The estimated covariance matrix using the Fourier methodology, when the Fejér kernel is used, has this important property. In particular, the following Fourier estimator of integrated volatility matrix (introduced in (3.9) for the univariate case and named the *Fourier-Fejér estimator*)

$$\frac{1}{N+1} \sum_{l=0}^{n_i-1} \sum_{r=0}^{n_j-1} F_N(t_l^i - t_r^j) \delta_{t_l^i}(p^i) \delta_{t_r^j}(p^j), \quad i, j = 1, \dots, d, \quad (3.31)$$

where $F_N(x)$ is the Fejér kernel defined in (3.10), is positive semi-definite.

Remark 3.8. When positive definiteness of the covariance matrix is required, the choice of the optimal cutting frequencies for the various volatility measures cannot be obtained independently for each entry and the same N must be used for all the entries. However, numerical experiments by Mancino and Sanfelici (2011b) show that the use of different optimal cutting frequencies N for variances and covariances does not spoil in general the positive definiteness property of the estimator.

Remark 3.9. The Fourier estimator of instantaneous volatility introduced in (2.11) may not preserve positive definiteness, due to the lack of symmetry in the definition. Akahori et al. (2016) proposed a modified Fourier estimator in order to overcome this problem.

Fourier-Malliavin Volatility Estimation

Theory and Practice

Mancino, M.E.; Recchioni, M.C.; Sanfelici, S.

2017, X, 138 p. 25 illus. in color., Softcover

ISBN: 978-3-319-50967-9