SPRINGER BRIEFS IN QUANTITATIVE FINANCE

Maria Elvira Mancino
Maria Cristina Recchioni
Simona Sanfelici

Fourier-Malliavin Volatility Estimation

Theory and Practice

Springer

SpringerBriefs in Quantitative Finance

More information about this series at http://www.springer.com/series/8784

Maria Elvira Mancino • Maria Cristina Recchioni
Simona Sanfelici

Fourier-Malliavin Volatility Estimation

Theory and Practice

Springer

Maria Elvira Mancino
Department of Economics and Management
University of Firenze
Firenze, Italy

Maria Cristina Recchioni
Department of Management
University Politecnica delle Marche
Ancona, Italy

Simona Sanfelici
Department of Economics
University of Parma
Parma, Italy

ISSN 2192-7006 ISSN 2192-7014 (electronic)
SpringerBriefs in Quantitative Finance
ISBN 978-3-319-50967-9 ISBN 978-3-319-50969-3 (eBook)
DOI 10.1007/978-3-319-50969-3

Library of Congress Control Number: 2016963594

Mathematics Subject Classification (2010): 42A38, 42B05, 62G05, 62F12, 62H12, 62P05, 62P20, 91G70

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To our families

Preface

The concept of *volatility* refers to any phenomenon presenting features of instability, unpredictability, and a likeliness to change frequently, often without apparent or cogent reason; in a word, a phenomenon that exhibits random variations. Therefore, it is an essential element of almost all branches of science, and the measurement of its impact and effects is of paramount importance. This book mainly focuses on the measurement of the statistical parameter which Bachelier (1900) called “nervosité” (the coefficient of nervousness) of a market price and which nowadays is referred as variance or volatility in the context of financial applications. Nevertheless, many of the methods and results presented here could be applied to other disciplines (from turbulence to chemistry, from physics to computer science and even medicine).

Ideally, we start from the book chapter “Volatility Estimation by Fourier Expansion” by Malliavin and Thalmaier (2006) and follow the rapid development of the Fourier-Malliavin estimation theory over the last decade. The purpose of this book is to give a picture of the state of the art concerning this theory and to suggest new directions for its application in the study of financial markets. We aim to give the interested reader a clear, comprehensive, and self-contained book on the use of the Fourier-Malliavin technique for volatility estimation, providing the theoretical and numerical tools needed to understand and apply the methodology to real cases. Specifically, readers are given examples and instruments to implement this methodology in various financial settings, and some new applications to real data are proposed. Detailed bibliographic references are pointed out to permit a study in depth. This book will appeal to the financial econometrics and quantitative finance community and, in particular, to PhD students, researchers, and practitioners in these fields.

Chapter 1 briefly introduces the main elements, namely, various concepts of volatility, the peculiar characteristics of market (high-frequency) data, and the Fourier analysis for financial time series. In Chapter 2, the reader is introduced to the basic idea underlying the Fourier-Malliavin method, and some intuitions on the method are anticipated. Chapter 3 mainly focuses on estimating integrated volatility and cross-volatility on a fixed time horizon, e.g. a day, while in Chapter 4, the Fourier estimation of instantaneous volatility is studied. In Chapter 5, the efficiency of the estimation method is analyzed when the observed asset prices are contaminated by market microstructure noise effects, as it happens when high-frequency data are employed. Chapter 6 gives some examples of the potential of the Fourier method to deal with the real-time use of the volatility estimates. The essentials of the mathematical background are presented in Appendix A, which enables the non-expert reader to follow the theory presented in the book. Furthermore, Appendix B provides a collection of MATLAB® codes useful for reproducing the numerical results contained in the book.

Acknowledgments

This book could not have existed without Professor Malliavin's initial interest in mathematical finance applications and without contribution from our direct collaborators and all those who explored and tested the Fourier estimation theory in their research. We are indebted to all them.

Particular thanks go to the organizers and participants of the Conference on Modeling High-Frequency Data in Finance at Steven Institute of Technology in 2010, for their stimulating feedback which led to the survey by Mancino and Sanfelici (2011c) that is the germ of this book project. Moreover, we wish to thank Joseph Teichmann and Christa Cuchiero who kindly contributed to Section 4.3 with some of their codes and to Fabrizio Laurini for his useful comments.

Finally, we would also like to thank the editorial board of "SpringerBriefs in Quantitative Finance"; the Springer staff, in particular Donna Chernyk, for her support; and the anonymous referees for their valuable comments and suggestions.

Firenze, Italy — Maria Elvira Mancino
Ancona, Italy — Maria Cristina Recchioni
Parma, Italy — Simona Sanfelici

Contents

Chapter 1
Introduction

Labitur occulte fallitque volatilis aetas
(Ovidio, Metamorfosi, Liber X v. 519–520)

Measurement of the volatility/covariance of financial-asset returns plays a central role in many issues in finance, e.g., risk and investment management, hedging strategies, forecasting. In connection with financial markets the word *volatility* is usually associated with the concepts of *risk* and *opportunity*, thus referring to a measure (as well as a feeling) of the movements and uncertainty in the markets. As a matter of fact, the constant-volatility assumption prescribed by the Black & Scholes model (Black and Scholes (1973)) does not account for some stylized facts such as variance heteroscedasticity, predictability, volatility smile, covariance between asset returns and volatility (the so-called leverage effect). Therefore, a wide set of time-dependent (stochastic) volatility models have been proposed to model asset-price evolution and to price options coherently with this evidence. Nevertheless, the volatility process is unobservable and its latency leads to the difficult task of developing efficient methods to measure it.

1.1 Implied and Historical Volatility

To measure volatility, both forward- and backward-looking methods are adopted: the *implied* and the *historical volatility* approaches. The former infers volatility levels by using options markets and has been privileged by practitioners for the purpose of forecasting. The implied volatility of an option is the measure of volatility that, when used in an option-valuation model, equates the theoretical value and the market value. If option pricing models are valid, implied volatilities express the market expectation about future volatility. The main reason for using implied volatility is the assumption that the market as a whole “may know some things about the future volatility in the stock that we don’t know,” with Black (1975). Interested readers will

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_1

find empirical and theoretical studies in Rubinstein (1994), Dupire (1994), Derman and Kani (1994) along with many others. More recently, a model-free measure of implied volatility that equals the market risk-neutral expectation of the total return variation has been introduced (see Britten-Jones and Neuberger (2000), Bollerslev et al. (2009, 2011)). On the contrary, the historical volatility measure is based on the magnitude of recent (past) moves of the prices, namely the (annualized) standard deviation of the log-returns. Volatility can be computed through *parametric* or *nonparametric* methods (see, for instance, the insightful review by Andersen et al. (2010)). In the first case, the expected volatility is modeled through a functional form of market or latent variables. In contrast, nonparametric methods address the computation of historical volatility without assuming any functional form of the volatility. The method studied in this book belongs to the second class. Finally, filtering methods have been applied to infer the volatility as well as its empirical distribution from historical asset-price observations, obtaining predictive distributions for multistep forecasts of volatility (among many, relevant contributions are Jacquier et al. (1994), Cvitanic et al. (2006), Chronopoulou and Viens (2012)).

1.2 High-Frequency Data

In the stochastic modeling of financial markets, the instantaneous volatility is described by the diffusion coefficient of a continuous time process. Measuring the diffusion coefficient from the observed asset prices is a challenging task, since data are not available continuously, but only on a discrete time grid. As volatility changes over time, its computation through nonparametric methods concentrates on a small time window (a day, a week), and high-frequency data are employed. In fact, the recent availability of time observations for all quotes and transactions, named *ultra-high-frequency* data by Engle (2000), has improved the capability of computing volatility efficiently, giving us new fundamental instruments and additional information about variation in return volatility, i.e., in the second moments of returns. Early recognition of this potential gain endowed by the use of high-frequency data has been noted by Nelson (1990, 1991), Andersen and Bollerslev (1998). Sophisticated technological tools and computer algorithms to rapidly trade securities have contributed to make high-frequency trading strategies more widely used by practitioners. Whereas at the turn of the twenty-first century, high-frequency trades had an execution time of several seconds, this had decreased to milliseconds and even microseconds by 2010.

At the same time, this fact poses new challenges to researchers both from the empirical and the theoretical sides, as observed early on by O'Hara (1995), Hasbrouck (1996), Goodhart and O'Hara (1997). In fact, the behavior of observed asset prices departs from what is prescribed by theoretical models (frictionless price), being affected by *noise microstructure* effects deriving from bid-ask bounce, infrequent trading, and price discreteness, among others. Furthermore, when computing covariances between returns recorded at the highest available observation frequency, returns are obviously asynchronous across different assets. Thus, the estimation of covariances suffers from a downward bias as the sampling interval is reduced (known as the *Epps effect* by Epps (1979)).

Most often, all these sources of microstructure effects are modeled as a nuisance component, in the form of additive noise components or rounding errors; this is the main approach followed in the present book. However, a very recent line of research on high-frequency data pursues a more modeling-based approach. Cartea and Jaimungal (2015) describe a model of the limit order book where agents solve a combined optimal stopping and control problem. Kercheval and Zhang (2015) propose a machine learning framework to capture the dynamics of limit order books. Other examples include the artificial "zero-intelligence" order-driven market model of Gatheral and Oomen (2010) and the Markovian queueing model of Cont and De Larrard (2013), proposing simple and tractable stochastic models for the dynamics of a limit order book in which orders to buy and sell are centralized and executed against the best available offers in the limit order book. These equilibrium models of limit order markets provide a glimpse into the dynamics of supply and demand and their role in price formation and are an attempt to describe the complex mechanisms producing microstructure effects.

1.3 Fourier Analysis for Volatility Measurement

Considering these specific characteristics of high-frequency data, a number of alternative volatility/covariance estimators have been proposed in the academic literature in the last twenty years. Most of them rely on the *quadratic covariation* formula, a classical result essentially due to Wiener, which permits the volatility in a time interval (integrated volatility) to be recovered from the observed asset prices. The *realized volatility-quadratic variation* estimators have been intensively studied and used for financial-econometrics purposes in a series of papers, and modifications of the realized volatility have been proposed to correct the bias due to microstructure noise (see Aït-Sahalia and Jacod (2014) for an updated presentation).

This book is devoted to studying an alternative nonparametric method originally proposed in Malliavin and Mancino (2002a) to compute the instantaneous multivariate volatility based on Fourier series. Owing to the book by Fourier (1822), Fourier analysis has been used in many fields because it allows one to represent a set of data as a sum of sinusoidal functions. A function of time, which is called *the signal*, is decomposed into the frequencies that constitute it. Therefore, the Fourier transform is frequently called the *frequency domain representation* of the original signal. Fourier analysis has been extensively applied to inference of processes in time-series analysis. However, these methods mainly hinge on the availability of a very long series of data and on the stationary or ergodic properties which are crucial for long time asymptotics. This fact contrasts with the approach of high-frequency data, where a finite horizon is considered and infill asymptotics (i.e., the time between two observations goes to zero) is performed, which exploits tick-by-tick data.

On the other hand, the underlying financial models fail to have stationary or ergodic properties, unlike the usual time series asymptotics prescribes. Regarding this point, the Fourier-Malliavin estimation approach does not assume any long range stationary condition as usually done in the statistical study of time series when using the ergodic theorem to compute a spectral measure or some other invariant from a single realization of the process. However, the fact that we need to construct an estimator of the desired quantity using only a single realization of the process is peculiar to financial experiments because, in contrast to other physical experiments, averaging the quantities obtained in each time window, e.g. one day, is meaningless.

Chapter 2
A First Glance at Fourier Method

Before tackling the estimation procedure in details, this chapter introduces the basic idea underlying the Fourier-Malliavin method, that is a general identity relating the Fourier transform of the (multivariate) volatility function with the Fourier transform of the log-returns. Moreover, some peculiar features of the method are briefly presented which will be more deeply addressed in the next chapters. The method has been originally proposed by Malliavin and Mancino (2002a) to estimate instantaneous multivariate volatilities from high-frequency observations of diffusion processes in a non-parametric way and without any stationarity assumption. The authors aimed at overcoming some difficulties arising from the application of the quadratic variation formula in the commonly used realized covariation methods with financial data.

2.1 Main Convolution Formula

The very first idea which led to the construction of the Fourier-Malliavin volatility estimator consists in the mathematical link between the Fourier transform of the *observed* asset prices and the Fourier transform of the *unobservable* volatility process. This section starts with an illustration of this main result.

From a theoretical viewpoint, suppose for the moment that the prices of d assets $p(t) = (p^1(t), \ldots, p^d(t))$ are observed in *continuous time* over a time interval $[0,T]$ and described by d continuous processes satisfying the following Itô stochastic differential equations[1]

$$dp^j(t) = \sum_{k=1}^{l} \sigma_k^j(t)\, dW^k(t) + b^j(t)\, dt, \quad j = 1, \ldots, d, \tag{2.1}$$

[1] The reader eventually unfamiliar with these dynamics for the price process can find a short introduction in the Appendix A.1.1.

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_2

where $W = (W^1, \ldots, W^l)$ are independent Brownian motions and σ_k^j and b^j are random processes satisfying mild regularity conditions which will be specified in the following sections. From the representation (2.1) the *(time-dependent) volatility*[2] *matrix* is defined as the matrix $\Sigma(t)$, whose (stochastic) entries are

$$\Sigma^{i,j}(t) = \sum_{k=1}^{l} \sigma_k^i(t)\sigma_k^j(t), \quad i, j = 1, \ldots, d. \tag{2.2}$$

For both functions, the asset return and the volatility matrix, the Fourier transform can be defined as follows, for any integer k and $i, j = 1, \ldots, d$, let

$$\mathscr{F}(dp^i)(k) := \frac{1}{2\pi} \int_0^{2\pi} e^{-\mathrm{i}kt}\, dp(t)$$

and

$$\mathscr{F}(\Sigma^{i,j})(k) := \frac{1}{2\pi} \int_0^{2\pi} e^{-\mathrm{i}kt}\, \Sigma^{i,j}(t)dt\,. \tag{2.3}$$

Note that by rescaling the unit of time[3] we can always reduce ourselves to the case where the time window $[0, T]$ becomes $[0, 2\pi]$.

First Step: for any integer k, compute the Fourier coefficients $\mathscr{F}(\Sigma^{i,j})(k)$ of the spot volatilities $\Sigma^{i,j}(t)$ by means of the Fourier coefficients $\mathscr{F}(dp^i)(k)$ of the price process $p(t)$.

Theorem 2.1. *Consider a process $p(t)$ satisfying (2.1). Then, for any $i, j = 1, \ldots, d$, it holds*

$$\frac{1}{2\pi}\mathscr{F}(\Sigma^{i,j}) = \mathscr{F}(dp^i) * \mathscr{F}(dp^j), \tag{2.4}$$

where the convolution product *which appears in (2.4) is defined as follows: for any i, j and for all integers k*

$$(\mathscr{F}(dp^i) * \mathscr{F}(dp^j))(k) := \lim_{N\to\infty} \frac{1}{2N+1} \sum_{|s|\le N} \mathscr{F}(dp^i)(s)\mathscr{F}(dp^j)(k-s). \tag{2.5}$$

The convergence of the convolution product (2.5) is attained in probability.[4]

We give a sketch of the proof that can be found in Malliavin and Mancino (2009). A preliminary step shows that the drift $b := (b^1, \ldots, b^d)$ of (2.1) gives no contribution to the formula (2.4). Therefore, we can assume $b = 0$. For any integer k and $i = 1, \ldots, d$, we set

[2] In the econometric literature the term *volatility* is often used as a synonym of *variance*.

[3] The analogous expressions for the Fourier transforms on $[0, T]$ are given in Appendix A.2, which contains a short introduction to Fourier analysis.

[4] See Definition A.4.

$$\Gamma_k^i(t) := \frac{1}{2\pi}\int_0^t e^{-\mathrm{i}k\tau}\,dp^i(\tau).$$

Then, by definition it holds $\Gamma_k^i(2\pi) = \mathscr{F}(dp^i)(k)$. For any integer $N \geq 1$ and any $|k| \leq N$, define

$$\gamma_k^{i,j}(N) := \frac{1}{2N+1}\sum_{|s|\leq N}\Gamma_s^i(2\pi)\Gamma_{k-s}^j(2\pi). \tag{2.6}$$

Note that the limit of (2.6) for $N \to \infty$ is equal to (2.5). By Itô formula (A.8), it follows that

$$\gamma_k^{i,j}(N) = \frac{1}{2\pi}\mathscr{F}(\Sigma^{i,j})(k) + R_N^{i,j}(k), \tag{2.7}$$

where

$$R_N^{i,j}(k) := \frac{1}{2N+1}\int_0^{2\pi}\sum_{|s|\leq N}\Gamma_s^i(t)d\Gamma_{k-s}^j(t) + \Gamma_{k-s}^j(t)d\Gamma_s^i(t)\,.$$

Therefore, the result holds true if we prove that, for any fixed k, $R_N^{i,j}(k)$ converges to 0 in probability, as $N \to \infty$. By writing $R_N^{i,j}(k)$ more explicitly, it is evident that it is equal to the sum of two analogous terms, each having the following expression

$$\frac{1}{(2\pi)^2}\int_0^{2\pi} dp^j(t_2)\int_0^{t_2} e^{\mathrm{i}kt_1}D_N(t_1-t_2)dp^i(t_1), \tag{2.8}$$

where $D_N(t)$ is the rescaled Dirichlet kernel

$$D_N(t) := \frac{1}{2N+1}\sum_{|s|\leq N} e^{\mathrm{i}st} = \frac{1}{2N+1}\frac{\sin[(N+\frac{1}{2})t]}{\sin\frac{t}{2}}. \tag{2.9}$$

By Itô energy identity (A.5) and some stochastic calculus, the variance of (2.8) is proved to be less or equal to

$$C\int_0^{2\pi} D_N^2(u)du = C\frac{2\pi}{2N+1},$$

where C is a constant, not depending on k. For the last identity, see, e.g., Malliavin (1995). Therefore, letting $N \to \infty$ in (2.7), the proof is completed. □

As soon as all the Fourier coefficients of the volatility matrix's entries have been computed, it suffices to apply an inversion formula to obtain the time-dependent volatility function.

> *Second Step*: Reconstruct the spot volatility matrix $\Sigma(t)$ using the Fourier-Fejér inversion formula.

The reconstruction of the stochastic function of time $\Sigma^{i,j}(t)$ from its Fourier coefficients can be obtained as follows: for $i = 1,\dots,d$ and $|s| \leq 2N$, compute the Fourier coefficients of prices $\mathscr{F}(dp^i)(s)$ and, for any $|k| \leq N$, $i, j = 1,\dots,d$, define

$$\mathscr{F}(\Sigma_N^{i,j})(k) := \frac{2\pi}{2N+1} \sum_{|s|\leq N} \mathscr{F}(dp^i)(s)\,\mathscr{F}(dp^j)(k-s). \tag{2.10}$$

If the volatility matrix has continuous paths, namely the function $t \to \Sigma^{i,j}(t)$ is continuous, then the Fourier-Fejér summation (see Appendix A.2, formula (A.13)) gives almost surely[5] and uniformly in time

$$\Sigma^{i,j}(t) = \lim_{N\to\infty} \sum_{|k|<N} (1 - \frac{|k|}{N})\mathscr{F}(\Sigma_N^{i,j})(k)\,e^{\mathrm{i}kt}\,, \quad \text{for all } t \in (0,2\pi). \tag{2.11}$$

Remark 2.1. In view of the issues we are going to study, we emphasize that all the volatility Fourier coefficients (2.3) are obtained by the formula (2.4). In particular, the 0-th Fourier coefficient

$$\mathscr{F}(\Sigma^{i,j})(0) := \frac{1}{2\pi}\int_0^{2\pi} \Sigma^{i,j}(t)dt$$

is computed. When multiplied by 2π, this coincides with a financially relevant quantity, that is the *integrated cross-volatility*.

Remark 2.2. It is possible to implement the Fourier method by expanding the volatility function $\Sigma^{i,j}(t)$ as a series of sines and cosines, as it has been originally done by Malliavin and Mancino (2002a). This result is a direct consequence of Remark A.2.

2.2 Specific Features of the Fourier Approach

In this section we highlight a few peculiar features of the Fourier estimation approach which result from the application of (2.4) and (2.11) with discretely observed asset price. These properties will be further studied throughout the book, even in comparison with other estimators.

Define the discrete analogue of the quantities introduced in Theorem 2.1. For notational simplicity, let us consider the case of two assets, which trade, respectively, on discrete grids $\{0 = t_0^j < t_1^j < \ldots < t_{n_j}^j = 2\pi\}$, with $j = 1,2$. It is worth noting that we allow irregularly spaced observation times and even asynchronous observations across different assets, as is usually the case with real transaction prices.

For any integer k, $|k| \leq 2N$, let us define the discrete Fourier transform for each asset return

$$c_k(dp_{n_j}^j) := \frac{1}{2\pi}\sum_{l=0}^{n_j-1} e^{-\mathrm{i}kt_l^j}\,\delta_{I_l^j}(p^j), \tag{2.12}$$

where $I_l^j := [t_l^j, t_{l+1}^j[$ and $\delta_{I_l^j}(p^j) := p^j(t_{l+1}^j) - p^j(t_l^j)$, $l = 0,\ldots,n_j-1$ with $j = 1,2$. For any $|k| \leq N$ and $i,j = 1,2$, let us consider the discrete analogue of the convolution (2.5), given by

[5] See Definition A.5.

$$\frac{1}{2N+1} \sum_{|s| \le N} c_s(dp^i_{n_i}) c_{k-s}(dp^j_{n_j}).$$

In virtue of the identity (2.4), the last term, when multiplied by 2π, is the candidate as estimator of the k-th Fourier coefficient of $\Sigma^{i,j}$. Therefore, we define

$$c_k(\Sigma^{i,j}_{n_i,n_j,N}) := \frac{2\pi}{2N+1} \sum_{|s| \le N} c_s(dp^i_{n_i}) c_{k-s}(dp^j_{n_j}). \tag{2.13}$$

Finally, the random function of time

$$\widehat{\Sigma}^{i,j}_{n_i,n_j,N,M}(t) := \sum_{|k| \le M} \left(1 - \frac{|k|}{M}\right) c_k(\Sigma^{i,j}_{n_i,n_j,N})\, e^{\mathrm{i}kt} \tag{2.14}$$

will be called the *Fourier estimator* of the instantaneous volatility matrix $\Sigma^{i,j}(t)$.

We highlight here some particular features of the just described estimation procedure that will be extensively studied in the following chapters.

> The definition of the Fourier spot volatility estimator (2.14) relies on the *integration* of the price observations rather than on a differentiation procedure.

This property is peculiar of the Fourier approach, as opposed to the realized volatility-type estimators (see the recent book by Aït-Sahalia and Jacod (2014) for a comprehensive treatment of these estimators). To be more specific, let us recall the procedure leading to the realized spot volatility-type estimators.

Consider the univariate case, that is the stochastic process p is defined by (2.1) with $d = l = 1$. Firstly, volatility is computed over finite time intervals $[0,t]$ (integrated volatility), relying upon the *quadratic variation* formula defined by

$$\langle p, p \rangle_t := \lim_{n \to \infty} \sum_{0 \le k < t2^n} \left(p((k+1)2^{-n}) - p(k2^{-n})\right)^2. \tag{2.15}$$

In fact, a classical result, essentially due to Wiener, states that the following identity holds almost surely

$$\langle p, p \rangle_t = \int_0^t \sigma^2(s)\, ds, \tag{2.16}$$

where σ^2 is the volatility function (denoted $\Sigma^{1,1}$ in the notation of (2.2)). Then, the spot volatility is derived from (2.16) by differentiation

$$\sigma^2(t) = \lim_{h \to 0} \frac{\int_0^{t+h} \sigma^2(s)ds - \int_0^t \sigma^2(s)ds}{h} = \lim_{h \to 0} \frac{\int_t^{t+h} \sigma^2(s)ds}{h}. \tag{2.17}$$

As a consequence, the realized volatility-type estimators measure the spot volatility at t as (weighted) sample averages of increasingly finer sampled squared (or absolute) returns over smaller and smaller $[t, t+h]$ intervals. The procedure involves a double asymptotics (for $n \to \infty$ and $h \to 0$) in order to perform both the numerical derivative (2.17) and the discretization procedure (2.15). This immediately raises important issues of efficiency and numerical instability, a critical point being the choice of the length of the time interval h.

The computation of the Fourier coefficients for each asset price (2.12) and the Fourier spot cross-volatility estimator (2.14) requires neither equally spaced price observations nor preliminary synchronization of the observed data.

The Fourier estimator uses all the available data through (2.12): the possibility of using all data avoiding any preliminary manipulation of them translates into the direct use of unevenly sampled returns and even asynchronous data in the multivariate case. In fact, when recorded at the highest available observation frequency, asset returns are asynchronous across different assets. On the contrary, the realized covariance-type estimators rely on the *quadratic covariation* formula, which states that, for $i \neq j$,

$$\langle p^i, p^j \rangle_t := \lim_{n\to\infty} \sum_{0\le k<t2^n} \left(p^i((k+1)2^{-n}) - p^i(k2^{-n})\right)\left(p^j((k+1)2^{-n}) - p^j(k2^{-n})\right) \tag{2.18}$$

is equal to

$$\int_0^t \Sigma^{i,j}(s)ds.$$

The definition of quadratic covariation (2.18) requires the data to be synchronous, thus these estimators suffer from a downward bias when applied to asynchronous intraday data.[6]

The effectiveness of the Fourier spot volatility estimator (2.14) is obtained by balancing three parameters: the numbers of data n_j, the cutting frequency N in the convolution formula, and the number M of estimated Fourier coefficients of volatility to be used in the inversion formula. It must hold $M \le N \le n_j,\ j = 1,\ldots,d$. Choosing these parameters according with specific market characteristics guarantees the efficiency of the Fourier estimator with high-frequency market data.

[6] This behavior is known as *Epps effect*, by Epps (1979).

In Section 5 we will show 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; on the other side, the high-frequency noise or short-run noise is ignored by the Fourier estimator by cutting the highest frequencies. In other words, when efficiently implemented, the Fourier estimator uses as much as possible of the available sample path without being excessively biased by the impact of market frictions.

The Fourier estimator is defined as a global estimator, that is an estimator of the path $t \to \Sigma^{i,j}(t)$ over the whole time interval of interest.

This property is manifest in the fact that the convergence of the random function (2.14) to the covariance function holds uniformly in time. Differently from local estimators, which require the bandwidths to be tuned with the specific time t considered, in the case of the Fourier estimator it is possible to choose the cutting frequencies N and M independently of the specific instant of time, still obtaining accurate spot volatility estimates inside the whole observed time range.

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

$$dp(t) = \sigma(t)\, dW(t) + b(t)\, dt, \tag{3.1}$$

where W is a Brownian motion on a filtered probability space $(\Omega, \mathscr{F}, \mathbf{P})$. Let σ and b be random processes, adapted to the filtration $(\mathscr{F}_t)_{t\geq 0}$ 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

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

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_3

Remark 3.1. Model (3.1) is referred to as a Brownian semimartingale[1] 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 $\mathscr{S}_n := \{0 = t_{0,n} \leq \cdots \leq t_{k_n,n} = 2\pi\}$ 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 $\mathscr{S}_n$, 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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

where

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

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 $\rho(n) \to 0$ as $n \to \infty$. The following asymptotic results hold in probability.

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

$$\lim_{n,N\to\infty} c_k(\sigma^2_{n,N}) = \mathscr{F}(\sigma^2)(k). \tag{3.11}$$

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

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

Central Limit Results. The asymptotic error distribution is Gaussian, with optimal rate[2] 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 $(1/2)k,\ k = 1,2,\ldots$. In this case, the following stable convergence in law[3] holds

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

Remark 3.2. The value of the asymptotic variance is linked to the ratio between n and N; more precisely, if we assume that $N/n \to c > 0$ as $N,n \to \infty$, 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 $N << n$ 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 $N/n \to 0$, the following limit theorem with slightly suboptimal rate holds:

$$\rho(n)^{-1/(2\gamma)}\left(\widehat{\sigma}^2_{n,N} - \int_0^{2\pi} \sigma^2(t)dt\right) \to \mathscr{N}\left(0, 2\int_0^{2\pi} \sigma^4(t)dt\right) \tag{3.14}$$

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

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

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 follows[5]

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

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 $\widehat{\sigma}^2_{n,N}$ 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 $\widehat{\sigma}^2_{n,N}$ 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

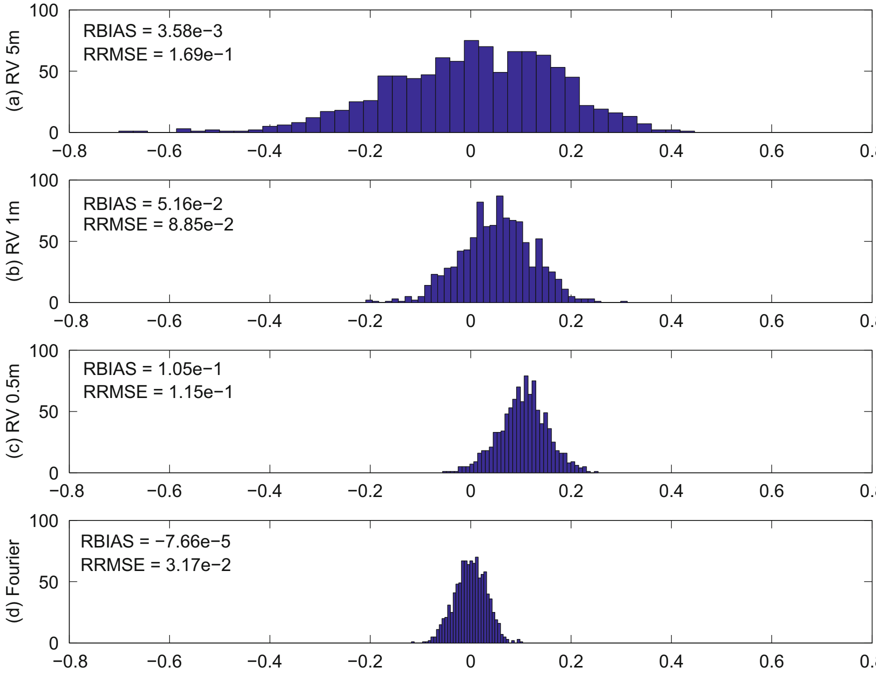

Fig. 3.1 Distribution of $\frac{\int_0^T \sigma^2(t)dt - \widehat{\sigma}^2}{\int_0^T \sigma^2(t)dt}$, where $\widehat{\sigma}^2$ are three different estimators of the integrated volatility: (a) estimator (3.16) with $\Delta t = 5$-minute; (b) estimator (3.16) with $\Delta t = 1$-minute; (c) estimator (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

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

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 $\mathscr{F}(\sigma^2)(k)$ 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

$$\mathscr{F}(\sigma^4)(k) = \sum_{s+h=k}\mathscr{F}(\sigma^2)(s)\,\mathscr{F}(\sigma^2)(h). \tag{3.18}$$

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

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

where the $c_s(\sigma^2_{n,N})$ 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:

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

Under the bandwidth conditions $\rho(n)NQ \to 0$ and $Q^2/N \to 0$ as $n,N,Q \to \infty$, 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 $c_s(\sigma^2_{n,N})$ 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

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

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.

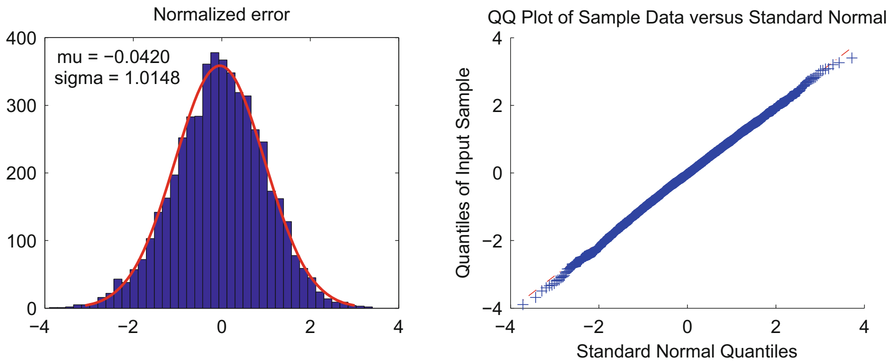

Fig. 3.2 Distribution of $\rho(n)^{-1/2} \frac{\widehat{\sigma}^2_{n,N} - \int_0^{2\pi} \sigma^2(t)dt}{(2\widehat{\sigma}^4_{n,N,M})^{1/2}}$. Parameter values: $\theta = 0.035$, $\omega = 0.6365$, $\lambda = 0.2962$. The distribution is computed with 5000 "daily" replications.

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

$$0 = t^j_{0,n_j} < t^j_{1,n_j} < \ldots < t^j_{k_{n_j},n_j} = 2\pi \;, \quad j = 1,\ldots,d.$$

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

For any $|k| \leq N$ and $i, j = 1,\ldots,d$, 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

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

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

$$\sum_{l=0}^{n_i-1} \sum_{r=0}^{n_j-1} D_N(t^i_l - t^j_r)\delta_{I^i_l}(p^i)\delta_{I^j_r}(p^j), \tag{3.23}$$

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 $\sigma^i_j(t)$ in (2.1) are continuous, e.g. the volatilities are driven by a second diffusion process. Let $\rho(n) := \rho(n_1) \vee \rho(n_2) \to 0$ as $n \to \infty$. The following results hold.[7]

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

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

$$\lim_{n,N\to\infty} c_k(\Sigma^{1,2}_{n_1,n_2,N}) = \mathscr{F}(\Sigma^{1,2})(k). \tag{3.24}$$

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

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

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

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

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 $N\rho(n) \to c > 0$; 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 Asyncronicity 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

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

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 $\overline{p}^1, \overline{p}^2$ 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 $\delta_i(\overline{p}^1)\delta_i(\overline{p}^2)$ 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

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

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 $\rho(n)N \to 0$ as $n,N \to \infty$, 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 $N \le n/2$ 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 $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,1}$ and $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,2}$ as a function of the number of the Fourier coefficients are shown. It is evident that the estimator $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,1}$ 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 $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,2}$. 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

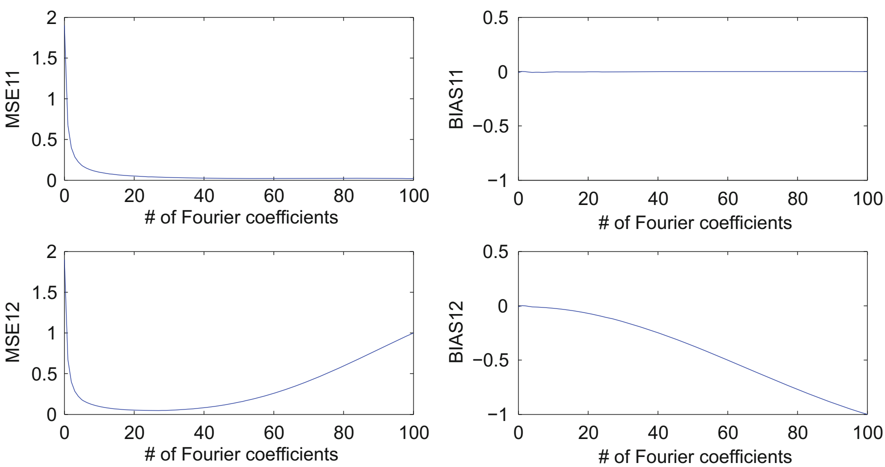

Fig. 3.3 MSE and BIAS of the Fourier estimators $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,1}$ and $(2\pi)^{-1}\widehat{\Sigma}_{n,N}^{1,2}$ as functions of the number of the Fourier coefficients.

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	
$n \times$ Variance	3.68		5.20		6.03		9.60	
$N \times$ Variance	0.92		0.94		0.96		0.94	
	Variance	BIAS	Variance	BIAS	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 $argmin\,MSE(N)$ 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 $\rho(n)N \to 0$ 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 $n \times Variance = Var(\sqrt{n}(\widehat{\Sigma}^{1,2}_{n,N} - 1))$ is increasing, while $N \times Variance = Var(n^{1/(2\gamma)}(\widehat{\Sigma}^{1,2}_{n,N} - 1))$ 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 subsampling 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

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

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

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

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

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 $\widehat{\Sigma}^{1,2}_{n_1,n_2,N}$, 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 $RC^{12}_{0.5min}$ (resp. RC^{12}_{1min} and RC^{12}_{5min}) 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
$\widehat{\Sigma}^{1,2}_{n_1,n_2,N}$	2.39e-3	-1.54e-2	3.65e-3	-3.88e-2
$RC^{12}_{0.5min}$	2.78e-2	-1.61e-1	3.13e-2	-1.71e-1
RC^{12}_{1min}	9.29e-3	-8.32e-2	1.01e-2	-8.87e-2
RC^{12}_{5min}	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*)

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

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.

Chapter 4
Estimation of Instantaneous Volatility

Unlike the integrated volatility, the nonparametric estimation of instantaneous volatility is a relatively recent topic. In the case of deterministic volatility function, Genon-Catalot et al. (1992) proposed a first approach through wavelet series, while Florens-Zmirou (1993), Jacod (2000) developed functional methods, which are local in space, for estimating the volatility as function of the underlying state variable level. Under the stochastic volatility paradigm, Foster and Nelson (1996) first proposed a local estimator of spot volatility from which many refinements have been derived in the subsequent literature. It consists in using a double asymptotics in order to perform both the discretization procedure contained in (2.15) and the numerical derivative involved in formula (2.17). In this scenario, Malliavin and Mancino (2002a) suggested to compute the instantaneous multivariate volatility function through its expansion in trigonometric polynomials, whose coefficients depend on the log-return processes. Other contributions to this field are given, among many others, by Comte and Renault (1998), Andreou and Ghysels (2002), Fan and Wang (2008), Mykland and Zhang (2008), Muller et al. (2011), Ogawa and Sanfelici (2011), Alvarez et al. (2011), Todorov and Tauchen (2012), Zu and Boswijk (2014).

In this chapter, the Fourier estimator of instantaneous multivariate volatility is defined for discrete, unevenly spaced and asynchronously sampled asset prices. Both the asymptotic and the finite sample properties of the estimator are studied. Finally, directions are provided to efficiently implement the estimator with real market data.

4.1 Univariate Estimator

Consider the asset price model (3.1) and notations from Section 3.1. Recalling that any k-th Fourier coefficient of the volatility process can be consistently estimated by (3.4) and using the Fourier-Fejér inversion formula, we define the *Fourier estimator of spot volatility* as follows:

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_4

$$\widehat{\sigma}^2_{n,N,M}(t) := \sum_{|k|\le M} \left(1 - \frac{|k|}{M}\right) e^{\mathrm{i}tk} c_k(\sigma^2_{n,N}), \quad t \in [0, 2\pi]. \tag{4.1}$$

The definition of the estimator (4.1) depends on three parameters, the number of data n and the two frequencies N, M. The choice of the relative growth rate between them is a relevant issue and will be discussed in the following sections.

By elementary calculus, by substituting formula (3.3) into (3.4), the Fourier estimator of spot volatility (4.1) can be expressed as follows:

$$\frac{1}{2\pi} \sum_{j=0}^{n-1} \sum_{j'=0}^{n-1} F_M(t - t_{j'}) D_N(t_{j'} - t_j) \delta_j(p) \delta_{j'}(p), \tag{4.2}$$

where F_M is the Fejér kernel defined by (3.10) and D_N is the rescaled Dirichlet kernel defined by (2.9). We stress the point that the estimator (4.2) contains two terms: the quadratic part

$$\frac{1}{2\pi} \sum_{j=0}^{n-1} F_M(t - t_j)(\delta_j(p))^2 \tag{4.3}$$

and the cross terms

$$\frac{1}{2\pi} \sum_{j=0}^{n-1} \sum_{\substack{j'=0 \\ j' \neq j}}^{n-1} F_M(t - t_{j'}) D_N(t_{j'} - t_j) \delta_j(p) \delta_{j'}(p). \tag{4.4}$$

The quadratic term (4.3) behaves like the Kernel-based spot volatility estimators considered by Fan and Wang (2008), Kristensen (2010). Nevertheless, the second addend (4.4) plays a crucial role in terms of the estimator robustness to microstructure noise effects by means of a suitable choice of the cutting frequency N. This point will be addressed in Chapter 5.

4.1.1 Asymptotic Results

The consistency of the spot volatility estimator (4.1) results from the preliminary proof of the convergence in probability of (3.4) to the k-th Fourier coefficient of the volatility stated by (3.11), for any k. Then, if the volatility path is continuous, the function $\sigma^2(t)$ can be determined in the sup-norm in virtue of the Fejér Convergence Theorem A.13. In fact, the Fourier coefficients represent the building blocks used to get all information about the latent variable, the *volatility*.

Consistency. Let $\widehat{\sigma}^2_{n,N,M}(t)$ be defined in (4.1). The following uniform convergence in probability holds

$$\lim_{n,N,M\to\infty} \sup_{t\in(0,2\pi)} |\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t)| = 0. \tag{4.5}$$

The proof can be found in Malliavin and Mancino (2009). The uniform convergence (4.5) highlights the fact that the Fourier estimator is a *global estimator*, in the sense that it estimates the whole path $t \to \sigma^2(t)$ in the interval of interest. Apart from results on the uniform convergence of the Fourier estimator, very few results are available for the global estimators, such those for the kernel-based estimators by Fan and Wang (2008) and the wavelet-based estimator by Hoffmann (1999) . In this respect, a relevant open problem consists in determining the rate of convergence in (4.5) (as well as in L^p-norm).

Remark 4.1. The Fourier spot volatility estimator works also as a pointwise estimator inside the time interval, but it slightly loses its accuracy near the boundaries. In fact, in order to assume that the price process (3.1) satisfies $p(0) = p(2\pi)$, the process p is eventually modified into the process $\widehat{p}(t) = p(t) - t\,(p(2\pi) - p(0))/(2\pi)$, which has the same volatility as p and $\widehat{p}(0) = \widehat{p}(2\pi)$. However, from the computational point of view, this periodization procedure affects the boundary precision of the estimator, as the Table 4.2 in Section 4.2.2 will show. With the purpose of avoiding this artificial periodization subjacent to Fourier series method, Curato et al. (2016) define an estimator based on the Laplace transform, which has similar features with the Fourier estimator but it is statistically efficient both inside the interval of observations and near the boundary.

Central Limit Result. Under the conditions $N/n \to c$ with $c = 1/2$ as $N, n \to \infty$ and $M/n \to 0$ as $M, n \to \infty$, the following stable convergence in law[1] holds

$$\sqrt{\frac{n}{M}}\ \left(\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{4}{3}\sigma^4(t)\right). \tag{4.6}$$

Note that $N = n/2$ is the *Nyquist frequency*[2] and it provides the optimal[3] asymptotic variance $(4/3)\sigma^4(t)$. Different choices of c are discussed in Mancino and Recchioni (2015), where (4.6) is proved. However, in Chapter 5 we will see that the possibility of choosing the cutting frequency N growing at a lower rate than n (i.e., $N/n \to 0$) is an important feature of the Fourier estimator, when dealing with high-frequency data in the empirical applications, as yet highlighted in Remark 3.6.

With respect to the localizing frequency M, the convergence (4.6) precisely requires $M = O(n^\beta)$ with $(2\nu + 1)^{-1} < \beta < 1$, where $\nu \in (0, 1/2)$ is the Hölder-continuity parameter of the volatility path $\sigma(t)$. Thus, for β close to $1/2$ the rate of convergence becomes $1/4$, which is the optimal rate of convergence for a non-parametric spot volatility estimator. Note that the Hölder-continuity assumption is not restrictive. For instance, it holds if the volatility process σ is driven by a

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

[2] A simple introduction to this important concept can be found in Section A.2.3.

[3] The localized realized volatility estimators has variance equal to $2\sigma^4(t)$, see Chapter 8 of the book by Aït-Sahalia and Jacod (2014) and the discussion by Cuchiero and Teichmann (2015).

second Brownian semi-martingale with bounded[4] coefficients, as it is for well-known stochastic volatility models. This result essentially follows by using the Kolmogorov continuity theorem (A.1) and Lévy's theorem (A.2). In particular, the choice $M = c_M \sqrt{n} \log n$, for a suitable constant c_M, is suggested by (A.2).

4.1.2 Finite Sample Properties

When volatility estimates are needed for empirical purposes, the main concern relies on the finite sample properties of the estimator, as suggested by Griffin and Oomen (2011). Thus, in this section some finite sample properties of the Fourier spot volatility estimator are presented to the reader.

Firstly, we show numerical evidence of the accuracy of the Fourier spot volatility estimator in approximating the volatility path. This is done by comparing the true (simulated) volatility path with the estimated one at different times by means of the respective standardized returns defined by

$$z(t) := \frac{r(t)}{\sigma(t)\sqrt{\Delta t}}, \tag{4.7}$$

where $r(t) := p(t + \Delta t) - p(t)$ is the log-return on the interval Δt. The standardized returns (4.7) are frequently used to study the performance of volatility estimators, because, for any t, the standardized returns are random variables normally distributed with zero mean and variance equal to one for small sampling intervals in the absence of microstructure noise (e.g., see Andersen et al. (2001b), Zu and Boswijk (2014)).

The proposed analysis is carried out using the following stochastic volatility model

$$dp(t) = \mu\, dt + \sigma(t)\, dW_1(t), \tag{4.8}$$

$$\sigma(t) = \exp(\beta_0 + \beta_1\, \tau(t)), \tag{4.9}$$

$$d\tau(t) = \beta_2\, \tau(t)\, dt + dW_2(t), \tag{4.10}$$

where $W_1(t)$ and $W_2(t), t \in [0,T]$, are dependent Brownian motions with correlation λ. The model parameters are: $\mu = 0.03$, $\beta_1 = 0.125$, $\beta_2 = -0.025$, $\lambda = -0.3$, $\beta_0 = \beta_1/(2\beta_2)$. The random variable $\tau(0)$ has distribution $\mathcal{N}(0, -1/(2\beta_2))$, while the initial log-price is $p(0) = \log 9$. We generate second-by-second return and variance paths for a total of $n = 21600$ observations per day. The simulation is carried out using the explicit Euler discretization scheme and the time horizon is $T = 1$ day.

[4] See Fisher and Nappo (2010) for a study of the modulus of continuity of a stochastic process with possibly unbounded coefficients and Fan and Wang (2008) for a proof of the Hölder continuity in the case of many common volatility models.

The standardized returns (4.7) are evaluated on the regular time grid $t_j = 0.5(2j-1)\Delta t$, $j = 1,2,\dots,\lfloor T/\Delta t \rfloor$, where $\lfloor \cdot \rfloor$ denotes the integer part operator,[5] using both the true volatility and the volatility estimated by the Fourier method. We denote the true and estimated standardized returns with $z(t)$ and $\widehat{z}_{N,M}(t)$, respectively.

The cutting frequency N has been selected equal to $n/2$ according to the analysis in Chapter 3 and M equal to $\frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$ in order to fulfill the assumptions required for the asymptotic normality (4.6). More specifically, setting $N = n/2$ and writing the cutting frequency M in the form $M = b\frac{1}{2\pi}\sqrt{n}\log n$, we select b in order to minimize the average MSE.

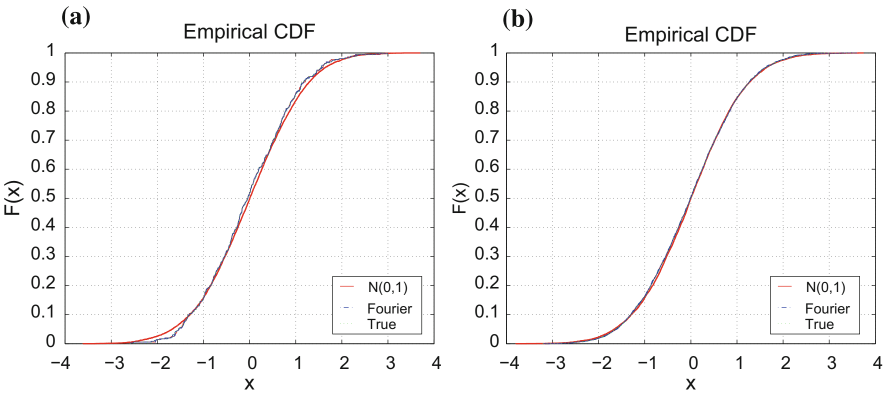

Fig. 4.1 Comparison between the cumulative density functions of a standard normal sample (red solid line), of the standardized returns using the true volatility (green dotted line) and using the Fourier spot volatility estimates (blue dash-dot line) when $\Delta t = 1$ minute (a), $\Delta t = 5$ seconds (b).

Figure 4.1 shows a comparison between the cumulative density functions of the two samples $\{z(t_j)\}_{j=1,\dots,T/\Delta t}$, $\{\widehat{z}_{N,M}(t_j)\}_{j=1,\dots,T/\Delta t}$ and the theoretical $\mathcal{N}(0,1)$, when $\Delta t = 1$ minute (Figure 4.1(a)) and $\Delta t = 5$ seconds (Figure 4.1(b)). Figure 4.1 shows that the theoretical cumulative density function is approximated with sufficient accuracy when $\Delta t = 1$ minute. Moreover, the quality of the approximation substantially improves when $\Delta t = 5$ seconds, where the cumulative density function of $\widehat{z}_{N,M}$ perfectly fits the probability distribution $\mathcal{N}(0,1)$.

It is worth noting that the volatility estimates used in $\widehat{z}_{N,M}$ have been computed using only one log-price path sampled at 1-second (i.e., $n = 21600$) and choosing the same cutting frequencies N and M for any t. This fact highlights the *global* character of the Fourier estimator, namely, the fact that it is designed to estimate the volatility path as a process over the entire time interval. Confirming this point, Figure 4.2 shows a realization of the true spot variance (dotted line) and the corresponding Fourier estimates (solid line) obtained with the same choice of N and

[5] For the sake of simplicity, hereafter we omit the integer part symbol $\lfloor \cdot \rfloor$.

M as in Figure 4.1 and recording the volatility path at the scale $\Delta t = 1$ minute. A comparative study of the performance of the Fourier spot volatility estimator with different local estimators can be found in Mancino and Recchioni (2015).

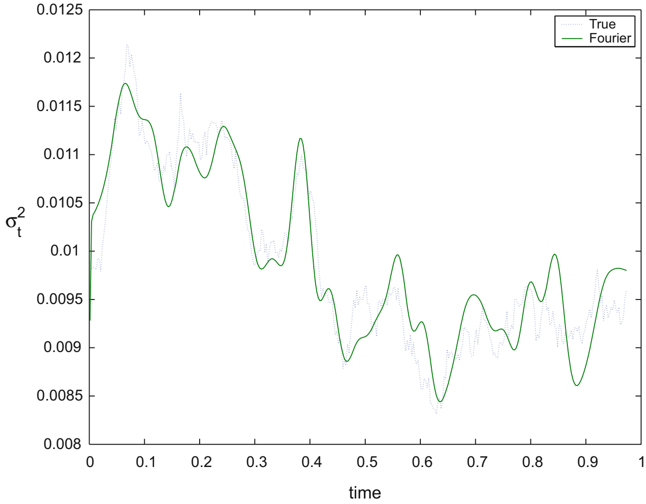

Fig. 4.2 True variance $\sigma^2(t)$ (dotted line) and Fourier estimated variance $\widehat{\sigma}^2_{n,N,M}(t)$ (solid line) as a function of time, recorded at a 1 minute scale for one realization of (4.8)–(4.10).

Figures 4.1 and 4.2 are obtained using a single realization of the log-price process. In order to investigate whether the fitting does not depend on the specific path, we proceed generating 500 replications of the standardized return paths $z(t)$ and $\widehat{z}_{N,M}(t)$, for four different values of the scale at which the volatility is reconstructed, that is $\Delta t = 5$, 30 seconds, 1, 3 minutes. We apply the Kolmogorov-Smirnov (KS) and the Jarque-Bera (JB) tests at the 5% significance level to determine whether the 500 random samples could have the hypothesized standard normal (KS-test) or a normal (JB-test) cumulative density function with unspecified mean and variance. The two tests are employed since it is known that the standardized returns may fail to be standard normal random variables when the scale Δt increases up to 3 minutes.

The upper panel in Table 4.1 shows the scale Δt, the percentage of the KS test rejections, the corresponding average p-values, the percentage of the JB test rejections and the corresponding average p-values obtained using the true standardized returns $z(t)$, while the lower panel shows the same quantities for the estimated standardized returns $\widehat{z}_{N,M}(t)$. The percentages of rejection obtained for the KS test are slightly larger than those obtained for the JB test. This result is due to the non-negligible role played by the drift when Δt is greater than 1 minute. We emphasize that a satisfactory approximation of the standardized returns is significant for an accurate volatility estimate, since the evaluation of the standardized return (4.7) at time t is based on the volatility evaluation at the same time t.

Table 4.1 Comparison of the true standardized returns $z(t)$ and the Fourier estimated standardized returns $\widehat{z}_{N,M}(t)$ using Kolmogorov-Smirnov and Jarque-Bera tests.

True standardized returns $z(t)$				
Δt	KS rejections (%)	KS-pvalue	JB rejections (%)	JB-pvalue
5 secs	7.4%	0.48	4.0%	0.49
30 secs	7.8%	0.45	4.4%	0.51
60 secs	8.8%	0.46	4.6%	0.52
180 secs	8.2%	0.45	4.2%	0.53
Estimated standardized returns $\widehat{z}_{N,M}(t)$				
Δt	KS rejections (%)	KS-pvalue	JB rejections (%)	JB-pvalue
5 secs	7.4%	0.49	3.8%	0.49
30 secs	7.8%	0.45	3.8%	0.51
60 secs	9.0%	0.45	4.0%	0.52
180 secs	7.6%	0.46	4.0%	0.52

Finally, the empirical distribution of $\sqrt{n/M}(\widehat{\sigma}^2_{n,N,M}(t)-\sigma^2(t))/\sigma^2(t)$ for the 500 replications of tick-by-tick data is computed. Figure 4.3 shows the empirical and theoretical distribution $\mathcal{N}(0,4/3)$ at time $t=0.09$ (panel (a)), $t=0.5$ (panel (b)) and $t=0.94$ (panel (c)). The empirical distributions shown in each panel are tested for normality using the Jarque-Bera test at the 5% significance level. The test shows the null hypothesis is not rejected. The p-values are shown in the panels of Figure 4.3.

4.2 Multivariate Estimator

The genuine advantage of the Fourier method becomes manifest when facing the issue of estimating the matrix $\Sigma(t)$ in (2.2) as a *stochastic function of time* for any t in $[0,2\pi]$. This is the topic of the present section, where the *Fourier estimator* of the multivariate spot volatility $\Sigma^{i,j}(t)$ defined in (2.14) is studied. Notations from Section 3.3 are in force.

4.2.1 Asymptotic Results

Confirming the fact that Fourier estimator is essentially defined to deal with the multivariate case, the asymptotic results for the Fourier estimator (2.14) follow along the same line as the univariate case. However, the (possible) asynchronicity between recorded data introduces new issues, which can be addressed with a suitable choice of the cutting frequencies N,M.

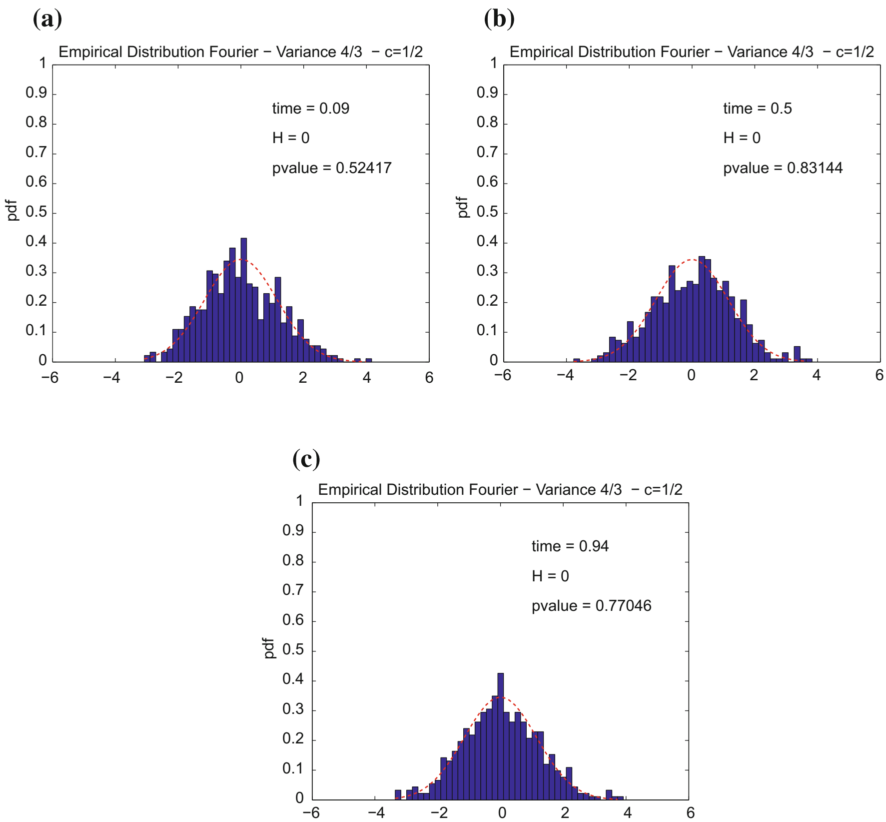

Fig. 4.3 Empirical distribution of $\sqrt{n/M}(\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t))/\sigma^2(t)$ at $t = 0.09$ (a), $t = 0.5$ (b), $t = 0.94$ (c), with $N = n/2$ and $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$.

Suppose that the processes $\sigma^i_j(t)$ in (2.1) are continuous, e.g., the volatilities are driven by a second diffusion process. Asymptotic conditions required for the irregular/asynchronous time grids are stated in Malliavin and Mancino (2009) Theorem 4.4. In order to simplify the notations, we consider two assets, labeled 1 and 2 and we focus on the covariance terms of the volatility matrix $\Sigma(t)$.

Consistency. Recall that $c_k(\Sigma^{1,2}_{n_1,n_2,N})$ defined in (2.13) converges to the Fourier coefficient of the cross-volatility function $\mathscr{F}(\Sigma^{1,2})(k)$ in virtue of (3.24). Let $\widehat{\Sigma}^{1,2}_{n_1,n_2,N,M}(t)$ be defined by (2.14) and suppose that $N\rho(n) \to 0$ and $M/N \to 0$, as $N, M, n \to \infty$, then the following uniform convergence holds in probability

$$\lim_{n,N,M\to\infty} \sup_{t\in(0,2\pi)} |\widehat{\Sigma}^{1,2}_{n_1,n_2,N,M}(t) - \Sigma^{1,2}(t)| = 0. \tag{4.11}$$

The assumption $N\rho(n) \to 0$ is required for the consistency of the estimator with asynchronous data. On the other hand, in the case when observed data of the two assets are synchronous, the best choice is $N = n/2$, as in the univariate case. The situation is different if data are contaminated by microstructure noise: when dealing with noisy high-frequency data, the choice of N being an infinite of lower growth rate than the number of data n is recommended (see also the discussion in Section 3.3.2). As for the localizing parameter M, its growth rate is to be related with the Holdër-continuity parameter of the volatility paths, as in the univariate case. The rate of convergence of the global estimator in (4.11) is still an open problem.

Central Limit Result. Regarding the central limit result for the multivariate spot volatility Fourier estimator, some partial results are known. The pointwise asymptotic normality in the multivariate synchronous setting holds given the assumptions already considered in Section 4.1.1. The weak convergence is obtained in Malliavin and Mancino (2009) for asynchronous data when some rather technical conditions are satisfied. In Mancini et al. (2015) the authors consider the quadratic part of the Fourier estimator, thus reducing it to a kernel-based estimator in the spirit of Kristensen (2010), and prove the pointwise asymptotic normality for this estimator with the optimal rate. However, as we will see in Section 5.4, such a modification of the Fourier estimator while allowing one to get this asymptotic result, destroys its efficiency in practical relevant cases, such as in the presence of high-frequency data.

4.2.2 Bandwidth and Scale Selection

This section studies the finite sample efficiency of the Fourier estimator of the spot volatility matrix. Being a rather unexplored field, we obtain here some preliminary original results, with the primary goal to give some directions for an efficient implementation of the estimator.

We simulate the following continuous time bivariate GARCH model

$$\begin{bmatrix} dp^1(t) \\ dp^2(t) \end{bmatrix} = \begin{bmatrix} \beta_1 \sigma_1^2(t) \\ \beta_2 \sigma_4^2(t) \end{bmatrix} dt + \begin{bmatrix} \sigma_1(t) & \sigma_2(t) \\ \sigma_3(t) & \sigma_4(t) \end{bmatrix} \begin{bmatrix} dW_5(t) \\ dW_6(t) \end{bmatrix} \tag{4.12}$$

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

where $\{W_i(t)\}$, $i = 1, 2, \ldots, 6$ are independent Brownian motions. The parameters of the model are: $\alpha_1 = 0.1$ $\alpha_2 = 0.1$, $\alpha_3 = 0.2$, $\alpha_4 = 0.2$, $\beta_1 = 0.02$, $\beta_2 = 0.01$, $\omega_1 = 0.1$, $\omega_2 = 0.1$, $\omega_3 = 0.2$, $\omega_4 = 0.2$, $\theta_1 = 0.1$, $\theta_2 = 0.1$, $\theta_3 = 0.1$, $\theta_4 = 0.1$, $\alpha = 0.1$; further, $\sigma_1(0) = 0.5$, $\sigma_2(0) = 0.1$, $\sigma_3(0) = 0.9$, $\sigma_4(0) = 0.25$ and $p^1(0) = \log 9$, $p^2(0) = \log 11$.

Firstly, we show empirical evidence that the performance of the Fourier estimator in the multivariate *synchronous* case is the same as in the univariate one, when the cutting frequency N and M are chosen as suggested in Section 4.1.2, that is $N = \frac{n}{2}$ and $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$.

To this end, we simulate synchronous evenly spaced 1-second returns, $p^1(t_l)$, $p^2(t_l)$, $t_l = lT/n$, $l = 0,1,\dots,n$, $T = 1$ day (6 hours), $n = 21600$, using the explicit Euler discretization scheme. We compute the true volatility matrix entries $\Sigma^{i,j}(t)$ and their Fourier estimates, $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$ as a function of time, running 500 daily replications of the 1-second log-returns.

In this synchronous setting, we investigate the empirical distributions of the true volatility matrix and its Fourier estimate by comparing the distribution of the 500 realizations of $\Sigma^{i,j}(t)$ and of $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$ at times $t = 0.05 + 0.1k$, with $k = 0,1,\dots,9$ and $t = 0.99$. The analysis is carried out using the Kolmogorov-Smirnov (KS) test to determine whether the two random samples are drawn from the same underlying continuous population. The significance level is set equal to 5%. Table 4.2 shows the time t considered, the result of the hypothesis test, H, (i.e., $H = 0$ the null hypothesis is not rejected at 5% level, $H = 1$ the null hypothesis is rejected at 5% level) and the corresponding p-value for the cross-volatility $\Sigma^{1,2}(t)$ and the volatilities $\Sigma^{1,1}(t)$, $\Sigma^{2,2}(t)$. Table 4.2 shows that the Fourier estimator's performance is excellent inside

Table 4.2 Synchronous data. Comparison between the true spot volatilities $\Sigma^{i,j}(t)$ under model (4.12) and the estimated ones $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$, using the two-sample K-S goodness-of-fit hypothesis test.

$\Sigma^{1,1}(t)$			$\Sigma^{1,2}(t)$			$\Sigma^{2,2}(t)$		
t	H	KS-pvalue	t	H	KS-pvalue	t	H	KS-pvalue
0.05	1	0.0000121	0.05	1	0.0006	0.05	1	0.0471
0.15	1	0.1931	0.15	0	0.3438	0.15	0	0.4431
0.25	0	0.0691	0.25	0	0.5361	0.25	0	0.8937
0.35	0	0.5560	0.35	0	0.9611	0.35	0	0.6766
0.45	0	0.5660	0.45	0	0.5560	0.45	0	0.5560
0.55	0	0.4431	0.55	0	0.7942	0.55	0	0.9921
0.65	0	0.8937	0.65	0	0.9921	0.65	0	0.9921
0.75	0	0.9921	0.75	0	0.9921	0.75	0	0.9999
0.85	0	0.04431	0.85	0	0.4431	0.85	0	0.8937
0.95	0	0.0993	0.95	0	0.6766	0.95	0	0.9610
0.99	1	0.0009	0.99	1	0.0004	0.99	1	0.0082

the time interval. Furthermore, the p-values for the estimated volatility $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$ $(i,j = 1,2)$ are similar, indicating that the cross-volatility estimation does not require a specific treatment in the synchronous case.

Consider now the empirically more relevant case of *asynchronous* data. The analyzed example consists of two asynchronous samples with the same number of data (i.e., $n_1 = n_2 = n = 21600$). The process p^1 is observed at t^1_k, $k = 1, 2, \ldots, n$, where t^1_k are evenly spaced, while p^2 is observed at t^2_k, $k = 1, \ldots, n$, where $t^1_0 = t^2_0$ and t^2_k is drawn out from a uniformly distributed random sample in the interval $[t^1_{k-1}, t^1_k]$, $k = 1, 2, \ldots, n$.

We compute the average relative errors of the Fourier estimates $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$ on the grid $\tau_\nu = (\nu + 1/2)/360$, $\nu = 0, 1, \ldots, 359$, that is at a 1 minute scale, as follows:

$$e^{i,j}(\tau_\nu) = \frac{1}{500} \sum_{l=1}^{500} \frac{|\Sigma^{i,j}_l(\tau_\nu) - \widehat{\Sigma}^{i,j}_{n,N,M,l}(\tau_\nu)|}{|\Sigma^{i,j}_l(\tau_\nu)|}, \; \nu = 0, 1, \ldots, 359, \tag{4.13}$$

where the subscript l refers to the l-th realization out of 500 in the Monte Carlo experiment. Figure 4.4(a)–(b) shows the average relative errors $e^{1,1}(t)$ (solid line), $e^{1,2}(t)$ (dotted-line), and $e^{2,2}(t)$ (dashed line) as a function of time $t \in (0, 1)$. In panel (a) the Fourier estimates are computed by choosing N and M as in the synchronous case; we note that the relative error sensibly increases for the estimates of the cross-volatilities. Panel (b) shows the same relative errors of $\Sigma^{1,1}(t)$ and $\Sigma^{2,2}(t)$ as those in panel (a), while the relative errors of the instantaneous cross-volatility $\Sigma^{1,2}(t)$ are obtained using the estimates $\widehat{\Sigma}^{1,2}_{n,N,M}(t)$ corresponding to the cutting frequencies $N = 0.85\,n^{3/4}$, and $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n^{3/4}}\log n^{3/4}$. The choice of the cutting frequencies N, M follows by the numerical studies in Sections 3.3.2 and 4.1.2, respectively.

This result confirms the two main findings illustrated in the case of the integrated volatility matrix and resumed in the following box.

First, the choice $N = n/2$ is not the proper one when asynchronous data are processed. Second, the Fourier methodology provides efficient estimates of the cross-volatility by suitably cutting the frequencies N and M.

Remark 4.2. The accuracy of the Fourier estimates could be further improved by choosing M depending on the time t, as it is usually done for the local volatility estimators (see Mancino and Recchioni (2015) for further details).

The last simulation exercise analyses the point-wise accuracy of the volatility matrix estimation over the entire time interval, with the aim of showing that

the Fourier method provides *global* estimation of the spot volatility, that is the volatilities are estimated with similar accuracy at any t in the interior of the domain by choosing the cutting frequencies N and M constant in time.

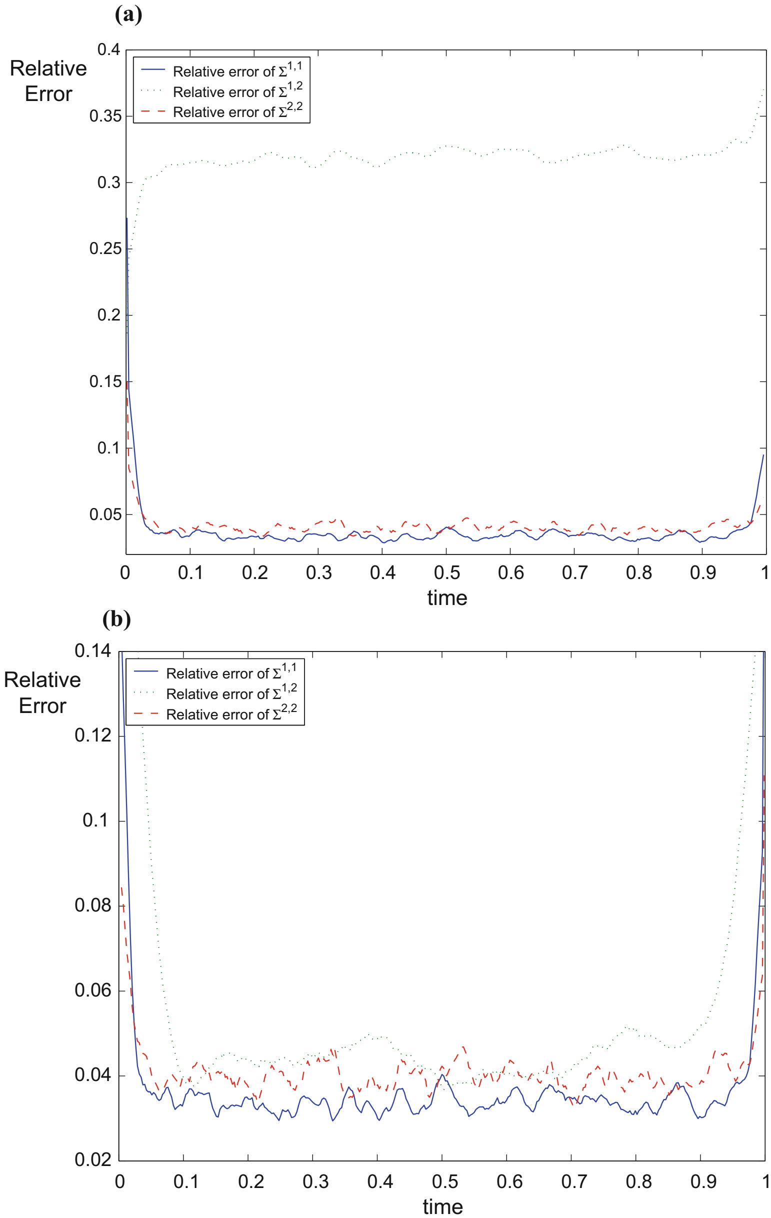

Fig. 4.4 Asynchronous data. Average relative errors of the Fourier estimates $\widehat{\Sigma}^{1,1}_{n,N,M}(t)$, $\widehat{\Sigma}^{1,2}_{n,N,M}(t)$ and $\widehat{\Sigma}^{2,2}_{n,N,M}(t)$ as a function of time. Panel (a): $n_1 = n_2 = n$, $N = \frac{n}{2}$, $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$. Panel (b): $n_1 = n_2 = n$ and $N = 0.85\ n^{3/4}$, $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n^{3/4}}\log n^{3/4}$ for the instantaneous covariance. Note that the y-scale in panel (a) is larger than in panel (b) due to the large values of the relative errors.

Figure 4.5 displays the true values of the volatility matrix $\Sigma^{i,j}(t)$ (dotted line) and the estimated ones $\widehat{\Sigma}^{i,j}_{n,N,M}(t)$ at a scale of 1 minute in the synchronous case.[6] The estimated values are obtained using one realization of the 1-second returns (i.e., $n = 21600$), while choosing $N = n/2$ and $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$. Notice that the reconstruction is quite satisfactory in all the cases.

Figure 4.6 (upper panel) displays the true values of the cross volatility $\Sigma^{1,2}(t)$ (dotted line) and the estimated one $\widehat{\Sigma}^{1,2}_{n,N,M}(t)$, in the non-synchronous case at a scale of 1 minute. The estimated values of the instantaneous covariance are obtained using one realization of the 1-second returns (i.e., $n = 21600$) with $N = 0.85\, n^{3/4}$ and

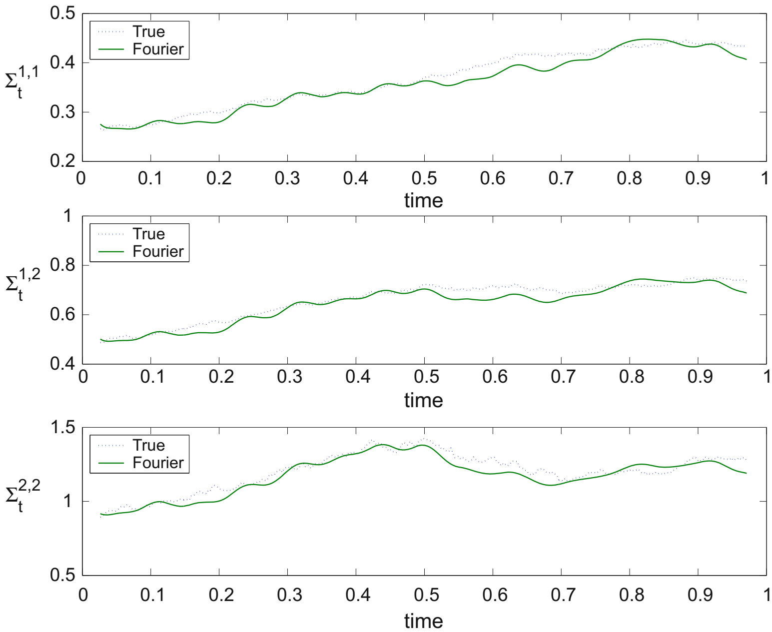

Fig. 4.5 Synchronous data. The upper panel shows $\Sigma^{1,1}(t)$ (dotted line) and the Fourier estimates (solid line), the middle panel shows $\Sigma^{1,2}(t)$ (dotted line) and the Fourier estimates (solid line), and the lower panel shows $\Sigma^{2,2}(t)$ (dotted line) and the Fourier estimates (solid line) as a function of time. The estimated spot volatility matrix is reconstructed at a time scale of 1 minute from one realization of the 1-second log-returns ($n = 21600$, $N = \frac{n}{2}$, $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n}\log n$).

$M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n^{3/4}}\log n^{3/4}$. As already noticed above (cfr. Figure 4.4(b)), this choice makes the estimates of the instantaneous cross volatility of quality comparable with those of the instantaneous volatilities. The estimated cross volatility $\widehat{\Sigma}^{1,2}_{n,N,M}(t)$ in the asynchronous case turns out to be a very smooth function. This is a consequence of the fact that the cutting frequency N and M have been reduced in order to eliminate

[6] Here the volatility is estimated at the same scale used to evaluate the average relative errors $e^{i,j}$ $(i,j = 1,2)$ defined in (4.13).

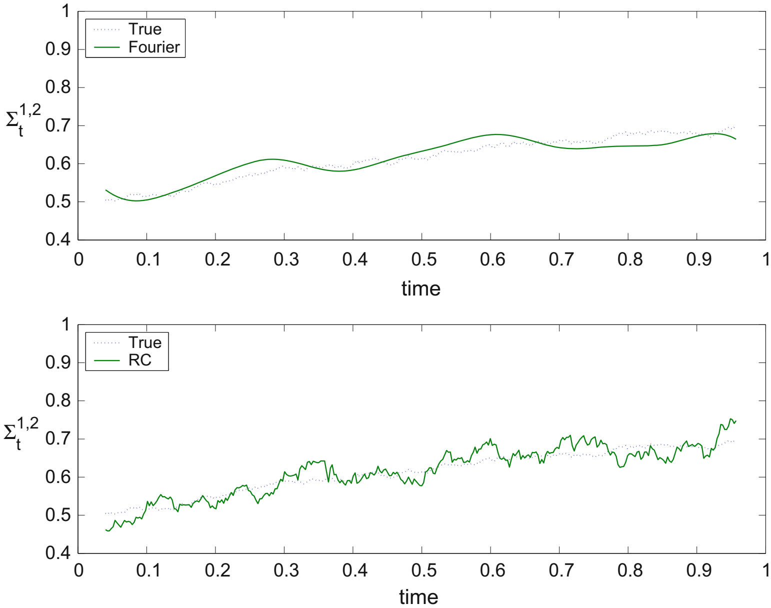

Fig. 4.6 Asynchronous data. True (dotted line) and estimated (solid line) cross volatility $\Sigma^{1,2}(t)$ as a function of time obtained by the Fourier estimator (upper panel) and by the Realized Covariance (lower panel). The estimated cross volatility is reconstructed at a scale of 1 minute from one realization of the 1-second log-returns ($n = 21600$, $N = 0.85\,n^{3/4}$, $M = \frac{1}{2\pi}\frac{1}{8}\sqrt{n^{3/4}}\,\log n^{3/4}$).

the bias caused by the asynchronous data. Nevertheless, despite the smoothness, the estimated $\widehat{\Sigma}^{1,2}_{n,N,M}$ is able to capture the major oscillations of the true simulated cross volatility $\Sigma^{1,2}$. As a benchmark, in the lower panel of Figure 4.6 we show the performance of the spot Realized Covariance estimator, obtained by localizing the estimator (3.27). The localizing bandwidth must be carefully chosen in order to reduce oscillations and minimize the discretization errors due to last tick interpolation. Due to the particular asynchronicity structure at hands, here the bandwidth is 7 minute large. It is evident that the Fourier estimates are much smoother and hence reliable than the ones obtained by the Realized Covariance.

Remark 4.3. A preliminary attempt to investigate the impact of the cutting frequencies on the time scale used to reconstruct the volatility path is due to Mattiussi and Iori (2010). By modifying the Fejér kernel appearing in the Fourier expansion of the volatility function with the dependence on a further positive parameter δ, the authors suggest to choose the Nyquist frequencies $N = n/2$, $M = N/2$ and then adapting the trajectory to the desired time scale by an appropriate selection of the parameter δ. Furthermore, they investigate whether an optimal time scale exists at which the instantaneous volatility matrix should be reconstructed. This is done by letting $N = n/2$ and $M = N/2$ and then choosing δ by minimizing the mean squared error. In the univariate case, a simulation study is also carried in Mancino and Recchioni

(2015). However, a theoretical result for the optimal choice of the scale of volatility estimates is still under investigation. The relevance of this issue becomes manifest when, for instance, we aim at estimating a second order quantity as the volatility of volatility (see Kanaya and Kristensen (2015) and discussions in Section 6).

4.3 Fourier Method in the Presence of Jumps

So far we have intentionally left out of the discussion the more sophisticated jump-diffusion models. Actually, the price model (2.1) can be generalized to allow for a jump component in addition to the continuous Brownian factor, usually described through the sum of non-zero random variables, whose sum is controlled by a Poisson process. In this case, the quadratic variation (2.15) includes both the continuous path variation (i.e., the integrated variance) studied so far and the contribution of the jump variation.[7] As a consequence, the classical Realized Volatility is no longer consistent.

The existing volatility and covariance estimators mainly focus on the estimation of integrated quantities. The most common approaches employ the Bi-power and Multi-power Variation estimators as proposed by Barndorff-Nielsen and Shephard (2004), later generalized by replacing the power function with different specifications (see, e.g., Jacod (2008), Todorov and Tauchen (2012)) and the threshold method by Mancini (2009).

The Fourier methodology has been extended by Cuchiero and Teichmann (2015) to estimate the path of the instantaneous volatility and covariance process in the presence of jumps. We only give a brief outline of the method, addressing the interested reader to the cited paper.

The procedure has two steps. First, it obtains an estimate of the Fourier coefficients of a continuous invertible function $\rho(\sigma^2)$ of the instantaneous volatility (or of the covariance) by using a jump robust estimator (like the ones cited above). Let $[0,2\pi]$ be the time horizon and consider the uniform time grid $\{0 = t_0 < \ldots < t_n = 2\pi\}$ with step size $\rho(n) = 2\pi/n$. The estimator of the k-th Fourier coefficient takes the form

$$\sum_{j=1}^{n} \frac{1}{n} e^{-\mathrm{i}kt_{j-1}} g(\sqrt{n}\delta_j(p)), \tag{4.14}$$

where the function g can assume different specifications. Second, it uses the Fourier-Fejér inversion formula as in (4.1) to reconstruct the path of the process $\rho(\sigma^2)$. This can thus be translated into an estimator of the volatility by inverting the function $\rho(\cdot)$.

The estimator of instantaneous volatility obtained so far is consistent and the Central Limit Theorem holds with rate of convergence equal to $n^{(1-\nu)/2}$, where ν

[7] A precise statement of the Itô formula for general semimartingale is beyond the scope of this book; we refer the interested reader to the book by Protter (1992).

is the Hölder continuity (between two jumps) of the volatility path. So, the rate of convergence lies in $(0,1/4)$ approaching the optimal rate $1/4$ as ν is close to $1/2$.

We illustrate the method by considering the case $\rho_g(\sigma^2(t)) = e^{-\sigma^2(t)/2}$, that is we choose $g(x) = \cos x$. Thus, by (4.14) the estimator of $e^{-\sigma^2(t)/2}$ is

$$\frac{1}{2\pi}\sum_{j=1}^{n}\frac{1}{n}F_M(t-t_{j-1})\, g(\sqrt{n}\delta_j(p)), \tag{4.15}$$

where $F_M(x)$ is the Fejér kernel (3.10). The performance of the estimator is studied by simulating a one-dimensional Bates-type model

$$dp(t) = -\left(\frac{\sigma^2(t)}{2} - (e^{v_{J,p}^2/2}-1)i_p\right)dt + \sigma(t)dW_1(t) + J_p dN_t \tag{4.16}$$

$$d\log\sigma(t) = -\left(\frac{\alpha^2}{2} + (e^{v_{J,\sigma}^2/2}-1)i_p\right)dt + \alpha\, dW_2(t) + J_\sigma\, dN_t, \tag{4.17}$$

where W_1 and W_2 are correlated Brownian motions with correlation λ, N_t is a Poisson process with intensity i_p responsible for jumps occurring simultaneously in price and volatility, while J_p and J_σ are normally distributed with zero mean and variances $v_{J,p}^2$, $v_{J,\sigma}^2$, respectively. The model parameters are chosen as follows: $\alpha = 0.5$, $\lambda = -0.5$, the jump rate per day is $i_p = 20/250$, the volatility of normal jumps are $v_{J,p} = 0.2$, $v_{J,\sigma} = 0.01$. We choose the number of the grid points, n, equal to 21600 that corresponds to one day ($T = 1$) of 1-second data while the cutting frequency is $M = 128$ (approximately a time scale of 1 minute and half). As suggested in Cuchiero and Teichmann (2015), the true (simulated) and reconstructed trajectories are evaluated at $2M+1$ points. Figure 4.7 illustrates that even in the case of jumps in the log-price and in the variance the trajectory of the log-price is satisfactorily reconstructed.

Remark 4.4. The jump-robust Fourier method shares the main feature of the Fourier estimation method, namely, the fact that it allows one to reconstruct the volatility as a *stochastic function of time* in the univariate and multivariate case. This property makes it possible to iterate the volatility estimation procedure and to compute second order quantities like the multivariate volatility of the volatility (see Chapter 6), when the involved processes have jumps.

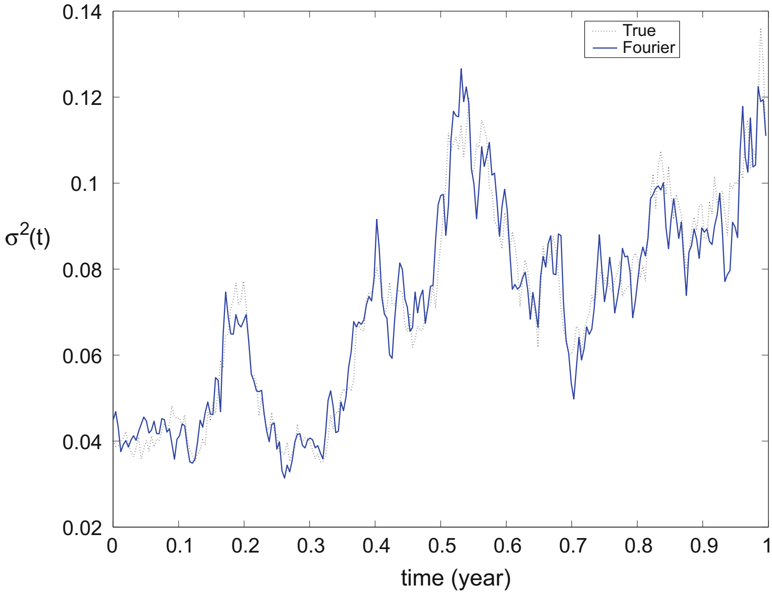

Fig. 4.7 Jump-diffusion model (4.16)–(4.17). Comparison between true (dotted line) and Fourier estimated volatility (solid line) as a function of time.

Chapter 5
High Frequency Analysis: Market Microstructure Noise Issues

The model-free measurement of volatility has recently received strong impulse by the availability of high-frequency financial data. Nevertheless, the efficiency of all the methodologies proposed for accurately estimating the volatility builds on the observability of the true price process, while observed asset prices diverge from their efficient values, being contaminated by market microstructure effects.

This section explores the *robustness* of the Fourier estimator of volatility to microstructure noise, more precisely, at which extent applying the Fourier method without any sophisticated *ad hoc* procedures provides a reliable value of the asset volatility. A feasible procedure for designing an optimal mean squared error-based Fourier estimator is presented which relies on the computation of analytic expressions for the bias and the mean squared error of the noise-affected estimator. These formulae provide the tool to optimize the finite sample performance of the estimator: the number of frequencies to be included in the Fourier series are selected with the aim of minimizing the mean squared error, for a given number of intra-daily observations. Further, Monte Carlo experiments and empirical exercises confirm that the Fourier estimator is at the same time statistically efficient and robust to some types of market frictions present in high-frequency data.

5.1 What Is the Noise Effect on Fourier Estimator?

The theory presented in Chapters 3 and 4 contrasts with the nonobservability of the true price process. In fact, when the price is sampled over small intervals such as few seconds, the observed price deviates from the efficient/latent price due to the imperfection of the trading process, as illustrated in Roll (1984), Glosten and Milgrom (1985), Harris (1991), O'Hara (1995). The econometricians do 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.

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_5

When the asset prices are observed without errors, both Realized Volatility and Fourier method provide us with a consistent estimate of integrated volatility and the two estimators have comparable statistical properties. On the contrary, the Realized Volatility calculated from high-frequency returns turns out to be an estimate of the variation of the noise, rather than of the latent true price process, as it has been empirically observed by Andersen et al. (1999a) and theoretically analyzed in Zhang et al. (2005), Bandi and Russell (2008). 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 (see, e.g., Zhou (1996), Andersen et al. (2001a), Zhang et al. (2005), Hansen and Lunde (2006b), Barndorff-Nielsen et al. (2008)). A study of these methods in comparison with the Fourier approach is conducted in Section 5.2.3.

It will be clear after reading this chapter that the Fourier estimator needs no correction in order to be statistically efficient and robust to various kinds 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; on the other side the high-frequency noise or short-run noise is ignored by cutting the highest frequencies in the construction of the Fourier 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.

We describe a feasible procedure to optimize the finite sample performance of the Fourier estimator of integrated volatility and covariance by minimizing the mean squared error (MSE hereafter) as a function of the number of frequencies, N, for a given number, n, of intra-daily observations. This procedure for the choice of a convenient cut-off frequency allows us to filter out a great portion of variation in the integrated volatility estimates which is attributed to the noise and can be applied as a rule-of-thumb in empirical cases. The method to find the optimal cutting frequency suggests that the optimal MSE-based estimator should be designed using quote-to-quote returns.

5.2 The Case of Integrated Volatility

In this section we study the performance of the Fourier estimator of integrated volatility when the asset price is contaminated by microstructure noise effects. The analysis starts by considering a simple but well-consolidated additive model where the microstructure noise displays an MA(1) structure with a negative first order autocorrelation. The MA(1) model is typically justified by bid-ask bounce effects (see Roll (1984)). It is known to be a realistic approximation in decentralized markets where trades arrive in a random fashion with idiosyncratic price setting behavior, the foreign exchange market being a valid example (see Zhang et al. (2005), Bandi and Russell (2006), Hansen and Lunde (2006b) for additional discussions on this point). However, as observed by Aït-Sahalia and Jacod (2014), some data on log-returns may be inconsistent with a simple MA(1) structure. Therefore, we study also a more general form of additive models, where the noise is correlated with the efficient returns, and the rounding error model, where the measurement error is mainly due to the fact that transaction prices are multiples of a tick size. We provide analytical formulae of the bias and of the MSE of the Fourier estimator under MA(1) 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.

5.2.1 Starting from the Additive MA(1) Model

Consider a given time interval (e.g., a trading day), scaled to be $[0,2\pi]$, as usual. Suppose that the process is observed at a discrete unevenly spaced grid $\{0 = t_0 \leq t_1 \leq \cdots \leq t_n = 2\pi\}$ for any $n \geq 1$, and that the logarithm of the observed price process can be split into the sum of two terms

$$\widetilde{p}(t_j) = p(t_j) + \eta(t_j), \quad j = 0,\ldots,n, \tag{5.1}$$

where p is the efficient log-price process and η is the microstructure noise component. 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 following hypotheses hold:

(A) The (latent) price process $p(t)$ satisfies the stochastic differential equation (3.1).

(M.I) The random shocks $\eta(t_j)$, for $0 \leq j \leq n$ and for all n, are independent and identically distributed with mean zero and bounded fourth moment.

(M.II) The true return process $\delta_j(p) := p(t_{j+1}) - p(t_j)$ is independent of $\eta(t_j)$ for any j,n.

To simplify the notation, in the sequel we will write η_j instead of $\eta(t_j)$.

Remark 5.1. The hypothesis that the η_j's are independent of the increments $\delta_j(p)$ is discussed in Hansen and Lunde (2006b). Their empirical work suggests that the independence assumption is not too damaging statistically, when we analyze data in tickly traded stocks recorded every minute.

> Aim: estimation of the integrated volatility by means of the Fourier estimator defined in (3.5), given the observations of the *contaminated process* $\widetilde{p}$ defined in (5.1).

In the sequel, we assume that $2\pi/n$ is the time distance between adjacent logarithmic prices and denote the integrated volatility by V.

Firstly, consider the Realized Volatility RV_n defined in (3.8): it 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. In fact, it is easy to prove (see Zhang et al. (2005), Bandi and Russell (2008)) that the bias of the Realized Volatility estimator diverges as the number n of observations increases and is given by

$$E[RV_n - V] = 2nE[\eta^2]. \tag{5.2}$$

Consider now the Fourier estimator defined in (3.5). The bias is computed as follows:

$$E[\widehat{\sigma}^2_{n,N} - V] = 2n\, E[\eta^2] \left(1 - D_N\left(\frac{2\pi}{n}\right)\right), \tag{5.3}$$

where D_N is the rescaled Dirichlet kernel defined in (2.9). The proof is given in Mancino and Sanfelici (2008) and can be easily obtained using the representation (3.7). Note that, under the condition $N^2/n \to 0$, the right-hand side of (5.3) tends to zero. Therefore, if we choose N "small" with respect to the number of observations 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.

> In a finite sample, given the number of data n, a suitable choice of the Fourier frequency N allows for lower bias with respect to the Realized Volatility estimator. Further, the Fourier estimator is asymptotically unbiased under the condition N^2/n goes to 0.

We compare now the MSE of the Realized Volatility with that of the Fourier estimator. The MSE of the Realized Volatility is the following (for the proof, see Hansen and Lunde (2006b), Bandi and Russell (2008))

$$MSE(RV_n) = 2\frac{2\pi}{n}(Q + o(1)) + n^2\alpha + n\beta + \gamma, \tag{5.4}$$

where Q is the integrated quarticity $\int_0^{2\pi} \sigma^4(s)ds$, $o(1)$ is a term which goes to zero as n goes to infinity and

$$\alpha := 4E[\eta^2]^2, \;\; \beta := 4E[\eta^4], \;\; \gamma := 8E[\eta^2]V + 2E[\eta^2]^2 - 2E[\eta^4]. \tag{5.5}$$

Therefore, while the addend $2\frac{2\pi}{n}(Q + o(1))$ is asymptotically vanishing, the polynomial $n^2\alpha + n\beta + \gamma$ diverges when the number n of observations increases.

Moving to the Fourier volatility estimator, for any given n, N, the MSE takes the form

$$MSE(\widehat{\sigma}^2_{n,N}) = 2\frac{2\pi}{n}(Q+o(1)) + n^2\widehat{\alpha}(n,N) + n\widehat{\beta}(n,N) + \widehat{\gamma}(n,N), \tag{5.6}$$

where

$$\widehat{\alpha}(n,N) := \alpha\left(1 - D_N\left(\frac{2\pi}{n}\right)\right)^2,$$

$$\widehat{\beta}(n,N) := \beta\,\left(1 - D_N\left(\frac{2\pi}{n}\right)\right)^2,$$

$$\widehat{\gamma}(n,N) := \gamma + 4Q\frac{2\pi}{2N+1} + 4(E[\eta^2]^2 + E[\eta^4])\left(2D_N\left(\frac{2\pi}{n}\right) - D_N^2\left(\frac{2\pi}{n}\right)\right), \tag{5.7}$$

with α, β, γ as in (5.5) and where D_N is the rescaled Dirichlet kernel defined in (2.9). The proof of the identity (5.6) is given by Mancino and Sanfelici (2008). Letting $N^2/n \to 0$, then it holds

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

and

$$\lim_{n,N\to\infty} \gamma(n,N) = 8E[\eta^2]V + 2E[\eta^4] + 6E[\eta^2]^2. \tag{5.8}$$

It follows that, conveniently tuning the parameter N, the MSE of the Fourier estimator does not diverge and it is given by (5.8), which is small in magnitude. The following conclusion can be drawn.

When microstructure effects are introduced in the model it is no longer true that the MSE decreases as the sampling frequency increases, as it happens in the absence of microstructure noise. Nevertheless, while in the presence of microstructure effects the MSE of the Realized Volatility diverges as n increases, due to the presence of the terms of order n^2 and n, the MSE of the Fourier estimator does not diverge if condition $N^2/n \to 0$ is met.

As an important by-product of the MSE computation, a feasible procedure for selecting the cutting frequency N_{cut} is obtained. In fact, given the analytical expression (5.6) of the MSE for the noise-affected volatility estimator as a function of the sampling frequency and of the number of Fourier coefficients, then an optimal cutting frequency N_{cut} can be chosen depending on the number of observations, the noise moments, and the quarticity.

Formulae (5.3) and (5.6) allow us to measure the bias and MSE of the volatility estimates also in the case of empirical market data, where the efficient price and the volatility are not available. The practical calculation of (5.3) and (5.6) hinges on the estimation of the relevant noise moments as well as on the preliminary identification of the integrated volatility V and quarticity Q. Since the noise moments do not vary across frequencies under the MA(1) model, in computing the MSE estimates we use sample moments constructed using quote-to-quote return data in order to estimate the relevant population moments of the noise components. Details on how it is possible to estimate these quantities and the corresponding Matlab® codes are given in Appendix B.2.[1] Preliminary estimates of V and Q are obtained by computing $\widehat{\sigma}^2_{n,N}$ or RV_n and the estimator defined by (3.20) or (3.17) for the integrated quarticity using sparse sampled data.

Feasible procedure: use quote-to-quote returns and minimize the MSE formula (5.6) as a function of the cutting frequency N_{cut}. For any given sample size n, the optimal cutting frequency N_{cut} is obtained by direct minimization of the estimated MSE by comparing the computed MSE values over distinct integer-valued N, having the Nyquist frequency as upper bound. (5.6) can be minimized on average using intra-day returns over many days or, alternatively, for every daily ex-post variation measure so that N_{cut} could vary from day to day.

Recently, a different mathematical framework for the derivation of the optimal cutting frequency has been presented by Wang (2014). Given the sample size n, an explicit asymptotic expression for the optimal cutting frequency is provided. Under the condition $N^2/n \to 0$ as $n,N \to \infty$, the MSE-optimal cutting frequency is given by the formula

$$N_{opt} = -b + (-b^3 - d + \sqrt{d(d+2b^3)})^{1/3} + (-b^3 - d - \sqrt{d(d+2b^3)})^{1/3}, \quad (5.9)$$

where the constants are: $b = (5+2\rho)/12$, $d = -(3n\rho(\rho+2) + 2\pi^2(1+\rho))/(16\pi^2)$ and $\rho = E[V]/(2E[\eta^2])$. We have performed the following simulation exercise assuming the GARCH diffusion model (3.15). Intra-day noisy prices are affected by a Gaussian noise η with mean 0 and variance $E[\eta^2] = 0.000142$. We simulate 500 daily replications for 24 hours of trading with a total of $n = 86400$ second-by-second returns per day. In Figure 5.1, the true and estimated bias and MSE of the Fourier estimator, as given by (5.3) and (5.6), are plotted as a function of the number of the Fourier coefficients. The minimum of the true MSE is 6.29e-4 and is attained for $N_{cut} = 793$ which, at least theoretically, corresponds to a sampling frequency of $24 \cdot 60/(2 \cdot 793) = 0.91$ minutes. The MSE of the estimator obtained by feasible minimization of the estimated MSE (5.6) is 6.41e-4 and the corresponding optimal cutting frequency $N_{cut} = 823$. On the other side, the optimal cutting frequency as

[1] Other possible estimators of these quantities are discussed in Barndorff-Nielsen et al. (2008), although the statistical gains are minor.

specified by formula (5.9) yields $N_{opt} = 2383$, with a corresponding MSE value of 1.50e-2, that is suboptimal.

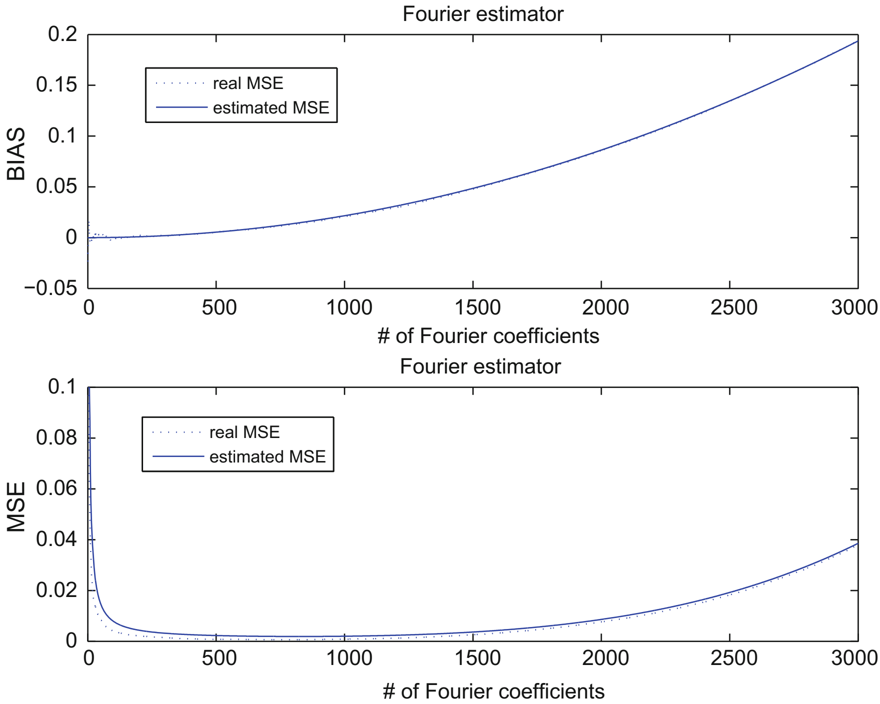

Fig. 5.1 True (dotted line) and estimated (solid line) bias and MSE of the Fourier estimator as a function of the number of the Fourier coefficients. Quote-to-quote returns. Parameter values: $\theta = 0.035$, $\omega = 0.6365$, $\lambda = 0.2962$, $p(0) = \log 100$, $\sigma^2(0) = 0.6365$.

5.2.2 Moving to Alternative Microstructure Noise Models

In the previous section a practical way to efficiently implement Fourier estimation method with high-frequency data has been proposed. Here, Monte Carlo evidence is given of the fact that such (intentionally) elementary rule of selecting the cutting frequency still works under more general noise provision.

To this end, two alternative models for market microstructure noise are considered. First, we relax the assumption **(M.II)** by considering the noise correlated with the efficient returns, following an example in Hansen and Lunde (2006b). More precisely, we assume that **(M.I)** and $(\mathbf{M.II})'$ hold, where

$(\mathbf{M.II})'$ the random shocks are defined as $\widetilde{\eta}_j := \zeta \delta_j(p) + \eta_j$, for any j, being ζ a real constant[2] and η_j as in **(M.I)**.

The second situation we study is the case of measurement errors due to the fact that transaction prices are multiples of a tick size. More precisely, we assume:

$(\mathbf{MR})$ $\widetilde{p}(t) := [\frac{p(t)}{l_\alpha}]\, l_\alpha$, being l_α the tick size and $[x]$ the integer closest to x.

Then, the noise η is defined by $\eta(t_i) = \widetilde{p}(t_i) - p(t_i)$ and can be modeled as a rounding off problem (see Aït-Sahalia and Jacod (2014), Li and Mykland (2014)).

Remark 5.2. The rounding noise is very different from the additive white noise, in many respects. In particular, it is not independent from p but, differently from $(\mathbf{M.II})'$, it is a deterministic and known function of p. Rounding results in autocorrelated returns with negative autocorrelation at lags 1 and 2, see Figure 5.2. Markets often specify a minimum price increment, also known as a tick size; this results in prices which are often unchanged for a few consecutive observations, a property which is not compatible with a semimartingale model. In fact, one implication of

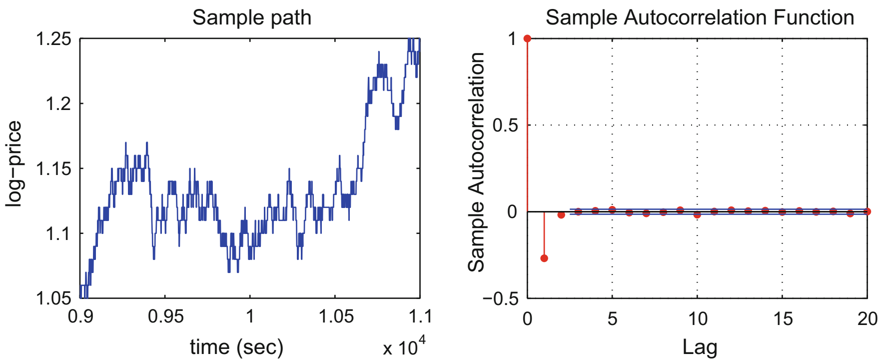

Fig. 5.2 Portion of trajectory of a rounded log-price process $\tilde{p}$ and sample autocorrelation function of log-returns of observed prices at a fixed 1 second frequency. For each lag $j = 0, 1, 2, \ldots, 20$ on the x-axis, the correlation of $\delta_i(\tilde{p})$ and $\delta_{i-j}(\tilde{p})$, estimated from observations $i = j+1, \ldots, n$, is reported.

rounding is that both observed returns and volatility can be zero over short intervals, an outcome that has zero probability of occurrence in any model that contains a Brownian semimartingale component and with non-noisy observations.

A realistic treatment of rounding effects would require that we operate on price levels instead of log returns.[3] However, rounding at price level has smaller impact on log-returns even if we assume an initial price of an order of magnitude of 2 dollars and a tick size of one cent. Therefore, in our simulations we operate rounding directly on the log-price $p(t)$.

[2] $\zeta = 0$ corresponds to the case with independent noise assumption.

[3] An example of rounding at price level is provided in Section 5.4.

We test the performance of the Fourier estimator of integrated variance under microstructure noise of the kind $(\mathbf{M.I}) - (\mathbf{M.II})'$ and $(\mathbf{MR})$ through a Monte Carlo simulation. The infinitesimal variation of the true log-price process and spot volatility is given by the square-root model by Cox et al. (1985)

$$\begin{aligned} dp(t) &= \sigma(t)\, dW_1(t) \\ d\sigma^2(t) &= \gamma(\beta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t), \end{aligned} \tag{5.10}$$

where W_1, W_2 are independent Brownian motions. The parameter values used in the simulations reflect the features of IBM time series: $\gamma = 0.01$, $\beta = 1.0$, $\nu = 0.05$. The initial value of σ^2 is set equal to one, while $p(0) = \log 2$ to make the effect of rounding more evident. The simulations are run for 500 daily replications.

Figure 5.3 shows the true and estimated bias and MSE for the integrated volatility computed from 1 second returns contaminated by dependent noise and by rounding errors, respectively. Dependent noise is defined by $(\mathbf{M.II})'$, with $\zeta = 0.1$ and $\eta_j \sim \mathcal{N}(0, \xi^2)$, $\xi^2 = 0.000142$, while rounding is operated with a tick size $l_\alpha = 0.01$. We notice that in both cases the estimated curves are very close to the true ones, although the formulae (5.3) for the bias and (5.6) for the MSE were obtained under different noise structure. In the case of dependent noise the optimal cutting frequency obtained by minimization of the true MSE is rather small $N_{cut} = 271$ and

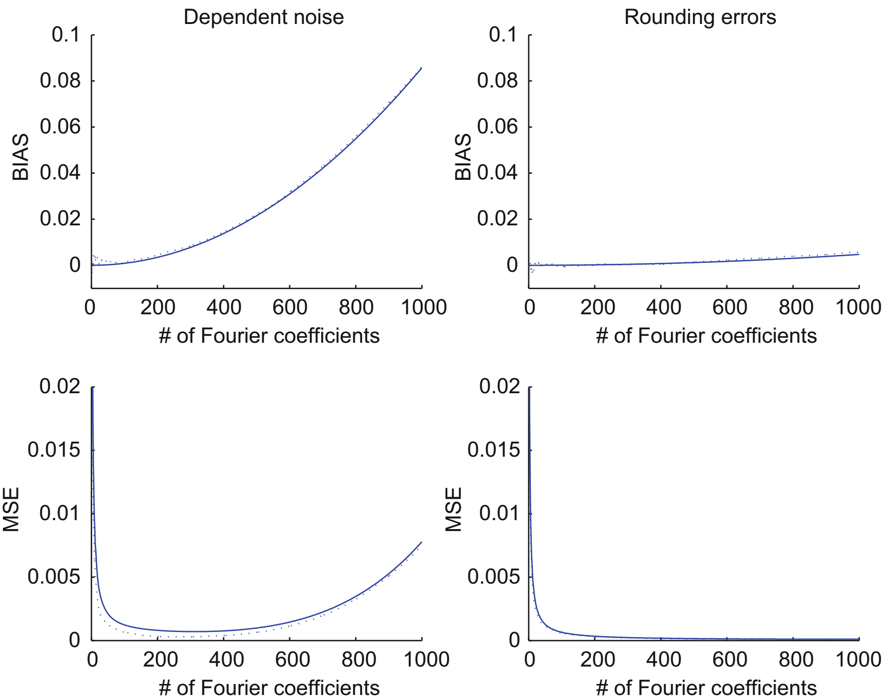

Fig. 5.3 True (dotted line) and estimated (solid line) bias and MSE of the Fourier estimator as a function of the number of the Fourier coefficients. Second-by-second returns over a daily trading period of $T = 6$ hours, for a total of 500 days.

yields an MSE value of 2.90e-4. The MSE achieved by the feasible procedure, i.e. by minimization of estimate (5.6), is 2.99e-4 for $N_{cut} = 313$. Note that the cutting frequency (5.9) selects the frequency $N_{cut} = 825$ that gives an MSE larger by an order of magnitude. Finally, the effect of rounding is rather small if compared to other forms of microstructure effects. This reflects on a higher optimal cutting frequency $N_{cut} = 768$ at which MSE$= 9.33$e-5. The MSE achieved by the feasible procedure is 9.74e-5. These values are resumed in Table 5.1 for the convenience of the reader.

Table 5.1 Optimal MSE based Fourier estimator characteristics under $(\mathbf{M.I}) - (\mathbf{M.II})'$ and $(\mathbf{MR})$ microstructure noise. Unfeasible optimization of the true MSE versus feasible optimization of estimate (5.6).

MSE optimization	UNFEASIBLE		FEASIBLE	
Noise structure	MSE	N_{cut}	MSE	N_{cut}
$(\mathbf{M.I}) - (\mathbf{M.II})'$	2.90e-4	271	2.99e-4	313
$(\mathbf{MR})$	9.33e-5	768	9.74e-5	934

5.2.3 Comparison with Other Estimators

This section studies the Fourier estimator performance in comparison with other estimators that have been specifically proposed in the literature to handle microstructure noise effects.

Alternative methods, which aim at controlling the microstructure noise effects, are essentially based on three techniques: *sub-sampling*, *bias-correction*, and *pre-averaging*. The simplest strategy to reduce the impact of noise is given by the sparse sampling of the available data, that is using a sampling interval of some minutes (5 minutes are suggested by Barndorff-Nielsen and Shephard (2002)). However, this methodology ignores a lot of information, which is not statistically efficient. For a more efficient implementation of the sparse sampling method, Zhang et al. (2005) proposed a sub-sampling approach, namely the Two-Scales Realized Volatility, that averages lower frequency realized volatilities. Bias-correction makes use of various order auto-covariances to correct the spurious noise-induced autocorrelation of observed log-returns, see, e.g., the bias-corrected estimator by Hansen and Lunde (2006b) and the Realized Kernels estimator by Barndorff-Nielsen et al. (2008, 2011b). The pre-averaging technique has been proposed by Jacod et al. (2009) and is based on the idea that if one averages a number of observed log-prices, one is closer to the latent process p.

Besides the well-known *Realized Volatility* defined in (3.8), hereafter denoted by RV, we consider the following estimators of the integrated volatility belonging to the above-mentioned classes: the *bias-corrected estimator*

$$HL := RV + 2\frac{n}{n-1}\sum_{j=1}^{n-1} \delta_j(p)\delta_{j+1}(p), \tag{5.11}$$

the *flat-top Realized Kernels*

$$RK := \sum_{h=-H}^{H} k\left(\frac{h}{H+1}\right) \sum_{j=|h|+1}^{n} \delta_j(p)\delta_{j-|h|}(p), \tag{5.12}$$

with kernels $k(\cdot)$ of *Bartlett*, *Cubic*, and *Tukey-Hunning* (hereafter TH_2) type.[4] The Realized Kernels may be considered as unbiased corrections of the Realized Volatility by means of the first H autocovariances of the returns. In particular, when H is selected to be zero the Realized Kernels become the Realized Volatility. Our analysis includes also the *Two-Scales estimator*

$$TSRV := \frac{S}{S-1}\left(\frac{1}{S}\sum_{s=1}^{S} RV(G^{(s)}) - \frac{1}{S}RV\right). \tag{5.13}$$

The Two-Scales (subsampling) estimator is a bias-adjusted average of lower frequency realized volatilities $RV(G^{(s)})$ computed on S non-overlapping observation subgrids $G^{(s)}$ containing n_S observations.

The *Pre-Averaging estimator* is defined as

$$PA := \frac{12}{k_n}\sum_{s=0}^{n-k_n+1}(\bar{\delta}_s(p))^2 - \frac{6}{k_n^2}\sum_{s=1}^{n}(\delta_s(p))^2, \tag{5.14}$$

where the pre-averaged returns are

$$\bar{\delta}_s(p) := \frac{1}{k_n}\left(\sum_{j=k_n/2}^{k_n-1} p_{t_{s+j}} - \sum_{j=0}^{k_n/2-1} p_{t_{s+j}}\right), \tag{5.15}$$

and k_n is a bandwidth parameter going to ∞ as $n \to \infty$ at the proper rate (see Jacod et al. (2009)).

The proposed Monte Carlo exercise simulates discrete data from the continuous time stochastic volatility model (5.10) with microstructure contaminations. In particular, we consider the microstructure noise model (5.1), assuming **(M.I)** - **(M.II)**: the logarithmic noises η_j are i.i.d. Gaussian with zero mean and variance equal to $E[\eta^2]$ and independent from p. The simulations are run for 500 daily replications, starting from the initial values $\sigma^2(0) = 1$ and $p(0) = \log 100$. 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-

---

[4] Bartlett kernel: $k(x) = 1-x$; Cubic kernel: $k(x) = 1-3x^2+2x^3$; TH_2 kernel: $k(x) = \sin^2[\pi/2(1-x)]$.

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_j. For instance, the choice $n = 360$ corresponds to a sampling period of $\rho(360) = 1$ minute.

Feasible optimal rules for choosing the bandwidth-parameters employed by the considered estimators are discussed in the cited papers and resumed in Table 5.2 for the reader's convenience. For the Fourier estimator, the optimal cutting frequency N can be obtained by direct minimization of the estimated MSE given by (5.6). Note that Q is the integrated quarticity estimated by means of low frequency returns.

Table 5.3 shows the performance of the different estimators when the key parameters are obtained by the feasible rules described in Table 5.2. For the Fourier and *TSRV* estimators the feasible procedure allows us to obtain levels of MSE that are close to the real optimum. On the contrary, the feasible optimization of the *RK* and

Table 5.2 Optimal bandwidth-parameters.

Estimator	Optimal bandwidth
RV (3.8)	$n^* = (TQ/4E[\eta^2]^2)^{1/3}$
HL (5.11)	n^* = number of price observations
RK (5.12)	$H = c^* \xi^{4/5} n^{3/5},\quad c^* = (144/0.269)^{1/5}$ $\xi^2 = E[\eta^2]/\sqrt{Q}$
TSRV (5.13)	$S = c^* n^{2/3},\quad c^* = (TQ/48E[\eta^2]^2)^{-1/3}$
PA (5.12)	$k_n = c^* \xi^{4/5} n^{3/5},\quad \xi^2 = E[\eta^2]/\sqrt{Q}$

Table 5.3 Comparison of optimized integrated volatility estimators. Feasible optimization.

	MSE				BIAS			
	1 sec	30 sec	1 min	5 min	1 sec	30 sec	1 min	5 min
RV	3.76e+1	4.12e-2	1.13e-2	2.32e-3	6.13e+0	2.01e-1	1.03e-1	1.52e-2
Sparse samp. *RV*	2.48e-3	2.48e-3	2.20e-3	2.32e-3	2.63e-2	3.00e-2	2.78e-2	1.52e-2
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
Sparse samp. *HL*	4.76e-3	4.20e-3	4.11e-3	5.51e-3	-2.33e-3	-1.38e-3	-7.71e-4	-3.70e-3
Fourier	2.99e-4	1.11e-3	1.52e-3	3.55e-3	8.95e-3	1.66e-2	1.94e-2	3.27e-3
Bartlett *RK*	2.04e-4	1.17e-3	1.82e-3	6.45e-3	8.58e-4	3.97e-4	5.53e-5	-5.80e-3
Cubic *RK*	2.25e-4	1.23e-3	1.94e-3	6.45e-3	8.39e-4	4.91e-4	-8.41e-5	-5.80e-3
TH_2 *RK*	1.47e-4	8.78e-4	1.45e-3	5.56e-3	6.06e-4	3.19e-4	-1.66e-4	-4.32e-3
TSRV	1.01e-4	7.49e-4	1.34e-3	5.43e-3	-3.99e-5	-1.67e-3	-1.49e-3	-1.11e-2
PA	1.67e-4	1.18e-3	2.03e-3	8.56e-3	-3.65e-4	-1.29e-2	-2.07e-2	-7.42e-2

PA estimators is not able to capture the actual minimum of the MSE and leads to slightly less efficient results so that, in practice, especially at the highest frequencies, their performance is comparable to the other ones. We notice that, at a sampling frequency of 5 minutes the effects of the microstructure noise are not evident and the *RV* and *HL* estimators are as efficient as the others. For frequencies higher than 1 minute, the noise-induced autocorrelation of returns becomes effective and the *RV* starts to strongly overestimate the underlying integrated volatility. The first order bias-correction of estimator HL compensates the microstructure effects and reduces the bias of the RV. Nevertheless, for each data set $(t_j,\ \tilde{p}(t_j),\ j = 0, 1 \ldots n)$, with $n = 21600, 720, 360, 72$, we tried to improve the performance of the RV and of the bias-corrected estimator HL by sparse sampling, according to the rule-of-thumb mentioned above giving the optimal sampling frequency T/n^*. Obviously, sparse sampling has a positive effect on RV, while it has negative effect on HL. Nevertheless, both estimators are not competitive with the others at the highest frequencies.

At the highest frequency, the $TSRV$ estimator provides 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 estimator provides very good results, even in comparison with methods specifically designed to handle market microstructure contaminations, and is practically unaffected by noise, having only a slightly higher MSE and bias for quote-to-quote returns.

Remark 5.3. The Fourier-Fejér estimator (3.9) slightly improves the behavior of the Fourier estimator with Dirichlet kernel for very high frequencies (see Mancino and Sanfelici (2008)).

5.2.4 An Empirical Application

This section contains a merely illustrative example of Fourier estimation method with empirical data. The reader can find the relevant codes employed in the Appendix B.

The data set is composed by quote-to-quote logarithmic prices of the Italian stock index futures, named FIB30, on January 14, 2000. Only the prices of the next-to-expiration contracts, which are the most liquid ones, are employed, with the FIB30 expiring quarterly. 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. We have a total of 7170 quotes and on average a new quote arrives every 4.14 seconds. The smallest return is -0.43% and the largest is 0.43%. Figure 5.4 shows the time series plot of the tick-by-tick returns and the autocorrelation up to lag 20. The first-order autocorrelation is significantly negative and equal to -0.3592, with 95% confidence interval $[-0.0236, 0.0236]$, while after lag 1 they are insignificantly different from 0 (or marginally significant). Thus, the MA(1) approximation seems to capture the main economic effects in the data. Quotes prior to 10 a.m. have been removed to eliminate opening quotes from our sample so that $n = 6157$.

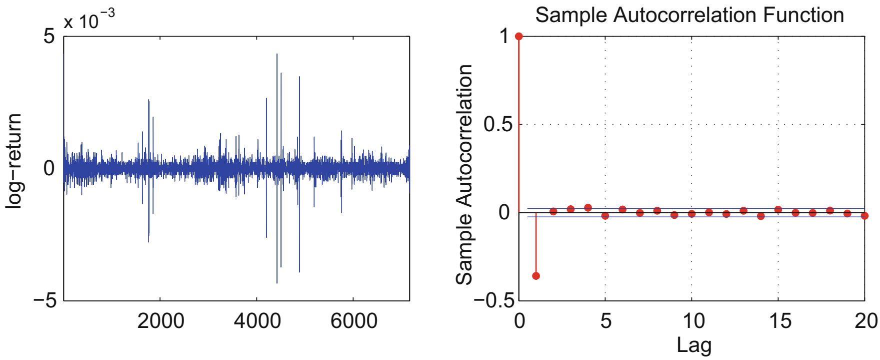

Fig. 5.4 Time plot of the tick-by-tick returns and the ACF for FIB30 on January 14, 2000.

In Figure 5.5 we plot the estimated conditional MSE and bias of the Fourier estimator based on quote-to-quote returns, for N ranging from 1 to $\lfloor n/2 \rfloor = 3078$. The

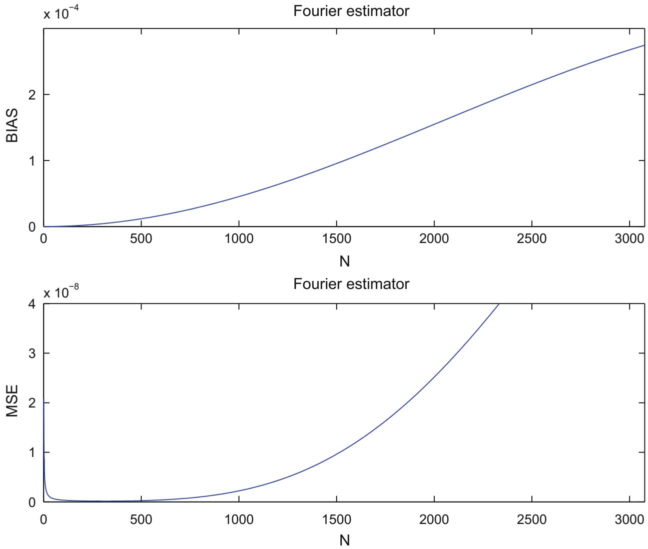

Fig. 5.5 Estimated conditional bias and MSE of the Fourier estimator as a function of the maximum number of Fourier coefficients.

MSE is estimated by implementing formula (5.6), 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 2-minute returns. The minimum of the MSE for the Fourier estimator is 1.565e-010 attained at $N = 319$. As suggested by the theory exposed in the previous sections, by choosing a suitable cutting frequency N it is possible to render the Fourier estimator invariant to short-run noise introduced by market microstructure effects, with consequent efficiency gains. Figure 5.6 plots the Fourier volatility estimates as a function of the maximum number of Fourier coefficients N based on tick-by-tick data (blue line)

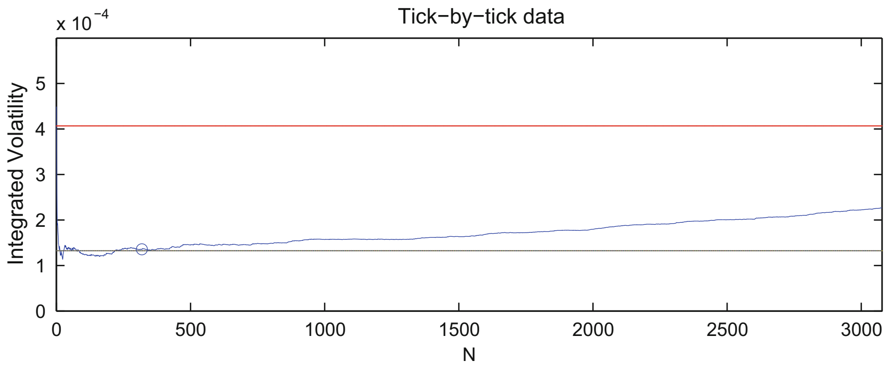

Fig. 5.6 Fourier volatility estimate as a function of the maximum number of Fourier coefficients N from tick-by-tick data (blue line); Realized Volatility based on tick-by-tick data (red line); Realized Volatility based on 2-minute returns (black line).

and the corresponding Realized Volatility estimate based on tick-by-tick data (red line). We clearly see that the microstructure effects contained in high-frequency data seriously spoil the Realized Volatility estimates and make sparse sampling strongly necessary. For instance, the Realized Volatility estimate 1.323e-4 based on 2-minute log-returns (black line) seems to filter out great portion of noise. The small circle indicates the Fourier estimate 1.355e-4 attained at $N = 319$, which is the cutting frequency minimizing the MSE estimated by (5.6). Since our theoretical results indicate that the Realized Volatility estimator is more biased than the Fourier estimator in the presence of market microstructure noise, the fact that the Realized Volatility estimate for commonly used sampling frequencies (e.g., 2 minutes) is 1.323e-4 indicates that the actual volatility might be higher than predicted by the commonly used Realized Volatility, as already noticed by Nielsen and Frederiksen (2008).

5.3 The Case of Integrated Covariance

The estimation of multivariate volatilities is a challenging task for the combined effects of microstructure noises and asynchronous trading times. In fact, when considering intraday financial data, trades on different assets are not likely to occur at the same time. In such situations, the addition of a moderate amount of independent and uncorrelated noise may not have great effect on the estimates and it may in some cases even compensate the effects due to non-synchronicity. On the other side, Griffin and Oomen (2011) find that the ordering of covariance estimators in terms of efficiency depends crucially on the level of microstructure noise; in particular for high level of noise an estimator which is consistent for asynchronous observations, like (3.29), can become less efficient than the standard Realized Covariance.

This chapter starts with the finite sample properties of the Fourier covariance estimator in the presence of microstructure effects. Under a benchmark price model, we analytically compute the bias and the MSE for given finite sample sizes of the different assets and given number of Fourier coefficients included in the series. Under suitable growth conditions for these parameters, it is possible to prove that

- the bias of the Fourier estimator asymptotically vanishes,
- the MSE of the Fourier estimator converges to a constant.

As a consequence, 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. Moreover, we provide a practical way to optimize the finite sample performance of the Fourier estimator as a function of the number of frequencies by the minimization of the MSE, for a given number of intra-daily observations.

For simplicity, we consider the case of two assets. Assume the following model for the observed log-prices

$$\widetilde{p}^i(t^i_j) := p^i(t^i_j) + \eta^i(t^i_j) \quad \text{for } i = 1,2, \tag{5.16}$$

where the processes p^i are driven by model (2.1) with $b^i \equiv 0$, and the noise model is specified by the assumptions[5]

(Mm.I) $p(t) := (p^1(t), p^2(t))$ and $\eta(t) := (\eta^1(t), \eta^2(t))$ are independent for any t;

(Mm.II) $\eta(t)$ and $\eta(s)$ are independent for $s \neq t$, $E[\eta(t)] = 0$ for any t and $E[\eta^i(t)\eta^k(t)] = \omega_{ik} < \infty$, for any t, with $i,k = 1,2$.

In order to obtain simple analytic formulae, the asset prices $(\widetilde{p}^1, \widetilde{p}^2)$ are observed on particular grids of *regular asynchronous trading*. The asset 1 trades at regular points: $\Pi^1 = \left\{ t^1_i \in [0,2\pi] : i = 1,\dots,n_1 \text{ and } t^1_{i+1} - t^1_i = \frac{2\pi}{n_1 - 1} \right\}$. Also asset 2 trades at regular points: $\Pi^2 = \left\{ t^2_j \in [0,2\pi] : j = 1,\dots,n_2 \text{ and } t^2_{j+1} - t^2_j = \frac{4\pi}{n_1 - 1} \right\}$, where $n_2 = n_1/2$, 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: $t^2_j = t^1_{2(j-1)+1} + \frac{\pi}{n_1 - 1}$ for $j = 1,\dots,n_2$. Moreover, suppose $t^1_1 = 0$ and $t^1_{n_1} = 2\pi$. For simplicity, denote $n := n_1$ and assume n is even. We consider the Fourier-Fejér estimator (3.31) of the covariance between asset 1 and 2, according to Remark 5.3, and we will denote it by $\widehat{\Sigma}^{1,2}_{N,n}$.

The bias of the Fourier covariance estimator under microstructure noise satisfying **(Mm.I)–(Mm.II)**, neglecting minor end-effects, is computed by Mancino and Sanfelici (2011b) and is equal to

$$E[\widehat{\Sigma}^{1,2}_{N,n} - \int_0^{2\pi} \Sigma^{1,2}(t)dt] = \sum_{j=1}^{\frac{n}{2}-1} \sum_{i=2(j-1)+1}^{2(j-1)+3} (F_N(t^1_i - t^2_j) - 1) E[\int_{t^1_i}^{t^1_{i+1}} \Sigma^{1,2}(t)dt]. \tag{5.17}$$

From (5.17) the following result can be stated

[5] This noise structure only allows for contemporaneous and not serial correlation in the noise.

the Fourier covariance estimator is asymptotically unbiased in the presence of microstructure noise, under the condition $N/n \to 0$ as $n,N \to \infty$.

Remark 5.4. Apparently, the microstructure noise plays no role in equation (5.17) which corresponds to (3.28) under the specific regular asynchronous trading setting. Actually, under regular asynchronous trading, trades on different assets never occur at the same time. As a consequence, the potential bias of the Fourier estimator is not affected by the presence of microstructure noise satisfying **(Mm.I)–(Mm.II)**, but is exclusively caused by the asynchronicity, as $F_N(t_i^1 - t_j^2) \neq 1$ in formula (5.17). This fact motivates the growth rate condition $N/n \to 0$ instead of $N^2/n \to 0$ as $n,N \to \infty$.

On the contrary, the presence of microstructure noise has effects on the Fourier estimator's MSE which is given by

$$
\begin{aligned}
E[(\widehat{\Sigma}^{1,2}_{N,n} - \int_0^{2\pi} \Sigma^{1,2}(t)dt)^2] &= o(1) + 2\omega_{22} \sum_{i=1}^{n-1} F_N^2(t_i^1 - t^2_{\frac{n}{2}-1}) E[\int_{t_i^1}^{t_{i+1}^1} \Sigma^{11}(t)dt] \\
&+ 2\omega_{11} \sum_{j=1}^{\frac{n}{2}-1} F_N^2(t^1_{n-1} - t_j^2) E[\int_{t_j^2}^{t_{j+1}^2} \Sigma^{22}(t)dt] + 4\,\omega_{22}\omega_{11}\, F_N^2(t^1_{n-1} - t^2_{\frac{n}{2}-1}),
\end{aligned}
\tag{5.18}
$$

where $o(1)$ is a term which goes to zero, for $N/n \to 0$ as $n,N \to \infty$. The $o(1)$ term in (5.18) is the MSE of the estimator for a pure diffusive process; it converges to zero, as the multivariate estimator is consistent under condition $N/n \to 0$ from (3.25). The other terms appear because of the microstructure noise components. The second and third terms are constant for increasing n. Finally, the term $4\omega_{11}\omega_{22}F_N^2(t^1_{n-1} - t^2_{\frac{n}{2}-1})$ converges to the constant $4\omega_{11}\omega_{22}$ as n,N increase at the proper rate $N/n \to 0$. We conclude as follows.

Under the condition $N/n \to 0$, the MSE of the Fourier multivariate volatility estimator: *(i)* converges to 0 (i.e., the estimator is consistent) under asynchronous observations, *(ii)* in the presence of microstructure noise (satisfying (**Mm**.**I**) – (**Mm**.**II**)) does not diverge at the highest data frequency.

Remark 5.5. In the case of *synchronous* trading times, the Fourier covariance estimator exhibits the same behavior with the univariate case analyzed in Section 5.2. In particular, if the two asset price processes are recorded on the same time grid (which for simplicity we suppose equally spaced with mesh size $2\pi/n$) and assumptions **(Mm.I) - (Mm.II)** hold, then it holds

$$
E[\widehat{\Sigma}^{1,2}_{N,n} - \int_0^{2\pi} \Sigma^{1,2}(t)dt] = 2n\,\omega_{1,2}\left(1 - F_N\left(\frac{2\pi}{n}\right)\right). \tag{5.19}
$$

Thus, in the synchronous case, the Fourier estimator of covariance is asymptotically unbiased in the presence of microstructure noise, if $N^2/n \to 0$ as $n, N \to \infty$.

5.3.1 Comparison with Other Estimators

In this section the performance of the Fourier covariance estimator is compared with that of different nonparametric estimators, taking into account two main intrinsic features of high-frequency data, namely the microstructure noise contamination and the asynchronicity.

The following covariance estimators will be considered in the simulation study. A first group of estimators requires a preliminary synchronization procedure, such as linear or piecewise constant interpolation, which yields the observations times $\{0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_n \leq 2\pi\}$ for both assets. The *Realized Covariance* estimator $RC^{1,2}$ defined in (3.27), is not consistent under asynchronous trading and highly biased in the presence of noise effects (see also Hayashi and Yoshida (2005), Zhang (2009)). Several modifications of the Realized Covariance have been proposed. We consider here the *Realized Covariance plus Leads and Lags*

$$RCLL^{1,2} := \sum_i \sum_{h=-l}^{L} \delta_{i+h}(p^1)\delta_i(p^2). \tag{5.20}$$

The estimator (5.20) has good properties under microstructure noise contaminations of the prices, but it is still not consistent for asynchronous observations.

The following two estimators employ the synchronization procedure known as *refresh time*, i.e. the first time when both posted prices are updated, setting the price of the quicker asset to its most recent value (last-tick interpolation). The first one is the *Multivariate Realized Kernel* estimator, introduced by Barndorff-Nielsen et al. (2011a) and defined as

$$K^{1,2} := \sum_{h=-n}^{n} k\left(\frac{h}{H+1}\right)\Gamma_h^{1,2}, \tag{5.21}$$

where $\Gamma_h^{1,2}$ is the h-th realized autocovariance of the two assets and $k(\cdot)$ belongs to a suitable class of kernel functions.[6] The second one, proposed and called *Modulated Realized Covariation* by Christensen et al. (2010), exploits the econometric technique of pre-averaging of the high-frequency data

$$MRC^{1,2} = \left(1 - \frac{6}{k_n^2}\right)^{-1}\left[\frac{n}{n-k_n+2}\frac{12}{k_n}\sum_{s=0}^{n-k_n+1}\bar{\delta}_s(p^1)\bar{\delta}_s(p^2) - \frac{6}{k_n^2}\sum_{s=1}^{n}\delta_s(p^1)\delta_s(p^2)\right], \tag{5.22}$$

[6] In our analysis, we will consider the Parzen weight kernel $k(x)$ defined as $1-6x^2+6x^3$ for $0 \leq x \leq 0.5$, $2(1-x)^3$ for $0.5 < x \leq 1$ and 0 otherwise.

where the pre-averaged return process is defined by (5.15) and k_n is a bandwidth parameter. In order to be a consistent estimator of the integrated covariance, however, it needs a bias-correction. Therefore, it is not guaranteed to be positive semi-definite. We will also consider the *All-Overlapping* estimator, $AO^{1,2}$, defined by (3.29), which is consistent under asynchronous prices observations, but not efficient in the presence of microstructure noise as follows from the studies by Voev and Lunde (2007), Griffin and Oomen (2011). Finally, the ideas of pre-averaging and All-Overlapping synchronization have been merged in the *Pre-averaged All-Overlapping estimator* by Christensen et al. (2010). This estimator can be implemented on the original data without prior alignment of prices and is defined as

$$PAO^{1,2} = \frac{16}{k_n^2} \sum_{i=0}^{n_1-k_n+1} \sum_{j=0}^{n_2-k_n+1} \bar{\delta}_i(p^1)\bar{\delta}_j(p^2)\, 1_{\{(t_i^1,t_{i+k_n}^1]\cap(t_j^2,t_{j+k_n}^2]\neq\emptyset\}}, \tag{5.23}$$

where the indicator function discards pre-averaged returns not overlapping in time.

Remark 5.6. (Optimal bandwidth selection rules). All the parameters involved in the design of the different estimators can be optimized on a daily basis. As suggested by Barndorff-Nielsen et al. (2011a), when implementing the multivariate Realized Kernel, on each day the univariate optimal MSE-based bandwidth selection applied to each asset price individually gives $H_i = c^*\xi_i^{4/5}n^{3/5}$, where $c^* = (144/0.269)^{1/5}$, $\xi_i^2 = \omega_{ii}/\sqrt{Q_{ii}}$ and Q_{ii} is the integrated quarticity of asset i estimated by means of low frequency returns. The two bandwidths are then averaged to obtain the global H value. In the case of the $MRC^{1,2}$ and the $PAO^{1,2}$ estimators, a recommended bandwidth parameter is $k_n = (k_n^{(1)} + k_n^{(2)})/2$, where $k_n^{(i)} = \theta_i n^{3/5}$ and $\theta_i = c^*\xi_i^{4/5}$. In the case of the Fourier estimator, we build the optimal MSE-based estimator by choosing the cutting frequency N_{cut} which minimizes the estimated MSE (5.18) instead of the true one over a finite range of N values.

We simulate discrete data from the continuous time bivariate GARCH model (3.30). Moreover, we assume that the logarithmic noises $\eta^1(t), \eta^2(t)$ are i.i.d. Gaussian, possibly contemporaneously correlated and independent from p. We set the *noise-to-signal ratio* ς, defined as the noise standard deviation over the total standard deviation for 1 second returns, equal to 5.5 which is in fact quite relevant. We also consider the case of dependent noise, assuming for simplicity $\eta_i^j = \zeta(p^j(t_i^j) - p^j(t_{i-1}^j)) + \bar{\eta}_i^j$, for $j = 1,2$, with $\bar{\eta}_i^j$ i.i.d. Gaussian and $\zeta = 0.1$.

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 $h = 6$ hours, for a total of 21600 observations per day. Then we sample the observations according to the regular non-synchronous trading sampling scheme, 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. From the simulated data, integrated covariance estimates can be compared to the value of the true variance quantities. The results are reported in Table 5.4 and have to be related to those of Table 3.2.

Table 5.4 Comparison of integrated volatility estimators. Noise ratio $\varsigma = 5.5$. $\rho_1 = 5$ sec, $\rho_2 = 10$ sec with a displacement of 2 seconds for Reg-NS trading. The low frequency returns necessary for the RC-type estimators are obtained by imputation on a uniform grid. The number of leads and lags for the $RCLL^{1,2}$ estimators is $l = L = 1$. All the other estimators have been optimized on a feasible and daily basis as indicated in Remark 5.6.

	Reg-NS + Unc		Reg-NS + Cor		Reg-NS + Dep	
	MSE	bias	MSE	bias	MSE	bias
$\widehat{\Sigma}^{1,2}_{n_1,n_2,N}$	2.40e-3	-1.08e-2	2.20e-3	-1.00e-2	9.85e-3	-1.77e-2
$RC^{1,2}_{0.5min}$	3.24e-2	-1.57e-1	3.16e-2	-1.59e-1	2.12e-1	-1.51e-1
$RC^{1,2}_{1min}$	1.28e-2	-8.21e-2	1.18e-2	-7.75e-2	1.37e-1	-7.69e-2
$RC^{1,2}_{5min}$	1.26e-2	-1.31e-2	1.31e-2	-1.04e-2	5.06e-2	-1.70e-2
$RCLL^{1,2}_{0.5min}$	7.83e-3	-1.47e-3	7.05e-3	1.98e-3	1.34e-1	-3.54e-3
$RCLL^{1,2}_{1min}$	1.05e-2	2.23e-3	9.22e-3	1.96e-3	8.97e-2	-1.33e-2
$RCLL^{1,2}_{5min}$	3.73e-2	4.36e-3	3.19e-2	1.56e-2	6.15e-2	6.22e-3
$AO^{1,2}$	1.08e-2	1.71e-5	9.15e-3	1.34e-3	4.02e-1	4.95e-2
$K^{1,2}$	6.56e-3	-1.80e-3	6.04e-3	-6.17e-4	2.06e-1	2.95e-2
$MRC^{1,2}$	4.71e-3	-1.14e-2	4.56e-3	-1.22e-2	7.38e-3	-1.11e-3
$PAO^{1,2}$	9.25e-3	-2.62e-3	9.55e-3	-1.61e-3	1.49e-2	-7.87e-3

Within each table entries are the values of the MSE and bias, using 500 Monte Carlo replications. Rows correspond to different estimators, while columns correspond to different type of noise, namely contemporaneously uncorrelated ($\omega_{ij} = 0$ for $i \neq j$), contemporaneously correlated and dependent on the price process, respectively. When the noise correlation matrix is not diagonal, the correlation is set to 0.5.

For any considered size of the synchronization grid (0.5, 1 and 5 minutes), the $RC^{1,2}$ estimator has poor performances; the lead/lag bias-correction partially compensates the effect of noise, at least in the cases of noise independent of the efficient price. As already found by Griffin and Oomen (2011), the performance of the $AO^{1,2}$ estimator is strongly affected by the extent of noise. The Fourier estimator always provides a valid alternative, even in the case of dependent noise although the MSE estimate (5.18) we use to select the cutting frequency does not account for dependence between noise and efficient price. The only estimator which is able to provide a good alternative to the Fourier estimator is the $MRC^{1,2}$ and, in the case of dependent noise, the $PAO^{1,2}$ estimator. The $K^{1,2}$ estimator provides acceptable estimate for low levels of noise but is rapidly swamped by the presence of large microstructure effects. Barndorff-Nielsen et al. (2011a) themselves in their simulations find out that there is not a great difference between the multivariate Realized (Parzen) Kernel and the sparse sampled Realized Covariance. This could be related to the synchronization procedure, which may result in excessive data reduction.

In agreement with our theoretical analysis, the proposed simulation study suggests the following conclusion.

The Fourier covariance estimator is not much affected by the presence of noise and asynchronicity, so that it becomes a very interesting alternative especially when microstructure effects are particularly relevant in the available data.

5.3.2 An Empirical Application

In this section, the estimator of the integrated covariance (3.22) is applied with high frequency market data which exhibit both asynchronicity and noise effects.

Specifically, we analyze four exchange rates with different liquidity (Forex data): AUD/JPY (Australian Dollar vs Japanese Yen), USD/CHF (US Dollar vs Swiss Franc), EUR/USD (Euro vs US Dollar) GBP/USD (British Pound vs US Dollar) observed on July 25, 2016. The sample sizes are: 1728 for AUD/JPY (50 seconds evenly sampled data), 8031 for USD/CHF (unevenly sampled data, one observation every 10.75 seconds on average), 38475 for EUR/USD (unevenly sampled data, one observation every 2.25 seconds on average), 57731 for GBD/USD (unevenly sampled data, one observation every 1.5 seconds on average). The USD/CHF exchange rate suffers from missing values with data gaps of one or two minutes while the GBD/USD and EUR/USD rates have only missing data with gaps of a few seconds. These exchange rates are also characterized by different trading volumes and, as a consequence, different liquidity. In fact, the trading volumes of AUD/JPY, USD/CHF, EUR/USD, and GBP/USD are 422355.5, 293248.4, 898692.5, and 397047.1, respectively. Therefore, the pairs (AUD/JPY, USD/CHF), (AUD/JPY, EUR/USD) and (AUD/JPY, GBP/USD) are those with the most severe non-synchronicity and the pair (AUD/JPY, EUR/USD) presents further criticism generated by the different liquidity. Finally, all data are affected by microstructure noise.

The integrated covariance of each pair of exchange rates is estimated by the Fourier estimator $\widehat{\Sigma}^{1,2}_{n_1,n_2,N}$ using the entire dataset and by the Realized Covariance estimates $RC^{1,2}$ using two different synchronization procedures. The first one provides the estimator $RC^{1,2}_{max,L}$ by last tick imputation over a grid with size equal to the average transaction interval of the least traded exchange rate in each considered pair. The other estimator, $RC^{1,2}_{RT}$, is obtained by using the refresh time procedure, employed in Section 5.3.1 to construct the Multivariate Realized Kernel estimator (5.21). In implementing the Fourier estimator, the cutting frequency N_{cut} of the Fourier estimator is chosen as indicated in Remark 5.6.

Table 5.5 shows that the $RC^{1,2}_{max,L}$ estimates deteriorate and a bias toward zero appears. This effect is particularly evident in the first three lines of the table, where the asynchronicity between the time series is more pronounced. Moreover, in the first three cases the imputation on a 50-second grid entails a larger discretization error. Finally, we observe that the $RC^{1,2}_{RT}$ estimates are generally affected by a large

Table 5.5 Integrated covariance estimation.

currency pair	$\widehat{\Sigma}^{1,2}_{n_1,n_2,N}$	$RC^{1,2}_{max,L}$	$RC^{1,2}_{RT}$
(AUD/JPY, USD/CHF)	-3.58e-6	-9.36e-7	-1.30e-6
(AUD/JPY, EUR/USD)	2.97e-6	1.42e-7	-1.85e-7
(AUD/JPY, GBP/USD)	1.12e-5	1.53e-6	-1.99e-6
(USD/CHF, EUR/USD)	-5.84e-6	-3.99e-6	-1.88e-6
(USD/CHF, GBP/USD)	-4.68e-6	-1.93e-6	-2.14e-7
(EUR/USD, GBP/USD)	7.99e-6	6.30e-6	4.47e-6

bias also for pairs with a low degree of asynchronicity, as a consequence of the considerable loss of data when implementing the refresh time procedure.

5.3.3 Asymptotic Results

The finite sample analysis conducted so far has shown that the Fourier estimator of integrated covariance can be efficiently employed in the joint presence of microstructure noise effects and asynchronicity of data. This section introduces consistency and asymptotic normality for a class of Fourier-type estimators of integrated covariance, named *Fourier Realized Kernel*, which has been recently established by Park et al. (2016). Notations from Section 3.3 are in force.

The Fourier Realized Kernel exploits the fact that the main convolution formula (2.5) leading to the estimator of the Fourier coefficients of the covariance given by (2.13) can be modified through a weighting function, in the same spirit of the Fourier-Fejér estimator (3.31). The Fourier Realized kernel estimator of the k-th Fourier coefficient of the covariance between asset i and j can be defined as follows

$$\sum_{|s|\leq N/2} K_H\left(\frac{2\pi s}{\overline{n}}\right) c_s(dp^i_{n_i}) c_{k-s}(dp^j_{n_j}) \tag{5.24}$$

where the kernel K_H satisfies suitable conditions.[7] The frequency N is defined as $N := \overline{n}/H \to \infty$ where $\overline{n}$ is the biggest number of sample sizes amongst all assets, while the bandwidth $H \to \infty$ with $H/n \to 0$, where n is the smallest number of sample sizes between all assets. The 0-th Fourier coefficient provides the estimator of the integrated covariance, as seen in the previous chapters.

The Central Limit theorem for any k-th Fourier coefficients holds under some general conditions that allow for microstructure noise effects and asynchronicity between different assets. Let the parameter β measure the degree of relative liquidity between assets; more precisely, it is defined as $\beta := \max_{i,j} \beta_{i,j}$, where

[7] These conditions are stated by Assumption 3′ in Park et al. (2016)

$$\beta_{i,j} = \lim_{n\to\infty} \frac{\log(n_i \vee n_j)}{\log(n_i \wedge n_j)}.$$

The rate conditions on the bandwidth imposed by Park et al. (2016) imply that $\beta \in [1,2)$. Then, the rate of convergence for the Fourier Realized kernel estimator of any k-th Fourier coefficient (and in particular for the integrated covariance) is $(n_i \wedge n_j)^\theta$ with $\theta = (2-\beta)/5 \in (0,1/5]$ if the sample sizes are of different order and $n^{1/5}$ (i.e. $\beta = 1$) for the volatility estimator (as well as for the covariance estimator with synchronous observations). As a consequence, the rate of convergence becomes slower if the degree of relative liquidity between assets increases. This result supplements the discussion in Remark 3.6.

It is worth noting that the advantage of the Fourier approach is that it does not require an explicit time alignment. Thus, even if the rate of convergence is not the optimal one, that is $O(n^{1/4})$, the sample size n is the sample size of the original dataset. The rate of convergence of the Multivariate Realized kernel (5.21) is $O(n^{1/5})$, closer to the optimal one, but the sample size n is the sample size of data after alignment. Another important aspect is that the Fourier approach guarantees the estimated covariance matrix to be positive defined. The Modulated Realized Covariation (5.22) and the Quasi-Maximum Likelihood estimator by Ait-Sahalia et al. (2010) attain the optimal rate $O(n^{1/4})$; however, in both cases the estimated covariance matrix is not guaranteed to be positive definite. These considerations show that the ranking between the estimators cannot rely only on the rate of convergence but one should take into consideration the specific application of interest.

Remark 5.7. The Fourier Realized kernel estimator coincide with the multivariate Realized kernel by Barndorff-Nielsen et al. (2008, 2011a) only in the special case when trading times are synchronized and equally spaced. Park et al. (2016) show that when the data is not synchronously observed, the Fourier Realized kernel has a superior performance using all the data.

5.4 The Case of Spot Volatility

As it concerns the robustness with respect to market microstructure effects, the Fourier estimator of spot volatility inherits the properties of the integrated volatility estimator studied in Sections 5.2 and 5.3. A consistent estimator of the spot volatility in the presence of noise effects and asynchronicity based on the estimated k-th Fourier coefficient (5.24) can be easily constructed through the inversion formula as in (2.14). The convergence is uniform in time, as in (4.11). The rate of convergence is still an open problem.

This section analyzes the performance of the (univariate) spot volatility estimator under different microstructure noise scenarios through a simulation study. We consider four microstructure noise models. The first one is the additive MA(1) model specified by the assumptions **(M.I)** - **(M.II)** (see Section 5.2). The second and third model are also additive models, but, for the second model, the assumption **(M.I)** is replaced by

(M.I)$'$ the random shocks $\eta(t_j)$ for any $j = 0, 1, \ldots, n$ are allowed for negative first order autocorrelation,

while for the third one, we consider the case when the noise is correlated with the efficient returns, as specified by assumption **(M.II)**$'$ in Section 5.2.2.

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

$$\widetilde{p}(t_j) = \log\left(\left[\frac{\exp(p(t_j))}{l_\alpha}\right] l_\alpha\right), \quad j = 0, \ldots, n, \tag{5.25}$$

where, as already mentioned, $[x]$ denotes the integer closest to x while l_α is the fixed rounding error level (i.e., the tick size). As highlighted in Section 5.2.2, given that stock prices are often rounded to the cent, the choice $l_\alpha = 0.01$ mimics the financial markets.

Let us now describe the considered dataset. The log-prices, $p(t_j)$, $j = 0, 1, \ldots, n$, are generated simulating the following stochastic volatility model

$$\begin{aligned} dp(t) &= (\mu - \sigma(t)^2/2)dt + \sigma(t)\, dW_1(t), \\ d\sigma^2(t) &= \gamma(\theta - \sigma^2(t))dt + \nu\sigma(t)\, dW_2(t), \end{aligned} \tag{5.26}$$

where $W_1(t)$ and $W_2(t)$ are correlated Brownian motions, being λ the correlation.[8] We set $\nu = 0.5/252$, $\gamma = 5/252$, $\theta = 0.1$, $\mu = 0.05/252$, $\lambda = -0.5$, $\sigma^2(0) = 1$, $p(0) = \log(9)$. The noise components $\eta(t_j)$, for any $j = 0, 1, \ldots, n$, have Gaussian distribution with mean zero and variance $\widetilde{\eta}^2$. We choose $\widetilde{\eta} = \varsigma\, std(r)$ where $std(r)$ is the standard deviation of the 1-second returns and $\varsigma = 0.8,\ 3.2$, while ζ appearing in **(M.II)**$'$ is set equal 0.1. The numerical simulations are obtained integrating numerically the stochastic differential equations (5.26) by the explicit Euler discretization scheme to compute second-by-second return and variance paths over a daily trading period of $T = 1$ day (6 hours trading). We simulate 500 trading days and $n = 21600$ observations per day. The volatility is estimated at every minute.

We measure the performance of the spot volatility estimator, $\widehat{\sigma}^2_{n,N,M}(t)$, defined in (4.1) over the entire interval $[0, T]$ by the relative mean squared error

$$RMSE(t) := E\left[(\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t))^2/\sigma^2(t)\right]$$

and the bias

$$BIAS(t) := E\left[\widehat{\sigma}^2_{n,N,M}(t) - \sigma^2(t)\right].$$

Specifically, the performance over the interval $[0, T]$ is evaluated using the integrated relative mean squared error $IRMSE := (1/T)\int_0^T RMSE(t)dt$ and the integrated bias, $IBIAS := (1/T)\int_0^T BIAS(t)dt$.

[8] This model, with the addition of jumps, has been considered by Li and Mykland (2014) to study the effect of rounding errors on integrated volatility estimators.

The study contained in Section 5.2 suggests that the microstructure noise is ignored by the Fourier estimator of integrated volatility by carefully selecting the cutting frequency N which appears in (4.2) through the Dirichlet kernel. In the case of spot volatility, a second parameter, M, has to be set. In particular, it is interesting to know whether it is possible to choose the cutting frequencies N and M independently of the specific point in time t, that is in a global manner, and still preserve the performance of the spot volatility Fourier estimator over the whole time interval.

The Fourier spot volatility estimates are computed using several values of the frequencies N and M in the form $N = n^{\alpha}/2$ and $M = \frac{1}{2\pi}\frac{1}{8}n^{\beta}$. More specifically, we use the values $\alpha = 1, 3/4, 2/3, 1/2, 1/3$ and $\beta = 3/4, 2/3, 1/2, 1/3, 1/4, 1/6$ to estimate the spot volatility; then, we select the pair (α,β) which minimizes the integrated relative mean squared error *IRMSE*. In the absence of noise we should choose $\alpha = 1$ and β such that $\frac{1}{2} < \beta < 1$, according to (4.6) and the numerical study in Section 4.1.2. However, in the present exercise different values of α and β are explored, in virtue of the fact that the highest frequencies must be cut in order to filter out microstructure noise effects arising from high-frequency data.

Table 5.6 Performance of the Fourier estimator in the absence of noise ($\varsigma = 0.0$) and under different microstructure noise effect models.

Noise model $(M.I)$-$(M.II)$				Noise model $(M.I)'$-$(M.II)$			
ς	(α,β)	*IRMSE*	*IBIAS*	ς	(α,β)	*IRMSE*	*IBIAS*
0.0		2.77e-4	-1.03e-3	0.0		2.77e-4	-1.03e-3
0.8	$(\frac{3}{4},\frac{1}{2})$	2.77e-3	1.31e-2	0.8	$(\frac{3}{4},\frac{1}{2})$	2.43e-3	3.12e-3
3.2	$(\frac{2}{3},\frac{1}{2})$	1.15e-2	4.87e-2	3.2	$(\frac{3}{4},\frac{1}{2})$	7.23e-3	6.10e-2
Noise model $(M.I)$-$(M.II)'$				Noise model (MR)			
ς	(α,β)	*IRMSE*	*IBIAS*	l_{α}	(α,β)	*IRMSE*	*IBIAS*
0.0		2.77e-4	-1.03e-3	0.0		2.77e-4	-1.03e-3
0.8	$(\frac{3}{4},\frac{1}{2})$	2.79e-3	1.57e-2	0.01	$(\frac{3}{4},\frac{1}{2})$	2.51e-3	1.99e-3
3.2	$(\frac{2}{3},\frac{1}{2})$	1.12e-2	4.71e-2	0.1	$(\frac{1}{2},\frac{1}{2})$	3.93e-2	3.31e-2

Table 5.6 shows the results of the simulation study. The first line of each panel contains the results in the absence of noise; in this case the cutting frequencies N and M are chosen to be equal to $n/2$ and $\frac{1}{2\pi}\frac{1}{8}\sqrt{n}\ln n$, respectively. In the sequent lines, from left to right, Table 5.6 shows the noise to signal ratio ς in the case of additive noise or the rounding level l_{α} in the case of rounding error, the pair (α,β) which minimizes the integrated relative mean squared error, the corresponding integrated relative mean squared error and bias. The results displayed in the table confirm a satisfactory performance of the Fourier estimator in the presence of both additive and rounding noises, due to its ability to filter out the noise by a suitable choice of the cutting frequencies N and M.

Remark 5.8. The global nature of Fourier estimator is shown by Table 5.6. In fact, the optimized estimator performs well over the entire time interval without being dependent on time (as well as on specific properties of the volatility process), differently from other spot volatility estimators whose defining parameters need to be tuned at each specific point in time. This property may be relevant when the estimated volatility is used to calibrate stochastic volatility models. Actually, if the value of M and N are independent of time, then the estimator $\widehat{\sigma}^2(t)$ is a continuous function of t and this can help the calibration process. For further discussion on this point see also Mancino and Recchioni (2015).

We conclude this section showing a final comparison with the Fejér Kernel-based realized spot volatility estimator (4.3), which can be erroneously identified with the Fourier estimator. The columns of Table 5.7 have the same format as those in Table 5.6 except for the pair (α,β) that is replaced by β, as the frequency N does not appear in the definition (4.3). Table 5.7 gives clear numerical evidence of the differences between the two estimators. The Fejér Kernel-based estimator provides accurate volatility estimates in the absence of noise but is highly biased in the presence of additive noises. In the case of rounding noise with low rounding level the performance of the Fejér Kernel-based estimator is satisfactory: this last finding confirms the observations made in Section 5.2.2 about the less damaging effect of rounding errors with respect to additive noises.

Table 5.7 Performance of the Fejér Kernel-based estimator (4.3) in the absence of noise and under microstructure noise effect models.

Noise model $(M.I)$-$(M.II)$				Noise model $(M.I)'$-$(M.II)$			
ς	β	*IRMSE*	*IBIAS*	ς	β	*IRMSE*	*IBIAS*
0.0	$\frac{1}{2}$	2.15e-4	-1.04e-3	0.0	$\frac{1}{2}$	2.15e-4	-1.04e-3
0.8	$\frac{1}{2}$	1.59e+0	1.26e+0	0.8	$\frac{1}{3}$	4.91e+0	2.21e+0
3.2	$\frac{1}{2}$	4.10e+2	2.02e+1	3.2	$\frac{1}{2}$	1.25e+3	3.45e+1
Noise model $(M.I)$-$(M.II)'$				**Noise model (MR)**			
ς	β	*IRMSE*	*IBIAS*	l_α	β	*IRMSE*	*IBIAS*
0.0	$\frac{1}{2}$	2.15e-4	-1.04e-3	0.0	$\frac{1}{2}$	2.15e-4	-1.04e-3
0.8	$\frac{1}{3}$	2.19e+0	1.48e+0	0.01	$\frac{1}{2}$	1.49e-2	3.21e-2
3.2	$\frac{2}{3}$	4.19e+2	2.04e+1	0.1	$\frac{1}{2}$	7.79e+0	1.51e+0

Remark 5.9. A more comprehensive study in Mancino and Recchioni (2015) shows that the Fourier estimator of spot volatility has a competitive performance on high-frequency data even in comparison to bias-adjusted estimators, such as the Two-Scales realized spot variance estimator which is proposed in Zu and Boswijk (2014) as a localization of the Two Scales estimator. The Fourier estimator satisfactorily performs in all the different scenarios illustrated in this section, without requiring any *ad hoc* adjustment.

Chapter 6
Getting Inside the Latent Volatility

This chapter introduces the reader into some recent financial applications of the Fourier estimator. We exploit here the ability of the method to reconstruct the volatility as a *stochastic function of time* in the univariate and multivariate case; in other words, we can handle the volatility function as an observable variable. This property makes it possible to have insights into various volatility related financial quantities, such as volatility of volatility and leverage. The chapter begins with an empirical exercise in which the latent volatility is estimated; we discuss in some extent the issue of the presence of jumps in the financial data. Then, in Sections 6.2 and 6.3 it is shown how to iterate the procedure for the purpose of parameter identification and calibration of stochastic volatility models and how to estimate in a model-free fashion a second order effect, known as price-volatility feedback rate. Finally, in Section 6.4 we analyze the forecasting power of the Fourier estimator of integrated volatility by a simple Monte Carlo experiment and an empirical application. Further directions for additional applications are given in Section 6.5.

6.1 Market Data Considerations

Since this chapter claims to apply the Fourier estimator on empirical data with the aim of testing the method by means of real financial applications, some considerations are needed whether the Fourier estimator should be used in an unadulterated way on real data.

As a matter of fact, both microstructure noise contaminations and jumps cannot be ignored when using high frequency data. In fact, when considering empirical data, the two issues of microstructure and jumps are strictly related and it may be difficult to distinguish between two unobservable components of observed data: large levels of noise and small jumps in efficient price (see, e.g., Lee and Mykland (2012)). Jumps have already been considered in Section 4.3, where we discussed the jump-robust Fourier estimator (4.14). However, this estimator is not robust to

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3_6

microstructure effects. At the same time, Chapter 5 showed the robustness of the (original) Fourier estimator (3.5) with respect to noise contaminations. In this section, we reconcile this dichotomy and supplement the analysis showing how the Fourier method could be easily and efficiently implemented with empirical data.

Firstly, to detect the presence of jumps in the considered sample, we employ the test statistics[1] proposed by Corsi et al. (2015). This test is based on the relative difference between the Realized Volatility and the Threshold Bipower Variation (hereafter, denoted by TBV). The latter is defined by

$$TBV := \mu_1^{-2} \sum_{j=1}^{n-1} |\delta_{j-1}(p)||\delta_j(p)| \; 1_{\{|\delta_{j-1}(p)|^2 \leq \vartheta_{j-1}\}} 1_{\{|\delta_j(p)|^2 \leq \vartheta_j\}},$$

where $\mu_1 \simeq 0.7979$ and ϑ_j is a suitable threshold vanishing with a slower rate than the modulus of continuity of the Brownian motion (see (A.2)) as $n \to \infty$. TBV provides a consistent estimator of the integrated variance in the presence of jumps. So, the difference between Realized Volatility and TBV affords a nonparametric estimator of the contribution to total price variation coming from the jump component.

After having detected the jumps in the sample, a simple *jump-robust* Fourier estimator denoted as FE_i^J can be defined as follows: if no jump has occurred on day i, the Fourier estimator $\widehat{\sigma}^2_{n,N}$ defined by (3.5) is employed, while if jumps have occurred on day i, we numerically integrate the jump-robust Fourier estimator of spot volatility obtained in Section 4.3. As a benchmark of our analysis we consider the estimator RV_i^J defined with the same approach as follows: if no jump has occurred on day i, the Realized Volatility is employed, while if jumps have occurred on day i, the jump-robust TBV estimator is considered.[2]

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. Only the prices of the next-to-expiration contracts are used, which are the most liquid ones, with the FIB30 expiring quarterly. We have a total of 1514523 quotes over the period and on average a new quote arrives every 5.67 seconds. Table 6.1 describes the main features of our data set.

Table 6.1 Summary statistics for the sample of the FIB30 index futures over the period January 11, 2000 to January 31, 2001 (1514523 trades). "Std. Dev." denotes the sample standard deviation of the variable.

Variable	Mean	Std. Dev.	Min	Max
FIB30	46182.98	2090.73	39182.50	51330.00
log-return (%)	7.9723e-6	2.6890e-2	-3.3731e+0	1.5150e+0

[1] In all the empirical experiments only one of the possible procedures to test for the presence of jumps in the data is considered, as the study of jumps goes beyond the scope of the present book.

[2] We use this definition of volatility measure because TBV is less efficient than the Realized Volatility if no jumps occur.

Figure 6.1 shows the time plot and the ACF of the tick-by-tick log-returns (upper panels) and of the 5 minute returns. Quotes prior to 9:30 a.m. are removed to eliminate opening quotes from our sample. We construct sparse sampled intraday returns using a sort of *tick time sampling* scheme, where the t_j's are chosen to be the time of the first transaction occurring a given period, say 5 minutes, after the previous one. We also exclude all overnight returns. Sparse sampling seems to eliminate a large portion of microstructure effects, although some statistically significant correlation in returns remains at lags larger than one.

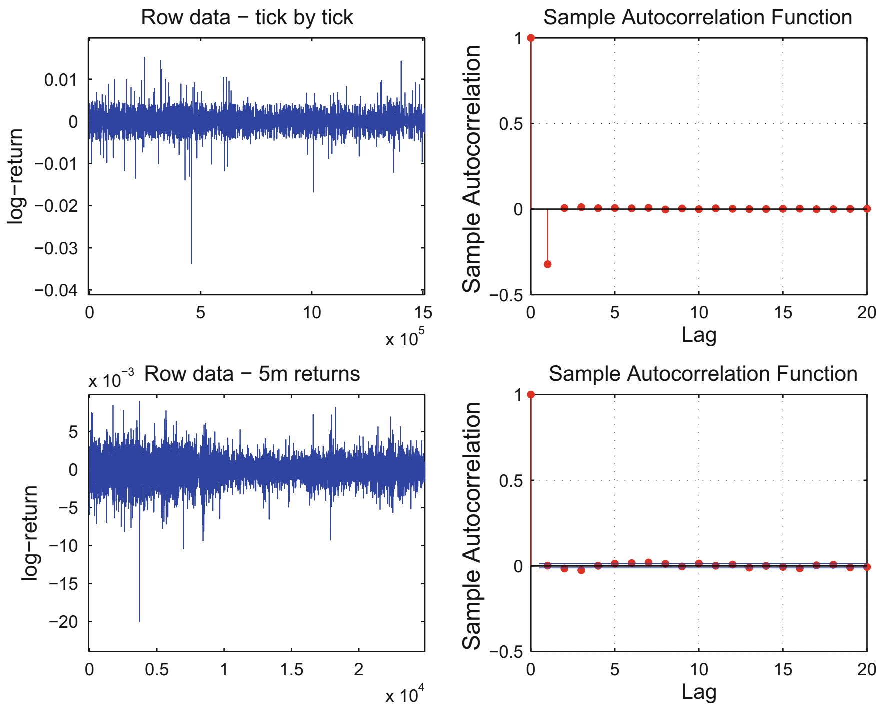

Fig. 6.1 Time plot of the log-returns and the ACF for FIB30 over the period January 11, 2000 to January 31, 2001.

The daily integrated volatility is estimated on each day of the sample by means of the Fourier and the Realized Volatility estimators. Since RV and TBV are not robust to microstructure effects, we construct both estimates from sparse sampled 5 minute returns. The Fourier estimator is instead built using all data and choosing a cutting frequency N that minimizes the estimated MSE as explained in Section 5.2.1.

The presence of jumps in our sample is revealed by the fact that the distribution of daily log-returns rescaled by the estimated daily volatility is not Gaussian. We apply the Jarque-Bera (JB) test at the 5% significance level to test the null hypothesis that the sample comes from a normal distribution with unknown mean and variance, against the alternative that it does not come from a normal distribution, but the test fails. To identify jumps in our sample, the jump-test by Corsi et al. (2015) is employed at the significance level of 99.9%. This allows us to identify a percentage of days with jumps around 12.27%.

Table 6.2 summarizes the unconditional distribution of the daily log-returns standardized by the integrated volatility estimated by the Fourier estimator $\hat{\sigma}^2_{n,N}$, the RV estimator and the jump robust versions of these estimators FE^J_i and RV^J_i, respectively, over the whole sample. As expected, the Fourier and RV standardized daily log-returns are slightly skewed and RV exhibits a standard deviation larger than one. Rescaling the daily returns by jump robust measures of integrated volatility, the return distributions are closer to the standard normal. In particular, the distribution obtained by the jump robust Fourier estimator FE^J_i is the closest to $\mathscr{N}(0,1)$.

Table 6.2 Standardized daily log-return distributions for the FIB30 index futures over the period January 11, 2000 to January 31, 2001 (whole sample, 269 days).

Volatility estimator	Mean	Std. Dev.	Skewness	Kurtosis
$\hat{\sigma}^2_{n,N}$	0.047418	1.077545	0.101114	2.761774
RV	0.061317	1.136009	0.164536	2.776498
Jump-robust FE^J_i	0.032659	1.095342	-0.025122	3.097388
Jump-robust RV^J_i	0.057457	1.123815	0.079632	2.829964

The following conclusion can be drawn out of this short analysis. When data contain jumps that, combined with other market microstructure frictions, contribute to render the basic assumption of a continuous semimartingale price invalid at the tick-by-tick level, then the jump robust measures of integrated volatility entail distributions that are closer to the standard normal. However, the Fourier estimator $\widehat{\sigma}^2_{n,N}$ seems to maintain its efficiency. This relevant issue should deserve further analysis.

6.2 Factor Identification for Stochastic Volatility Models

We consider a fairly general class of stochastic volatility models in continuous time, including classical models such as Hull and White (1987), Stein and Stein (1991). The log-price and variance processes satisfy

$$\begin{aligned} dp(t) &= \sigma(t)dW(t) + a(t)dt\,, \\ d\sigma^2(t) &= \gamma(t)dZ(t) + b(t)dt\,, \end{aligned} \tag{6.1}$$

where W and Z are correlated Brownian motions with instantaneous correlation $\rho(t)$ and $\sigma(t)$, $\gamma(t)$, $a(t)$, $b(t)$ are adapted random processes satisfying mild integrability conditions (see Barucci and Mancino (2010) for details). In this model, jumps in the price and in the volatility are ruled out, however, extensions in this direction can be obtained using the method presented in Section 4.3.

In Chapter 4 the Fourier estimation method has been efficiently used to compute pathwise the diffusion coefficient in (6.1), i.e. $\sigma^2(t)$. Here, we proceed a step further and we apply the Fourier approach in order to obtain accurate estimates of the volatility of the variance process $\gamma(t)$ and of the covariance between the price and the instantaneous variance, also in the case when the data are contaminated by microstructure noise.[3]

The intuition is the following. The knowledge of all Fourier coefficients $\mathscr{F}(\sigma^2)(k)$ of latent instantaneous variance process $\sigma^2(t)$ allows us to iterate the main convolution formula from Theorem 2.1, in its univariate version, in order to compute the *volatility of the volatility* process; analogously, using the multivariate expression, it is possible to compute the *leverage*, that is the covariance between the price and the variance. More precisely, the convolution formula applied to the variance process is specified through the following limit in probability

$$\mathscr{F}(\gamma^2)(k) = \lim_{M\to\infty} \frac{2\pi}{2M+1} \sum_{|s|\le M} \mathscr{F}(d\sigma^2)(s)\,\mathscr{F}(d\sigma^2)(k-s), \tag{6.2}$$

where we can use the integration by parts formula to write the Fourier coefficients of $d\sigma^2$, that is, for any integer s,

$$\mathscr{F}(d\sigma^2)(s) = \mathrm{i}s\,\mathscr{F}(\sigma^2)(s) + \frac{1}{2\pi}(\sigma^2(2\pi) - \sigma^2(0)). \tag{6.3}$$

Similarly, the following result, relating the Fourier coefficients of returns and variance process to the Fourier coefficients of the leverage process $\eta(t)$, holds in probability

$$\mathscr{F}(\eta)(k) = \lim_{M\to\infty} \frac{2\pi}{2M+1} \sum_{|s|\le M} \mathscr{F}(d\sigma^2)(s)\,\mathscr{F}(dp)(k-s). \tag{6.4}$$

6.2.1 *Volatility of Volatility*

In this section, we focus on the estimation of the *integrated stochastic volatility of volatility* using high-frequency data. Given the estimated Fourier coefficients of the volatility process (3.4) and relying on (6.2), the Fourier estimator of the second order quantity $\int_0^{2\pi} \gamma^2(t)dt$ is defined as follows[4]

$$\widehat{\gamma}^2_{n,N,M} := \frac{(2\pi)^2}{M+1} \sum_{|s|\le M} \left(1 - \frac{|s|}{M}\right) s^2\, c_s(\sigma^2_{n,N})c_{-s}(\sigma^2_{n,N}). \tag{6.5}$$

[3] An early attempt to use the Fourier method to identify the parameters of stochastic volatility models is present in Malliavin and Mancino (2002b), Barucci and Mancino (2010), Renò (2008), while a deep study is done in Sanfelici et al. (2015), Curato and Sanfelici (2015).

[4] In the convolution (6.5) a Barlett kernel has been added, which improves the behavior of the estimator for very high observation frequencies.

Notice that in order to define the estimator (6.5), the integration by parts formula (6.3) has been replaced with the following approximation

$$c_s(d\sigma^2_{n,N}) \cong \mathrm{i}s\, c_s(\sigma^2_{n,N}). \tag{6.6}$$

We highlight that, according to definition (6.5), the computation of the volatility of volatility needs only to pre-estimate the Fourier coefficients of the volatility process from the asset returns and does not require a preliminary estimation of the instantaneous volatility. In this respect, the Fourier estimator of the volatility of volatility is notably different from the other proposed estimators, which first estimate the volatility path using some consistent estimate of the instantaneous volatility and then estimate the volatility of volatility using the estimated volatility process as a proxy of the unknown paths. The rationale is that the reconstructed (estimated) path of the volatility is plugged into an estimator of integrated volatility, e.g., the Realized Volatility (see, for instance, Barndorff-Nielsen and Veraart (2013), Vetter (2015)). Therefore, a large number of observations for the price process is necessary, as it is statistically clear that the integrated variance of the volatility process can be estimated only on a larger time scale than the one used for estimating the volatility path from the observed prices. This yields a huge loss of information contained in the original dataset. On the other side, it is well known that spot volatility estimation is quite unstable, especially in the presence of microstructure effects as it happens with high-frequency data. This point is summarized in the following box.

The Fourier estimator reconstructs the volatility of volatility using as input the Fourier coefficients of the volatility and, ultimately, the Fourier coefficients of the observable log-returns. In other words, it uses only integrated quantities from the whole available dataset. This fact renders the method easily implementable and computationally stable.

In the context of the stochastic volatility model (6.1), the estimator (6.5) is consistent in probability and asymptotical unbiased in the presence of MA(1) microstructure noise, under the conditions that $N = n^{\alpha}$ $(0 < \alpha < 1/2)$ and $M = n^{\beta}$ $(0 < \beta < \alpha/4)$, being n the sample of the underlying price process. The proof can be found in Sanfelici et al. (2015). The rate of convergence of the estimator is still to be derived.

Remark 6.1. It should be clear that, based on (6.2) and approximation (6.6), a Fourier estimator of the *instantaneous volatility of volatility* can be defined by the same procedure adopted in the definition (4.1): in the first step estimate the k-th Fourier coefficients of the volatility of volatility; in the second step, use the Fourier-Fejér inversion formula to write the estimator of the spot quantity. Finally, we remark that the method proposed can be extended without any conceptual difficulties to the multidimensional setting.

6.2.2 Leverage

The so called *leverage effect* refers to the relationship between returns and the corresponding volatility which tend to be negatively correlated. One possible economic interpretation of this phenomenon was developed by Black (1976) and Christie (1982) and is connected with the concept of financial leverage (debt-to-equity ratio). As asset prices decline, companies become automatically more leveraged since the relative value of their debts rises relative to that of their equities. The probability of default rises and then their stocks become riskier, inducing a higher volatility. From a mathematical point of view, the no leverage hypothesis means that the process $\sigma(t)$ is independent from the Brownian motion W in model (6.1). This hypothesis simplifies the study of the properties of the volatility estimator, but is not realistic for the analysis of equity returns.

It appears evident from (6.4) that in the context of the Fourier approach the definition of an estimator of the leverage process can be based on the same approximation (6.6) as used for the volatility of volatility. Therefore, we define the Fourier estimator of the *integrated leverage* $\int_0^{2\pi} \eta(t)dt$ by

$$\widehat{\eta}_{n,N,M} := \frac{(2\pi)^2}{2M+1} \sum_{|s|\le M} \mathrm{i}\, s\, c_s(\sigma^2_{n,N}) c_{-s}(dp_n). \tag{6.7}$$

Similar to the volatility of volatility Fourier estimator, even the definition (6.7) does not require the preliminary estimation of the instantaneous volatility path, but only the estimated Fourier coefficients of the volatility.

Curato and Sanfelici (2015) prove that the estimator is consistent in probability under suitable growth conditions between the parameters, precisely if $N/n \to c > 0$ and $M^2/N \to 0$, in the absence of microstructure noise.[5] The rate of convergence of the estimator is currently under investigation. The authors also perform an extensive simulation study of its efficiency and show that, in finite sample, the Fourier estimator of the leverage effect turns out to be accurate in the presence of non-equidistant observations of the price process and microstructure noise contaminations. In the presence of noise, the sufficient bandwidth conditions for the asymptotic unbiasedness under MA(1) model become $N/n \to 0$ and $M^2/N \to 0$.

Remark 6.2. The same argument illustrated in Remark 6.1 applies to the leverage estimator. Thus, a *Fourier estimator of spot leverage* can be defined as follows

$$\widehat{\eta}_{n,N,M,L}(t) := \sum_{|k|\le L} \left(1 - \frac{|k|}{L}\right) e^{\mathrm{i}tk} c_k(\eta_{n,N,M}), \tag{6.8}$$

where

$$c_k(\eta_{n,N,M}) := \frac{2\pi}{2M+1} \sum_{|s|\le M} \mathrm{i}\, s\, c_s(\sigma^2_{n,N}) c_{k-s}(dp_n).$$

[5] The proof by Curato and Sanfelici (2015) require the differentiability in the sense of distributions of the drift and diffusion coefficients. We refer the interested reader to the paper for further details.

6.2.3 Empirical Analysis

In order to illustrate the efficiency of the estimators (6.5) and (6.7), we consider a case study based on 5-second returns of the S&P 500 index recorded at the Chicago Mercantile Exchange (CME) on March 4th, 2013. The sample contains 4921 observations. Table 6.3 describes the main features of our data set. High frequency returns

Table 6.3 Summary statistics for the sample of the traded CME S&P 500 index on March 4, 2013 (4921 trades). "Std. Dev." denotes the sample standard deviation of the variable.

Variable	Mean	Std. Dev.	Min	Max
S&P 500 index	1518.43	3.56	1512.29	1525.27
log-return (%)	9.35e-5	4.84e-3	-1.36e-1	5.19e-2

are contaminated by transaction costs, bid-and-ask bounce effects, etc., leading to biases in the variance measures. Figure 6.2 shows the time plot of the log-returns and the autocorrelation function for the row 5-second data (upper panels) and for the 5-minute aggregated data (lower panel). Row data exhibit a strongly significant positive first order autocorrelation and higher order autocorrelations remain significant up to lag 11. Sparse sampling at 5 minute frequency makes the data free from microstructure effects.

As a benchmark for the Fourier estimator of volatility of volatility (resp. of leverage), we use the *Pre-estimated Spot variance-based Realized Variance* of Barndorff-Nielsen and Veraart (2013), which we call *PSRV* (resp. the *Pre-estimated Spot variance-based Realized Leverage* of Mykland and Zhang (2009), Barndorff-Nielsen and Veraart (2013), which we call *PSRL*). These estimators are consistent in the absence of microstructure frictions. For the reader's convenience, we recall their construction. Hypothetically, let us assume that the variance process σ^2 is observed at equally spaced times $\{j\Delta_n : j = 0, 1, 2, \dots, \lfloor T/\Delta_n \rfloor\}$, for some $\Delta_n > 0$ such that $\Delta_n \to 0$, as $n \to \infty$. For any function f, denote $\Delta_j f := f(j\Delta_n) - f((j-1)\Delta_n)$.

The *PSRV* estimator is then defined as the sum of squared increments over the time interval $[0, T]$

$$PSRV_n := \sum_{j=1}^{\lfloor T/\Delta_n \rfloor} (\Delta_j(\sigma^2))^2. \tag{6.9}$$

However, since volatility is unobservable, we have to replace the variance process $\sigma^2(t)$ with a consistent spot variance estimator, such as the locally averaged realized variance defined for any $t \in (0, T)$ by

$$\widehat{\sigma}^2(t) := \frac{1}{K_n \delta_n} \sum_{j=\lfloor t/\delta_n \rfloor - K_n/2}^{\lfloor t/\delta_n \rfloor + K_n/2} (\delta_j(p))^2, \tag{6.10}$$

where $\delta_j(p) := p(j\delta_n) - p((j-1)\delta_n)$ is the j-th log-return computed on a different time scale at which we observe the logarithmic asset price p, with mesh size $\delta_n > 0$.

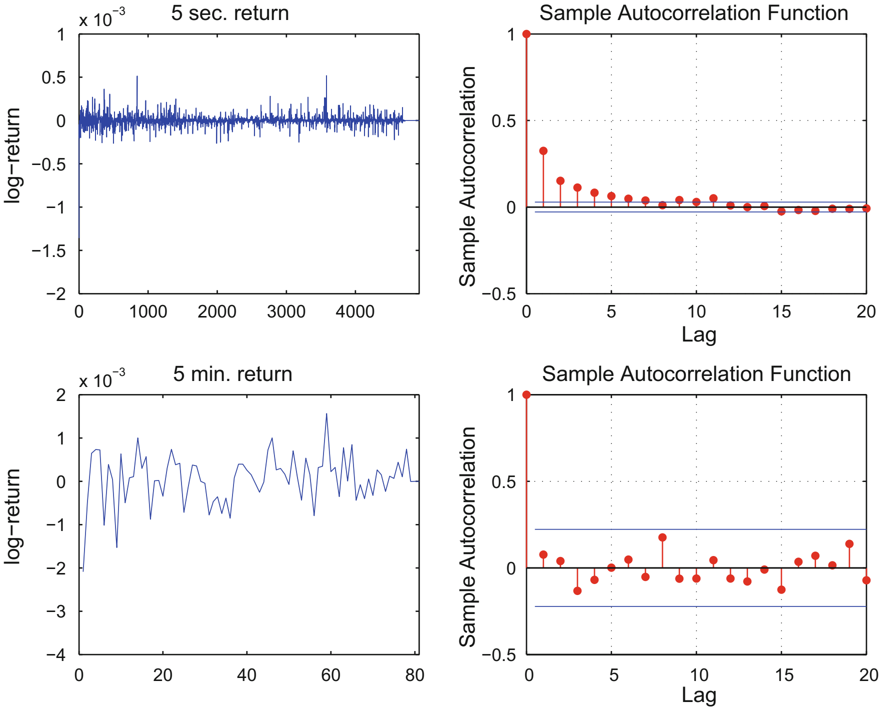

Fig. 6.2 Time plot of the tick-by-tick log-returns and ACF for S&P 500 index on March 4, 2013.

This estimator is constructed over a local window of size $K_n\delta_n$, where we require $K_n \to \infty$ such that $K_n\delta_n \to 0$. The spot volatility is estimated on a finer time scale than the one used in formula (6.9). Then, we must assume $\Delta_n > \delta_n$.

Similarly, the *PSRL* estimator is defined as the bias corrected sum of products of the estimated spot variance increments times the log-returns over the time interval $[0,T]$

$$PSRL_n := 2 \sum_{j=1}^{\lfloor T/\Delta_n \rfloor} \Delta_j(\hat{\sigma}^2)\,\Delta_j(p). \tag{6.11}$$

Remark 6.3. The condition $\Delta_n > \delta_n$ imposed on the choice of the two time scales for the estimators *PSRV* and *PSRL* represents a limit for the efficiency of such procedures. On one side, it requires using huge datasets of high-frequency returns, where market microstructure effects likely become manifest. On the other side, the choice of the second level time scale Δ_n implies a loss of the information contained in the original time series. The *PSRV* and *PSRL* estimators are not robust to microstructure noise. Therefore, in the presence of microstructure effects, the choice of the sampling interval δ_n is conditioned by the well known bias-variance trade off. However, we remark that sparse sampling may produce a loss of the rich information contained in the original high-frequency dataset.

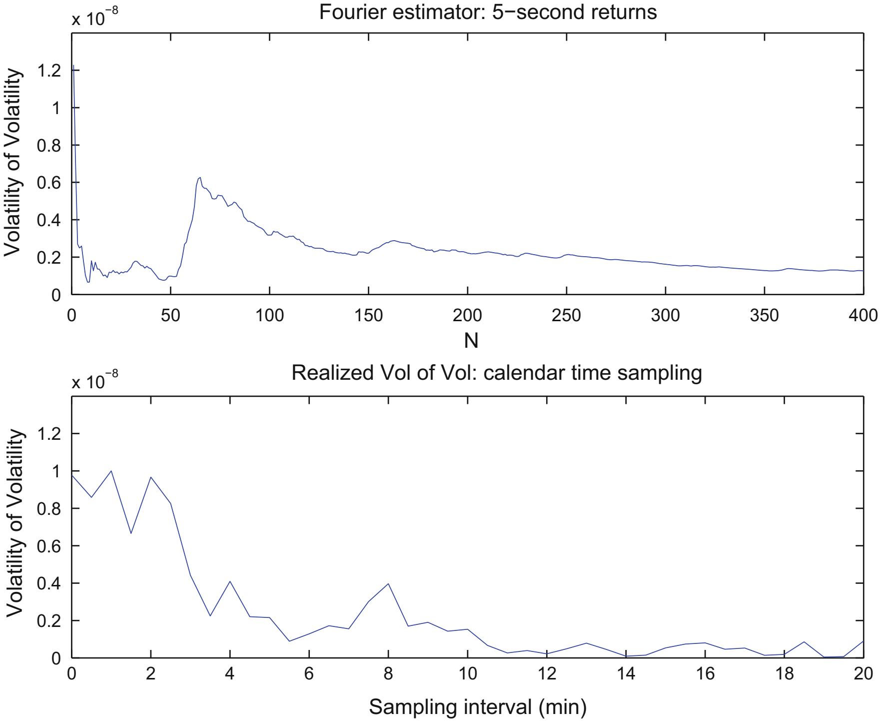

Fig. 6.3 S&P 500 index on March 4, 2013. Integrated volatility of volatility estimated by the Fourier method using 5-second returns (upper panel) and by the *PSRV* estimator (lower panel) as a function of the highest Fourier frequency N employed and of the sampling interval δ_n, respectively.

Figure 6.3 shows the integrated volatility of volatility estimated by the Fourier method (upper panel) and by the *PSRV* (lower panel), as a function of the highest frequency N of return coefficients employed and of the sampling interval δ_n, respectively. The volatility signature plots clearly indicate that the bias induced by market microstructure effects dies for sampling frequencies larger than 10 minutes. The impact of market microstructure effects on the 10-minute Realized Volatility measure for the S&P 500 index on this day can therefore be regarded as negligible. However, for low frequencies the *PSRV* estimator becomes downwards biased because sparse sampling has a severe impact on the cardinality of the database. In particular, for any value of the sample size n_S the *PSRV* is estimated at the frequency Δ_n corresponding to $K_n/2$ ticks, where the parameter K_n is chosen as $K_n = 2\sqrt{n_S}$. This implies that most of the data are neglected when estimating the second order quantities so that the volatility of volatility estimates are poor, especially when we start from sparse sampled data i.e. $n_S \ll n$. In the case of the Fourier estimator $\widehat{\gamma}^2_{n,N,M}$ defined in (6.5), the value of the parameter M is set to 3, according with the growth conditions ensuring both the consistency of the Fourier estimator and its asymptotically unbiasedness in the presence of microstructure noise, that is $N = O(n^{\alpha})$ and

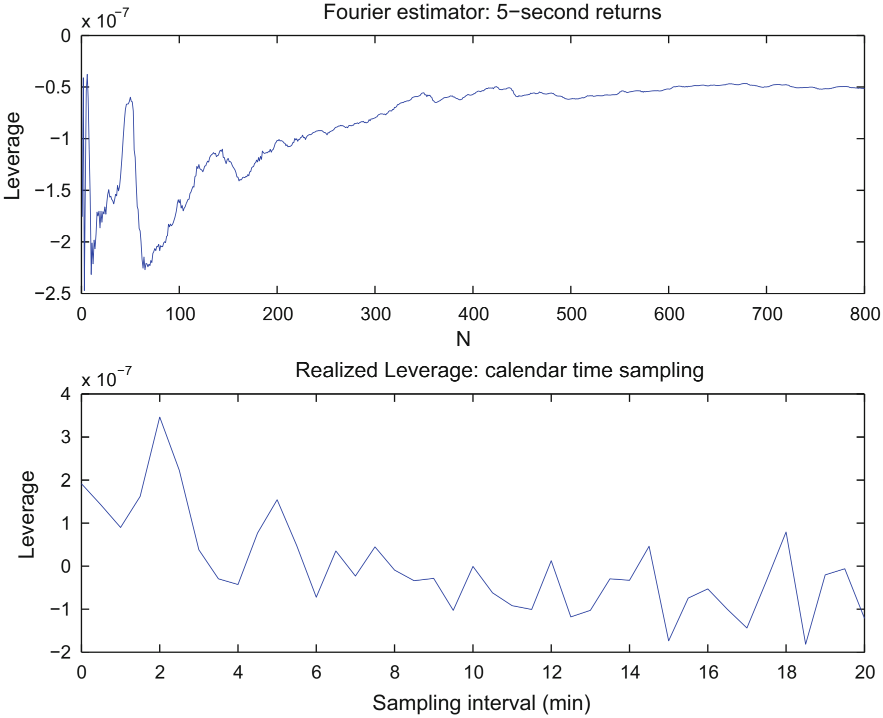

Fig. 6.4 S&P 500 index on March 4, 2013. Integrated leverage estimated by the Fourier method using 5-second returns (upper panel) and by the *PSRL* estimator (lower panel) as a function of the highest Fourier frequency N employed and of the sampling interval δ_n, respectively.

$M = O(n^{\beta})$ with $0 < \alpha < \frac{1}{2}$ and $0 < \beta < \frac{\alpha}{4}$. We remark that the Fourier estimator makes use of all the n observed prices, because it reconstructs the signal in the frequency domain and therefore it can filter out microstructure effects by a suitable choice of M and N instead of reducing the sampling frequency. We can see that for N larger than 250 the estimates become much stable.

Figure 6.4 shows the daily integrated leverage estimated by the Fourier method and by the *PSRL* estimator as a function of the highest frequency N of return coefficients employed and of the sampling interval δ_n, respectively. When estimating the leverage effect, a larger variability in the estimates can be observed if compared to other quantities such as volatility or quarticity. The cutting parameter M in (6.7) is set equal to 2, on the basis of the theoretical and empirical results obtained in Curato and Sanfelici (2015), while $K_n = 2\sqrt{n_S}$, where n_S is the total number of sparse sampled returns. For instance, when sampling returns at 10-minute frequency, the sample size is $n_S = 41$. From the plot, it is evident that the Fourier methodology provides quite stable estimates at cut-off frequencies N that turn out to be much smaller than the Nyquist frequency (i.e. $N \ll n/2 = 2460$), whereas the *PSRL* estimator is quite unstable as the estimates strongly depend on the sampling frequency.

6.3 Volatility Feedback Effects

Feedback effects of assets prices on volatilities are widely recognized in the financial market literature, both theoretically and empirically. Volatility feedback and leverage effects are related to the same phenomena: the leverage effect explains why a negative return causes an increase in the volatility; on the contrary, the notion of volatility feedback effect is based on the argument that volatility is priced and an increase in the volatility requires a higher rate of return from the asset, which can only be produced by a decline in the asset price as observed by French et al. (1987). Moreover, in Bekaert and Wu (1997) the interaction of these two effects is analyzed and used to explain the irregular behavior of volatility causing instability in financial markets such as excess volatility, volatility persistence and volatility smile.

The volatility feedback rate is a second order quantity which is supposed to describe the facility for the market to absorb small perturbations; therefore, it is related to the change through time in the rescaled variation of the stochastic process describing the asset price.

For simplicity, consider the univariate case and assume that the drift coefficient is zero[6] (i.e., we work with the discounted asset price under the risk neutral measure). Suppose the log-asset price $p(t)$ satisfies the model

$$dp(t) = \sigma(p(t))\, dW(t) - \frac{1}{2}\sigma^2(p(t))dt, \tag{6.12}$$

where the function σ is positive and twice differentiable. For shortness, denote $\sigma(t) := \sigma(p(t))$.

Remark 6.4. The volatility is modeled in (6.12) as a level dependent quantity (i.e. an unknown time independent function of asset price). There are different motivations for this type of dependence. First of all, this way to model asset price volatility is well suited to capture the relationship between volatility and asset price-return. The simplest way to model the negative relation between asset price-return and volatility is to assume a constant elasticity variance model, see Cox and Ross (1976). Level dependent volatility has been also conjectured to reproduce the implied volatility smile, see Derman and Kani (1994), Dupire (1994), Hobson and Rogers (1998). A different perspective to introduce level dependent volatility is to build a model with heterogeneous agents (e.g. fundamentalist, rational, portfolio insurance traders) deriving the asset price process in equilibrium, see Frey and Stremme (1997), Platen and Schweizer (1998).

[6] The extension to a finite number of assets can be found in Barucci et al. (2003). Further, the computation with non-zero drift is in Malliavin and Thalmaier (2006) Chapter 3.

Consider an infinitesimal perturbation $p(t)+\varepsilon\zeta(t)$ of the asset price. The pathwise sensitivity $\zeta(t)$ (i.e., *variation process*, see Kunita (1988)) is defined as the solution of the linearized stochastic differential equation[7]

$$d\zeta(t) = \zeta(t)\,(\sigma'(t)\,dW(t) - \sigma'(t)\sigma(t)dt\,).$$

We associate to $\zeta(t)$ the *rescaled variation* defined as

$$z(t) := \frac{\zeta(t)}{\sigma(t)}.$$

By applying Itô formula, in Malliavin and Mancino (2002b), Barucci et al. (2003) it is proved that the rescaled variation is a differentiable function with respect to t and, for any $s<t$, it can be expressed as

$$z(t) = \exp\Big(\int_s^t \lambda(\tau)\,d\tau\Big)\,z(s), \tag{6.13}$$

where

$$\lambda(t) := -\frac{1}{2}(\sigma(t)\sigma'(t)+\sigma(t)\sigma''(t)). \tag{6.14}$$

The function $\lambda(t)$ is called the *price-volatility feedback rate*.

Remark 6.5. In particular, in the Black-Scholes framework it holds $\lambda = 0$.

The price-volatility feedback rate can be understood as the appreciation rate of the rescaled variation. In this respect, large positive values of λ indicate market instability, while negative values corresponds to stable market directions:

- a negative λ would witness a period of stability, because[8] $z(t)\to 0$ as $t\to+\infty$;
- a positive λ would signal instability.

Furthermore, large positive values of the feedback rate usually anticipate a significant decrease in the price level and values of λ around zero imply that the price level is close to the equilibrium level (see also the empirical analysis in Inkaya and Yolcu Ocur (2014)). Thus, it would be important to estimate the volatility feedback rate without assuming the knowledge of an explicit expression of the volatility function.

> It is possible to obtain a non-parametric estimation of the feedback rate $\lambda(t)$ from the observation of a single price path using the Fourier method.

Define the following processes:

[7] The use of prime stands here for the first derivative with respect to the level $p(t)$.

[8] The proof is in Malliavin and Thalmaier (2006).

$$A(t) := \frac{d\langle p,p\rangle_t}{dt}, \;\; B(t) := \frac{d\langle A,p\rangle_t}{dt}, \;\; C(t) := \frac{d\langle B,p\rangle_t}{dt},$$

where $\langle x,y\rangle$ is the quadratic covariation defined in (2.18). Using Itô formula and (6.14), the feedback effect rate function λ can be expressed as

$$\lambda(t) = \frac{3}{8}\frac{B^2(t)}{A^3(t)} - \frac{1}{4}\frac{B(t)}{A(t)} - \frac{1}{4}\frac{C(t)}{A^2(t)}. \tag{6.15}$$

Note that the process $A(t)$ is the spot volatility, the process $B(t)$ is the spot leverage, while $C(t)$ is the quadratic variation process between the leverage and the asset price. Therefore, all the processes $A(t), B(t)$ and $C(t)$ can be obtained from the asset prices data through three subsequent iterations of the Fourier cross-volatility estimation procedure. More precisely, the estimator of $A(t)$ is given by (4.1), $B(t)$ by (6.8) and for $C(t)$ the estimator is obtained with the same procedure as for $B(t)$, by using the Fourier coefficients of $B(t)$ computed in the second step. It follows that the feedback rate can be empirically estimated from a single path of the asset price.

6.3.1 An Empirical Application: Market Stability

This section investigates whether the estimate of the feedback effect rate, λ, using high-frequency data observed in just one trading day gives insights into the market stability's assessment. Precise estimation of quadratic and higher order variation requires huge quantities of data. This is the reason why high-frequency data and the Fourier methodology are natural candidates to cope with this difficulty.

The data set consists of the 5-second observations of the EUR/GBP exchange rate on three different dates (September 30, 2008, May 21 and August 17, 2015) which represent different macroeconomic conditions in the financial scenario. The first date is close to the Lehman Brothers bankruptcy on September 15, 2008. May 21, 2015 follows the instability raised by the United Kingdom election on the 7, May. Finally, August 17, 2015, is chosen due to the calm climate in the financial market. One trading day consists of 24 hours with a total of 17280 observations (i.e. one observation every 5-seconds).

A preliminary analysis through the autocorrelation function plot of the 5-second returns is conducted in Figure 6.5 to detect the presence of market microstructure noise. In fact, a significant first-order autocorrelation in the high-frequency returns reveals the presence of noise. Figure 6.5 shows the 5-second returns on September 30, 2008 (left panel), on May 21, 2015 (middle panel) and on August 17, 2015 (right panel). The most noisy returns are those observed on September 30, 2008, while the microstructure present in the returns on the other two days behaves similarly.

The Fourier method has been iterated three times in order to compute the feedback effect rate λ according to (6.15). In the Fourier expansion of the volatility the frequencies N and M have been selected following the analysis in Chapter 5 in order to avoid the bias generated by the microstructure noise effect, i.e., we set $N = n^{2/3}$

and $M = n^{1/2}$, where $n = 17280$. This choice yields $N = 667$ and $M = 131$ for the estimation of $A(t)$. In the subsequent iterations, the two cutting frequencies are chosen dividing by two the values obtained in the previous step, e.g., for $B(t)$ the bandwidths are 333 and 65. This provides a reasonable approximation of the path $C(t)$ at a time scale of 10 minutes.

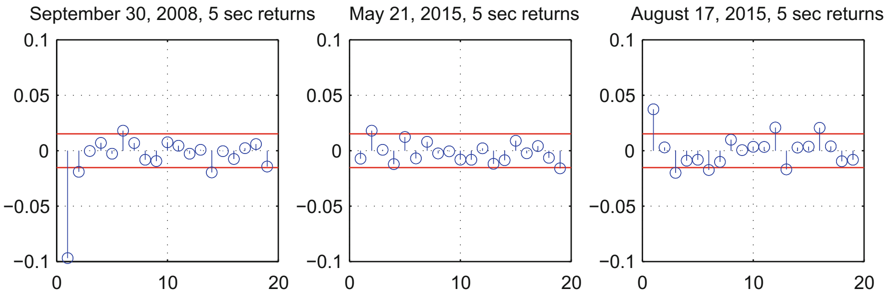

Fig. 6.5 Autocorrelation function plots of log-returns

Figure 6.6 compares the EUR/GBP exchange rate and the functions A, B, C and λ obtained on September 30, 2008 (left panels), May 21, 2015 (middle panels) and August 17, 2015 (right panels). The values of λ on September 30, 2008 are often large and positive despite the fact that the observed values of the volatility A (left panel second row) on the same date are smaller than those observed on May 21, 2015 (middle panel second row). Comparing the upper and lower left panels we can see that the highest positive values of the feedback rate observed on September 30, 2008 anticipate the fall in the exchange rate. This finding coheres with the interpretation of the sign of the feedback rate given by the theory.

An accurate look at the left panel in the second row of Figure 6.6 reveals that the volatility A displays a three-peaks structure with peaks corresponding to the opening of the major markets (Asia, Europe, North America). The most interesting finding is that, although we expect a higher market instability on September 30, 2008 than on May 21, 2015 or August 17, 2015, there is no evidence of this instability when looking at the values of EUR/GBP exchange rate or at the values of the volatility. Only by looking at the values of the feedback effect rate is it possible to detect the market instability. Roughly speaking, the feedback effect rate measures the sentiment of the market on the future behavior of the observed price dynamics.

Let us further analyze the EUR/GBP exchange rate by comparing the values of A, B, C and λ on May, 21, 2015 and August 17, 2015. The lower right panel of Figure 6.6 shows that the values of the feedback rate λ on August 17, 2015 oscillate around zero (often being negative) for almost all the day. This stabilization of λ around zero combined with very low values of the volatility A indicates that the price level is close to the equilibrium level. On May 21, 2015 the feedback rate shows strong fluctuations in the morning with a large positive peak which is followed by a fall in the exchange rate. Later, λ oscillates around the origin assuming also negative values which suggests that the market is coming back to a calm climate. During the

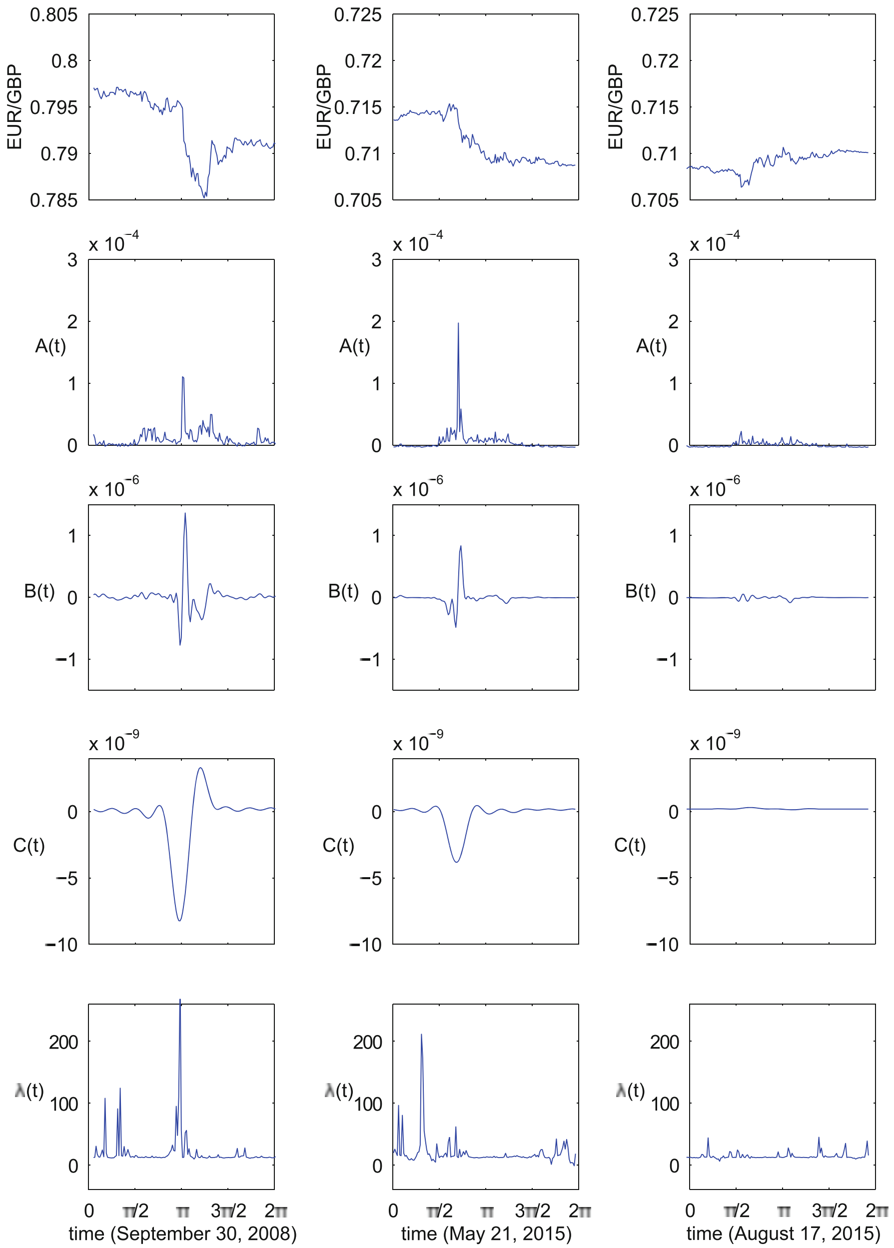

Fig. 6.6 EUR/GBP exchange rate and estimates of A, B, C, λ on September 30, 2008 (left panels), May 21, 2015 (middle panels) and August 17, 2015 (right panels). The time window $[0,2\pi]$ corresponds to one trading day (24 hours). The same y-scale is used in the left, middle and right panels, with the only exception for the panels showing the EUR/GBP exchange rate where a different y-scale is used in order to make the changes of EUR/GBP exchange rate dynamics more visible.

first two hours of trading A has about the same shape every day, while λ develops dramatically different shapes.

6.4 Volatility Forecasting Performance

Risk and asset management practices as well as derivative pricing rely upon a precise measure/forecast of volatility and covariance. By treating volatility as observed rather than latent, non parametric estimation methods improve forecasting performance using simple methods directly based on observable variables. Empirical analysis have shown that the forecasting performance of the Realized Volatility is superior to that of classical ARCH models (see, e.g., Andersen et al. (2003)). However, for high-frequency observations, the Realized Volatility is no longer a reliable measure, hence the need of exploring the forecasting performance of methods designed to handle market microstructure contaminations, such as the Fourier estimator.

Given a measure of the integrated volatility in the period $[t-1,t]$ obtained through the Realized Volatility or the Fourier methodology, we intend to evaluate its capability of forecasting the integrated volatility on day $[t,t+1]$: to this end the linear regression of one day ahead integrated volatility over today estimated volatility is considered. In this setting the forecasting performance can be evaluated through the R^2 of the linear regression.

Suppose that the logarithm of the observed asset price $\widetilde{p}(t)$ follows model (5.1), where the efficient log-price satisfies the stochastic differential equation (3.1). Integrability conditions stated in (3.2) hold.

Denote by $IV(t)$ the integrated volatility over one period, say $[t-1,t]$. Given n equally spaced observations in $[t-1,t]$, we denote by $\widehat{IV}(t)$ (resp. $\widetilde{IV}(t)$) the integrated volatility over the interval $[t-1,t]$ estimated by either the Realized Volatility or the Fourier estimator in the absence of microstructure noise effects (resp. in the presence of microstructure noise). Assuming that the spot volatility model belongs to the *Eigenfunction Stochastic Volatility* (ESV) models,[9] Andersen et al. (2011) and Barucci et al. (2012) proved that under the no leverage hypothesis it holds

$$Cov(IV(t+1),\widetilde{IV}(t)) = Cov(IV(t+1),\widehat{IV}(t)) = Cov(IV(t+1),IV(t)),$$

therefore,

$$R^2_{\widetilde{IV}} := \frac{Cov(IV(t+1),\widetilde{IV}(t))^2}{Var[IV(t)]Var[\widetilde{IV}(t)]} = \frac{Cov(IV(t+1),IV(t))^2}{Var[IV(t)]Var[\widetilde{IV}(t)]}. \tag{6.16}$$

[9] The ESV models introduced by Meddahi (2001) include most continuous-time stochastic volatility models. Roughly speaking, under these models the volatility process depends only on a single (latent) state variable and can be expressed as a linear combination of the eigenfunctions of the infinitesimal generator associated with this latent variable.

Formula (6.16) shows that maximizing the R^2 of the linear regression is equivalent to minimizing the variance of the considered estimator. Hence, minimum variance estimators should have better forecasting performances.

6.4.1 Monte Carlo Analysis

In this section the capability of the Fourier estimator to forecast the integrated volatility one step (day) ahead is evaluated, taking as benchmark the performance of the Realized Volatility.[10] Given the volatility process of the theoretical asset price, we calculate the exact integrated volatility. The comparison between the Realized Volatility and the Fourier estimation methods is accomplished through the R^2 associated with the Mincer-Zarnowitz style regression of the integrated volatility in $[t, t+1]$ $(IV(t+1))$ onto a constant and the estimated integrated volatility of the previous day, see formula (6.16).

6.4.1.1 In-sample Forecast

We consider the GARCH model

$$\begin{aligned} dp(t) &= \sigma(t)dW_1(t) \\ d\sigma^2(t) &= \theta(\omega - \sigma^2(t))dt + \overline{\sigma}\sigma^2(t)dW_2(t), \end{aligned} \tag{6.17}$$

with $\theta = 0.035$, $\omega = 0.636$, $\overline{\sigma} = 0.1439$, where $E[IV(t)] = \omega = 0.636$. High-frequency evenly sampled theoretical prices $p(t)$ and observed returns are generated by simulating second-by-second return and variance paths over 252 trading days (one trading year) using a Euler Monte Carlo discretization procedure with a trading day made of $T = 6$ hours for a total of 21600 observations. Then, we sample the observations by varying the uniform sampling interval $\rho(n) = T/n$ and obtaining data sets with different frequencies. The initial point of the simulation of the volatility process is set at ω. For each observation t_j, the observed asset price is obtained by adding i.i.d random variables $\eta(t_j)$ $(j = 0, 1, \ldots, n)$ with zero mean and constant variance to the theoretical price. Microstructure noise variance is set equal to a given percentage of the integrated volatility. In particular, we consider a model without microstructure noise and two different noise levels:

$$Var[\eta(t)] = \varsigma E[IV(t)] \quad \text{with} \quad \varsigma = 0\%,\ 0.1\%,\ 0.5\%.$$

In our analysis we consider the following sampling intervals:

[10] A comparison with methods specifically designed to handle market microstructure noise can be found in Barucci et al. (2012).

n	2160	1440	720	360	180	120	72
$\rho(n)$	10s	15s	30s	1min	2min	3min	5min

Table 6.4 provides the value of the R^2 for the Realized Volatility ($R^2(RV(n))$) and for the Fourier estimator ($R^2(F_N(n))$). The cutting frequency N is set equal to the Nyquist frequency, $N = n/2$, in the absence of microstructure noise (i.e., $\varsigma = 0$), while it is chosen by using the feasible minimization of the estimated MSE (5.6), when $\varsigma \neq 0$. Concerning the Realized Volatility estimator we observe that the R^2 increases monotonically as the sampling interval decreases only in the model without noise; if noise is added, then the R^2 reaches the highest value for a sampling interval between 1–5 minutes. For the Fourier estimator the R^2 increases with the sampling frequency even when microstructure noise effects are added.

Table 6.4 R^2 for integrated variance forecasts: linear regression of the integrated volatility at time $t+1$ onto a constant and the volatility at time t estimated by the Realized Volatility or the Fourier estimator for model (6.17). ς is the noise-to-signal ratio, n the number of observations and N the Fourier cutting frequency obtained with the feasible minimization of (5.6).

	ς =0%			$\varsigma = 0.1\%$			ς=0.5%		
n	$R^2(RV(n))$	$R^2(F_N(n))$	N	$R^2(RV(n))$	$R^2(F_N(n))$	N	$R^2(RV(n))$	$R^2(F_N(n))$	N
2160	0.9254	0.9254	1080	0.8020	0.8703	96	0.2014	0.8552	59
1440	0.9193	0.9131	720	0.8188	0.8708	83	0.2465	0.8372	50
720	0.9181	0.9181	360	0.8505	0.8544	64	0.4391	0.8078	39
360	0.9032	0.9030	180	0.8269	0.8411	52	0.5220	0.7477	31
180	0.8689	0.8685	90	0.8204	0.7807	37	0.6345	0.7237	27
120	0.8558	0.8565	60	0.8198	0.7615	28	0.6476	0.6768	25
72	0.8390	0.8355	36	0.7977	0.7033	18	0.6798	0.6480	17

The following conclusion can be drawn.

The forecasting performance of the Realized Volatility and Fourier estimators is quite similar in a model without noise. When noise is added the Fourier estimator outperforms the Realized Volatility in a significant extent for high-frequency observations and when the noise component is relevant.

In line with the results in Chapter 5, the Fourier cutting frequency becomes smaller, when the noise effect increases, even maintaining the same size of the grid.[11]

[11] The cutting frequency in this experiment differs from the one considered in Barucci et al. (2012), where N is selected in order to maximize the R^2. Moreover, the results in Table 6.4 are slightly less performing than those in Barucci et al. (2012) for low values of n due to the feasible minimization adopted here.

6.4.1.2 Out-of-Sample Forecast

The final simulation study compares the out-of-sample, one-day ahead forecasts of the integrated volatility obtained by the Fourier estimator or by the Realized Volatility with the true (simulated) values. Specifically, we consider a rolling window of 230 consecutive days. This choice of the window size allows us to forecast the last 22 days (i.e., a month). The steps of the procedure are the following:

Step (a) select 230 consecutive days (i.e., a time window) whose last date is t. For each day compute an estimate of the integrated volatility by the Fourier or the Realized Volatility estimator;

Step (b) using the 230 estimated integrated volatilities, regress one day ahead integrated variance over today estimated variance by

$$IV(t+1) = \phi_0 + \phi_1 \widehat{IV}(t) + \varepsilon_t, \tag{6.18}$$

where $t = 1, 2, \ldots, m$, ε_t is the error term and $\widehat{IV}(t)$ stands for either the Fourier estimator or Realized Volatility;

Step (c) use the coefficients of the regression in Step (b) to forecast the integrated volatility at $t+1$ day, denoted by $IV_{t+1|t}$:

$$IV_{t+1|t} = \phi_0 + \phi_1 \widehat{IV}(t);$$

Step (d) move the rolling window along the series discarding the first estimate and inserting the new one $IV(t+1)$ available at $t+1$ day, update t and repeat Steps (a)–(c).

Steps (a)–(d) provide a daily time series of one-day ahead forecasts $IV_{t+1|t}$ of the integrated volatility. These values are compared with the true (simulated) values of the integrated volatility.The two panels of Figure 6.7 show the true integrated volatility (dotted line) and the one-day ahead forecasts obtained using the Fourier (solid line) and Realized Volatility (dashed line) estimators. The forecast values are relative to the last month (i.e., the last twenty two days) of the trading year. In the left panel the prices are not affected by noise and the forecasts of both methodologies are obtained using 10 second returns. In the right panel the prices are affected by noise ($\varsigma = 0.5$) and the Fourier forecasts are still obtained using 10 second returns while the Realized Volatility forecasts are obtained from 5 minute returns in order to filter out microstructure effects. In the left panel (no noise) the mean squared errors of the forecasts obtained by the Fourier and by the Realized Volatility estimators are both equal to 0.002789, while in the right panel (presence of noise) the mean squared error of the Fourier and Realized Volatility forecasts are 0.0049 and 0.0104 respectively, thus the Fourier method allows us to halve the error of the forecasts. These results confirm that the Fourier methodology is strongly recommended in the presence of noise.

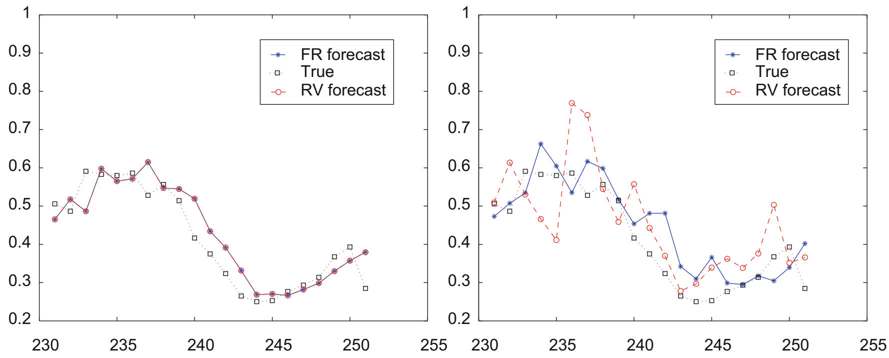

Fig. 6.7 Left Panel: true integrated daily volatility (black dotted line) and one-day ahead forecasts obtained with the Fourier method (blue solid line) and Realized Volatility using 10 second returns (no noise). Right Panel: true integrated daily volatility (black dotted line), one-day ahead forecasts obtained with the Fourier method (blue solid line) using 10 second returns and Realized Volatility (red dashed line) using 5 minute returns (with noise $\varsigma = 0.5\%$)

6.4.2 An Empirical Application

In this section the forecasting power of the Fourier estimator is studied using the S&P 500 index futures recorded at the Chicago Mercantile Exchange (CME). The sample covers the period from January 3, 2006 to December 31, 2007, a period of 500 trading days (1074825 tick-by-tick observations). Table 6.5 describes the main features of our dataset.

Table 6.5 Summary statistics for the sample of the traded CME S&P 500 index futures in the period from 3 January 2006 to 31 December 2007 (1074825 trades). Std. Dev. denotes the sample standard deviation of the variable.

Variable	Mean	Std. Dev.	Min	Max
S&P 500 index futures	1401.80	99.28	1.23e+3	1.59e+3
Number of trades per minute	5.6229	3.601	1	36

High-frequency returns are contaminated by transaction costs, bid-ask bounce effects, etc. leading to biases in the variance measures. Therefore, daily integrated volatility has been computed by the Fourier estimator using tick-by-tick data and the optimal cutting frequency N obtained by feasible minimization of the MSE estimate (5.6). As a benchmark, we consider daily estimates obtained by the Realized Volatility estimator from 5-minute returns. Jumps have been identified and measured using the Threshold Bipower Variation based method introduced in Section 6.1. The test is employed at the significance level of 99.9%. Days having trading period shorter than 5 hours have been eliminated. The cleaned samples contains 450 days.

We split our sample into two parts: the first one containing 50% of total estimates is used as a "burn-in" period to fit a univariate AR(1) model[12] for the estimated variance time series or equivalently to estimate the intercept and slope in the regression (6.18) where the left hand side is replaced by the corresponding estimate and then the fitted model is used to forecast integrated variance on the next day. The AR(1) models are separately estimated for both time series of integrated volatilities given by the Fourier and Realized Volatility estimators. The total number of out-of-sample forecasts m is equal to 225. Each time a new forecast is performed, the corresponding actual variance measure is moved from the forecast period to the first sample and the AR(1) parameters are re-estimated in real time. Figure 6.8 shows the estimated daily volatility (red line) versus one-day ahead forecasts (blue line) obtained by the Fourier and the Realized Volatility estimators over the forecast period (December 3, 2007 - December 31, 2007). Although not very different from a visual point of view, the mean squared error of the Fourier forecasts is 2.99e-9, while for the Realized Volatility it is 3.61e-9, i.e., a relative difference of 21%.

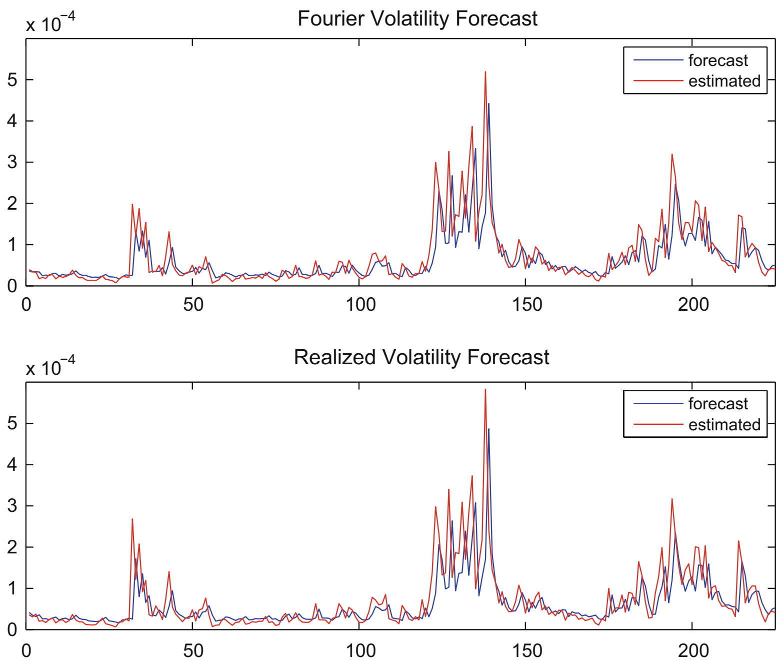

Fig. 6.8 Estimated daily volatility (red line) versus one-day ahead forecasts (blue line) obtained by the Fourier (tick-by-tick data) and the Realized Volatility (sparse sampled data) estimators over the forecast period (December 3, 2007 - December 31, 2007).

[12] Even if a simple AR(1) model cannot perfectly capture the dynamic of the integrated volatility, this model has been mainly chosen to make the empirical analysis comparable with the Monte Carlo analysis.

In order to better examine the informational content of forecasts and to assess the advantage of using the Fourier estimates from high-frequency data instead of Realized Volatility estimates from sparse sampled data, for each estimated series $\widehat{IV}(t)$ of integrated volatility, we project the estimated daily variance measure on day $[t, t+1]$ on a constant and the corresponding one-step-ahead forecast $IV_{t+1|t}$ obtained from the series of the Fourier estimates ($IV_{t+1|t}^{Fourier}$) and from the Realized Volatility estimates ($IV_{t+1|t}^{RV}$), respectively. The Mincer-Zarnowitz forecast evaluation regression takes the form

$$\widehat{IV}(t+1) = \phi_0 + \phi_1 IV_{t+1|t} + \varepsilon_{t+1}, \tag{6.19}$$

where $t = 0, 1, \ldots, m-1$. Conditioning upon the forecast, the prediction is unbiased only if $\phi_0 = 0$ and $\phi_1 = 1$. The R^2 from these regressions provides a direct assessment of the variability in the integrated variance that is explained by the particular estimates in the regressions. The R^2 can therefore be interpreted as a simple gauge of the degree of predictability in the volatility process and hence of the potential economic significance of the volatility forecasts.

Table 6.6 shows the OLS estimates from regressions (6.19) where the Fourier/Realized Volatility estimates are regressed on the forecasts based on the AR(1) model fitting the corresponding Fourier/RV estimates on the first year. We notice that the R^2 is higher for the Fourier forecasts than for Realized Volatility. In particular, the Fourier-based forecasts explain around 46% of the time series variability. All the estimated coefficient are significant at the 5% significance level, although the constant term ϕ_0 is very close to zero. The coefficient estimate for ϕ_1 is generally close to unity and it is higher for the Fourier estimator than for Realized Volatility. Moreover, the null hypothesis that $\phi_1 = 1$ cannot be rejected at 5% level, at least for the Fourier forecast. This seems to confirm a slightly higher forecasting power of the Fourier volatility estimates in comparison to Realized Volatility.

In cases where there are more than one forecasting models, additional forecasts are added to the right-hand side of (6.19) to check for incremental explanatory power. Therefore, we also consider the regression

$$\widehat{IV}(t+1) = \phi_0 + \phi_1 IV_{t+1|t}^{Fourier} + \phi_2 IV_{t+1|t}^{RV} + \varepsilon_{t+1}. \tag{6.20}$$

Table 6.6 OLS estimates from regressions (6.19) of Fourier/RV estimated integrated variance on day $t+1$ over a constant and each corresponding variance forecast over the forecast period.

Estimator	ϕ_0	ϕ_1	R^2
Fourier	0.000014	0.880960	0.458190
Std	(0.000005)	(0.064151)	
T-statistics	(2.648269)	(13.732580)	
RV	0.000016	0.862673	0.401689
Std	(0.000006)	(0.070504)	
T-statistics	(2.812579)	(12.235843)	

Table 6.7 shows the results for the cases where the left-hand side of the regression is given by the Fourier estimates and by Realized Volatility, respectively. In the first

Table 6.7 OLS estimates from regressions (6.20) of Fourier/RV estimated integrated variance on day $t+1$ over a constant and both forecasts over the forecast period.

Estimator	ϕ_0	ϕ_1	ϕ_2	R^2
Fourier	0.000014	1.403016	-0.527042	0.460974
Std	(0.000005)	(0.491791)	(0.492249)	
T-statistics	(2.579898)	(2.852868)	(-1.070681)	
RV	0.000014	1.323618	-0.450864	0.417799
Std	(0.000006)	(0.534044)	(0.534541)	
T-statistics	(2.422372)	(2.478484)	(-0.843461)	

case (upper panel), the R^2 is almost unchanged compared to the R^2 based solely on Fourier. Moreover, the coefficient estimate for ϕ_1 is close to unity and the null hypothesis that $\phi_1 = 1$ cannot be rejected at 5% level using the corresponding t tests. On the contrary, the coefficient ϕ_2 corresponding to the Realized Volatility estimates is not significantly different from zero at the 5% level. This means that the forecasts deriving from the Realized Volatility-based AR(1) model do not increment the explanatory power of the regression. When we regress the Realized Volatility series on both forecasts (lower panel), the R^2 is slightly increased compared to the R^2 based solely on Realized Volatility. This means that the forecasts deriving from the Fourier-based AR(1) model explain the sample variance of the Realized Volatility series better than $IV^{RV}_{t+1|t}$ itself. Moreover, unexpectedly, again the coefficient ϕ_2 corresponding to the Realized Volatility estimates is not significantly different from zero at the 5% level while the coefficient estimate for ϕ_1 is close to unity and the null hypothesis that $\phi_1 = 1$ cannot be rejected. This means that the Fourier forecasts have a larger explanatory power even when we regress the series of the Realized Volatility estimates.

These results confirm the higher informational content of forecasts based on the Fourier estimates from high-frequency data versus Realized Volatility estimates from sparse sampled data, mainly due to the higher accuracy and lower variability of the Fourier variance estimates which translate into superior forecasts of future variances.

6.5 Further Readings

The academic literature proposes many other interesting applications of the Fourier method which analyze the effects of the volatility estimates on other volatility-related quantities. Far from being exhaustive, we list below some applications to option pricing, principal component analysis, VaR estimation, term structure of interest rates study, credit risk and medicine: Renò and Rizza (2003), Precup and Iori (2004), Mancino and Renò (2005), Liu and Ngo (2014), Pasquale and Renò (2005), Malliavin et al. (2007), Liu and Mancino (2012), Papantonopoulos et al. (2013), Barsotti and Sanfelici (2014), Kenmoe and Sanfelici (2014), Han et al. (2014), Sanfelici and Uboldi (2014).

Appendix A
Mathematical Essentials

A.1 Stochastic Processes

We resume in this section a few fundamental concepts frequently used across the book. The reader can find a deeper and more rigorous treatment of the huge theory of stochastic processes in beautiful books such as Revuz and Yor (1991), Øksendal (1995).

A.1.1 Diffusion Processes

In simple words, we can say that a diffusion process is a process that can be locally described by the following stochastic difference equation

$$p(t+\Delta t) - p(t) = b(t)\Delta t + \sigma(t)\varepsilon(t)\sqrt{\Delta t}, \tag{A.1}$$

where $\varepsilon(t)$ are independent identically distributed random variables having standard Gaussian distribution and b and σ are deterministic functions. The first component, b, is called the *drift*, while the second one, σ, is named the *diffusion*.

More precisely, the diffusive component is described through a Brownian motion W, defined as follows:

Definition A.1. A Brownian motion $(W(t))_{t\in[0,T]}$ is a stochastic process such that the following properties hold:
i) $W(0) = 0$,
ii) for any $r < s \le t < v$, the increments $W(v) - W(t)$ and $W(s) - W(r)$ are independent random variables,
iii) for any $s < t$, the increment $W(t) - W(s)$ has Gaussian distribution with zero mean and variance $t - s$.

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3

In a natural way an *information structure* is associated with the Brownian motion, which is called the *natural filtration*. Loosely speaking, the filtration $\mathscr{F}_t^W$ denotes the information generated by the Brownian motion W on the interval $[0,t]$. More generally, we have the following:

Definition A.2. A filtration $(\mathscr{F}_t)_{t\geq 0}$ on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ is an increasing family of sub sigma-algebras of the sigma-algebra $\mathscr{F}$.

Definition A.3. Given a filtration $(\mathscr{F}_t)_{t\geq 0}$ on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ and a stochastic process $X = (X_t)_{t\geq 0}$ defined on the same space, the process X is adapted to the filtration $\mathscr{F}$ if X_t is measurable with respect to $\mathscr{F}_t$ for any $t \geq 0$.

For example, the Brownian motion $(W_t)_{t\geq 0}$ is (by definition) adapted to the natural filtration $(\mathscr{F}_t^W)_{t\geq 0}$.

Remark A.1. When studying asset price models, it is often necessary to consider a filtration $\mathscr{F}$ which is strictly bigger than $\mathscr{F}^W$, in the sense that $\mathscr{F}_t^W \subseteq \mathscr{F}_t$ for any $t \geq 0$. In this respect, it is enough that $W_t - W_s$ is independent of $\mathscr{F}_s$ whenever $0 \leq s < t$ in order that W is a Brownian motion also on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$.

The Brownian motion has continuous paths, that is $t \to W(t)$ is a continuous function. This is a consequence of the following fundamental result.

Theorem A.1. *(Kolmogorov Continuity Theorem) Given a stochastic process $X = (X_t)_{0\leq t\leq T}$ defined on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$, suppose that it satisfies the condition*

$$E[|X_t - X_s|^{\alpha}] \leq C\,|t-s|^{1+\beta}, \quad 0 \leq s,t \leq T$$

for some positive constants α, β, C, then (up to an equivalent modification[1]), X has locally Holdër continuous paths with exponent $\gamma \in (0, \beta/\alpha)$.

Actually, a remarkable result by Lévy (1937) gives the modulus of continuity of the Brownian paths on any bounded interval $[0,T]$. It states that almost surely

$$\limsup_{h\to 0} \frac{1}{\sqrt{2h\log(1/h)}} \max_{\substack{|t-s|\leq h \\ s,t\in[0,T]}} |W_t - W_s| = 1\,. \tag{A.2}$$

Therefore, diffusion models rule out the presence of jumps. Now, letting $\Delta t \to 0$ in (A.1), we can write formally the (univariate) Itô stochastic differential equation

$$dp(t) = b(t)dt + \sigma(t)dW(t), \qquad p(0) = p_0, \tag{A.3}$$

where the processes $b(t)$ and $\sigma(t)$ satisfy appropriate measurability and integrability conditions.

Consider now l independent Brownian motions $W^1, \ldots, W^l$. Then, we can define the d-variate stochastic differential system

$$dp^j(t) = b^j(t)dt + \sum_{i=1}^{l} \sigma_i^j(t)dW^i(t), \qquad p^j(0) = p_0^j, \;\; j = 1, \ldots, d. \tag{A.4}$$

[1] For the precise statement of *modification* of a stochastic process see, e.g., Protter (1992)

In this case, the drift is a d-dimensional vector and the diffusion is a $d \times l$ matrix

$$\begin{pmatrix} \sigma_1^1 & \cdots & \sigma_l^1 \\ \sigma_1^2 & \cdots & \sigma_l^2 \\ \vdots & \ddots & \vdots \\ \sigma_1^d & \cdots & \sigma_l^d \end{pmatrix}$$

Finally, recall the following types of convergence for random variables.

Definition A.4. Consider a sequence of real random variables $(X_n)_{n\geq 0}$ and a random variable X defined on the same probability space. We will say that the sequence X_n converges in probability to X if for any $\varepsilon > 0$ it holds

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0.$$

Definition A.5. Consider a sequence of real random variables $(X_n)_{n\geq 0}$ and a random variable X defined on the same probability space. We will say that the sequence X_n converges almost surely to X if it holds

$$P(\limsup_{n\to\infty} |X_n - X| > 0) = 0.$$

A.1.2 Itô Energy Identity

The following result, also known as *Itô isometry*, is a fundamental result in the study of the quadratic variation/volatility. Let X be a stochastic process adapted to the Brownian filtration and square integrable, then

$$E[(\int_0^t X(s)dW(s))^2] = E[\int_0^t X^2(s)ds], \quad \text{for any } t \in [0,T]. \tag{A.5}$$

A.1.3 Itô Formula

The fundamental goal which can be achieved through Itô formula is that, given the dynamics of the underlying factor (both univariate and multivariate), we can obtain the stochastic evolution of any (smooth) function of the underlying. An intuitive idea is given by applying a Taylor expansion

$$df = \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial x}\, dp + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dp)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x}\, dt \cdot dp \tag{A.6}$$

(where the dependence on (t,x) is usually omitted for ease of notation) and then using the following Itô table rules

$$(dW)^2 = dt, \;\; dW \cdot dt = 0, \;\; (dt)^2 = 0.$$

Theorem A.2. *Let $p(t)$ satisfy the dynamic (A.3) and $f(t,x)$ be a function, differentiable with respect to t and twice differentiable with respect to x, then*

$$df(t,p(t)) = \left(\frac{\partial f}{\partial t} + b(t)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2(t)\frac{\partial^2 f}{\partial x^2} \right) dt + \sigma(t)\frac{\partial f}{\partial x} dW(t). \tag{A.7}$$

With the same method, for the diffusion (A.4) the multidimensional Itô formula holds

Theorem A.3. *Let $p(t)$ satisfy the dynamics (A.4) and $f(t,x)$ be a function, differentiable with respect to t and twice differentiable with respect to x, then*

$$df(t,p(t)) = \left(\frac{\partial f}{\partial t} + \sum_{j=1}^{d} b^j(t)\frac{\partial f}{\partial x_j} + \frac{1}{2}\sum_{i,j=1}^{d} \Sigma^{i,j}(t)\frac{\partial^2 f}{\partial x_i x_j} \right) dt + \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}\sigma_j(t) dW(t), \tag{A.8}$$

where σ_j is the vector

$$\sigma_j = (\sigma_1^j, \ldots, \sigma_l^j)$$

and the entries of the matrix Σ are equal to

$$\Sigma^{i,j}(t) = \sum_{k=1}^{l} \sigma_k^i(t)\sigma_k^j(t), \quad i,j = 1,\ldots,d.$$

The matrix $\Sigma(t)$ is the variance-covariance matrix, also called *volatility matrix*.

The financial modeling also needs to consider the case of dependent Brownian motions. Consider now d Brownian motions such

$$dW^i \cdot dW^j = \rho^{i,j} dt, \quad i,j = 1,\ldots,d$$

where $\rho^{i,j}$ is the correlation between the two Brownian motions W^i, W^j and the system of stochastic differential equations

$$dp^j(t) = b^j(t)dt + \sigma^j(t)dW^j(t), \quad p^j(0) = p_0^j, \quad j = 1,\ldots,d. \tag{A.9}$$

Then, the multidimensional Itô formula is the following:

$$df(t,p(t)) = \left(\frac{\partial f}{\partial t} + \sum_{j=1}^{d} b^j(t)\frac{\partial f}{\partial x_j} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^i(t)\sigma^j(t)\rho^{i,j}\frac{\partial^2 f}{\partial x_i x_j} \right) dt + \sum_{j=1}^{d} \sigma^j(t)\frac{\partial f}{\partial x_j} dW^j(t). \tag{A.10}$$

A.2 Fourier Analysis

The *Fourier transform* is a mathematical tool which has been widely used in many applied fields, such as engineering, physics and, more recently, finance. It allows us to represent a possibly complicated periodic function as a linear combination of projections onto a trigonometric basis. The term Fourier transform refers to both the frequency domain representation and the mathematical operation that associates the frequency domain representation to a function of time. Linear operations done in one domain (time or frequency) have corresponding operations in the other domain, which are sometimes easier to perform. For instance, the operation of differentiation in the time domain corresponds to multiplication by the frequency and (which is of primarily interest for our theory) convolution in the time domain corresponds to ordinary multiplication in the frequency domain. Further, for many functions of practical interest the inverse Fourier transform can be defined. Therefore, after performing the desired operations, transformation of the result can be made back to the time domain. We recall here few definitions which are needed for reading this book. More exhaustive study can be found in many insightful books, e.g., Bloomfield (2000), Hannan (1970), Priestley (1983).

Given a function f defined and integrable on $[0,2\pi]$, with $f(0) = f(2\pi)$, the k-th Fourier coefficient is defined for any integer k as

$$\mathscr{F}(f)(k) := \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-\mathrm{i}kt}dt \tag{A.11}$$

being $\mathrm{i} = \sqrt{-1}$. Thus, we can consider its Fourier series

$$\sum_{|k|\leq N} \mathscr{F}(f)(k)\, e^{\mathrm{i}kt} \tag{A.12}$$

which, in the limit for $N \to \infty$ (in a suitable sense), gives $f(t)$.

Remark A.2. It is also possible to expand the function f using only real number, as a series of sine and cosine, in virtue of the *Euler identity* $e^{\mathrm{i}t} = \cos t + \mathrm{i}\,\sin t$; for any integer $k > 0$, set

$$a_k(f) := \frac{1}{\pi}\int_0^{2\pi} \cos(kt)f(t)dt, \qquad b_k(f) := \frac{1}{\pi}\int_0^{2\pi} \sin(kt)f(t)dt,$$

and

$$a_0(f) := \frac{1}{2\pi}\int_0^{2\pi} f(t)dt,$$

then it holds for any integer $k > 0$

$$\mathscr{F}(f)(k) = a_k(f) - \mathrm{i}b_k(f), \quad \mathscr{F}(f)(-k) = a_k(f) + \mathrm{i}b_k(f).$$

Remark A.3. If we liked to work on the interval $[0,T]$, the definition of the k-th Fourier coefficient (A.11) would change into

$$\mathscr{F}(f)(k) := \frac{1}{T}\int_0^T f(t)e^{-\mathrm{i}\frac{2\pi}{T}kt}dt,$$

and the trigonometric series (A.12) would become

$$\sum_{|k|\le N} \mathscr{F}(f)(k)\, e^{\mathrm{i}\frac{2\pi}{T}kt}.$$

A.2.1 Fejér's Convergence Theorem

It is not difficult to see that if f is square-integrable, then its Fourier series (A.12) converges in mean squared norm to the function f. However, this result is not enough for our estimation problem because we need to measure accurately the difference between the function f at some given point and its Fourier approximation at the same point.

In virtue of the *Fejér theorem*, by considering the Cesaro summation, it is possible to obtain the convergence in the stronger form of uniform convergence, which ensures that the rate at which the series converges is the same for any point in $[0,2\pi]$. More precisely, the trigonometric series

$$\sum_{|k|\le N}\left(1-\frac{|k|}{N}\right)\mathscr{F}(f)(k)\, e^{\mathrm{i}kt} \tag{A.13}$$

converges uniformly (and in mean squared norm) as $N\to\infty$ to $f(t)$ on $[0,2\pi]$ if the function f is continuous and periodic on $[0,2\pi]$.

If $f(t)$ has cadlag paths, then the limit of (A.13) gives $(f(t)+f(t^-))/2$. For a proof of these results see, e.g., Malliavin (1995).

A.2.2 Product Formula

Given two square integrable and periodic functions f and g on $[0,2\pi]$, we are interested in computing the Fourier coefficients of their product fg.

First, note that the product function fg is a periodic function, integrable on $[0,2\pi]$. Then, for any integer k, it holds that

$$\mathscr{F}(fg)(k) = \sum_{s+h=k}\mathscr{F}(f)(s)\mathscr{F}(g)(h).$$

A proof of this classical result can be found, e.g., in Malliavin (1995).

A.2.3 Nyquist Frequency

The concept of *Nyquist frequency* appears in the signal processing theory. Roughly speaking, the Sampling Theorem (see, e.g., Priestley (1983)) states that the sampling frequency should be at least twice the highest frequency contained in the original function. More precisely, when considering the discrete Fourier transform of a periodic function f sampled at n points $\{t_j,\ j = 0,\dots,n-1\}$, the Nyquist frequency is given by $N = n/2$.

The Nyquist frequency is related to the so-called *aliasing* phenomenon: $N = n/2$ is the critical frequency at which any frequency component k such that $|k| > n/2$ is *aliased* (i.e., falsely translated) into this range by the very act of discrete sampling. In fact, let us consider a discrete sample $f(t_j)$, $j = 0,\dots,n-1$ of a function f with $t_j = j2\pi/n$. The *discrete Fourier coefficients* of f are, for any integer k

$$c_k(f) = \sum_{j=0}^{n-1} e^{-\mathrm{i}\,\frac{2\pi}{n}\,k\,j} f(t_j).$$

It is easy to see that the frequencies k such that $n/2 \le |k| \le n$ can be expressed as $k = n-m$ or $k = -(n-m)$, $0 \le m \le n/2$ and the corresponding coefficients for $k = n-m$ are given by

$$c_{n-m}(f) = \sum_{j=0}^{n-1} e^{-\mathrm{i}\,2\frac{\pi}{n}\,(n-m)\,j} f(t_j),$$

that is

$$c_{n-m}(f) = \sum_{j=0}^{n-1} e^{\mathrm{i}\,2\frac{\pi}{n}\,m\,j} f(t_j) = c_{-m}(f)\,,$$

while for $k = -(n-m)$ we obtain $c_{-(n-m)}(f) = c_m(f)$. Finally, bearing in mind that the frequencies k with $|k| > n$ can be expressed as $|k| = l\,n + m$, $m = 0,1,\dots,n-1$, $l = 1,2,\dots$, it is easy to see that the corresponding coefficients satisfy $c_{l\,n+m}(f) = c_m(f)$ and $c_{-(l\,n-m)}(f) = c_{-m}(f)$.

This computation shows that the discrete Fourier coefficients with frequencies larger than $n/2$ are falsely translated into the frequency range $|k| \le n/2$.

Appendix B
Codes for the Fourier Estimator

This appendix contains the Matlab® implementation of some Fourier estimators illustrated in the previous chapters.

B.1 Integrated Volatility

The implementation of the Fourier estimator of integrated volatility can be easily obtained from formulae (3.3)–(3.5). It should be remarked that this form of the estimator is computationally more efficient than the mathematically equivalent (3.6). The algorithm is structured as a Matlab® function that returns the daily value of the integrated variance, recorded in the variable `ivol`. The input parameters are: the observed log-prices $p(t_i)$, $i = 0, 1, \dots, n$ collected in a vector `P`; the observation time vector `t` $= (t_0, t_1, \dots, t_n)$ spanning the interval $[0, 2\pi]$; the cutting frequency `N`. Observation times need not to be equally spaced.

In the Matlab® algorithm 1, `c_0` contains the zero-th Fourier coefficient of the log-returns and the k-th entry of vector `c_p` contains the k-th Fourier coefficient, $k = 1, 2, \dots, N$, up to the normalizing factor. Algorithm 1 exploits the fact that $c_{-s}(dp_n) = \overline{c_s(dp_n)}$.

Matlab® Algorithm 1

```
function ivol = FE(P,t,N)

% Computes the integrated volatility by the Fourier
% estimator with Dirichlet kernel

% Input variables:
%   P       vector of the observed log-prices
```


M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3

```
%   t        vector of the observation times
%   N        cutting frequency

% Output variables:
%   ivol     integrated variance

r=diff(P); % log-returns

c_p=zeros(N,1);
c_0=sum(r);

for k=1:N
    c_p(k)=sum(exp(1i*k*t(1:end-1)).*r);
end
ivol=(c_0.*conj(c_0)+2*sum(c_p.*conj(c_p)))/
      (2*N+1);
```

We can slightly modify the algorithm 1 by introducing the Fejér kernel in the convolution product as in (3.9)

Matlab® Algorithm 2

```
function ivol = FEker(P,t,N)

% Computes the integrated volatility by the Fourier
% estimator with Fejer kernel

% Input variables:
%   P        column vector of the observed
        log-prices
%   t        column vector of the observation times
%   N        cutting frequency

% Output variables:
%   ivol     integrated variance

r=diff(P); % log-returns

c_p=zeros(1,2*N+1); c_pp=zeros(1,2*N+1);
```

```
for k=1:(2*N+1)
    s=k-N-1;
    c_p(k)=sum(exp(-1i*s*t(1:end-1)).*r);
    c_pp(k)=sum(exp(1i*s*t(1:end-1)).*r);
    term(k)=(1-abs(s)/N)*c_p(k)*c_pp(k);
end
ivol=sum(term)/(N+1);
```

B.2 Estimated Bias and MSE

The practical calculation of (5.3) and (5.6) hinges on the estimation of the relevant noise moments as well as on the preliminary identification of the integrated volatility V and integrated quarticity Q. The function `moments.m` allows one to compute daily values for the sample moments, while V and Q are taken as input variables. Since the noise moments do not vary across frequencies under the MA(1) model, in computing the MSE estimates we use sample 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 (2008), so that for n sufficiently large we have

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

$$E[\eta^2] = \frac{E[\varepsilon^2]}{2}, \qquad E[\eta^4] = \frac{E[\varepsilon^4]}{2} - 3\frac{E[\varepsilon^2]^2}{4},$$

where ε is the noise return process. The outputs of the function are daily estimates for these moments, stored in the variables `E2`, `E4`, `Eeta2`, `Eeta4`, respectively, and daily values of α, β, and γ defined by (5.5).

Matlab® Algorithm 3

```
function [alpha,beta,gamma,E2,E4,Eeta2,Eeta4]
 = moments(P,n,V)

% Computes sample moments of the noise

% Input variables:
%   P       column vector of the observed
           log-prices
```

```
% n total number of intra-day returns
% V integrated variance

r=diff(P); % log-returns

E2=sum(r.^2)/n-V/n;
E4=sum(r.^4)/n-(6*E2*V)/n;
Eeta2=E2/2;
Eeta4=E4/2-3*E2^2/4;

alpha=E2^2;
beta=4*Eeta4;
gamma=8*Eeta2*V+alpha/2-2*Eeta4;
```

The following algorithm 4 implements the computation of estimates (5.3) and (5.6) starting from daily measurements of the sample moments of the noise. The function provides two row vectors (`BIAS`, `MSE`) containing the estimated values of the bias and MSE for the Fourier estimator of integrated volatility as a function of the maximum frequency `trunc` at which we decide to truncate the Fourier expansion. Estimates (5.3) and (5.6) allow us to measure the bias and MSE of the volatility estimators from observed prices also in the case of empirical market quote data, where the efficient price and volatility and the noise contaminations are not available. By direct comparison of the values `MSE(trunc)`, for `trunc` spanning from 1 to `N`, one can select the optimal cutting frequency minimizing the MSE of the Fourier estimator. The optimal cutting frequency, N_{opt} is equal the index of the minimum value component of the vector `MSE`.

Matlab® Algorithm 4

```
function [BIAS,MSE]
 = estimates(n,N,ND,Q,alpha,beta,gamma,E2,Eeta4,T)

% Input variables:
% n total number of intra-day returns
% N maximum Fourier frequency
% ND number of days in the sample
% Q daily integrated quarticity vector
% P column vector of the observed
 log-prices
```

```
%   alpha   daily estimations of alpha in (5.5)
%   beta    daily estimations of beta in (5.5)
%   gamma   daily estimations of gamma in (5.5)
%   E2      daily estimations of E2
%   Eeta4   daily estimations of Eeta4
%   T       trading period

% Output variables:
%   BIAS    row vector of bias estimates (5.3)
%   MSE     row vector of MSE estimates (5.6)

    h=2*pi/n;
    MSE1=zeros(ND,N); BIAS1=zeros(ND,N);
    alphaFE=zeros(ND,N); betaFE=zeros(ND,N);
    gammaFE=zeros(ND,N);

    for k=1:N
        trunc=min(n/2,k);
        for i=1:ND % i-th day
            BIAS1(i,k)=n*E2(i)*(1-D(trunc,h));

            alphaFE(i,k)=alpha(i)*(1+(D(trunc,h))^2
            -2*D(trunc,h));

            betaFE(i,k)=beta(i)*(1+(D(trunc,h))^2
            -2*D(trunc,h));

            gammaFE(i,k)=gamma(i)+4*(Eeta4(i)+
                   E2(i)^2)
            *(2*D(trunc,h)-(D(trunc,h))^2)+4*
                pi*Q(i)
            /(2*trunc+1);

            MSE1(i,k) = 2*Q(i)*h+betaFE(i,k)*n
            +alphaFE(i,k)*n^2+gammaFE(i,k);
        end
    end
    MSE=mean(F1);
    BIAS=mean(BIAS1);
```

Algorithm 4 calls the function `D.m` that provides the computation of the rescaled Dirichlet kernel.

Matlab® Algorithm 5

```
function d = D(N,t)

% Rescaled Dirichlet kernel:

d=1;

for s=1:N
    d=d+2*cos(s*t);
end

d=d/(2*N+1);
```

Analogous estimation of the bias and MSE for the Fourier-Fejér estimator (3.9) of integrated volatility can be obtained from algorithm 4 by substituting the function `D.m` with the function `V.m` of algorithm 6.

Matlab® Algorithm 6

```
function f = V(N,t)

% Rescaled modified Fejer kernel:

 f=sin(N*t).^2./((N*t).^2);
```

In algorithm 6 we have implemented a modified version of the Fejér kernel which is asymptotically equivalent to (3.10) but improves the performance of the estimators. Note that null elements in the input vector t are not allowed.

Writing in the Matlab® command windows the following lines:

```
x=[-pi:0.01:pi];
f=zeros(5,629); g=zeros(5,629);

for N=1:5
    f(N,:)=D(N,x); g(N,:)=V(N,x);
end
```

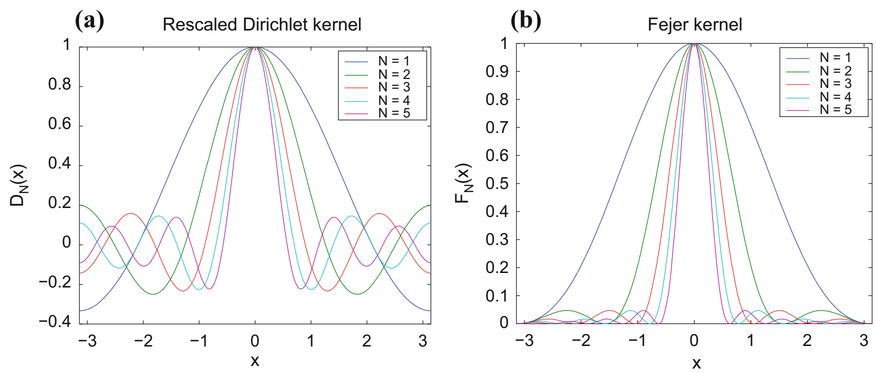

Fig. B.1 Rescaled Dirichlet kernel and rescaled Fejér kernel.

```
figure, plot(x,f)
xlabel('x')
ylabel('D_N(x)')
title('Rescaled Dirichlet kernel')
axis([-pi,pi,-0.4,1])

figure, plot(x,g)
xlabel('x')
ylabel('F_N(x)')
title('Fejer kernel')
axis([-pi,pi,0,1])
```

we obtain Figure B.1, showing the graphs of the rescaled Dirichlet kernel $D_N(x)$ and of the rescaled Fejér kernel $F_N(x)$, for $N = 1,2,3,4,5$.

B.3 Integrated Covariance

The implementation of the Fourier estimator of integrated co-volatility can be easily obtained from formula (3.22). We consider the case of two assets. The input parameters are: the observed log-prices $p^1(t_i^1)$, $i = 0,1,\ldots,n_1$ and $p^2(t_j^2)$, $j = 1,2,\ldots,n_2$, collected in two vectors `P1` and `P2`; the observation time vectors `t1` $= (t_0^1, t_1^1, \ldots, t_{n_1}^1)$ and `t2` $= (t_0^2, t_1^2, \ldots, t_{n_2}^2)$; the cutting frequency `N`. Observation times need not be either equally spaced or synchronous. Daily values of the integrated variances and covariance are recorded in the variables `iv11`, `iv22`, and `iv12` and are the output of algorithm 7.

The implementation of the Matlab® algorithm 7 is slightly different from algorithm 1. The vectors `c_p1` and `c_p2` contain the s-th Fourier coefficient of the log-returns on the two assets up to the normalizing factor, for $s = -N, -N+1, \ldots, N-1, N$, and the vectors `c_pp1` and `c_pp2` contain their conjugates.

Matlab® Algorithm 7

```
function [iv11,iv12,iv22]=FE2(P1,P2,t1,t2,N)

% Computes the integrated variances and
        covariances on
% two assets by the Fourier estimator with
        Dirichlet
% kernel

% Input variables:
%   P1, P2      vectors of the observed log-prices
%   t1, t2      vectors of the observation times
%   N           cutting frequency

% Output variables:
%   iv11, iv22     integrated variances on asset 1
    and 2
%   iv12           integrated covariance

r1=diff(P1); r2=diff(P2); % log-returns

c_p1=zeros(2*N+1,1); c_pp1=zeros(2*N+1,1);
c_p2=zeros(2*N+1,1); c_pp2=zeros(2*N+1,1);

for k=1:(2*N+1)
    s=k-N-1;
    c_p1(k)=sum(exp(-1i*s*t1(1:end-1)).*r1);
    c_pp1(k)=sum(exp(1i*s*t1(1:end-1)).*r1);
    c_p2(k)=sum(exp(-1i*s*t2(1:end-1)).*r2);
    c_pp2(k)=sum(exp(1i*s*t2(1:end-1)).*r2);
end
iv11=sum(c_p1.*c_pp1)/(2*N+1);
iv12=sum(c_p1.*c_pp2)/(2*N+1);
iv22=sum(c_p2.*c_pp2)/(2*N+1);
```

B.4 Spot Volatility

The implementation of the Fourier estimator of spot volatility can be easily obtained from formula (4.1) in Chapter 4. The Matlab® algorithm 8 implements the spot volatility estimator in the interval $[0,T]$. This is done suitably rescaling the time interval (see Appendix A.2).

The input parameters are: the observed log-prices $p(t_i)$, $i=0,1,\dots,n$ collected in a vector `P`; the observation time vector `t` $=(t_0,t_1,\dots,t_n)$; the cutting frequency `N`; the time vector `tau` $=(\tau_0,\tau_1,\dots,\tau_m)$ at which the spot volatility is evaluated; the time horizon `T` and the cutting frequency `M`. The values of the estimated spot volatility $\widehat{\sigma}^2_{n,N,M}(\tau_i)$, $i=0,1,\dots,m$ are recorded in the vector `spot` and the Fourier coefficients of the spot volatility are recorded in the vector `c_s`.

Observation times are not required to be equally spaced.

Matlab® Algorithm 8

```
function [spot,c_s] = FE_spot_vol(P,t,tau,T,N,M)

% Computes the spot variance by the Fourier
        estimator
% with Dirichlet kernel

% Input variables:
%   P     vector of the observed log-prices
%   t     vector of the observation times
%   tau   vector of the times where the
%         volatility is estimated
%   N     maximum Fourier frequency for price
          returns
%   M     maximum Fourier frequency for spot
          variance

% Output variables:
%   spot  vector of spot variance at the time
%         grid tv
%   c_s   Fourier coefficients of the spot variance

n=max(size(P)); nv=max(size(tau)); const=2*pi/T;
c_pp=zeros(N+M,1); c_p=zeros(2*N+2*M+1,1);
r=diff(P);
c_0=sum(r);

for k=1:N+M
```

```
        c_pp(k)=sum(exp(-1i*const*k*t(1:end-1)).*r);
    end
    for j=1:N+M
       c_p(j)=conj(c_pp(N+M+1-j))/T;
    end
    c_p(N+M+1)=c_0/T;
    for j=1:N+M
       c_p(N+M+1+j)=c_pp(j)/T;
    end

    % Fourier coefficients of the spot variance
          in [0,T]
    fact=T/(2*N+1);
    nshift=N+M+1;
    for k=-M:M
       c_s(k+M+1)=0.0;
       for l=-N:N
         c_s(k+M+1)=c_s(k+M+1)+fact*(c_p(l+nshift)*
                                 c_p(k-l+nshift));
       end
    end
    for it=1:nv
       spot(it)=0.0;
       for k=-M:M
          spot(it)=spot(it)+(1-abs(k)/M)*c_s(k+M+1)*
                             exp(1i*tau(it)*const*k);
       end
    end
    spot=real(spot);
```

B.5 Using Fast Fourier Transform Algorithm

The Fourier spot volatility estimator can also be implemented using the *Fast Fourier Transform algorithm* (FFT). The advantage of using FFT is mainly computational. In fact, the FFT reduces the complexity of computing the discrete Fourier transform from $O(n^2)$, which arises if one simply applies the basic algorithm, to $O(n \log n)$, where n is the data size.

Matlab® includes built-in routines `fft(x)` and `ifft(x)` which implement discrete Fourier and inverse transforms, using the algorithm in Cooley and Tukey (1965).

The reference interval is again $[0,T]$ as for algorithm 8. However, due to the specific structure of the FFT algorithm, the time grid at which the price is sampled is $t_i = i(T/n)$, $i = 0,1,\dots,n$, while the time grid at which the spot volatility is reconstructed is given by $\tau_j = (j-1)T/(2M+1)$, $j = 1,2,\dots,2M+1$, where, as in Section B.4, M is the cutting frequency in the spot volatility reconstruction.

The input parameters are: the observed log-prices $p(t_i)$, $i = 0,1,\dots,n$ collected in a vector `P`; the cutting frequencies `N` and `M`; the time horizon `T`. The values of the estimated spot volatility $\widehat{\sigma}^2_{n,N,M}(\tau_j)$ are recorded in the vector `spot` and the Fourier coefficients of the spot volatility are recorded in the vector `C`.

The k-th discrete Fourier transform of the log-return vector `r` is obtained by the Matlab® routine `fft` and recorded in the variables

$$\texttt{fft\_v(k)} = \sum_{j=1}^{n} r(j)\omega_n^{(j-1)(k-1)}, \text{ where } \omega_n = e^{-2\pi \mathrm{i}/n}.$$

Then, the useful Fourier coefficients of the log-returns are collected in the vector `fft_def`.

The spot volatility estimate at time τ_j, $j = 1,\dots,2M+1$, can be obtained by the Matlab® inverse Fourier transform function `ifft(f)` through the following steps

$$\texttt{Fsum(j)} = \sum_{k=-M}^{M} (1-\frac{|k|}{M})c_k(\sigma^2_{n,N})e^{\mathrm{i}\frac{2\pi}{T}k\tau_j}$$

$$= (2M+1)\left[\frac{1}{2M+1}\sum_{h=1}^{2M+1}(1-\frac{|h-M-1|}{M})c_{h-M-1}(\sigma^2_{n,N})\omega_{2M+1}^{-(j-1)(h-1)}\right]\omega_{2M+1}^{(j-1)M}$$

$$= (2M+1)\,\texttt{ifft(f)(j)}\;\omega_{2M+1}^{(j-1)M}$$

Matlab® Algorithm 9

```
function [spot] = FE_spot_vol_FFT(P,T,N,M)

% Computes the spot variance by the Fourier
        estimator
% with Dirichlet kernel and FFT

% Input variables:
%   P       vector of the observed log-prices
%   N       maximum Fourier frequency for price
            returns
```

```
%   M      maximum Fourier frequency for spot
            variance

% Output variables:
%   spot   vector of spot variance at the time grid
        tau

r=diff(P); fft_v=fft(r);
idx=M+N+1:-1:2; ff=fft_v(idx);
fft_def=[conj(ff) fft_v(1:M+N+1)];
fft_def=fft_def./T; % Fourier coeff. of log-returns

idxk=-N:1:N; nshift=M+N+1;
for kk=-M:M
    idxx=idxk+nshift+kk;
    Capp=fft_def(idxx);
    coeff(M+kk+1)=Capp*fft_def(nshift-N:nshift+N)';
end
C=coeff.*(T/(2*N+1)); % Fourier coeff. of variance

k=(-M:1:M);
f=C.*(1-abs(k)/M);
Fsum=(2*M+1)*ifft(f);
Fsum=exp(-1i*2*pi*M*(k+M)/(2*M+1)).*Fsum;
spot=real(Fsum);
```

It is worth noting that the best performance of the fast Fourier algorithm 9 is obtained choosing n, N and M to be a power of two. However, the function works even for values of n, N and M different from a power of two.

B.6 Volatility of Volatility

In this final section, we provide the codes used for the estimation of the volatility of volatility in Section 6.2.3 from empirical data. Observed asset log-prices should be uploaded as a column vector `p` of length `n`. Observation times, spanning the interval $[0,T]$, should be uploaded as a column vector `t`. The integrated volatility of volatility estimates $\hat{\gamma}^2_{n,N,M}$ for $N=1,\dots,N_{max}$ and $M=1,\dots,M_{max}$ are saved in a matrix `VoV(N,M)`. The main code is the following:

```
% Fourier as a function of N and M:
n=length(p); T=t(n); % trading period (Rmk. t(0)=0)
```

```
Nmax=floor(n/2); Mmax=floor(n^0.25); % maximum freq.
VoV=zeros(Nmax,Mmax); Ncut=[1:Nmax]; Mcut=[1:Mmax];
VoV=FEvov_NMfej(p,t,T,n-1,Nmax,Mmax);
```

and it calls the following function:

Matlab® Algorithm 10

```
function Q=FEvov_NMfej(P,t,T,n,Nmax,Mmax)

% Computes the integrated volatility of volatility
           by
% the Fourier estimator with Fejer kernel

% Input variables:
%   P      vector of observed log-prices of
           length n+1
%   t      observation times in [0,T]
%   Nmax   maximum Fourier frequency for price
            returns
%   Mmax   maximum Fourier frequency for spot
            variance

% Output variables:
%   VoV    integrated volatility of volatility
%          as a function of N and M

r=P(2:n+1)-P(1:n); % log-returns
c_p=zeros(1,2*Nmax+1+Mmax); % return Fourier coeff.
c_s=zeros(1,2*Mmax+1); % sigma squared Fourier
         coeff.
VoV=zeros(Nmax,Mmax); const=(2*pi)/T;

% Computes the Fourier coefficients of dp.
        (RMK.c_p(k)
% is the coefficient for s=k-N-1)
for k=1:(2*Nmax+1)+Mmax
    s=k-Nmax-1;
    c_p(k)=sum(exp(-1i*const*s*t(1:end-1)).*r)/T;
end

% Computes the sigma squared coefficients and the
% integrated volatility of volatility:
for N=1:Nmax
```

```
        for k=1:Mmax+1
            c_s(k)=sum(c_p(1+(Nmax-N):Nmax+N+1).*
               c_p(Nmax+N+1+(k-1):-1:(Nmax-N)+1+
                    (k-1)))*
               T/(2*N+1);
        end
        for k=1:Mmax+1
            c_s(k)=1i*(k-1)*c_s(k);
        end
        c_s(Mmax+2:2*Mmax+1)=conj(c_s(2:Mmax+1));

        for M=1:Mmax
            S=[1:M];
            VoV(N,M)=(T^2/(M+1))*(2*sum((1-S/M).*
               c_s(Mmax+2:Mmax+M+1).*c_s(2:M+1)));
            VoV(N,M)=real(VoV(N,M));
        end
    end
```

References

Aït-Sahalia Y, Jacod J (2014) High-Frequency Financial Econometrics. Princeton University Press

Ait-Sahalia Y, Fan J, Xiu D (2010) High-frequency covariance estimates with noisy and asynchronous financial data. Journal of the American Statistical Association 105(492):1504–1517

Akahori J, Liu NL, Mancino ME, Yasuda Y (2016) The Fourier estimation method with positive semi-definite estimators. Working Paper arXiv:14100112

Aldous DJ, Eagleson GK (1978) On mixing and stability of limit theorems. Annals of Probability 6:325–331

Alvarez A, Panloup F, Pontier M, Savy N (2011) Estimation of the instantaneous volatility. Statistical Inference for Stochastic Processes 15:27–59

Andersen T, Bollerslev T (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review 39(4):885–905

Andersen T, Bollerslev T, Diebold F, Labys P (1999a) (understanding, optimizing, using and forecasting) realized volatility and correlation. New York Univ, Sterne School Finance Dept Working paper pp 1–22

Andersen T, Bollerslev T, Lange S (1999b) Forecasting financial market volatility: sample frequency vis-à-vis forecast horizon. Journal of Empirical Finance 6(5):457–477

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

Andersen T, Bollerslev T, Diebold F, Labys P (2001b) The distribution of exchange rate volatility. Journal of the American Statistical Association 96:42–55

Andersen T, Bollerslev T, Diebold F, Labys P (2003) Modeling and forecasting realized volatility. Econometrica 71:579–625

Andersen T, Bollerslev T, Diebold FX (2010) Parametric and nonparametric volatility measurement. Handbook of Financial Econometrics pp 67–137

Andersen T, Bollerslev T, Meddahi N (2011) Realized volatility forecasting and market microstructure noise. Journal of Econometrics 160:220–234

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3

Andersen T, Dobrev D, Schaumburg E (2014) A robust neighborhood truncation approach to estimation of integrated quarticity. Econometric Theory 30:3–59

Andreou E, Ghysels E (2002) Rolling-sample volatility estimators: some new theoretical, simulation and empirical results. Journal of Business & Economic Statistics 20:363–375

Bachelier L (1900) Théorie de la spéculation. Annales Scientifiques de l'Ecole Normale Supérieure, 3e séries, 17

Bandi FM, Russell JR (2006) Separating market microstructure noise from volatility. Journal of Financial Economics 79(3):655–692

Bandi FM, Russell JR (2008) Microstructure noise, realized variance and optimal sampling. Review of Economic Studies 75(2):339–369

Bandi FM, Russell JR (2011) Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Journal of Econometrics 160(1):145–159

Bandi FM, Russel JR, Zhu Y (2008) Using high-frequency data in dynamic portfolio choice. Econometric Reviews 27(1-3):163–198

Barndorff-Nielsen OE, Schmiegel J (2008) Time change, volatility, and turbulence. In: Sarychev A, Shiryaev A, Guerra M, Grossinho MR (eds) Mathematical Control Theory and Finance, Springer, pp 29–53

Barndorff-Nielsen OE, 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 OE, Shephard N (2004) Power and bipower variation with stochastic volatility and jumps (with discussion). Journal of Financial Econometrics 2:1–48

Barndorff-Nielsen OE, Veraart AED (2013) Stochastic volatility of volatility and variance risk premia. Journal of Financial Econometrics 11(1):1–46

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N (2008) Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 6:1481–1536

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N (2011a) Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Journal of Econometrics 162(2):149–169

Barndorff-Nielsen OE, Hansen PR, Lunde A, Shephard N (2011b) Subsampling realised kernels. Journal of Econometrics 160(1):204–219

Barsotti F, Sanfelici S (2014) Firm's volatility risk under microstructure noise. In: Corazza M, Pizzi C (eds) Mathematical and Statistical Methods for Actuarial Sciences and Finance, Springer, pp 55–67

Barucci E, Mancino ME (2010) Computation of volatility in stochastic volatility models with high frequency data. International Journal of theoretical and Applied Finance 13(5):1–21

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

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

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

Barucci E, Magno D, Mancino ME (2012) Fourier volatility forecasting with high frequency data and microstructure noise. Quantitative Finance 12(2):281–293

Bekaert G, Wu G (1997) Asymmetric volatility and risk in equity markets. NBER Working Paper w6022

Black F (1975) Fact and Fantasy in the use of Options . Financial Analyst Journal 31(4):36–41

Black F (1976) Studies of stock market volatility changes. In Proceedings of the Business and Economic Statistic Section, American Statistical Association pp 177–181

Black F, Scholes M (1973) The price of options and corporate liabilities. Journal of Political Economy 81:637–659

Bloomfield P (2000) Fourier Analysis of Time Series. An Introduction. Wiley Series in Probability and Statistics

Bollerslev T, Zhang L (2003) Measuring and modeling systematic risk in factor pricing models using high-frequency data. Distribution of realized stock return volatility. Journal of Empirical Finance 10:533–558

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

Bollerslev T, Gibson M, Zhou H (2011) Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities. Journal of Econometrics 160:235–245

Bouchaud JP, Potters M (2003) Theory of financial risk and derivative pricing: from statistical physics to risk management. Cambridge University Press, Cambridge

Brandt MW, Diebold FX (2006) A no-arbitrage approach to range-based estimation of return covariances and correlations. Journal of Business 79:61–73

Britten-Jones M, Neuberger A (2000) Option prices, implied price processes, and stochastic volatility. Journal of Finance 55(2):839–866

Cartea A, Jaimungal S (2015) Optimal execution with limit and market orders. Quantitative Finance 15(8):1279–1291

Christensen K, Kinnebrock S, Podolskij M (2010) Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. Journal of Econometrics 159:116–133

Christie A (1982) The stochastic behavior of common stock variances. Journal of Financial Econometrics 10:407–432

Chronopoulou A, Viens F (2012) Estimation and pricing under long-memory stochastic volatility. Annals of Finance 8(2-3):379–403

Clement E, Gloter A (2011) Limit theorems in the Fourier transform method for the estimation of multivariate volatility. Stochastic Processes and Their Applications 121:1097–1124

Cohen KJ, Hawawini GA, Maier SF, Schwartz RA, Whitcomb DK (1983) Friction in the trading process and the estimation of systematic risk. Journal of Financial Economics 12:263–278

Comte F, Renault E (1998) Long memory in continuous-time stochastic volatility models. Mathematical Finance 8(4):291–323

Cont R, De Larrard A (2013) Price dynamics in a Markovian limit order market. SIAM Journal on Financial Mathematics 4(1):1–25

Cooley JW, Tukey JW (1965) An algorithm for the machine computation of the complex Fourier series. Mathematics of Computation 19:297–301

Corsi F, Audrino F (2010) Realized correlation tick-by-tick. Computational Statistics and Data Analysis 54(11):2372–2382

Corsi F, Pirino D, Renó R (2015) Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics 159:276–288

Cox J, Ross S (1976) The valuation of options for alternative stochastic processes. Journal of Financial Economics 3:145–166

Cox JC, Ingersoll JE, Ross SA (1985) A theory of the term structure of interest rates. Econometrica 53:385–408

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

Curato IV, Sanfelici S (2015) Measuring the leverage effect in a high frequency framework. In: Gregoriou GN (ed) The Handbook of High Frequency Trading, Elsevier, Amsterdam: North-Holland, pp 425–446

Curato IV, Mancino ME, Recchioni MC (2016) Spot volatility estimation using the Laplace transform. Forthcoming Econometrics and Statistics, available at http://ssrncom/abstract=2572340

Cvitanic J, Lipster R, Rozovskii B (2006) A filtering approach to tracking volatility from prices observed at random times. The Annals of Applied Probability 16(3):1633–1652

Dacorogna M, Gençay R, Müler UA, Olser RB, Pictet OV (2001) An introduction to high-frequency finance. Academic Press

De Jong F, Nijman T (1997) High frequency analysis of lead-lag relationships between financial markets. Journal of Empirical Finance 4:259–277

De Pooter M, Martens M, van Dijk D (2008) Predicting the daily covariance matrix for S&P100 stocks using intraday data: but which frequency to use? Econometric Reviews 27:199–229

Derman E, Kani I (1994) Riding on the smile. RISK 7:32–39

Dimson E (1979) Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7:197–226

Dupire B (1994) Pricing with a smile. RISK 7:18–20

Engle R (2000) The econometrics of ultra-high-frequency data. Econometrica 68(1):1–22

Engle R, Colacito R (2006) Testing and valuing dynamic correlations for asset allocation. Journal of Business & Economic Statistics 24(2):238–253

Epps T (1979) Comovements in stock prices in the very short run. Journal of the American Statistical Association 74:291–298

Fan J, Wang Y (2008) Spot volatility estimation for high frequency data. Statistics and its Interface 1:279–288

Fisher M, Nappo G (2010) On the moments of the modulus of continuity of it diffusions. Stochastic Analysis and Applications 28, 103–122, (2010) 28:103–122

Fleming J, Kirby C, Ostdiek B (2001) The economic value of volatility timing. The Journal of Finance LVI, 1:329–352

Fleming J, Kirby C, Ostdiek B (2003) The economic value of volatility timing using realized volatility. Journal of Financial Economics 67:473–509

Florens-Zmirou D (1993) On estimating the diffusion coefficient from discrete observations. Journal of Applied Probability 30:790–804

Foster DP, Nelson DB (1996) Continuous record asymptotics for rolling sample variance estimators. Econometrica 64:139–174

Fourier J (1822) Théorie analytique de la chaleur. Firmin Didot Père et Fils, Paris

French KR, Schwert GW, Stambaugh R (1987) Expected stock returns and volatility. Journal of Financial Economics 19:3–29

Frey R, Stremme A (1997) Market volatility and feedback effects from dynamic hedging. Mathematical Finance 7:351–374

Gatheral J, Oomen RCA (2010) Zero-intelligence realized variance estimator. Finance & Stochastic 14(2):249–283

Genon-Catalot V, Laredo C, Picard D (1992) Nonparametric estimation of the diffusion coefficient by wavelet methods. Scandinavian Journal of Statistics 19:317–335

Ghysels E, Sinko A (2011) Volatility forecasting and microstructure noise. Journal of Econometrics 160(1):257–271

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

Goodhart CAE, O'Hara M (1997) High frequency data in financial markets: Issue and applications. Journal of Empirical Finance 4:73–114

Griffin JE, Oomen RCA (2011) Covariance measurement in presence of non-synchronuous trading andmarket microstructure noise. Journal of Econometrics 160(1):58–68

Han CH, Liu W, Chen TY (2014) VaR/CVaR estimation under stochastic volatility models. International Journal of Theoretical and Applied Finance 17(2):1450,009

Hannan EJ (1970) Multiple Time Series. John Wiley and Sons

Hansen PR, Lunde A (2006a) Consistent ranking of volatility models. Journal of Econometrics 131:97–121

Hansen PR, Lunde A (2006b) Realized variance and market microstructure noise (with discussions). Journal of Business and Economic Statistics 24:127–161

Harris FHd, McInish TH, Shoesmith GL, Wood RA (1995) Cointegration, error correction, and price discovery on informationally linked security markets. Journal of Financial and Quantitative Analysis 30:563–579

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

Hasbrouck J (1996) Modeling market microstructure time series. In: Maddala GS, Rao CR (eds) Handbook of Statistics, Vol. 14, Elsevier, Amsterdam: North-Holland, pp 647–692

Hayashi T, Yoshida N (2005) On covariance estimation of nonsynchronously observed diffusion processes. Bernoulli 11(2):359–379

Hobson DG, Rogers LCG (1998) Complete models with stochastic volatility. Mathematical Finance 8:27–48

Hoffmann M (1999) L_p estimation of the diffusion coefficient. Bernoulli 5(3):447–481

Hoshikawa T, Kanatani T, Nagai K, Nishiyama Y (2008) Nonparametric estimation methods of integrated multivariate volatilities. Econometric Reviews 27(1):112–138

Hull J, White A (1987) The pricing of options on assets with stochastic volatilities. Journal of Finance 42:281–300

Inkaya A, Yolcu Ocur Y (2014) Analysis of volatility feedback and leverage effects on the ISE30 index using frequency data. Journal of Computational and Applied Mathematics 259:377–384

Jacod J (2000) Non-parametric kernel estimation of the coefficient of a diffusion. Scandinavian Journal of Statistics 27:83–96

Jacod J (2008) Asymptotic properties of realized power variations and related functionals of semimartingales. Stochastic Processes and their Applications 118(4):517–559

Jacod J, Rosenbaum M (2013) Quarticity and other functionals of volatility: efficient estimation. The Annals of Statistics 41:1462–1484

Jacod J, Shiryaev AN (2003) Limit Theorems for Stochastic Processes. 2nd ed. Springer-Verlag, New York

Jacod J, Li Y, Mykland PA, Podolskij M, Vetter M (2009) Microstructure noise in the continuous case: the pre-averaging approach. Stochastic Processes and their Applications 119:2249–2276

Jacquier E, Polson NG, Rossi PE (1994) Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12(4):371–389

Jiang GJ, Oomen RCA (2008) Testing for jumps when asset prices are observed with noise - a “swap variance” approach. Journal of Econometrics 144(2):352–370

Kanaya S, Kristensen D (2015) Estimation of stochastic volatility models by nonparametric filtering. Cemmap working paper CWP09/15

Kenmoe R, Sanfelici S (2014) An application of nonparametric volatility estimation to option pricing. Decisions Econ Finance 37(2):393–412

Kercheval A, Zhang Y (2015) Modelling high-frequency limit order book dynamics with support vector machines. Quantitative Finance 15(8):1315–1329

Kristensen D (2010) Nonparametric filtering of the realized spot volatility: a kernel-based approach. Econometric Theory 26:60–93

Kunita H (1988) Stochastic Flows and Stochastic Differential Equations. Cambridge University Press

Lee SS, Mykland PA (2012) Jumps in equilibrium prices and market microstructure noise. Journal of Econometrics 168:396–406

Lévy P (1937) Theorie de l'Addition des variables alatoires. Gauthier-Villars Paris

Li Y, Mykland PA (2014) Rounding errors and volatility estimation. Journal of Financial Econometrics 13(2):478–504

Liu NL, Mancino ME (2012) Fourier estimation method applied to forward interest rates. JSIAM Letters 4:17–20

Liu NL, Ngo HL (2014) Approximation of eigenvalues of spot cross volatility matrix with a view towards principal component analysis. Working paper available at https://arxivorg/pdf/14092214

Malliavin P (1995) Integration and Probability. Springer-Verlag

Malliavin P, Mancino ME (2002a) Fourier series method for measurement of multivariate volatilities. Finance and Stochastics 4:49–61

Malliavin P, Mancino ME (2002b) Instantaneous liquidity rate, its econometric measurement by volatility feedback. CRAS Paris 334:505–508

Malliavin P, Mancino ME (2009) A Fourier transform method for nonparametric estimation of volatility. The Annals of Statistics 37(4):1983–2010

Malliavin P, Thalmaier A (2006) Stochastic Calculus of Variations in Mathematical Finance. Springer

Malliavin P, Mancino ME, Recchioni MC (2007) A non parametric calibration of HJM geometry: an application of Itô calculus to financial statistics. Japanese Journal of Mathematics 2:55–77

Mancini C (2009) Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics 36(2):270–296

Mancini C, Mattiussi V, Renò R (2015) Spot volatility estimation using delta sequences. Finance and Stochastics 19:261–293

Mancino ME, Recchioni MC (2015) Fourier spot volatility estimator: asymptotic normality and efficiency with liquid and illiquid high-frequency data. PLoS ONE 10(9), URL e0139041.doi:10.1371/journal.pone.0139041

Mancino ME, Renò R (2005) Dynamic principal component analysis of multivariate volatility via Fourier analysis. Applied Mathematical Finance 12(2):187–199

Mancino ME, Sanfelici S (2008) Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise. Computational Statistics and Data Analysis 52(6):2966–2989

Mancino ME, Sanfelici S (2011a) Covariance estimation and dynamic asset allocation under microstructure effects via Fourier methodology. In: Gregoriou GN, Pascalau R (eds) Handbook of Econometrics, Palgrave-MacMillan, London, UK

Mancino ME, Sanfelici S (2011b) Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise. Journal of Financial Econometrics 9(2):367–408

Mancino ME, Sanfelici S (2011c) Multivariate volatility estimation with high frequency data using Fourier method. In: Florescu I, Mariani M, Viens F (eds) Handbook of Modeling High-Frequency Data in Finance, Wiley, New York

Mancino ME, Sanfelici S (2012) Estimation of quarticity with high frequency data. Quantitative Finance 12(4):607–622

Mandelbrot B, Van Ness J (1968) Fractional Brownian Motions, Fractional Noises and Applications . SIAM Review 10:422437

Martens M (2004) Estimating unbiased and precise realized covariances. EFA 2004 Maastricht Meetings Paper 4299

Mattiussi V, Iori G (2010) A nonparametric approach to estimate volatility and correlations dynamics. Working Paper City University London

Meddahi N (2001) An eigenfunction approach for volatility modeling. Working paper of University of Montreal available at https://gremaquniv-tlse1fr/perso/meddahi/29-2001-cahpdf

Muller HG, Sen R, Stadtmuller U (2011) Functional data analysis for volatility. Journal of Econometrics 165:233–245

Mykland PA (2012) A Gaussian calculus for inference from high frequency data. Annals of Finance 8:235–258

Mykland PA, Zhang L (2008) Inference for volatility-type objects and implications for hedging. Statistics and Its Interface 1:255–278

Mykland PA, Zhang L (2009) Inference for continuous semimartingales observed at high frequency. Econometrica 77:1403–1445

Nelson D (1990) Arch models as diffusion approximations. Journal of Econometrics 45:7–38

Nelson D (1991) Conditional heteroskedasticity in asset returns. a new approach. Econometrica 59:347–370

Nielsen MO, Frederiksen PH (2008) Finite sample accuracy and choice of sampling frequency in integrated volatility estimation. Journal of Empirical Finance 15(2):265–286

Ogawa S, Sanfelici S (2011) An improved two-step regularization scheme for spot volatility estimation. Economic Notes 40:107–134

O'Hara M (1995) Market Microstructure Theory. Blackwell

Øksendal B (1995) Stochastic Differential Equations. (4th Ed.). Springer Verlag. Berlin Heidelberg

Oya K (2005) Measurement of volatility of diffusion processes with noisy high frequency data. Proceedings of MODSIM05 available at wwwmssanzorgau/modsim05/papers/oyapdf

Papantonopoulos G, Takahashi K, Bountis T, Loos BG (2013) Mathematical modeling suggests that periodontitis behaves as a nonlinear chaotic dynamical process. Journal of Periodontology 84:e29–e39

Park S, Hong SY, Linton O (2016) Estimating the quadratic covariation matrix for an asynchronously observed continuous time signal masked by additive noise. Journal of Econometrics 191(2):325–347

Pasquale M, Renò R (2005) Statistical properties of trading volume depending on size. Physica A 346:518–528

Platen E, Schweizer M (1998) On feedback effects from hedging derivatives. Mathematical Finance 8:67–84

Precup OV, Iori G (2004) A comparison of high-frequency cross-correlation measures. Physica A 344:252–256

Precup OV, Iori G (2007) Cross-correlation measures in the high-frequency domain. European Journal of Finance 13(4):319–331

Priestley MB (1983) Spectral Analysis and Time Series. Academic Press

Protter P (1992) Stochastic Integration and Differential Equations – A new Approach. Springer Verlag

Renò R (2008) Nonparametric estimation of the diffusion coefficient of the stochastic volatility models. Econometric Theory 24:1174–1206

Renò R, Rizza R (2003) Is volatility lognormal? Evidence from Italian futures. Physica A 322:620–628

Revuz D, Yor M (1991) Continuous Martingales and Brownian Motion. Springer Verlag, Berlin Heidelberg

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

Rubinstein M (1994) Implied binomial trees. Journal of Finance 69(3):771–818

Sanfelici S, Uboldi A (2014) Assessing the quality of volatility estimators via option pricing. Studies in Nonlinear Dynamics & Econometrics 18(2):103–124

Sanfelici S, Curato IV, Mancino ME (2015) High frequency volatility of volatility estimation free from spot volatility estimates. Quantitative Finance 15(8):1–15

Scholes M, Williams J (1997) Estimating betas from nonsynchronous data. Journal of Financial Economics 5:309–327

Stein E, Stein J (1991) Stock price distributions with stochastic volatility: an analytic approach. Review of Financial Studies 4:727–752

Todorov V, Tauchen G (2012) The realized Laplace transforms of volatility. Econometrica 80:1105–1127

Vetter M (2015) Estimation of integrated volatility of volatility with applications to goodness-of-fit testing. Bernoulli

Voev V, Lunde A (2007) Integrated covariance estimation using high-frequency data in the presence of noise. Journal of Financial Econometrics 5(1):68–104

Wang F (2014) Optimal design of Fourier estimator in the presence of microstructure noise. Computational Statistics & Data Analysis 76:708–722

Zhang L (2009) Estimating covariation: Epps effect, microstructure noise. Journal of Econometrics 160(1):33–47

Zhang L, Mykland P, 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 B (1996) High frequency data and volatility in foreign-exchange rates. Journal of Business and Economic Statistics 14(1):45–52

Zu Y, Boswijk HP (2014) Estimating spot volatility with high-frequency financial data. Journal of Econometrics 18:117–135

Index

M.E. Mancino et al., *Fourier-Malliavin Volatility Estimation*,
SpringerBriefs in Quantitative Finance, DOI 10.1007/978-3-319-50969-3