SCUOLA
NORMALE
SUPERIORE

Classe di Scienze

Corso di perfezionamento in
Metodi computazionali e modelli matematici per le scienze e la finanza

35° ciclo

Spot volatility and covariance estimation on high frequency data using the Fourier methodology: microstructure noise, Limit Order Book and positive semi-definiteness

Settore Scientifico Disciplinare **SECS-S/06**

Candidato
dr. Tommaso Mariotti

Relatrice
Prof.ssa Maria Elvira Mancino

Supervisione interna
Prof. Fabrizio Lillo

Anno accademico 2022–2023

Contents

Chapter 1

Introduction

The field of non-parametric volatility estimation[1] attracted a huge interest in the recent literature on financial econometrics, in particular since the availability of high frequency data paved the way to the study of daily measures of volatility and intra-day volatility patters (see, e.g., Andersen and Bollerslev (1997)), focusing on quantities such as the integrated and, more recently, the spot (or instantaneous) volatility. The interest was accompanied by the discovery that, while offering new opportunities, high-frequency data also pose new challenges to the well-studied topic of volatility estimation. In particular the presence of market microstructure noise, i.e., a disturbance component originated by the features of trading in financial markets that renders traditional non-parametric estimators inconsistent (see, e.g., Bandi and Russell (2008)). While the market microstructure has been extensively studied in the last decades of the past century focusing on centralized forms of exchange with different competing classes of market participants (see Hasbrouck (2007) for a review), with the advent of electronic exchanges the interest shifted towards the widespread Limit Order Book structure, i.e. a double auction mechanism in which orders are queued and executed according to a price priority. Independently of the different approaches, both the classical agent-based models and the newer models for the LOB dynamics are characterize by the presence of features that drive the observed price away from the usual assumptions of stochastic modeling, being influenced for example by bid-ask spread, rounding and sequential and strategic trading of market participants. In addition to this microstructure friction, the presence of irregularly spaced and, most importantly, in a multivariate setting, possibly asynchronous observations, creates additional difficulties that need to be solved. To address the main issues raised by the use of high-frequency data[2], a variety of methods have been proposed in the literature (see, e.g., Jacod (2019) for a review), many of whom are briefly recalled in this work, in particular in Section 3.2, while other issues, such as the positive semi-definiteness of estimated spot volatility matrix, have not been fully addressed so far.

Among the several techniques proposed to reconstruct the intra-day volatility path, this work focus in particular on the Fourier method for spot volatility estimation, originally presented in Malliavin and Mancino (2002) and Malliavin and Mancino (2009). In comparison to other proposed estimators, the Fourier one may offer some advantages: its theoretical foundation lies on well known

[1] See, e.g., Andersen et al. (2010) for an introduction to the topics of integrated and spot volatility, and for the distinction between the parametric and non-parametric approach.

[2] Other than microstructure noise, another aspect that attracted the attention of the literature is the presence of jumps, i.e. discontinuities in the price and/or volatility process, but this aspect will not be addressed in the present work. Moreover, a new strand of literature is born, that tends to question the presence of intra-day discontinuities at ultrahigh-frequencies, see, e.g. Christensen et al. (2014).

results of Fourier analysis, empirical evidence shows that it does not require pre-manipulation of the data to be applied to noisy price series, it is able to manage irregularly spaced observations without any adjustment and it is expression of a more general method that may be applied not only to volatility, but also to estimate other relevant financial quantities, such as the quarticity, the volatility of volatility and the leverage (see Mancino et al. (2017) for an in dept discussion of the topic). The Fourier estimator reconstructs the instantaneous volatility as a series expansion, with coefficients estimated from the Fourier coefficients of the log-price increment. Specifically, for a given sample size n, the efficient implementation of the Fourier spot volatility estimator requires the balancing of only two parameters: N, the cutting frequency in the convolution formula that reconstructs the Fourier coefficients of the volatility, and M, the number of Fourier coefficients of the volatility to be used in the Fourier inversion formula, which yields the estimated spot volatility path.

This work aims at further developing the literature on the Fourier estimator, analyzing in particular the spot estimator with high-frequency data which are contaminated by microstructure noise, an aspect that was only partially addressed by previous works, such as Mancino and Recchioni (2015). The advances that this work presents with respect to the existing literature are related to three outstanding issues, involving both theoretical and practical aspects of spot volatility estimation. The first point considered is the need, from a theoretical point of view, of providing a Central Limit Theorem in presence of noise for the Fourier spot estimator, thus obtaining a convergence rate and an asymptotic error, which are still unknown, for the estimator under specific assumptions for the noise. A side quest related to the findings of the above mentioned CLT is to establish a feasible procedure to optimize the parameters involved in the estimator itself, together with the assessment of its performance. As second point, considering the needs of the applications, we stress the relevance, in presence of a vast literature on volatility estimation, of a broad analysis of the finite sample performance of both the Fourier and competing volatility estimators proposed in other works, with the accuracy of estimations evaluated in a setting that should be as realistic as possible and in relation to applications that might be of real interest for finance practitioners. Finally, as third pillar of this work, we also underline the interest of considering a multivariate setting for spot volatility estimation, that could be of paramount importance in the rising field of principal-component analysis with high-frequency data, focusing in particular on the need, both theoretical and practical, of obtaining, in the context of irregularly spaced and asynchronous observations, positive semi-definite estimations of the spot volatility matrix using a Fourier-type estimator. These three open problems, and the methods and techniques employed to solve them, are at the core of the three main chapters in which this work is divided.

In Chapter 2 first the definition of the Fourier spot estimator is recalled, then the results of our asymptotic study are presented. In particular the asymptotic normality of the estimator is derived in presence of i.i.d. noise when the underlying price follows a continuous Itô semimartingale, with an optimal convergence rate of $n^{1/8}$, together with a Central Limit Theorem in absence of noise, where the optimal convergence rate of $n^{1/4}$ is attained, refining the results of Mancino and Recchioni (2015); moreover an additional version of the previous results, with slightly suboptimal convergence rate but smaller asymptotic variance, is provided. The analysis shows that, in the presence of noise, efficiency is achieved by carefully reducing N, that is, by cutting out the highest frequencies in the construction of the Fourier coefficients of the volatility. This way, the Fourier

method allows obtaining a noise-robust estimator of the instantaneous volatility without the need for any manipulation of the original observations or any bias correction. Building on the asymptotic result, a feasible procedures to fine tune the frequencies N and M involved in the computation of the estimator is presented, based on a minimization of the asymptotic variance of the estimation error, which can be deemed observable as it depends on quantities that can be consistently estimated with high-frequency prices, namely, the integrals of the volatility, the quarticity and the volatility of volatility (all of which can be estimated with Fourier-type estimators) along with the variance of the noise. While the introduction of a feasible estimation strategy, rooted in asymptotic theory, represents the main innovation of the finite-sample study conducted in this chapter, the study is extended by a detailed finite-sample study of the performance of the spot Fourier estimator when the proposed tuning procedure is applied, both with simulated and real data. In both of the cases the results show the satisfactory accuracy of the estimation using price series with different microstructure characteristics, and it is shown that the estimator may have an edge with respect to some competitors proposed in the literature.

To extend and deepen the assessment of the finite-sample performances of alternatives volatility estimators in the presence of microstructure noise is one of the main aims of Chapter 3. Here, however, we notice that having the best asymptotic properties, such as the optimal rate of convergence and the minimum asymptotic error variance, do not guarantee the best performance of the estimator in finite-sample applications. To gain a more realistic comparison of the performances of the estimators we chose to use a Limit Order Book generator as a model to produce simulations, and under this point of view we extend the study by Gatheral and Oomen (2010) in two directions: first, by considering the more realist queue-reactive model for the LOB, proposed by Huang et al. (2015), and second by including in the analysis an enlarged collection fo both integrated and spot volatility estimators. For what concerns the first point indeed we show that, when calibrated to real data, the queue-reactive model is able to better reproduce the feature of high-frequency data with respect to the zero-intelligence model used in Gatheral and Oomen (2010), in particular when looking at the average spread and at the noise accumulation pattern produced. For the second point it is worth underlying that we consider three price series for the exercise: the mid-price, that is, the average between the best bid and best ask quotes; the micro-price, i.e., the volume-weighted average of the best bid and best ask quotes; the trade price, namely the price at which a market order is executed, all extracted from the simulated LOB. The analysis is conducted on different scenarios, in which the novelty of time-varying parameters for the model is introduced, and shows that for the spot volatility the Fourier estimators seems confirm its edge in accuracy, while for integrated volatility the best estimator may be different depending on the considered price series, with the Fourier still performing satisfactory. The second goal of the chapter is to study the impact of the availability of efficient volatility estimates on optimal execution. Specifically, being able to include the strategy of execution directly in the simulation of the LOB, we show the significance of the deviations between the predictions of the variance of a VWAP execution when different spot estimator are employed.

In Chapter 4 the focus of the analysis is enlarged, shifting from the estimation of the volatility of a single asset to the estimation of a volatility matrix in a multivariate setting that, considering also the correlations between assets, is of great interest in applications with high-frequency data. To guarantee symmetry and positive semi-definiteness of the estimated matrix in presence of asynchronous observations a new estimator, called PDF (positive definite Fourier), is introduced, and its

positive semi-definiteness is proved. The introduction of the new estimator is necessary, since the Fourier spot estimator, in this setting, has problems of symmetry, thus leading to possibly complex eigenvalues of the volatility matrix. The difficulties of obtaining positive semi-definite estimated spot volatility matrices is not limited to Fourier estimator, but, on the contrary of what established for integrated volatility (see, e.g., Mancino and Sanfelici (2011)), is a widespread issue for spot volatility estimator (see, e.g., Bibinger et al. (2019)). The finite-sample performances are evaluated starting with a sensitivity analysis of the estimation error to the choice of N and M, and it is found out that the preferable level of the cutting frequency N is lower than in the original case, this being influenced by the asynchronicity in the observation scheme. The accuracy of the estimator and its ability to produce positive estimations is then evaluated in an extensive numerical study, that include a variety of noise specification and different models for the efficient price process. In this study, its seems that the PDF estimator is able to obtain a gain in efficiency with respect to competing estimators able to manage asynchronous observations, while always producing the desired positive semi-definite estimations; at the same time, the influence of noise shape and intensity on the ability of the other estimators to produce positive estimations is analyzed. The chapter is concluded by introducing an alternative implementation of the estimator that, being based on a randomization technique, may reduce the required computational time, while maintaining positivity. Analyzing the impact of the different parameters on the performances of the two algorithms is found out that, however, this alternative approach entails, in the conducted simulated exercise, a trade-off between execution speed and accuracy, where the former can be increased at the sacrifice of the latter.

Finally, Chapter 5 summarizes the results of the work and concludes, while supplementary and auxiliary materials are presented, when needed, in the appendices of each chapter.

The results that constitute Chapter 3 have been published on *Quantitative Finance*, see Mariotti et al. (2023), while the results presented in Chapters 2 and 4 have been collected in two papers that are both currently submitted to peer-reviewed journals.

Chapter 2

Asymptotic normality and finite-sample robustness of the Fourier spot volatility estimator in the presence of microstructure noise

Even though the non-parametric estimation of the time-varying volatility of financial assets has long been recognized as a relevant topic (see, e.g., Andersen et al. (2003) and the references therein) in the last couple of decades, the increasing availability of high-frequency market data has given an enormous impulse to the investigation of novel estimation methodologies and the exploration of their applications. Besides, the focus has shifted from the estimation of the volatility accumulated over a given time horizon, i.e., the integrated volatility, to the reconstruction of the trajectory of the volatility process, i.e., the instantaneous (or spot) volatility. We refer to Aït-Sahalia and Jacod (2014) for an extensive review of the literature.

The benchmark for estimating the integrated volatility of an asset on a fixed time interval with high-frequency prices is given by the sum of the squared log-returns, i.e., the realized volatility. In the limit as the time between two consecutive observations tends to zero, the realized volatility converges to the quadratic variation of the asset log-price process, and its derivative provides the instantaneous volatility thereof. However, the differentiation of the quadratic variation may generate appreciable numerical instabilities. Moreover, high-frequency prices are affected by the presence of the noise due to market microstructure, which causes a discrepancy between asset pricing theory, based on semimartingales, and the actual price observations, sampled at very fine intervals.

In order to reduce the numerical instabilities due to the differentiation of the quadratic variation, authors have extensively adopted kernel-smoothing procedures, see Kristensen (2010) for a study of the asymptotic properties of kernel-based spot estimators. Moreover, modifications of realized spot estimators have been proposed to correct for the bias introduced by the presence of microstructure noise. Such modifications, which rely on the pre-processing of price observations, include the Two-Scale sub-sampling method by Zu and Boswijk (2014) and the Pre-Averaging method for kernel-based estimators by Aït-Sahalia and Jacod (2014) and Figueroa-López and Wu (2022). This chapter studies the asymptotic error normality and the finite-sample robustness, in the presence of microstructure noise, of an alternative non-parametric estimator of the instantaneous volatility, namely, the Fourier

estimator introduced by Malliavin and Mancino (2002).

The asymptotic properties of the Fourier volatility estimator in the absence of noise have been investigated by several papers. The weak convergence of the asymptotic error distribution of the Fourier spot covariance estimator was first derived in Malliavin and Mancino (2009), with a sub-optimal rate. Clement and Gloter (2011) then proved the asymptotic normality of the Fourier integrated covariance estimator with the optimal rate of convergence and asymptotic-error variance, under the assumption of irregular and non-synchronous observations. A Central Limit Theorem for the Fourier spot volatility estimator was proved in Mancino and Recchioni (2015), with a slightly sub-optimal rate of convergence. Besides, Cuchiero and Teichmann (2015) proposed a spot covariance estimator based on a modification of the classical Fourier method and studied its asymptotic normality of in the presence of jumps.

In the presence of microstructure noise, several studies have investigated the finite-sample performance of the Fourier method. For what concerns the Fourier estimator of the integrated variance, see, e.g., Barucci and Reno (2002), Hansen and Lunde (2006), Nielsen and Frederiksen (2008), Mancino and Sanfelici (2008). As for the Fourier spot volatility estimator, its robustness to noise was first investigated in Mancino and Recchioni (2015), where the authors performed a thorough simulation study that included auto-correlated and price-dependent noise patterns.

The previous studies have explored the finite-sample efficiency and robustness of the Fourier estimator of the volatility in the presence of noise, without investigating the asymptotic properties of the estimator, namely the rate of convergence and the asymptotic-error variance. This work fills this gap and provides novel results on the asymptotic normality of the Fourier estimator, which show that the latter reaches the optimal rate of convergence and asymptotic-error variance in the presence of additive microstructure noise. Specifically, in the novel asymptotic study presented in this chapter, under the assumption that the volatility process is also a continuous Itô semimartingale, we prove that the Fourier spot volatility estimator attains the optimal rate of convergence in the absence of noise, that is, $n^{1/4}$. Then, we prove that the Fourier estimator attains the optimal rate of convergence also in the presence of additive microstructure noise, that is, $n^{1/8}$. This is the convergence rate of the spot volatility estimators in Aït-Sahalia and Jacod (2014, Section 8.7) and Figueroa-López and Wu (2022, Theorem 2.2), while the Two-scale estimator in Zu and Boswijk (2014, Theorem 2) attains the convergence rate $n^{1/12}$. However, while all these estimators require a preliminary manipulation of price observations and a bias correction, it is worth remarking that the Fourier estimator does not need any modification to be efficient in the presence of noise. Indeed, as mentioned, it only requires a proper reduction of the cutting frequency N. Precisely, in the absence of noise, the choice of N that optimizes the rate of convergence and the asymptotic-error variance is the Nyquist frequency $n/2$ (see Theorem 2.2.1); instead, in the presence of noise, the cutting frequency N has order $n^{1/2}$ (see Theorem 2.2.7). The optimal convergence rate is obtained by choosing $M = O(N^{1/2})$ both in the absence and the presence of noise.

The convolution formula (2.3) exploits the contribution of weighted realized auto-covariances to deal with the presence of noise, with N determining the weight. The role of realized auto-covariances has been studied in the early work by Zhou (1996) and, later, in Barndorff-Nielsen et al. (2008). In this spirit, the Fourier spot volatility estimator can be related to a local version of the infinite-lag realized kernel estimator, see Barndorff-Nielsen et al. (2008). A different but related study is present in Park et al. (2016), where the authors prove the consistency and asymptotic normality for a

class of Fourier-type estimators of the Fourier coefficients of the covariance, named Fourier Realized Kernels. The Central Limit Theorem for any Fourier coefficient is derived under some general conditions that allow for microstructure noise effects and asynchronicity. The rate of convergence depends on the relative liquidity between assets and reaches $n^{1/5}$ if the assets are synchronous. The rate of convergence of the spot Fourier Realized Kernel estimator is not obtained. Moreover, the estimator studied in Mancini et al. (2015), Proposition 4.2, even if referred to by the authors as a Fourier-type estimator, substantially differs from the Fourier spot volatility estimator, in that it is actually a Féjèr-kernel based realized estimator. In fact, the authors get rid of the cross products of the log-returns (which are the main by-product of the original convolution formula, see Malliavin and Mancino (2009)), thereby loosing the robustness to microstructure noise.

For the selection of the cutting frequencies N and M in the presence of noise, we introduce a feasible adaptive method that minimizes the asymptotic-error variance of the Fourier estimator. Specifically, this method relies on the gradient-descent algorithm and optimizes the conditional mean integrated squared error, which is approximated, at high-frequencies, by the integrated asymptotic-error variance (see Zu and Boswijk (2014, Section 3.4)). This improves the finite-sample analysis provided in Mancino and Recchioni (2015, Section 4), where the authors adopt an indirect inference procedure for selecting N and M, based on the testing of the empirical distribution of the log-returns standardized by volatility estimates.

We evaluate the robustness of the feasible adaptive method for selecting N and M in four simulated scenarios, which generate incrementally more realistic noise patterns. Additionally, in the different simulated scenarios, we compare the finite-sample performance of the Fourier estimator with that of alternative noise-robust spot volatility estimators. The results of the simulation study support the robustness of the feasible method for selecting N and M and the relative finite-sample efficiency of the Fourier estimator. We begin by simulating a stylized scenario, which assumes an i.i.d. Gaussian noise, a constant leverage effect, a perfect positive correlation between the volatility of volatility and the volatility, and an equally-spaced price sampling grid. Then we progressively allow for stochastic leverage and volatility of volatility, auto-correlated and price-dependent noise with Student-t innovations and, finally, random sampling. obtaining numerical evidence that suggests that the feasible procedure is satisfactorily accurate and robust to the more realistic scenarios. Finally, we test of the feasible method for selecting N and M in an empirical exercise, conducted with tick-by-tick prices of three US stocks with different liquidity, sampled over the period December 1, 2022 - December 30, 2022. The results of the test suggest that the feasible method is satisfactorily accurate with real market prices, for the different liquidity levels considered, and the same results are obtained using a stock index, namely the S&P500 index.

The chapter is organized as follows. Section 2.1 contains the definition of the Fourier estimator of the instantaneous volatility. Novel asymptotic results are detailed in Section 2.2. Section 2.3 illustrates the simulation study, while Section 2.4 contains the empirical exercise. The Appendix 2.A contains some auxiliary results on the Dirichlet and Féjèr kernels.

2.1 Fourier estimator of the spot volatility

The Fourier estimator of the spot volatility was originally proposed in Malliavin and Mancino (2002). Its implementation consists in three steps. As a first step, one computes the discrete Fourier coeffi-

cients of the log-price increment. The second step then entails computing the convolution of these coefficients, which yields the Fourier coefficients of the volatility. Finally, the last step involves the reconstruction of the volatility path via the Fourier-Féjèr inversion formula.

For simplicity, we consider the estimation interval $[0, 2\pi]$. By scaling the unit of time, one can immediately obtain the arbitrary interval $[0, T]$. Suppose that the asset log-price, denoted by p, is observed at discrete, irregularly-spaced points in time: $\{0 = t_{0,n} \leq \dots t_{i,n} \dots \leq t_{n,n} = 2\pi\}$. For brevity, we will omit the second index n. Define $\rho(n) := \max_{0 \leq h \leq n-1} |t_{h+1} - t_h|$ and suppose that $\rho(n) \to 0$ as $n \to \infty$.

Consider the following interpolation formula

$$p_n(t) := \sum_{i=0}^{n-1} p_{t_i} I_{[t_i, t_{i+1}[}(t)$$

and define the discrete Fourier coefficients of the log-price increment as

$$c_k(dp_n) := \frac{1}{2\pi} \sum_{j=0}^{n-1} e^{-\mathrm{i} t_j k} \delta_j(p), \tag{2.1}$$

with $\delta_i(p) := p(t_{i+1}) - p(t_i)$. According to Malliavin and Mancino (2009), for any $t \in (0, 2\pi)$, the spot volatility estimator is defined as

$$\widehat{\sigma}^2_{nNM}(t) := \sum_{|k| \leq M} \left(1 - \frac{|k|}{M+1}\right) e^{\mathrm{i} t k} c_k(\sigma^2_{nN}), \tag{2.2}$$

where $c_k(\sigma^2_{nN})$ is an unbiased estimator of the k-th Fourier coefficient of the volatility process, obtained through the following convolution formula:

$$c_k(\sigma^2_{nN}) := \frac{2\pi}{2N+1} \sum_{|h| \leq N} c_h(dp_n) c_{k-h}(dp_n). \tag{2.3}$$

The estimator in (2.2) is a consistent estimator of $\sigma^2(t)$ for $t \in (0, 2\pi)$, as proved in Theorem 2.2.1. On the contrary, for $t = 0$ or $t = 2\pi$, which represent fixed times of discontinuity, the estimator converges to $(\sigma^2(t^-) + \sigma^2(t))/2$.

By means of the rescaled Dirichlet kernel, defined as

$$D_N(x) := \frac{1}{2N+1} \sum_{|k| \leq N} e^{\mathrm{i} k x} = \frac{1}{2N+1} \frac{\sin((2N+1)x/2)}{\sin(x/2)}, \tag{2.4}$$

it is possible to express (2.3) as

$$c_k(\sigma^2_{nN}) = \frac{1}{2\pi} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} D_N(t_j - t_i) e^{-\mathrm{i} k t_j} \delta_i(p) \delta_j(p).$$

Therefore, the Fourier spot volatility estimator (2.2) can be rewritten, using two kernels, as

$$\widehat{\sigma}^2_{nNM}(t) = \frac{1}{2\pi} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} F_M(t - t_j) D_N(t_j - t_i) \delta_i(p) \delta_j(p), \tag{2.5}$$

where

$$F_M(x) := \sum_{|k|\leq M} \left(1 - \frac{|k|}{M+1}\right) e^{\mathrm{i}kx} = \frac{1}{M+1} \left(\frac{\sin((M+1)x/2)}{\sin(x/2)}\right)^2 \tag{2.6}$$

is the Féjèr kernel.

Remark 2.1.1. *If the auto-covariances in the convolution formula (2.3) are ignored, the Fourier estimator in (2.5) becomes a kernel-type spot volatility estimator (with the kernel specified as the one by Féjèr), namely it reduces to*

$$\widehat{\sigma}^2_{nM}(t) = \frac{1}{2\pi} \sum_{j=0}^{n-1} F_M(t - t_j)(\delta_j(p))^2. \tag{2.7}$$

As the inversion formula in definition (2.2) can be obtained with different kernels, the Fourier spot volatility estimator generalizes kernel-type estimators for the joint presence of two kernels. Kernel-type spot volatility estimators are studied, in the absence of noise, in Kristensen (2010), Mancini et al. (2015), Figueroa-López and Li (2020). However, kernel-type estimators are not robust in the presence of noise, as shown in Mancino and Recchioni (2015): therefore, the authors need to employ the pre-averaging technique and introduce a bias correction, see Zu and Boswijk (2014), Figueroa-López and Wu (2022).

On the contrary, in Section 2.2.3 we will prove that the presence of the cross-products, arising from the convolution formula in the original definition of the Fourier estimator (2.2), is crucial to ensure the robustness of the estimator in the presence of microstructure noise, without the need of any manipulation of data, such as sparse sampling or preaveraging, and without resorting to any bias correction. The contribution of auto-covariances has early been considered by Zhou (1996) and, later, by Barndorff-Nielsen et al. (2008) to correct for the bias of realized-variance-type estimators of the integrated volatility in the presence of noise. The Fourier estimator remains unbiased and efficient with an appropriate choice of the cutting frequency N, which controls the convolution that estimates the Fourier coefficients of the volatility.

Remark 2.1.2. *The definition of the Fourier spot volatility estimator does not require to consider equidistant observations. However, for simplicity, in Section 2.2 we will refer to the latter sampling scheme to derive the asymptotic theory. The general case with irregular observations can be obtained as in Malliavin and Mancino (2009).*

2.2 Asymptotic theory

2.2.1 Asymptotic Normality in the absence of microstructure noise

In Mancino and Recchioni (2015), the authors show that the estimator defined in (2.2) achieves a slightly sub-optimal rate of convergence, under the hypothesis that the volatility process is a random function with Hölder-continuous paths. In this section, we fine-tune this result by proving that the estimator can reach the optimal rate of convergence, $n^{1/4}$, as well as the efficient asymptotic-error variance, under the assumption that the volatility process is a Brownian semimartingale.

The following assumptions are made.

(A.I) The price process p is a continuous Itô semimartingale satisfying

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

where W is a standard Brownian motion on a filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,2\pi]}, P)$, which satisfies the usual conditions.

(A.II) The spot volatility[1] process σ^2 is an Itô semimartingale satisfying

$$d\sigma^2(t) = \gamma(t)dZ_t + b_v(t)dt,$$

where Z is a standard Brownian motion adapted to the filtration $\mathcal{F}$ and such that $d\langle W, Z\rangle_t = \phi\, dt$, with constant ϕ.

(A.III) The processes σ, b, γ and b_v are continuous and adapted stochastic processes defined on the same probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and such that, for any $p \geq 1$,

$$E\left[\int_0^T \sigma^p(t)dt\right] < \infty\ ,\ E\left[\int_0^T b^p(t)dt\right] < \infty\ ,\ E\left[\int_0^T \gamma^p(t)dt\right] < \infty\ ,\ E\left[\int_0^T b_v^p(t)dt\right] < \infty.$$

The processes are specified in such a way that σ and γ are a.s. positive.

Theorem 2.2.1. *Let assumptions (A.I), (A.II) and (A.III) hold. Moreover, assume that* $\lim_{n,N\to\infty} N/n = c > 0$. *Then, for any fixed* $t \in (0, 2\pi)$, *the following stable convergence in law holds, as* $n, N, M \to \infty$:

- *if* $\lim_{n,M\to\infty} Mn^{-1/\tau} = a > 0$, *for* $1 < \tau < 2$,

$$n^{1/2}M^{-1/2}\left(\widehat{\sigma}^2_{nNM}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{4}{3}(1 + 2K(2c))\,\sigma^4(t)\right); \tag{2.8}$$

- *if* $\lim_{n,M\to\infty} Mn^{-1/2} = a > 0$,

$$n^{1/2}M^{-1/2}\left(\widehat{\sigma}^2_{nNM}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{4}{3}(1 + 2K(2c))\sigma^4(t) + \frac{2\pi}{3a^2}\,\gamma^2(t)\right), \tag{2.9}$$

with

$$K(c) := \frac{1}{2c^2} r(c)(1 - r(c)), \tag{2.10}$$

where $r(x) = x - [x]$ *and* $[x]$ *is the integer part of* x.

Remark 2.2.2. *(Convergence Rate)*
The asymptotic result in (2.8) is in line with the one obtained in Mancino and Recchioni (2015), where the rate is $n^{1/4}(\log n)^{-1/2}$, *which is optimal up to a logarithmic correction. In Mancino and Recchioni (2015) it is assumed that* $N/n \sim c > 0$ *and* $Mn^{-1/\tau} \sim a > 0$, *for* $1 < \tau < 2$. *Note that in this case* $Mn^{-1/2} \to \infty$. *It is worth noting that the Central Limit Theorem in Mancino and Recchioni (2015) has been proved for a volatility process* σ^2 *belonging to a more general class, that is, under the assumption that* σ^2 *is a.s. Hölder continuous in* $[0, 2\pi]$ *with parameter* $\alpha \in (0, 1/2)$. *Therefore, rough volatility models are also included.*

[1] We follow the relevant econometric literature and use the term volatility as a synonym of variance, thus we refer to $\sigma^2(t)$ as the volatility process.

Letting $\beta = \sqrt{2\pi}/a$, compared to Theorem 8.6 in Aït-Sahalia and Jacod (2014), the asymptotic result in (2.8) corresponds to the case $\beta = 0$, where the rate is sub-optimal, while the asymptotic result in (2.9) corresponds to the case $\beta \in (0, \infty)$, which instead attains the optimal rate $n^{1/4}$. An analogous result is obtained in Theorem 2.1 of Figueroa-López and Wu (2022), where alternative localization kernels are considered.

Remark 2.2.3. *(Asymptotic Variance)*
The function $K(c)$ is obtained in Clement and Gloter (2011) and arises from the behavior of the discretized Dirichlet kernel, in the case when N has the same order of the sample size n (see Lemma 2.A.4). The choice $c = 1/2$ in (2.10) gives $K(2c) = 0$. Note that $K(2c)$ is nonnegative for any positive c and equal to zero if $c = (1/2)k$, $k = 1, 2, \ldots$. If $k = 1$, i.e., $c = 1/2$, then $N = n/2$, that is, N is equal to the Nyquist frequency, which reppresents the natural choice for N. Moreover, if $N = n/2$, one achieves the optimal asymptotic variance with regards to the term $(4/3)\sigma^4(t)$, which is smaller than its counterpart $2\sigma^4(t)$ in Theorem 8.6 by Aït-Sahalia and Jacod (2014); besides, letting $\beta = \sqrt{2\pi}/a$, even the second addend of the asymptotic variance in the efficient-rate case (2.9), that is, $2\pi/(3a^2)\gamma^2(t)$, is smaller than the corresponding term $2\pi/a^2\gamma^2(t)$. This is due to the presence of the Féjèr kernel in the Fourier-Féjèr inversion formula, see also Cuchiero and Teichmann (2015).

Note that, with the choice $c = 1/2$ (equivalently, $N = n/2$), the Fourier estimator has the same rate of convergence and asymptotic variance of the Féjèr-kernel based spot volatility estimator in (2.7), studied in Kristensen (2010), Mancini et al. (2015). However, the estimator in (2.7) is not robust to market microstructure and thus a pre-averaging procedure and a bias correction are needed. In summary, with an appropriate choice of N/n, the effect of adding the cross terms in (2.5), which is essential in order to get an noise-robust estimator, is also non-detrimental in view of the asymptotic efficiency.

Based on the previous remarks, the following result easily follows.

Corollary 2.2.4. *Let assumptions $(A.I)$, $(A.II)$ and $(A.III)$ hold. Moreover, assume that $\lim_{n,N\to\infty} N/n = 1/2$ and $\lim_{n,M\to\infty} Mn^{-1/\tau} = a > 0$, for $1 < \tau < 2$. Then, for any fixed $t \in (0, 2\pi)$, the following stable convergence in law holds, as $n, N, M \to \infty$:*

$$n^{1/2}M^{-1/2}\left(\widehat{\sigma}^2_{nNM}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{4}{3}\,\sigma^4(t)\right).$$

2.2.2 Proof of Theorem 2.2.1

Remark 2.2.5. *As every continuous process is locally bounded, thus all processes appearing are. Moreover, standard localization procedures (see, e.g., Aït-Sahalia and Jacod (2014)) allow to assume that any locally bounded process is actually bounded and almost surely positive processes can be considered as bounded away from zero.*

Remark 2.2.6. *Given the discrete time observations $0 = t_0 \leq \ldots t_j \ldots \leq t_n = 2\pi$, we use the notation in continuous time by letting $\varphi_n(t) := \sup\{t_j : t_j \leq t\}$ for the sake of simplicity.*

The proof follows the lines of Theorem 3.1 in Mancini et al. (2015), thus we highlight only the main differences.

According to Lemma 2.2 Malliavin and Mancino (2009), we can assume that $b = 0$ and $b_v = 0$. Using the Itô formula we can rewrite (2.5) as

$$\widehat{\sigma}^2_{nNM}(t) = \frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\sigma^2(v)dv + F_1(t) + F_2(t), \tag{2.11}$$

where

$$F_1(t) := \frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\int_0^v D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\ \sigma(v)dW_v, \tag{2.12}$$

$$F_2(t) := \frac{1}{2\pi}\int_0^{2\pi}\int_0^v F_M(t-\varphi_n(u))D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\ \sigma(v)dW_v, \tag{2.13}$$

and using the decomposition in (2.11) we write

$$\sqrt{\frac{n}{M}}(\widehat{\sigma}^2_{nNM}(t) - \sigma^2(t))$$

$$= \sqrt{\frac{n}{M}}\ \left(\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\sigma^2(v)dv - \sigma^2(t)\right) + \sqrt{\frac{n}{M}}\ (F_1(t)+F_2(t)).$$

The proof follows Jacod (1997), Jacod and Protter (1998) and is divided into three steps.

I) In the first step we consider the term:

$$\sqrt{\frac{n}{M}}\ \left(\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\sigma^2(v)dv - \sigma^2(t)\right). \tag{2.14}$$

The term (2.14) is decomposed into two addends:

$$\sqrt{\frac{n}{M}}\ \left(\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\sigma^2(v)dv - \frac{1}{2\pi}\int_0^{2\pi} F_M(t-v)\sigma^2(v)dv\right) \tag{2.15}$$

$$+\sqrt{\frac{n}{M}}\ \left(\frac{1}{2\pi}\int_0^{2\pi} F_M(t-v)\sigma^2(v)dv - \sigma^2(t)\right). \tag{2.16}$$

Observe that the term (2.15) converges to zero in probability, in fact:

$$\sqrt{\frac{n}{M}}\,E\left[\left|\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\sigma^2(v)dv - \frac{1}{2\pi}\int_0^{2\pi} F_M(t-v)\sigma^2(v)dv\right|\right]$$

$$= \sqrt{\frac{n}{M}}\,E\left[\left|\frac{1}{2\pi}\int_0^{2\pi}\left(F_M(t-\varphi_n(v)) - F_M(t-v)\right)\sigma^2(v)dv\right|\right]$$

$$\leq \operatorname{ess\,sup}\|\sigma^2\|_\infty\sqrt{\frac{M}{n}} \to 0,$$

by Lemma 2.A.1 i).

Consider now the term (2.16), which we denote by $(n/M)^{1/2}G(t)$. $G(t)$ splits into two terms:

$$G_1(t) := \frac{1}{2\pi}\int_{[0,2\pi]\cap\{|t-v|\leq\frac{2\pi}{M+1}\}} F_M(t-v)\left(\sigma^2(v)-\sigma^2(t)\right)dv \tag{2.17}$$

and

$$G_2(t) := \frac{1}{2\pi} \int_{[0,2\pi]\cap\{|t-v|>\frac{2\pi}{M+1}\}} F_M(t-v)\,(\sigma^2(v)-\sigma^2(t))dv. \tag{2.18}$$

As for the term $G_2(t)$ in (2.18), it is shown that

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

as in Mancino and Recchioni (2015), eq. (31). So, it converges to zero in probability.
Consider $G_1(t)$ in (2.17), which is equal to:

$$\sum_{|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{1}{2\pi}\int_{[0,2\pi]\cap\{|t-v|\leq\frac{2\pi}{M+1}\}} e^{-\mathrm{i}kv}(\sigma^2(v)-\sigma^2(t))dv. \tag{2.19}$$

First, we prove that the term corresponding to $k=0$ in (2.19) is asymptotically negligible. It splits into

$$\frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t}(\sigma^2(v)-\sigma^2(t))\,dv + \frac{1}{2\pi}\int_{t}^{t+\frac{2\pi}{M+1}}(\sigma^2(v)-\sigma^2(t))\,dv.$$

For the first addend, applying the integration by parts, it holds:

$$\frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t}(\sigma^2(v)-\sigma^2(t))dv = \frac{1}{2\pi}\left[v\,(\sigma^2(v)-\sigma^2(t))\right]_{t-\frac{2\pi}{M+1}}^{t} - \frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t} v\,d\sigma^2(v)$$

$$= \frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t}\left(t-\frac{2\pi}{M+1}-v\right)d\sigma^2(v).$$

Then, we compute the convergence in probability of[2]:

$$\frac{n}{M}\left\langle\frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t}\left(t-\frac{2\pi}{M+1}-v\right)d\sigma^2(v), \frac{1}{2\pi}\int_{t-\frac{2\pi}{M+1}}^{t}\left(t-\frac{2\pi}{M+1}-v\right)d\sigma^2(v)\right\rangle$$

$$= \frac{n}{M}\,\frac{1}{(2\pi)^2}\int_{t-\frac{2\pi}{M+1}}^{t}\left(t-\frac{2\pi}{M+1}-v\right)^2\gamma^2(v)dv$$

$$= \frac{n}{M}\,\frac{1}{(2\pi)^2}\int_{t-\frac{2\pi}{M+1}}^{t}\left(t-\frac{2\pi}{M+1}-v\right)^2 dv\left(\gamma^2\left(t-\frac{2\pi}{M+1}\right)+O_p\left(\frac{1}{M+1}\right)\right)$$

$$= \frac{n}{M}\,\frac{1}{(2\pi)^2}\,\frac{1}{3}\left(\frac{2\pi}{M+1}\right)^3\left(\gamma^2\left(t-\frac{2\pi}{M+1}\right)+O_p\left(\frac{1}{M+1}\right)\right) = O_p\left(\frac{n}{M^4}\right)\to 0.$$

The second addend is analogous.

Thus, we consider (2.19) for $k\neq 0$. Using the integration by parts, it is equal to:

$$\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{1}{2\pi}\left[\frac{\mathrm{i}}{k}\,e^{-\mathrm{i}kv}(\sigma^2(v)-\sigma^2(t))\right]_{t-\frac{2\pi}{M+1}}^{t+\frac{2\pi}{M+1}} \tag{2.20}$$

[2]The notation $\langle,\rangle$ denotes the quadratic variation on $[0,2\pi]$.

$$+\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{1}{2\pi}\frac{\mathrm{i}}{k}\int_{[0,2\pi]\cap\{|t-v|\leq\frac{2\pi}{M+1}\}}e^{-\mathrm{i}kv}d\sigma^2(v). \tag{2.21}$$

Consider (2.20). It is equal to:

$$\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{1}{2\pi}\frac{\mathrm{i}}{k}\left(e^{-\mathrm{i}k(t+\frac{2\pi}{M+1})}\left(\sigma^2\left(t+\frac{2\pi}{M+1}\right)-\sigma^2(t)\right)-e^{-\mathrm{i}k(t-\frac{2\pi}{M+1})}\left(\sigma^2\left(t-\frac{2\pi}{M+1}\right)-\sigma^2(t)\right)\right)$$

$$=\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\,\frac{1}{2\pi}\frac{\mathrm{i}}{k}\left(e^{-\mathrm{i}k\frac{2\pi}{M+1}}\int_t^{t+\frac{2\pi}{M+1}}d\sigma^2(v)+e^{\mathrm{i}k\frac{2\pi}{M+1}}\int_{t-\frac{2\pi}{M+1}}^{t}d\sigma^2(v)\right). \tag{2.22}$$

For the first integral (the second one is analogous), we study the convergence in probability of:

$$\frac{n}{M}\frac{1}{(2\pi)^2}\left\langle\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\,\frac{\mathrm{i}}{k}\,e^{-\mathrm{i}k\frac{2\pi}{M+1}}\int_t^{t+\frac{2\pi}{M+1}}d\sigma^2,\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\,\left(-\frac{\mathrm{i}}{k}\right)\,e^{\mathrm{i}k\frac{2\pi}{M+1}}\int_t^{t+\frac{2\pi}{M+1}}d\sigma^2\right\rangle.$$

The dominant term is studied as follows

$$\frac{n}{M}\frac{1}{(2\pi)^2}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\,\frac{1}{k^2}\int_t^{t+\frac{2\pi}{M+1}}\gamma^2(s)ds$$

$$=\frac{1}{2\pi}\frac{1}{M+1}\,\frac{n}{M}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\,\left(\gamma^2(t)+O_p\left(\frac{1}{M+1}\right)\right). \tag{2.23}$$

It is easily seen that

$$\lim_M\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\,\frac{1}{k^2}=\lim_M\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\,\frac{1}{k^2}=\lim_M\sum_{0<|k|\leq M}\frac{1}{k^2}=\frac{\pi^2}{3}.$$

Further, by the condition $Mn^{-1/2}\sim a$, and the continuity of $\gamma(t)$, we conclude that (2.23) converges to

$$\frac{1}{a^2}\frac{\pi}{6}\gamma^2(t).$$

Finally, noting that the covariance between the two stochastic integrals in (2.22) is zero, the contribution of (2.20) is equal to

$$\frac{1}{a^2}\frac{\pi}{3}\gamma^2(t).$$

Similarly, we deal with the term (2.21), which is equal to:

$$\frac{1}{2\pi}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{\mathrm{i}}{k}\,\left(\int_{t-\frac{2\pi}{M+1}}^{t}e^{-\mathrm{i}kv}d\sigma^2(v)+\int_t^{t+\frac{2\pi}{M+1}}e^{-\mathrm{i}kv}d\sigma^2(v)\right).$$

The convergence in probability for the bracket of the first addend, that is,

$$\frac{n}{M}\frac{1}{(2\pi)^2}\left\langle\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\,\frac{\mathrm{i}}{k}\int_{t-\frac{2\pi}{M+1}}^{t}e^{-\mathrm{i}kv}d\sigma^2(v),\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)e^{-\mathrm{i}kt}\,\left(-\frac{\mathrm{i}}{k}\right)\int_{t-\frac{2\pi}{M+1}}^{t}e^{\mathrm{i}kv}d\sigma^2(v)\right\rangle,$$

reads as follows:

$$\frac{n}{M}\frac{1}{(2\pi)^2}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\int_{t-\frac{2\pi}{M+1}}^{t}\gamma^2(v)dv$$

$$=\frac{n}{M}\frac{1}{(2\pi)^2}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\frac{2\pi}{M+1}\left(\gamma^2\left(t-\frac{2\pi}{M+1}\right)+O_p\left(\frac{1}{M+1}\right)\right)\to\frac{1}{a^2}\frac{\pi}{6}\gamma^2(t),$$

given the condition $Mn^{-1/2}\sim a$ and the fact that $\gamma(t)$ is continuous. The second addend is analogous.
Finally these four terms contribute to the variance with

$$\frac{2\pi}{3}\frac{1}{a^2}\gamma^2(t).$$

We remark that these terms converge to 0 if $Mn^{-1/\tau}\sim a$ for $1<\tau<2$ (thus $n/M^2\to 0$). Therefore, under this condition, the volatility of volatility component does not appear in the asymptotic variance in (2.8).

II) The study of the bracket

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

is analogous to the one in Mancini et al. (2015). It is composed of four terms and each of them leads to the same limit. We give details of the proof only for the first one.

Using the Itô formula, it holds that

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

$$=\frac{n}{M}\frac{1}{(2\pi)^2}\int_0^{2\pi}F_M^2(t-\varphi_n(v))\left(\int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\right)^2\sigma^2(v)dv$$

$$=\frac{n}{M}(F_{11}(t)+F_{12}(t)),$$

where

$$F_{11}(t):=\frac{1}{(2\pi)^2}\int_0^{2\pi}F_M^2(t-\varphi_n(v))\int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))\sigma^2(s)ds\,\sigma^2(v)dv, \tag{2.24}$$

$$F_{12}(t):=\frac{1}{(2\pi)^2}\int_0^{2\pi}F_M^2(t-\varphi_n(v))\int_0^v\left(\int_0^s D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\right) \tag{2.25}$$

$$\times D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\,\sigma^2(v)dv.$$

We prove that, for any $t\in(0,2\pi)$ fixed, in probability it holds:

$$\frac{n}{M}F_{11}(t)\to\frac{1}{3}(1+2K(2c))\sigma^4(t), \tag{2.26}$$

where $K(c)$ is equal to (2.10), and

$$\frac{n}{M}F_{12}(t)\to 0. \tag{2.27}$$

We begin with (2.26). Let $V := \pi(1 + 2K(2c))$. We have:

$$\left| \frac{n}{M} F_{11}(t) - V \, \frac{2}{3} \frac{1}{2\pi} \, \sigma^4(t) \right|$$

$$= \left| \frac{n}{M} \frac{1}{(2\pi)^2} \int_0^{2\pi} F_M^2(t - \varphi_n(v)) \int_0^v D_N^2(\varphi_n(v) - \varphi_n(s)) \sigma^2(s) ds \, \sigma^2(v) dv - V \frac{2}{3} \, \frac{1}{2\pi} \sigma^4(t) \right|$$

$$\leq \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \varphi_n(v)) \left(n \int_0^v D_N^2(\varphi_n(v) - \varphi_n(s)) \sigma^2(s) ds - V \sigma^2(v) \right) \sigma^2(v) dv \right| \tag{2.28}$$

$$+ V \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \varphi_n(v)) \sigma^4(v) dv - \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right|. \tag{2.29}$$

Consider (2.28): it is less than

$$\frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \varphi_n(v)) \left| n \int_0^v D_N^2(\varphi_n(v) - \varphi_n(s)) \sigma^2(s) ds - V \sigma^2(v) \right| \, \sigma^2(v) dv$$

which is $o_p(1)$ by virtue of Lemma 2.A.1 iv) and Lemma 2.A.4 ii).
Consider now (2.29): by Lemma 2.A.1 iv), it holds, in probability, that

$$\lim_{n, M \to \infty} \left| \frac{1}{(2\pi)^2} \int_0^{2\pi} \frac{1}{M} F_M^2(t - \varphi_n(v)) \sigma^4(v) dv - \frac{2}{3} \frac{1}{2\pi} \sigma^4(t) \right| = 0.$$

Finally, observe that $V \frac{2}{3} \frac{1}{2\pi} = \frac{1}{3}(1 + 2K(2c))$.
We prove now (2.27). Using the Itô isometry, we have:

$$n^2 E \left[\left(\int_0^v \int_0^s D_N(\varphi_n(v) - \varphi_n(u)) \sigma(u) dW_u \, D_N(\varphi_n(v) - \varphi_n(s)) \sigma(s) dW_s \right)^2 \right]$$

$$\leq \|\sigma^2\|_\infty^2 \, n^2 \int_0^v \int_0^s D_N^2(\varphi_n(v) - \varphi_n(u)) du \, D_N^2(\varphi_n(v) - \varphi_n(s)) ds.$$

Therefore, it is enough to observe that, as $N/n \sim c$,

$$n \int_0^v \left(n \int_0^s D_N^2(\varphi_n(v) - \varphi_n(u)) du \right) D_N^2(\varphi_n(v) - \varphi_n(s)) ds = o(1),$$

by combining Lemma 2.A.4 ii) and iii). Finally, we remark that $\frac{1}{M} \int_0^{2\pi} F_M^2(t - \varphi_n(v)) dv = O(1)$, by Lemma 2.A.1 iii).

III) The last step of the proof concerns the orthogonality. The first result requires to prove the convergence in probability of

$$\left\langle \sqrt{\frac{n}{M}} \, (F_1(t) + F_2(t) + G_1(t)), W \right\rangle \to 0,$$

where F_1 and F_2 are defined, respectively, in (2.12) and (2.13), while G_1 is defined in (2.17).

Consider $F_1(t)$ (the term $F_2(t)$ is analogous). It holds that:

$$\left\langle \sqrt{\frac{n}{M}}\, F_1(t), W \right\rangle = \sqrt{\frac{n}{M}} \frac{1}{2\pi} \int_0^{2\pi} F_M(t-\varphi_n(v)) \int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\sigma(v)dv.$$

Consider the L^1-norm:

$$E\left[\left|\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\sigma(v)dv\right|\right]$$

$$\leq \|\sigma\|_\infty \frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v)) E\left[\left(\int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\right)^2\right]^{1/2} dv.$$

Moreover, we have

$$E\left[\left(\int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\right)^2\right] \leq \|\sigma^2\|_\infty \int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))ds.$$

Thus

$$\sqrt{\frac{n}{M}} E\left[\left|\frac{1}{2\pi}\int_0^{2\pi} F_M(t-\varphi_n(v))\int_0^v D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\sigma(v)dv\right|\right]$$

$$\leq C\sqrt{\frac{1}{M}}\int_0^{2\pi} F_M(t-\varphi_n(v))\left(n\int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))ds\right)^{1/2} \leq C\sqrt{\frac{1}{M}} \to 0,$$

using Lemma 2.A.4 i).

Consider now the term $G_1(t)$ and note that it reduces to study the contribution of (2.20) and (2.21). We give details of the proof for (2.21). It requires to prove the following convergence in probability:

$$\sqrt{\frac{n}{M}}\frac{1}{2\pi}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\frac{\mathrm{i}}{k}\int_{t-\frac{2\pi}{M+1}}^t \gamma(v)\phi\, dv \to 0.$$

Consider:

$$\left|\sqrt{\frac{n}{M}}\frac{1}{2\pi}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\frac{\mathrm{i}}{k}\int_{t-\frac{2\pi}{M+1}}^t \gamma(v)\phi\, dv\right|$$

$$\leq C\sqrt{\frac{n}{M}}\frac{1}{2\pi}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)\frac{1}{|k|}\frac{2\pi}{M+1}\left(\gamma\left(t-\frac{2\pi}{M+1}\right)+O_p\left(\frac{1}{M+1}\right)\right)$$

$$\leq C\sqrt{\frac{n}{M}}\frac{1}{M+1}\sum_{k=1}^M\left(1-\frac{k}{M+1}\right)\frac{1}{k} \sim \frac{1}{\sqrt{M}}\left(\ln M+\overline{\gamma}+O\left(\frac{1}{M}\right)\right)\to 0,$$

where $\overline{\gamma}$ is the Euler-Mascheroni constant. The proof is complete.

2.2.3 Asymptotic Normality in the presence of microstructure noise

We study the asymptotic normality for the estimator (2.2) when the price process is contaminated by microstructure noise. The main result of this section shows that the spot volatility estimator (2.2) reaches the optimal rate of convergence and asymptotic variance.

The following assumption on the noise process η is made.

(N) $(\eta_{t_i})_{i\geq 0}$ is a family of i.i.d. random variables, independent of the log-price process p, with zero mean and finite second and fourth moments, equal to ξ and ω, respectively.

Denote by $\widetilde{\sigma}^2_{nNM}(t)$ the Fourier spot volatility estimator in the presence of microstructure noise, that is,

$$\widetilde{\sigma}^2_{nNM}(t) := \sum_{|k|\leq M} \left(1 - \frac{|k|}{M+1}\right) e^{\mathrm{i}tk} c_k(\widetilde{\sigma}^2_{nN}), \tag{2.30}$$

where $c_k(\widetilde{\sigma}^2_{nN})$ is the estimator of the k-th Fourier coefficient of the volatility process, obtained via (2.3) using the discrete Fourier coefficients of the increments of the observed noisy log-price process

$$\widetilde{p}_{t_i} := p_{t_i} + \eta_{t_i}. \tag{2.31}$$

In the rest of the work, we will refer to p as the efficient log-price.

Theorem 2.2.7. *Let assumptions (A.I), (A.II), (A.III) and (N) hold. Moreover, assume that* $\lim_{N,n\to\infty} Nn^{-1/2} = c > 0$. *Then, for any fixed* $t \in (0, 2\pi)$, *the following stable convergence in law holds, as* $n, N, M \to \infty$*:*

- *if* $\lim_{N,M\to\infty} M\, N^{-1/\tau} = a > 0$, *for* $1 < \tau < 2$,

$$n^{1/4} M^{-1/2} \left(\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{1}{c}\frac{2}{3}\sigma^4(t) + c\frac{8\pi}{9}\sigma^2(t)\xi + c^3\frac{4\pi^2}{15}\xi^2\right); \tag{2.32}$$

- *if* $\lim_{N,M\to\infty} M\, N^{-1/2} = a > 0$,

$$n^{1/4} M^{-1/2} \left(\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right) \to \mathcal{N}\left(0, \frac{1}{c}\frac{2}{3}\sigma^4(t) + \frac{1}{a^2 c}\,\frac{2\pi}{3}\,\gamma^2(t) + c\frac{8\pi}{9}\sigma^2(t)\xi + c^3\frac{4\pi^2}{15}\xi^2\right). \tag{2.33}$$

Remark 2.2.8. *(Convergence rate)*
In the presence of microstructure noise, the convergence rate attained by the Fourier estimator in (2.33) is the optimal one, that is, $n^{1/8}$. *This rate is also found for the spot volatility estimators by Aït-Sahalia and Jacod (2014, Section 8.7) and Figueroa-López and Wu (2022, Theorem 2.2). However, while in these cases pre-averaging and a bias correction are needed, the Fourier estimator in (2.2) does not need any modification. The case in (2.32), which assumes* $MN^{-1/\tau} \sim a > 0$, *for* $1 < \tau < 2$, *reaches a slightly sub-optimal rate; however, the resulting asymptotic variance is smaller and does not depend on the volatility of volatility.*

Remark 2.2.9. *(Asymptotic Variance)*
The asymptotic variance in (2.33) depends on the spot quarticity, the spot volatility of volatility, and the variance of the noise. All these factors can be estimated; thus, in principle, a feasible Central Limit Theorem is achievable. The same addends appear in the asymptotic variance of the pre-averaging and bias-corrected kernel-type estimators by Aït-Sahalia and Jacod (2014), Figueroa-López and Wu (2022). Notice that, as the condition $N/n \to 0$ *is in force, the quantity* $K(2c)$ *in (2.10), which arises from the discretization of the Dirichlet kernel, disappears (see Lemma 2.A.5).*

It is worth noticing the role of the cutting frequency N, which is evident by comparing Theorems 2.2.7 and 2.2.1. In Theorem 2.2.7, N is chosen smaller than in the noise-less case, in line with the

findings of Mancino and Sanfelici (2008). In fact, the role of N is that of filtering out the noise: the high-frequency, or short-run, noise is ignored by cutting the highest frequencies in the construction of the Fourier coefficients of the volatility $c_k(\sigma^2_{nN})$.

2.2.4 Proof of Theorem 2.2.7

The notation introduced in Remark 2.2.6 for the proof of Theorem 2.2.1 still holds in the following.

For brevity, let $\eta_i := \eta_{t_i}$. Moreover, let $\varepsilon_i := \eta_{i+1} - \eta_i$. For any $t \in (0, 2\pi)$ we rewrite the estimator error $\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)$ as:

$$\frac{1}{2\pi}\int_0^{2\pi} \left(F_M(t-\varphi_n(v)) - F_M(t-v)\right)\, \sigma^2(v)dv + F_1(t) + F_2(t) + G(t) + E_1(t) + E_2(t) + E_3(t), \quad (2.34)$$

where $F_1(t), F_2(t)$ and $G(t)$ are defined, respectively, in (2.12), (2.13) and (2.16), while $E_1(t)$, $E_2(t)$ and $E_3(t)$ are defined as follows:

$$E_1(t) := \frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t-t_j) D_N(t_j - t_i)\delta_i(p)\varepsilon_j, \quad (2.35)$$

$$E_2(t) := \frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t-t_j) D_N(t_j - t_i)\delta_j(p)\varepsilon_i, \quad (2.36)$$

$$E_3(t) := \frac{1}{2\pi}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t-t_j) D_N(t_j - t_i)\varepsilon_i\varepsilon_j. \quad (2.37)$$

Using the decomposition in (2.34), the proof consists in the study of the following four components, which are addressed in the next sections:

$$n^{\frac{1}{4}}M^{-\frac{1}{2}} \int_0^{2\pi} \left(F_M(t-\varphi_n(v)) - F_M(t-v)\right)\, \sigma^2(v)dv \quad (2.38)$$

$$+ n^{\frac{1}{4}}M^{-\frac{1}{2}}\, \left(G(t) + F_1(t) + F_2(t)\right) \quad (2.39)$$

$$+ n^{\frac{1}{4}}M^{-\frac{1}{2}}\, \left(E_1(t) + E_2(t)\right) \quad (2.40)$$

$$+ n^{\frac{1}{4}}M^{-\frac{1}{2}}\, E_3(t). \quad (2.41)$$

No-noise component

The study of the terms (2.38) and (2.39) closely follows the study of the corresponding terms in Theorem 2.2.1. However, due to the slower convergence rate considered here, some different constants appear when computing the asymptotic variance contribution given by $F_1(t)$, $F_2(t)$ and $G(t)$. Thus, we argue only this point.

Consider the contribution to the asymptotic variance given by

$$\left\langle n^{\frac{1}{4}}M^{-\frac{1}{2}}\, F_1(t), n^{\frac{1}{4}}M^{-\frac{1}{2}}\, F_1(t)\right\rangle.$$

Let $F_{11}(t)$ and $F_{12}(t)$ be defined as in (2.24) and (2.25). We prove that the following convergence

in probability holds:

$$\frac{\sqrt{n}}{M}F_{11}(t) \to \frac{1}{6}\frac{1}{c}\,\sigma^4(t).$$

Noting that $\sqrt{n}/M \sim N/(cM)$, we have:

$$\left|\frac{1}{c}\frac{N}{M}F_{11}(t) - \frac{1}{6c}\,\sigma^4(t)\right| = \left|\frac{1}{c}\frac{N}{M}\frac{1}{(2\pi)^2}\int_0^{2\pi} F_M^2(t-\varphi_n(v))\int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))\sigma^2(s)ds\,\sigma^2(v)dv - \frac{1}{c}\frac{1}{6}\,\sigma^4(t)\right|$$

$$\leq \frac{1}{c}\frac{1}{(2\pi)^2}\int_0^{2\pi}\frac{1}{M}F_M^2(t-\varphi_n(v))\left|N\int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))\sigma^2(s)ds - \frac{\pi}{2}\sigma^2(v)\right|\sigma^2(v)dv \tag{2.42}$$

$$+\frac{1}{c}\left|\frac{\pi}{2}\frac{1}{(2\pi)^2}\int_0^{2\pi}\frac{1}{M}F_M^2(t-\varphi_n(v))\sigma^4(v)dv - \frac{1}{6}\,\sigma^4(t)\right|. \tag{2.43}$$

Using Lemma 2.A.5, as $N/n \to 0$ (recall that $N/\sqrt{n} \sim c$), then

$$\lim_{n,N\to\infty} N\int_0^v D_N^2(\varphi_n(v)-\varphi_n(s))\sigma^2(s)ds = \frac{\pi}{2}\sigma^2(v).$$

Moreover, by Lemma 2.A.1 iv), as $M/n \to 0$, it holds that

$$\lim_{n,M\to\infty}\int_0^{2\pi}\frac{1}{M}F_M^2(t-\varphi_n(v))\sigma^2(v)dv = \frac{4\pi}{3}\sigma^4(t).$$

Therefore, (2.42) goes to 0. Analogously, (2.43) converges to 0.

We prove now the following convergence in probability:

$$\frac{\sqrt{n}}{M}F_{12}(t) \to 0. \tag{2.44}$$

It is enough to study

$$\frac{N}{M}\int_0^{2\pi} F_M^2(t-v)\int_0^v\int_0^s D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\,D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\,\sigma^2(v)dv.$$

Using the Itô isometry, we have that

$$N^2E\left[\left(\int_0^v\int_0^s D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\,D_N(\varphi_n(v)-\varphi_n(s))\sigma(s)dW_s\right)^2\right]$$

$$\leq \|\sigma^2\|_\infty N^2\int_0^v E\left[\left(\int_0^s D_N(\varphi_n(v)-\varphi_n(u))\sigma(u)dW_u\right)^2\right]D_N^2(\varphi_n(v)-\varphi_n(s))ds$$

$$\leq \|\sigma^2\|_\infty^2 N^2\int_0^v\int_0^s D_N^2(\varphi_n(v)-\varphi_n(u))du\,D_N^2(\varphi_n(v)-\varphi_n(s))ds.$$

Therefore, it is enough to observe that

$$N\int_0^v N\int_0^s D_N^2(\varphi_n(v)-\varphi_n(u))du\;D_N^2(\varphi_n(v)-\varphi_n(s))ds = o(1),$$

by combining Lemmas 2.A.4 and 2.A.5 (remember that $N/n \to 0$). Finally, $(1/M)\int_0^{2\pi}F_M^2(t-v)dv = O(1)$, which concludes the proof of (2.44).

Further, note that the contribution to the variance of all these components remains the same under the condition $M/N^{-1/\tau} \sim a > 0$, for $\tau > 1$.

Finally, as for the contribution of $G(t)$ defined in (2.16), the proof follows in the same way as in Theorem 2.2.1. It is enough to point out that, when studying $G_1(t)$, it holds:

$$\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\int_t^{t+\frac{2\pi}{M+1}}\gamma^2(s)ds$$

$$=\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\frac{2\pi}{M+1}\left(\gamma^2(t)+O_p\left(\frac{1}{M+1}\right)\right)$$

$$=\frac{1}{2\pi}\frac{1}{M+1}\frac{\sqrt{n}}{M}\sum_{0<|k|\leq M}\left(1-\frac{|k|}{M+1}\right)^2\frac{1}{k^2}\left(\gamma^2(t)+O_p\left(\frac{1}{M+1}\right)\right). \tag{2.45}$$

Further, by the conditions $MN^{-1/2} \sim a$ and $Nn^{-1/2} \sim c$, and the continuity of $\gamma(t)$, we conclude that (2.45) converges to

$$\frac{1}{a^2c}\frac{\pi}{6}\gamma^2(t).$$

Noise component

In this section we compute the contribution to the asymptotic variance from the pure noise term (2.37). We observe that, neglecting end-effects that will be studied later, the term

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_i-t_j)\varepsilon_i\varepsilon_j$$

$$=\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_i-t_j)(\eta_{i+1}\eta_{j+1}-\eta_i\eta_{j+1}-\eta_{i+1}\eta_j+\eta_i\eta_j)$$

can be expressed as follows:

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_i-t_j)\eta_{i+1}\eta_{j+1}-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_{i+1}-t_j)\eta_{i+1}\eta_{j+1} \tag{2.46}$$

$$-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_{j+1})D_N(t_i-t_{j+1})\eta_{i+1}\eta_{j+1}+\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_{j+1})D_N(t_{i+1}-t_{j+1})\eta_{i+1}\eta_{j+1}, \tag{2.47}$$

plus the remaining end points equal to the following:

$$-\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_0-t_j)\eta_0\eta_{j+1}+\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_n-t_j)\eta_n\eta_{j+1} \tag{2.48}$$

$$-\sum_{j=0}^{n-1}F_M(t-t_0)D_N(t_j-t_0)\eta_0\eta_{j+1}+\sum_{j=0}^{n-1}F_M(t-t_n)D_N(t_j-t_n)\eta_n\eta_{j+1} \tag{2.49}$$

$$+\sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_0-t_j)\eta_0\eta_j - \sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_n-t_j)\eta_n\eta_j \tag{2.50}$$

$$+\sum_{j=0}^{n-1} F_M(t-t_0)D_N(t_j-t_0)\eta_j\eta_0 - \sum_{j=0}^{n-1} F_M(t-t_n)D_N(t_j-t_n)\eta_j\eta_n. \tag{2.51}$$

We begin by studying the terms (2.46) and (2.47), that can be written as follows:

$$-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t-t_j)\left(D_N(t_{i+1}-t_j)-D_N(t_i-t_j)\right)\eta_{i+1}\eta_{j+1} \tag{2.52}$$

$$+\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} F_M(t-t_{j+1})\left(D_N(t_{i+1}-t_{j+1})-D_N(t_i-t_{j+1})\right)\eta_{i+1}\eta_{j+1}. \tag{2.53}$$

With $\rho(n) = t_{i+1} - t_i$, using the Taylor formula, we express (2.52) and (2.53) as

$$\sum_i\sum_j \left[F_M(t-t_{j+1})D'_N(t_i-t_{j+1}) - F_M(t-t_j)D'_N(t_i-t_j)\right]\,\rho(n)\,\eta_{i+1}\eta_{j+1}\, + o(\rho(n))$$

$$= \sum_i\sum_j \left[F'_M(t-t_j)D'_N(t_i-t_j) + F_M(t-t_j)D''_N(t_i-t_j)\right]\,\rho(n)^2\,\eta_{i+1}\eta_{j+1}\, + o(\rho(n)^2).$$

We study the contribution to the asymptotic variance of these two terms:

$$E_{31}(t) := \frac{1}{2\pi}\sum_i\sum_j F'_M(t-t_j)D'_N(t_i-t_j)\,\rho(n)^2\,\eta_{i+1}\eta_{j+1},$$

$$E_{32}(t) := \frac{1}{2\pi}\sum_i\sum_j F_M(t-t_j)D''_N(t_i-t_j)\,\rho(n)^2\,\eta_{i+1}\eta_{j+1}.$$

Consider $E_{31}(t)$. Given the assumptions that the noise and price processes are independent and that the noise is zero-mean, conditionally on the price process, it holds:

$$E[(n^{1/4}M^{-1/2}\frac{1}{2\pi}\sum_i\sum_j F'_M(t-t_j)D'_N(t_i-t_j)\,\rho(n)^2\,\eta_{i+1}\eta_{j+1})^2]$$

$$= \frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i\sum_j \left(F'_M(t-t_j)\right)^2\left|D'_N(t_i-t_j)\right|^2\rho(n)^4E[\eta_{i+1}^2\eta_{j+1}^2] \tag{2.54}$$

$$+\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i\sum_{i'\neq i}\sum_j\sum_{j'\neq j} F'_M(t-t_j)F'_M(t-t_{j'})\,D'_N(t_i-t_j)D'_N(t_{i'}-t_{j'})\,\rho(n)^4E[\eta_{i+1}\eta_{i'+1}\eta_{j+1}\eta_{j'+1}]. \tag{2.55}$$

First, we study the term (2.54), which is equal to:

$$\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i \left|F'_M(t-t_i)\right|^2\left|D'_N(0)\right|^2\rho(n)^4E[\eta_{j+1}^4]$$

$$+\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i\sum_{j\neq i} \left|F'_M(t-t_j)\right|^2\left|D'_N(t_i-t_j)\right|^2\rho(n)^4E[\eta_{i+1}^2\eta_{j+1}^2].$$

By adopting the continuous-time notation and using the fact that $|D'_N(0)|^2 = O(N)$, and that the η's are i.i.d. with fourth moment equal to ω, we can study:

$$\frac{\sqrt{n}}{M}\rho(n)^3\, N\, \frac{1}{(2\pi)^2}\int_0^{2\pi} |F'_M(t-\varphi_n(s))|^2 ds\ \omega$$

$$+\frac{\sqrt{n}}{M} 2\,\rho(n)^2\, \frac{1}{(2\pi)^2}\int_0^{2\pi} |F'_M(t-\varphi_n(s))|^2 \int_0^s |D'_N(\varphi_n(s)-\varphi_n(u))|^2\ du\, ds\ \xi^2$$

$$= \frac{1}{c} N^2 \rho(n)^3\, M^2\, \frac{1}{(2\pi)^2}\frac{1}{M^3}\int_0^{2\pi} |F'_M(t-\varphi_n(s))|^2\, ds\ \omega \tag{2.56}$$

$$+\frac{1}{c} N^2\, \rho(n)^2\, M^2\, 2\frac{1}{(2\pi)^2}\frac{1}{M^3}\int_0^{2\pi} |F'_M(t-\varphi_n(s))|^2\, \frac{1}{N}\int_0^s |D'_N(s-\varphi_n(u))|^2 du\, ds\ \xi^2. \tag{2.57}$$

Consider the order of the two addends by using Lemmas 2.A.3 and 2.A.6. For (2.56), the order is:

$$N^2\rho(n)^3\, M^2\, \omega \sim \frac{N^2}{n}\frac{M^2}{n^2} = o(1).$$

For (2.57), we obtain:

$$N^2\rho(n)^2\, M^2 \xi \sim \frac{N^2}{n}\frac{M^2}{n} = o(1).$$

Therefore (2.54) goes to zero in probability.

Consider now (2.55). We distinguish four cases:

$$(I)\quad i=j,\ \ i'\neq i,\ \ j'\neq j,\ \ i'=j',$$

$$(II)\quad i=j,\ \ i'\neq i,\ \ j'\neq j,\ \ i'\neq j',$$

$$(III)\quad i\neq j,\ \ i'\neq i,\ \ j'\neq j,\ \ i'\neq j',$$

$$(IV)\quad i\neq j,\ \ i'\neq i,\ \ j'\neq j,\ \ i'=j'.$$

Observe that in the cases $(II), (III)$ and (IV) the conditional expectation is null because of the independence of the η's. Consider the case (I). Using that $|D'_N(0)|^2 = O(N)$ and the independence property of the noise process, it is equal to

$$\frac{1}{c}\frac{N^2}{M}\,\rho(n)^4\,\frac{1}{(2\pi)^2}\,2\sum_j\sum_{j'<j} F'_M(t-t_j)F'_M(t-t_{j'})E[\eta^2_{j+1}]E[\eta^2_{j'+1}]. \tag{2.58}$$

Then it is equivalent to study

$$\frac{N^2}{M}\,\rho(n)^2\,\left|\int_0^{2\pi} F'_M(t-\varphi_n(s))\int_0^s F'_M(t-\varphi_n(u))\,du\,ds\right|\ \xi^2$$

$$\leq C\frac{N^2}{M}\rho(n)^2\int_0^{2\pi}|F'_M(t-\varphi_n(s))|^2\,ds = CN^2M^2\rho(n)^2 = o(1),$$

therefore (2.55) goes to zero in probability.

Now we study the contribution to the asymptotic variance from $E_{32}(t)$. It holds:

$$E[(n^{1/4}M^{-1/2}\frac{1}{2\pi}\sum_i\sum_j F_M(t-t_j)D''_N(t_i-t_j)\,\rho(n)^2\,\eta_{i+1}\eta_{j+1})^2]$$

$$=\frac{\sqrt{n}}{M}\rho(n)^4\,\frac{1}{(2\pi)^2}\sum_i\sum_j F_M^2(t-t_j)|D''_N(t_i-t_j)|^2\,E[\eta_{i+1}^2\eta_{j+1}^2] \tag{2.59}$$

$$+\frac{\sqrt{n}}{M}\,\rho(n)^4\,\frac{1}{(2\pi)^2}\sum_i\sum_{i\neq i'}\sum_j\sum_{j\neq j'}F_M(t-t_j)D''_N(t_i-t_j)F_M(t-t_{j'})D''_N(t_{i'}-t_{j'})\,E[\eta_i\eta_j\eta_{i'+1}\eta_{j'+1}]. \tag{2.60}$$

First, consider (2.59). It is equal to

$$\frac{1}{c}\frac{N}{M}\,\rho(n)^4\frac{1}{(2\pi)^2}\sum_j F_M^2(t-t_j)|D''_N(0)|^2\,E[\eta_{j+1}^4] \tag{2.61}$$

$$+\frac{2}{c}\frac{N}{M}\rho(n)^4\,\frac{1}{(2\pi)^2}\,\sum_j\sum_{i<j}F_M^2(t-t_j)|D''_N(t_i-t_j)|^2\,E[\eta_{i+1}^2\eta_{j+1}^2]. \tag{2.62}$$

Consider the first addend (2.61). Given that $|D''_N(0)|^2=O(N^4)$, it reduces to:

$$\frac{N}{M}N^4\rho(n)^3\int_0^{2\pi}F_M^2(t-\varphi_n(s))\,ds\,\omega\sim\frac{N^4}{n^2}\frac{N}{n}\to 0.$$

Consider the second addend (2.62). Thanks to Lemma 2.A.6, it holds:

$$\frac{2}{c}\frac{N}{M}\,\rho(n)^2\,\frac{1}{(2\pi)^2}\int_0^{2\pi}F_M^2(t-\varphi_n(s))\int_0^s|D''_N(\varphi_n(s)-\varphi_n(u))|^2\,du\,ds\;\xi^2$$

$$=\frac{2}{c}\,N^4\,\rho(n)^2\,\frac{1}{(2\pi)^2}\frac{1}{M}\int_0^{2\pi}F_M^2(t-\varphi_n(s))\frac{1}{N^3}\int_0^s|D''_N(\varphi_n(s)-\varphi_n(u))|^2\,du\,ds\,\xi^2\to c^3\,\frac{4\pi^2}{15}\,\xi^2.$$

We now study (2.60). Using the properties of the noise process, it is enough to observe that

$$\frac{N}{M}\rho(n)^4\sum_j\sum_{j'\neq j}F_M(t-t_j)|D''_N(0)|^2F_M(t-t_{j'})\,E[\eta_{j+1}^2\eta_{j'+1}^2]$$

$$=2\frac{N}{M}N^4\rho(n)^2\int_0^{2\pi}F_M(t-\varphi_n(s))\int_0^sF_M(t-\varphi_n(u))\,du\,ds\;\xi^2.$$

$$=4\frac{N^5}{M}\rho(n)^2\int_t^{2\pi}F_M(t-\varphi_n(s))\int_0^sF_M(t-\varphi_n(u))\,du\,ds\;\xi^2$$

$$\leq C\frac{N^5}{M}\rho(n)^2\int_t^{t+\epsilon_n}F_M(t-\varphi_n(s))\,ds=C\frac{N^4}{n^2}\frac{N}{M}\epsilon_n,\quad\text{for any sequence}\,\epsilon_n\to 0.$$

Therefore, to have convergence to 0, it is sufficient to choose a sequence $\epsilon_n\to 0$ s.t. $\frac{N}{M}\epsilon_n\to 0$.

It remains to prove the asymptotic negligibility of the end points' contribution. The terms are

grouped as follows:

$$-\sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_j)\eta_{j+1}\eta_0 + \sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_j)\eta_j\eta_0 \tag{2.63}$$

$$+\sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_j)\eta_{j+1}\eta_n - \sum_{j=0}^{n-1} F_M(t-t_j)D_N(t_j)\eta_j\eta_n \tag{2.64}$$

$$-\sum_{j=0}^{n-1} F_M(t)D_N(t_j)\eta_{j+1}\eta_0 + \sum_{j=0}^{n-1} F_M(t)D_N(t_j)\eta_j\eta_0 \tag{2.65}$$

$$+\sum_{j=0}^{n-1} F_M(t)D_N(t_j)\eta_{j+1}\eta_n - \sum_{j=0}^{n-1} F_M(t)D_N(t_j)\eta_j\eta_n. \tag{2.66}$$

We first consider (2.65). It can be written as

$$\begin{aligned}
&-\sum_{j=0}^{n-1} F_M(t)D_N(t_j)\eta_{j+1}\eta_0 + \sum_{j=0}^{n-1} F_M(t)D_N(t_{j+1})\eta_{j+1}\eta_0 \\
&\qquad +F_M(t)D_N(t_0)\eta_0^2 - F_M(t)D_N(t_n)\eta_n\eta_0 \\
&= F_M(t)\sum_{j=0}^{n-1}\left(D_N(t_{j+1}) - D_N(t_j)\right)\eta_{j+1}\eta_0 + F_M(t)\eta_0^2 - F_M(t)\eta_n\eta_0,
\end{aligned}$$

where we used that $D_N(t_0) = D_N(0) = 1 = D_N(2\pi) = D_N(t_n)$. Notice that, for the i.i.d property of the noise, it holds:

$$\begin{aligned}
&n^{1/2}M^{-1}E\left[\left(F_M(t)\sum_{j=0}^{n-1} D'_N(t_j)\rho(n)\,\eta_{j+1}\eta_0 + F_M(t)\eta_0^2 - F_M(t)\eta_n\eta_0\right)^2\right] \\
&= n^{1/2}M^{-1}\left(F_M^2(t)\sum_{j=0}^{n-1}|D'_N(t_j)|^2\rho(n)^2\,\xi^2 + F_M^2(t)\omega + F_M^2(t)\,\xi^2\right).
\end{aligned}$$

Finally, observe that

$$n^{1/2}M^{-1}\left(F_M^2(t)\rho(n)\int_0^{2\pi}|D'_N(\varphi_n(s))|^2 ds\,\xi^2 + F_M^2(t)(\omega+\xi^2)\right) = O(n^{-1/2}M^{-3}N) + O(n^{1/2}M^{-3}) = o(1), \tag{2.67}$$

where we used the fact that, for any $\varepsilon > 0$ and $\varepsilon \le t \le 2\pi - \varepsilon$, $F_M(t) \le M^{-1}\left(\sin(\varepsilon/2)\right)^{-2}$, and Lemma 2.A.5.

Consider now (2.63). With a similar decomposition as for (2.65), it can be written as

$$\sum_{j=0}^{n-1}[\,(F'_M(t-t_j)D_N(t_j) + F_M(t-t_j)D'_N(t_j))\;\rho(n) + o(\rho(n))]\eta_{j+1}\eta_0 + F_M(t)\eta_0^2 - F_M(t)\eta_0\eta_n,$$

where we used that $D_N(t_0) = D_N(0) = 1 = D_N(2\pi) = D_N(t_n)$. Then, for the i.i.d property of the

noise, it holds:

$$n^{1/2}M^{-1}E\left[\left(\sum_{j=0}^{n-1}\left(F'_M(t-t_j)D_N(t_j)+F_M(t-t_j)D'_N(t_j)\right)\rho(n)\eta_{j+1}\eta_0+F_M(t)\eta_0^2-F_M(t)\eta_0\eta_n\right)^2\right]$$

$$=n^{1/2}M^{-1}\left(\sum_{j=0}^{n}\left(|F'_M(t-t_j)|^2D_N^2(t_j)+F_M^2(t-t_j)|D'_N(t_j)|^2\right)\rho(n)^2\xi^2+F_M^2(t)(\omega+\xi^2)\right).$$

The first two addends are studied as follows. For the first one, we have:

$$n^{1/2}M^{-1}\rho(n)\int_0^{2\pi}|F'_M(t-\varphi_n(s))|^2D_N^2(\varphi_n(s))ds\sim n^{1/2}M^{-1}\rho(n)M^3\,D_N^2(t),$$

where we have used Lemma 2.A.3. Moreover, using that for $t\in(0,2\pi)$, then $D_N^2(t)=O(N^{-1})$, this term is

$$O(n^{-1/2}M^2N^{-1})=o(1).$$

For the second term, we study:

$$n^{1/2}M^{-1}\rho(n)\int_0^{2\pi}F_M^2(t-\varphi_n(s))|D'_N(\varphi_n(s))|^2ds\,\xi^2.$$

Using Lemmas 2.A.1 and 2.A.6, this gives, for $t\neq 0,2\pi$, that

$$n^{1/2}M^{-1}\rho(n)M|D'_N(t)|^2=O(n^{-1/2}N)o(1)=o(1).$$

Finally, consider the remaining term, which is equal to

$$n^{1/2}M^{-1}F_M^2(t)(\omega+\xi^2)=O(n^{1/2}M^{-1}M^{-2})=O(n^{-1/4}).$$

The terms (2.64) and (2.66) are analogous.

Mixed noise-price component

Consider now the asymptotic variance contribution from the term (2.40). It comprises four analogous terms. We focus on $E_1(t)$, defined in (2.35). We observe that the term $E_1(t)$ can be written as:

$$\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_j)D_N(t_j-t_i)\delta_i(p)\eta_{j+1}-\sum_{i=0}^{n-1}\sum_{j=0}^{n-1}F_M(t-t_{j+1})D_N(t_{j+1}-t_i)\,\delta_i(p)\eta_{j+1}\tag{2.68}$$

plus the remaining end-terms equal to

$$-\sum_{i=0}^{n-1}F_M(t-t_0)D_N(t_i-t_0)\,\delta_i(p)\eta_0+\sum_{i=0}^{n-1}F_M(t-t_n)D_N(t_i-t_n)\,\delta_i(p)\eta_n.\tag{2.69}$$

Consider (2.68). It is equal to

$$\sum_i \sum_j F'_M(t-t_j) D_N(t_j - t_i)\, \delta_i(p) \eta_{j+1}\, \rho(n) + o(\rho(n)) \tag{2.70}$$

$$+ \sum_i \sum_j F_M(t-t_j) D'_N(t_i - t_j)\, \delta_i(p) \eta_{j+1}\, \rho(n) + o(\rho(n)). \tag{2.71}$$

Thus we first consider the term (2.70). It holds:

$$\left(n^{\frac{1}{4}} M^{-\frac{1}{2}} \frac{1}{2\pi} \sum_i \sum_j F'_M(t-t_j) D_N(t_j - t_i)\, \rho(n)\, \delta_i(p)\, \eta_{j+1} \right)^2$$

$$= \frac{\sqrt{n}}{M} \frac{1}{(2\pi)^2} \sum_i \sum_j |F'_M(t-t_j)|^2 D_N^2(t_j - t_i)\, \rho(n)^2\, \delta_i^2(p)\, \eta_{j+1}^2 \tag{2.72}$$

$$+ \frac{\sqrt{n}}{M} \frac{1}{(2\pi)^2} \sum_i \sum_{i' \neq i} \sum_j \sum_{j' \neq j} F'_M(t-t_j) D_N(t_j - t_i) F'_M(t-t_{j'}) D_N(t_{j'} - t_{i'})\, \rho(n)^2\, \delta_i(p) \eta_{j+1} \delta_{i'}(p) \eta_{j'+1}. \tag{2.73}$$

Consider the conditional (with respect to the price process p) expectation of (2.72). Using the fact that the noise process is independent from the price process, then it is equal to

$$\frac{1}{c} \frac{N}{M} \frac{1}{(2\pi)^2} \sum_j |F'_M(t-t_j)|^2 D_N^2(0)\, \rho(n)^2\, \delta_j^2(p)\, E[\eta_{j+1}^2] \tag{2.74}$$

$$+ \frac{2}{c} \frac{N}{M} \frac{1}{(2\pi)^2} \sum_j \sum_{i<j} |F'_M(t-t_j)|^2 D_N^2(t_j - t_i)\, \rho(n)^2\, \delta_i^2(p)\, E[\eta_{j+1}^2]. \tag{2.75}$$

Because $D_N^2(0) = 1$, then the first addend, (2.74), is equal to

$$\frac{1}{c} \frac{N}{M} M^3\, \rho(n)^2\, \frac{1}{(2\pi)^2} \left(\frac{1}{M^3} \int_0^{2\pi} |F'_M(t - \varphi_n(s))|^2 \sigma^2(s) ds + o_p(1) \right) \xi,$$

which converges to 0 by virtue of Lemma 2.A.3.

Consider the term (2.75). It is equal to

$$\frac{1}{c} \frac{N}{M}\, \rho(n) \frac{1}{(2\pi)^2} \left(\int_0^{2\pi} \int_0^s |F'_M(t - \varphi_n(s))|^2 D_N^2(\varphi_n(s) - \varphi_n(u)) \sigma^2(u) du\, ds + o_p(1) \right) \xi. \tag{2.76}$$

Using Lemmas 2.A.3 and 2.A.5, the following convergence in probability holds:

$$\int_0^{2\pi} \frac{1}{M^3} |F'_M(t - \varphi_n(s))|^2\, N \int_0^s D_N^2(\varphi_n(s) - \varphi_n(u)) \sigma^2(u) du\, ds \to C\sigma^2(t).$$

Therefore (2.76) has order $\rho(n) M^2$, which converges to 0.

Then consider the term (2.71). It holds:

$$\left(n^{\frac{1}{4}} M^{-\frac{1}{2}} \frac{1}{2\pi} \sum_i \sum_j F_M(t-t_j) D'_N(t_i - t_j)\, \rho(n)\, \delta_i(p)\, \eta_{j+1} \right)^2$$

$$= \frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i\sum_j F_M^2(t-t_j)|D'_N(t_i-t_j)|^2\,\rho(n)^2\,\delta_i^2(p)\,\eta_{j+1}^2 \tag{2.77}$$

$$+\frac{\sqrt{n}}{M}\frac{1}{(2\pi)^2}\sum_i\sum_{i'\neq i}\sum_j\sum_{j'\neq j} F_M^2(t-t_j)|D'_N(t_i-t_j)|^2\,\rho(n)^2\,\delta_i(p)\delta_{i'}(p)\,\eta_j\eta_{j'+1}. \tag{2.78}$$

Consider (2.77). The conditional variance with respect to p gives

$$\frac{1}{c}\frac{N}{M}\rho(n)^2\frac{1}{(2\pi)^2}\sum_j F_M^2(t-t_j)|D'_N(0)|^2\,\delta_j^2(p)\,\xi$$

$$+\frac{2}{c}\frac{N}{M}\rho(n)^2\frac{1}{(2\pi)^2}\sum_j\sum_{i<j} F_M^2(t-t_j)|D'_N(t_i-t_j)|^2\,\delta_i^2(p)\,\xi.$$

The first addend converges to zero. Indeed, it holds that

$$N^2\frac{1}{n^2}\left(\frac{1}{M}\int_0^{2\pi} F_M^2(t-\varphi_n(s))\sigma^2(s)ds + o_p(1)\right)\xi = O_p\left(\frac{N^2}{n^2}\right) = o_p(1).$$

Consider then the second one. Based on Lemmas 2.A.1 and 2.A.6, it holds:

$$\frac{2}{c}\frac{N}{M}\rho(n)\frac{1}{(2\pi)^2}\left(\int_0^{2\pi} F_M^2(t-\varphi_n(s))\int_0^s |D'_N(\varphi_n(s)-\varphi_n(u)|^2\,\sigma^2(u)du\,ds + o_p(1)\right)\xi \to c\,\frac{2\pi}{9}\,\sigma^2(t)\xi.$$

As for the term (2.78), this gives zero contribution for the i.i.d property of the noise process and the fact that it is zero-mean.

Consider now the end-points term in (2.69) and prove their asymptotic negligibility. Consider the first addend:

$$\left(n^{1/4}M^{-1/2}\sum_{i=0}^{n-1} F_M(t)D_N(t_i)\delta_i(p)\eta_0\right)^2$$

$$= n^{1/2}M^{-1}\sum_{i=0}^{n-1} F_M^2(t)D_N^2(t_i)(\delta_i(p))^2\eta_0^2 \tag{2.79}$$

$$+n^{1/2}M^{-1}2\sum_i\sum_{j<i} F_M^2(t)D_N(t_i)D_N(t_j)\delta_i(p)\delta_j(p)\eta_0^2. \tag{2.80}$$

The conditional expectation with respect to p of (2.79) gives

$$n^{1/2}M^{-1}F_M^2(t)\left(\int_0^{2\pi} D_N^2(\varphi_n(u))\sigma^2(u)du + o_p(1)\right)\xi.$$

Noting that, for any $\varepsilon > 0$ and $\varepsilon \leq t \leq 2\pi - \varepsilon$, $F_M(t) \leq M^{-1}\,(\sin(\varepsilon/2))^{-2}$, and using Lemma 2.A.5, then (2.79) has order $n^{1/2}M^{-3}N^{-1} = o(1)$.

Consider now (2.80). Given the independence between the noise and price processes, we get:

$$n^{1/2}M^{-1}\,2F_M^2(t)\int_0^{2\pi}\int_0^s D_N(\varphi_n(u))\sigma(u)dW_u D_N(\varphi_n(s))\sigma(s)dW_s\,\xi.$$

It holds:

$$E\left[\left(\int_0^{2\pi}\int_0^s D_N(\varphi_n(u))\sigma(u)dW_u D_N(\varphi_n(s))\sigma(s)dW_s\right)^2\right]^{1/2} \leq C\int_0^{2\pi} D_N^2(u)du = C\frac{2\pi}{2N+1}.$$

Therefore (2.80) has order $O_p(n^{1/2}M^{-2}N^{-1}) = o_p(1)$.

For the second addend in (2.69) it is enough to observe that $t_n = 2\pi$ and use the periodicity of the kernels and the properties of the noise process.

Asymptotic orthogonality

The last step needed to ensure the stable convergence requires proving the following convergence in probability:

$$\left\langle n^{1/4}M^{-1/2}\left(F_1(t)+F_2(t)+G(t)\right), W\right\rangle \to 0.$$

The proof is analogous to *Step III)* in Theorem 2.2.1, taking into account the relative rates of N and M and using the fact that, by Lemma 2.A.4, it holds:

$$\lim_{n,N\to\infty} N\int_0^v D_N^2(\varphi_n(v)-\varphi_n(u))\sigma^2(u)du = \frac{\pi}{2}\sigma^2(v).$$

Finally, with a conditional expectation argument and the fact that at least one of the terms in the product has expectation equal to zero, it is easily seen that the covariance of $F_1(t)+F_2(t)+G(t)$ and $E_1(t)+E_2(t)+E_3(t)$ is asymptotically zero. The proof is now complete.

2.3 Finite-sample properties

We investigate the finite-sample properties of the Fourier estimator of the spot volatility defined in (2.30) in four simulated scenarios, which generate incrementally more realistic microstructure noise patterns. We begin with a stylized scenario (scenario I) where the additive noise is i.i.d. Gaussian and independent of the price, the leverage parameter is constant, the volatility of volatility is a linear function of the volatility and price observations are sampled on an equally-spaced grid. We then progressively allow for stochastic leverage and volatility of volatility (scenario II), auto-correlated and price-dependent noise, with t-distributed innovations (scenario III) and, finally, a random price-sampling scheme (scenario IV). Firstly, we study the sensitivity of the estimator performance to the cutting frequencies N and M, with a focus on the relationship between the (unfeasible) optimal choice of the cutting frequencies and the level of the noise-to-signal ratio. Secondly, we introduce a feasible adaptive method for selecting N and M, which optimizes the integrated asymptotic-error variance in the rate-efficient case, see result (2.33) in Theorem 2.2.7, and test its robustness when the hypotheses of Theorem 2.2.7 are partially violated. Finally, we compare the finite-sample performance of the Fourier estimator with those of alternative noise-robust spot volatility estimators.

2.3.1 Simulation design

For each of the following scenarios, we simulate 10^4 daily trajectories of noisy log-price observations. As we use the trading day as unit of time, the horizon T of the simulated trajectories is equal to 1.

For simulating scenarios I, II and III we consider the equally-spaced price sampling grid with mesh $\rho(n) = T/n$ equal to 1 second, which implies $n = 23400$, as we assume the trading day to be 6.5-hour long; for the simulation of scenario IV, we employ a random sampling scheme, which is detailed in Section 2.3.1.

Scenario I

In the first scenario, we simulate the efficient log-price process using the classical model by Heston (1993), whose dynamics satisfy Assumptions (A.I), (A.II) and (A.III). The model reads as

$$\begin{cases} dp(t) & = \sigma(t)dW_t + \left(\mu - \frac{1}{2}\sigma^2(t)\right)dt \\ d\sigma^2(t) & = \gamma\sigma(t)dZ_t + \theta\left(\alpha - \sigma^2(t)\right)dt \end{cases}, \tag{2.81}$$

where W and Z are two Brownian motions with constant correlation ρ and $\mu, p(0) \in \mathbb{R}, \gamma, \theta, \alpha, \sigma^2(0) > 0$. For the simulation, we set $(\mu, \theta, \alpha, \gamma, \rho) = (0.001, 1, 0.3, 0.2, -0.8)$. Further, we set $p(0) = \ln(100)$, while $\sigma^2(0)$ is drawn, trajectory-wise, from the stationary distribution of σ^2, that is, from $\Gamma\left(u_1 = \frac{2\theta\alpha}{\gamma^2}, u_2 = \frac{\gamma^2}{2\theta}\right)$, where u_1 and u_2 denote, respectively, the shape and scale parameters.

For simulating the additive noise, we assume that, for any i, η_{t_i} is independent of p and such that $\eta_{t_i} \sim_{i.i.d.} \mathcal{N}(0, \xi)$ and $\sqrt{\xi} = \zeta\sqrt{\alpha\rho(n)}$, where ζ denotes the noise-to-signal ratio, expressed as the quotient between the standard deviation of the noise, i.e., $\sqrt{\xi}$, and the steady-state standard deviation of the efficient log-return sampled at the highest available frequency, i.e., $\sqrt{\alpha\rho(n)}$, with $\rho(n)$ fixed and equal to 1/23400. In the simulation, we adopt three increasing values of ζ, namely, $\zeta = 1, 2, 3$.

Scenario II

In the second scenario, we simulate efficient log-price observations from the following model:

$$\begin{cases} dp(t) & = \sigma(t)dW_t + \left(\mu - \frac{1}{2}\sigma^2(t)\right)dt \\ d\sigma^2(t) & = \gamma(t)dZ_t + \theta_1(\alpha_1 - \sigma^2(t))dt \\ d\gamma^2(t) & = \beta\gamma(t)dY_t + \theta_2(\alpha_2 - \gamma^2(t))dt \end{cases}, \tag{2.82}$$

where W, Z and Y are three Brownian motions and $\mu, p(0) \in \mathbb{R}$, $\theta_1, \alpha_1, \theta_2, \alpha_2, \beta, \sigma^2(0), \gamma^2(0) > 0$. The correlations of the pairs (W, Z) and (Z, Y) satisfy the model by Veraart and Veraart (2012), that is,

$$\begin{cases} dW_t & = \phi_1(t)dZ_t + \sqrt{1 - \phi_1^2(t)}dZ_t^{\perp} \\ d\phi_1(t) & = \kappa_1\sqrt{1 - \phi_1^2(t)}dB_t^{(1)} + \lambda_1(\nu_1 - \phi_1(t))dt \end{cases} \tag{2.83}$$

and

$$\begin{cases} dZ_t & = \phi_2(t)dY_t + \sqrt{1 - \phi_2^2(t)}dY_t^{\perp} \\ d\phi_2(t) & = \kappa_2\sqrt{1 - \phi_2^2(t)}dB_t^{(2)} + \lambda_2(\nu_2 - \phi_2(t))dt \end{cases}, \tag{2.84}$$

where: $Z^{\perp}$ and $Y^{\perp}$ denote two Brownian motions that are independent of, respectively, Z and Y; $B^{(1)}$ and $B^{(2)}$ are two additional Brownian motions which are independent of each other and independent of W, Z and Y; for $i = 1, 2$, $k_i \in \mathbb{R}, \lambda_i > 0, \nu_i \in [-1, 1]$. For the simulation, we set

$(\mu, \theta_1, \alpha_1, \beta, \theta_2, \alpha_2) = (0.001, 8, 0.3, 0.5, 10, 0.1)$ and $(\kappa_1, \lambda_1, \nu_1, \kappa_2, \lambda_2, \nu_2) = (0.1, 0.2, -0.6, 0.6, 1, 0.5)$. The initial conditions are set as $p(0) = \ln(100), \sigma^2(0) = \alpha_1, \gamma^2(0) = \alpha_2, \phi_1(0) = \nu_1$ and $\phi_2(0) = \nu_2$.

Compared to the Heston model in (2.81), the model in (2.82-2.84) is capable of generating more realistic dynamics of the efficient log-price. Specifically, it allows for a stochastic volatility of volatility γ^2, a stochastic leverage ϕ_1 and a stochastic correlation between the volatility and the volatility of volatility ϕ_2. Empirical evidence supporting the existence of a stochastic leverage can be found, e.g., in Kalnina and Xiu (2017). Moreover, the empirical study by Toscano et al. (2022) supports the existence of a mean-reverting stochastic volatility of volatility that displays a time-varying positive correlation with the volatility.

The additive noise is simulated using the same specification adopted in scenario I, with, again, $\zeta = 1, 2, 3$.

Scenario III

In the third scenario, the efficient price process is simulated via the same model employed in scenario II, that is, the model in (2.82)(2.83)-(2.84). Instead, for simulating the additive noise we use a specification of the model introduced by Jacod et al. (2017). For any i, the specification reads as

$$\eta_{t_i} = \psi_{t_i} \chi_{t_i}, \tag{2.85}$$

where: ψ_{t_i} is sampled from the continuous-time process $\psi(t)$, whose dynamics satisfy

$$d\psi(t) = \theta_\psi(\alpha_\psi - \psi(t))dt + \sigma_\psi dW_t,$$

with $\psi(0), \alpha_\psi, \theta_\psi, \sigma_\psi > 0$; χ_{t_i} is a discrete-time process, independent of W, whose dynamics satisfy an AR(1) model, that is, $\chi_{t_i} = a\chi_{t_{i-1}} + e_{t_i}$, where $|a| < 1$ and e_{t_i} is an i.i.d. sequence of t-distributed random variables with $v > 4$ degrees of freedom.
For the simulation of χ, we set $a = 0.6$ and $v = 5$. As for the simulation of ψ, we consider three alternative choices for the parameter vector $(\alpha_\psi, \theta_\psi, \sigma_\psi)$, namely: $\mathrm{v}_1 := (0.0001, 10, 0.01)$, $\mathrm{v}_2 := (0.0002, 10, 0.02)$ and $\mathrm{v}_3 := (0.0003, 10, 0.03)$. The initial condition is set as $\psi(0) = \alpha_\psi$, independently of the parameter vector chosen for ψ. Let $\bar{\xi} := \frac{\nu}{\nu-2}\frac{1}{1-a^2}\left(\alpha_\psi^2 + \frac{\sigma_\psi^2}{2\theta_\psi}\right)$ denote the steady-state variance of η. Hence, by imposing that $\sqrt{\bar{\xi}} = \zeta\sqrt{\alpha_1 \rho(n)}$, with α_1 denoting the steady-state variance in (2.82) and $\rho(n) = 1/23400$, then $\mathrm{v}_1, \mathrm{v}_2$ and v_3 correspond to a noise-to-signal ratio ζ (approximately) equal to $1, 2$ and 3, respectively.

The specification in (2.85) generates more realistic noise patterns, compared to the one adopted in scenarios I and II, in that it allows for auto-correlation in the noise through χ and for dependence between the efficient log-price and the noise through ψ, which is driven by the same Brownian motion as the efficient log-price.

Scenario IV

The fourth scenario differs from the third only for the different price sampling scheme employed. As outlined in Remark 2.1.2, the Fourier estimation methodology does not require equidistant observations as inputs. Accordingly, in the last scenario we adopt an irregular random sampling scheme which assumes that observation times follow an non-homogeneous Poisson process. Denoting by ι

the time-dependent rate, we set

$$\iota(t) = 23400\left(1 + \frac{1}{2}\cos\left(\frac{2\pi}{T}t\right)\right). \tag{2.86}$$

The specification in (2.86), which is also adopted in the simulation study by Li and Linton (2022), mimics the clustering of trades in the beginning and the end of a trading day, which generates the typical intra-day volatility U-shape.

2.3.2 Performance sensitivity to N and M

For the different noise-to-signal ratio ζ considered, we investigate the sensitivity of the finite-sample performance of the Fourier spot volatility estimator to the cutting frequencies N and M. For the implementation of the estimator, we use the entire sample of available high-frequency prices, corresponding to $n = 23400$ in the case of regular sampling. As for N and M, based on the conditions of the rate-efficient Central Limit Theorem, see (2.33) in Theorem 2.2.7, we set $N = \lfloor c\sqrt{n} \rfloor$ and $M = \lfloor a\sqrt{N} \rfloor$. We then evaluate the sensitivity of the finite-sample performance to changes in the constants c and a. Specifically, we allow c and a to vary, respectively, in the range $[1, 10]$ and $[0.1, 0.5]$, with a step size equal to, respectively, 1 and 0.1. Following Mancino and Recchioni (2015, Section 4.2), we reconstruct the spot volatility path on the sparser equally-spaced grid with mesh equal to 1 minute.

The finite-sample performance is assessed based on the mean integrated squared error (MISE), which is defined as $MISE := E\left[\int_0^T \left(\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right)^2 dt\right]$. Figure 2.1 shows the MISE as a function of the pair (c, a) in the four scenarios, for the different noise intensities considered. Further, Table 2.1 illustrates the values of the pair (c, a) that minimize the MISE, along with the resulting optimized MISE values.

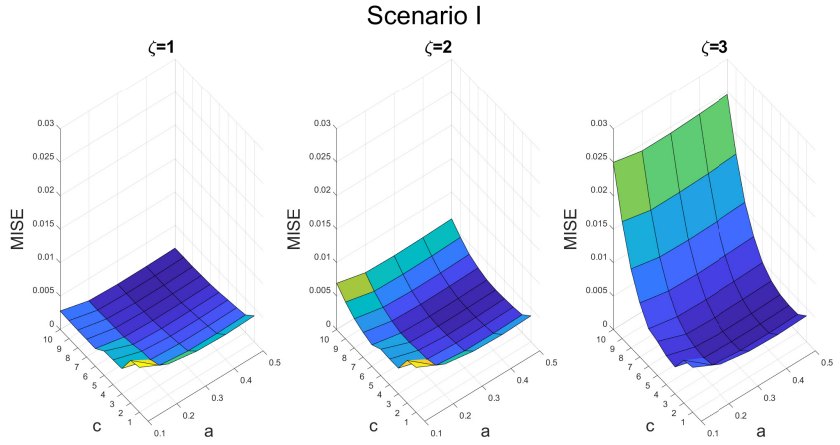

Figure 2.1: Scenario I: MISE as a function of (c, a) for the different values of the noise-to-signal ratio ζ.

Based on Figure 2.1, the finite-sample efficiency of the estimator appears to be much more sensitive to changes of c for a fixed a than viceversa. Moreover, the sensitivity to c appears to

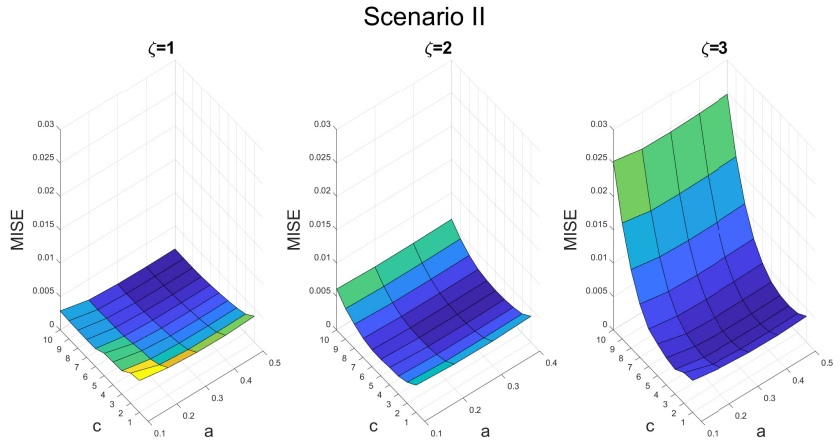

Figure 2.2: Scenario II: MISE as a function of (c, a) for the different values of the noise-to-signal ratio ζ.

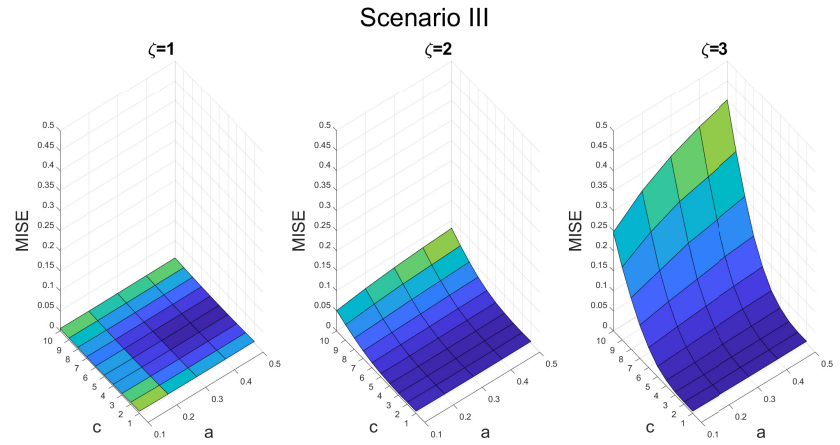

Figure 2.3: Scenario III: MISE as a function of (c, a) for the different values of the noise-to-signal ratio ζ.

be amplified with the increase of the noise-to-signal ratio. In other words, in all the scenarios considered, a sub-optimal selection of c appears to determine a larger performance deterioration when the intensity of the noise is higher. Instead, the lack of sensitivity to a seems to be independent of the noise intensity and the scenario.

Furthermore, Table 2.1 suggests that the optimal choice of c is smaller in correspondence of a larger noise-to-signal ratio. Additionally, it is worth noting that the optimal values of c are approximately halved when moving from an i.i.d. noise specification (scenarios I and II) to an auto-correlated, price-dependent noise specification (scenarios III and IV). Instead, the optimal choice of a appears to be noise- and scenario-insensitive, oscillating between 0.3 and 0.4. The results illustrated in Table 2.1 are in line with the fact that, when the amount of noise present in the data is larger, a smaller value of the cutting frequency N is needed, in the convolution formula (2.3), in order to

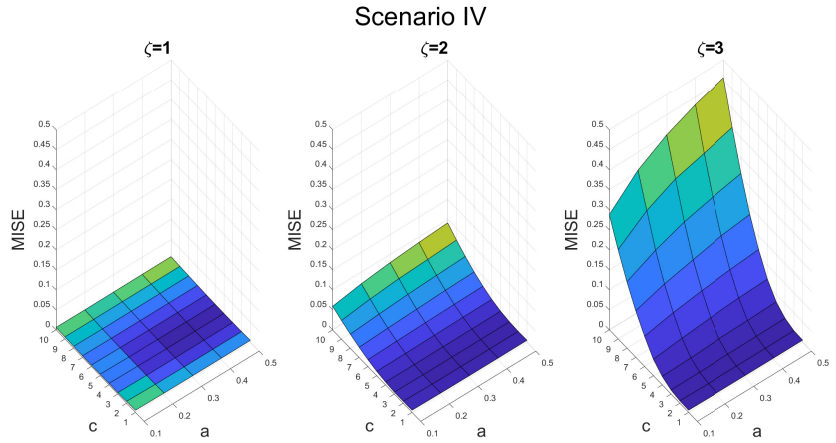

Figure 2.4: Scenario IV: MISE as a function of (c, a) for the different values of the noise-to-signal ratio ζ.

scenario	ζ	optimal (c, a)	MISE
I	1	$(8, 0.3)$	$1.4 \cdot 10^{-3}$
	2	$(5, 0.3)$	$2.1 \cdot 10^{-3}$
	3	$(3, 0.3)$	$2.5 \cdot 10^{-3}$
II	1	$(8, 0.4)$	$1.5 \cdot 10^{-3}$
	2	$(5, 0.4)$	$2.2 \cdot 10^{-3}$
	3	$(3, 0.4)$	$2.6 \cdot 10^{-3}$
III	1	$(4, 0.4)$	$2.2 \cdot 10^{-3}$
	2	$(2, 0.4)$	$3.3 \cdot 10^{-3}$
	3	$(2, 0.4)$	$3.8 \cdot 10^{-3}$
IV	1	$(4, 0.3)$	$2.5 \cdot 10^{-3}$
	2	$(2, 0.3)$	$3.3 \cdot 10^{-3}$
	3	$(1, 0.3)$	$4.0 \cdot 10^{-3}$

Table 2.1: MISE-optimal selection of (c, a) and resulting MISE values in the different scenarios and for the different values of the noise-to-signal ratio ζ.

efficiently filter out the noise and reconstruct the volatility coefficients.

2.3.3 Feasible method for selecting N and M

The feasible method for selecting the pair (N, M) that we propose aims at optimizing the conditional MISE (c-MISE) of the Fourier spot volatility estimator under the constraint $M < N < n$, for the given available sample size n. Define the c-MISE as

$$c\text{-}MISE := E^{\sigma}\left[\int_0^T \left(\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right)^2 dt\right],$$

where E^σ denotes the conditional expectation with respect to the filtration generated by $(\sigma_t)_{t\in[0,T]}$. As shown in Zu and Boswijk (2014, Section 3.4), the following decomposition holds:

$$c\text{-}MISE = \int_0^T E^\sigma \left[\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right]^2 dt + \int_0^T Var^\sigma \left[\widetilde{\sigma}^2_{nNM}(t) - \sigma^2(t)\right] dt,$$

where Var^σ denotes the conditional variance with respect to the filtration generated by $(\sigma_t)_{t\in[0,T]}$. Under the hypotheses that yield the rate-efficient result in Theorem 2.2.7, Equation (2.33), for a sufficiently large n, the c-MISE is approximated by the conditional asymptotic integrated mean squared error (c-AMISE). After scaling the unit of time, so that the estimation horizon becomes $[0,T]$, the c-AMISE reads as

$$c\text{-}AMISE = \frac{M}{N}\frac{2}{3}\int_0^T \sigma^4(t)\,dt + \frac{1}{M}\frac{T}{3}\int_0^T \gamma^2(t)\,dt + \frac{M\,N}{n}\frac{4T}{9}\,\xi\int_0^T \sigma^2(t)\,dt + \frac{N^3M}{n^2}\frac{T^3}{15}\,\xi^2. \quad (2.87)$$

The c-AMISE in (2.87) depends on the integrated volatility $IV_T := \int_0^T \sigma^2(t)dt$, the integrated quarticity $IQ_T := \int_0^T \sigma^4(t)dt$, the integrated volatility-of-volatility $IVV_T := \int_0^T \gamma^2(t)dt$ and the noise variance ξ. All these quantities can be estimated with high-frequency prices and thus the c-AMISE can be treated as observable. Specifically, for estimating IV_T, IQ_T and IVV_T we propose to use Fourier-transform based estimators. We refer the reader to, respectively, Mancino and Sanfelici (2008), Mancino and Sanfelici (2012) and Sanfelici et al. (2015) for the study of the finite-sample efficiency of the Fourier estimators of IV_T, IQ_T and IVV_T in the presence of noise and guidance on the practical implementation thereof with high-frequency prices. Moreover, for estimating ξ, one can use the estimator proposed in Bandi and Russell (2006, Section 3).

For brevity, denote the c-AMISE in (2.87) as a function of (N,M) by $\Psi(N,M)$. Moreover, let $\widehat{\Psi}(N,M)$ indicate the estimator obtained by plugging the Fourier estimators of IV_T, IQ_T and IVV_T, along with the estimator of ξ by Bandi and Russell (2006), into (2.87), in place of the respective true latent quantities. The feasible selection of (N,M) for a given n amounts to solving the constrained optimization problem

$$\min_{(N,M)\in S} \widehat{\Psi}(N,M),$$

where $S = \left[\underline{N},\overline{N}\right]\times\left[\underline{M},\overline{M}\right] \subset \mathbb{R}^+\times\mathbb{R}^+$ is such that $M < N < n$. To do so, we propose to use the following adaptive method, based on the gradient-descent algorithm (see, e.g., Ruder (2017)).

Adaptive method for the selection of the pair (N,M)
Given the initial condition $(N_0 = \underline{N}, M_0 = \underline{M})$, for $k = 1,...,K$, follow the update rule[a]

$$N_k = \left(\underline{N} \vee N_{k-1} - \lambda\frac{\partial\,\widehat{\Psi}}{\partial N}(N_{k-1},M_{k-1})\right)\wedge\overline{N},$$

$$M_k = \left(\underline{M} \vee M_{k-1} - \lambda\frac{\partial\,\widehat{\Psi}}{\partial M}(N_{k-1},M_{k-1})\right)\wedge\overline{M},$$

where λ is a positive constant, typically referred to as the *learning rate* parameter. The

optimal value (N^*, M^*) is achieved as soon as the marginal absolute change of the objective function falls below a given threshold ϑ, i.e., in correspondence of the smallest k such that

$$\delta_k := \frac{\left|\widehat{\Psi}(N_k, M_k) - \widehat{\Psi}(N_{k-1}, M_{k-1})\right|}{\widehat{\Psi}(N_{k-1}, M_{k-1})} < \vartheta.$$

Instead, if $\delta_k > \vartheta \ \forall k$, make the final selection $(N^*, M^*) = (N_K, M_K)$.

[a]The notation $\frac{\partial \widehat{\Psi}}{\partial x}(u, v)$ is shorthand for $\frac{\partial \widehat{\Psi}(N,M)}{\partial x}\Big|_{N=u, M=v}$.

For a given n, the implementation of the proposed adaptive method for selecting (N, M) requires the selection of the constraint region S and the vector of optimization parameters (λ, ϑ, K). In this regard, for the largest n available[3], we set

$$S = \left[\lfloor 2^{-1} n^{1/4} \rfloor, \lfloor 10 n^{1/4} \rfloor\right] \times \left[\lfloor n^{1/2} \rfloor, \lfloor 20 n^{1/2} \rfloor\right] \tag{2.88}$$

and

$$(\lambda, \vartheta, K) = (\propto \hat{\xi}^{-1}, 10^{-4}, 10^4). \tag{2.89}$$

The choice of S is rather conservative, as it allows for a fairly large range of variation for N and M. The selection of ϑ and K is also conservative, since it ensures that the stopping rule $\delta_k < \vartheta$ is always satisfied for some $k \leq K$ in the simulation study. In general, the choice of the learning rate λ is known to play a crucial role in determining the accuracy of the gradient-descent algorithm, see Ruder (2017) and the references therein. With regard to the specific optimization algorithm presented in this section, numerical evidence suggests that selecting a value for λ which is inversely proportional to the (estimated) variance of the noise (that is, to the order of magnitude of the tick-by-tick variance of the noisy log-returns) may yield satisfactory estimation accuracy. Specifically, after setting

$$\lambda = c_\lambda \, \hat{\xi}^{-1}, \tag{2.90}$$

we numerically evaluated the efficiency of the adaptive method for different values of c_λ in the region $[0.1, 10]$ and found that the MISE is optimized if c_λ is chosen in a neighborhood of $c_\lambda = 5$ in the case of scenarios I and II, or in a neighborhood of $c_\lambda = 1$ in the case of scenarios III and IV. Based on the higher degree of realism of scenarios III and IV, for empirical applications (see Section 2.4) we thus recommend the choice $c_\lambda = 1$. As for robustness, it is worth noting that MISE values appear to be rather stable for $c_\lambda \in [3, 7]$ in scenarios I and II, while in scenarios III and IV stability is achieved only on a much smaller window, that is, only if $c_\lambda \in [0.5, 1.5]$.

The efficiency of the feasible adaptive method for selecting N and M is summarized in Table 2.2. Compared to the unfeasible results reported in Table 2.1, the feasible results illustrated in Table 2.2 appear to be satisfactory. Indeed, based on the values of Δ_{MISE}, which denotes the relative change in the MISE compared to the corresponding unfeasible MISE reported in Table 2.1, the loss of efficiency ranges between 4% and 18%. The average c produced by the feasible parameter selection method ($\bar{c}$ in Table 2.2) is smaller than the optimal value of c reported in Table 2.1, except

[3]Recall that n is equal to 23400 in scenarios I, II and III, in which we adopt the equally-spaced one-second sampling grid. Instead, in scenario IV, n is random, with expected value equal to 23400, see (2.86).

in scenario II, for $\zeta = 2$, and in scenario III, for $\zeta = 2, 3$. The average value of the feasible optimal a, denoted by $\bar{a}$, is instead close to the unfeasible optimal value, as it stays in the range $[0.3, 0.4]$. Finally, note that the feasible method for selecting N and M provides a satisfactory performance also when the assumptions of Theorem 2.2.7 are partially violated, that is, in scenarios II, III and IV, thereby showing robustness in view of empirical applications.

scenario	ζ	$(\bar{c}, \bar{a})$	MISE	Δ_{MISE}
I	1	$(5.3, 0.4)$	$1.6 \cdot 10^{-3}$	+14%
	2	$(3.2, 0.4)$	$2.3 \cdot 10^{-3}$	+10%
	3	$(2.6, 0.4)$	$2.6 \cdot 10^{-3}$	+4%
II	1	$(5.7, 0.3)$	$1.7 \cdot 10^{-3}$	+13%
	2	$(3.3, 0.3)$	$2.5 \cdot 10^{-3}$	+14%
	3	$(2.7, 0.4)$	$2.9 \cdot 10^{-3}$	+12%
III	1	$(3.4, 0.4)$	$2.4 \cdot 10^{-3}$	+9%
	2	$(2.7, 0.4)$	$3.6 \cdot 10^{-3}$	+9%
	3	$(1.7, 0.3)$	$4.3 \cdot 10^{-3}$	+13%
IV	1	$(3.4, 0.3)$	$2.6 \cdot 10^{-3}$	+4%
	2	$(2.3, 0.3)$	$3.9 \cdot 10^{-3}$	+18%
	3	$(1.7, 0.3)$	$4.7 \cdot 10^{-3}$	+18%

Table 2.2: MISE values produced by the feasible adaptive method for selecting N and M, in the different scenarios and for the different values of the noise-to-signal ratio ζ considered. Letting $N = c\sqrt{n}$ and $M = a\sqrt{N}$, $\bar{c}$ and $\bar{a}$ denote the sample averages of, respectively, the values of c and the values of a produced by the feasible procedure. Moreover, Δ_{MISE} denotes the relative change in the MISE compared to its unfeasible values reported in Table 2.1.

2.3.4 Performance comparison

We compare the finite-sample performance of the Fourier spot volatility estimator with that of three alternative noise-robust estimators, namely the Smoothed Two-Scale estimator (STS) by Mykland et al. (2019), the Pre-Averaging Kernel estimator by Figueroa-López and Wu (2022) (PAK) and the Pre-Averaging ReMeDI (PAR) estimator by Li and Linton (2022). Differently from the Fourier estimator, these estimators are local estimators based on the numerical differentiation of the estimated integrated variance. Moreover, they all rely on price pre-averaging and a bias correction. We briefly recall their definitions.

The STS estimator is defined as

$$\widetilde{\sigma}^2_{STS}(t) := \frac{1}{s_n}\left(\widehat{\langle p,p\rangle}_{t+s_n} - \widehat{\langle p,p\rangle}_t\right),$$

where $\{s_n\}_{n\geq 1}$ denotes a sequences of positive numbers such that $s_n \to 0$ as $n \to \infty$ and $\widehat{\langle p,p\rangle}_t$ indicates the integrated STS estimator on $[0,t]$. Specifically, consider the grid $\mathcal{S} = \{\tau_{n,0} = 0 < \tau_{n,1} < ... < \tau_{n,n_b} = T\}$ and let $n_b^*(t) = \sup\{1 \leq i \leq n_b : \tau_{n,i} \leq t\}$ and $M_{n,i} = \#\{j : t_j \in (\tau_{n,i-1}, \tau_{n,i}]\}$. It holds that

$$\widehat{\langle p,p\rangle}_t = \frac{1}{(1-(K+J)/n_b)(K-J)}\left(K\widetilde{[\bar{P},\bar{P}]}_t^{(K)} - J\widetilde{[\bar{P},\bar{P}]}_t^{(J)}\right),$$

where

$$K\widetilde{[\bar{P},\bar{P}]}_t^{(K)} = \left(\frac{1}{2}\sum_{i=1}^{J} + \sum_{i=J+1}^{n_b^*(t)-(K+J)} + \frac{1}{2}\sum_{i=n_b^*(t)-(K+J)+1}^{n_b^*(t)-K}\right) \times \left(\bar{P}_{\tau_{i+K}} - \bar{P}_{\tau_i}\right)^2, \tag{2.91}$$

$$\bar{P}_{\tau_{n,i}} = \frac{1}{M_{n,i}} \sum_{t_j \in (\tau_{n,i-1}, \tau_{n,i}]} \tilde{p}_{t_j}$$

and $J\widetilde{[\bar{P},\bar{P}]}_t^{(JK)}$ is obtained by switching K and J in (2.91). For the implementation of the STS estimator, we set $\mathcal{S} = \{iT/n_b, i = 0,1,...,n_b\}$ and select n_b, s_n, J and K using as guidance the analysis presented in Section 6.3 of Chen et al. (2020).

The PAK estimator reads

$$\widetilde{\sigma}^2_{PAK}(t) := \frac{1}{\phi_{k_n}(g)} \sum_{i=1}^{n-k_n+1} \frac{1}{m_n\rho(n)}\, K\left(\frac{t_i - t}{m_n\rho(n)}\right)\left(\bar{P}^2_{t_i} - \frac{1}{2}\widehat{P}_{t_i}\right),$$

where $k_n = \lfloor 1/\left(c_k\sqrt{\rho(n)}\right)\rfloor$, $m_n = c_m\rho(n)^{-3/4}$, $K(x)$ is a kernel function satisfying Assumption 2 of Figueroa-López and Wu (2022) and

$$\bar{P}_{t_i} = -\sum_{j=1}^{k_n}\left(g\left(\frac{j}{k_n}\right) - g\left(\frac{j-1}{k_n}\right)\right)\widetilde{p}_{t_{i+j-2}}, \tag{2.92}$$

$$\widehat{P}_{t_i} = \sum_{j=1}^{k_n}\left(g\left(\frac{j}{k_n}\right) - g\left(\frac{j-1}{k_n}\right)\right)^2 \delta_{i+j-2}(\widetilde{p})^2, \quad \phi_{k_n}(g) = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)^2,$$

where g denotes a real-valued weight function on $[0,1]$ which is continuous and piece-wise C^1, with piece-wise Lipschitz derivative g' and such that $g(0) = g(1) = 0$, $\int_0^1 g^2(s)ds = 1$. In order to apply the PAK estimator, we select $K(x) = \frac{1}{2}\mathrm{e}^{-|x|}$, based on the discussion in Figueroa-López and Wu (2022, Section 3). As for the function g, we select $g(x) = x \wedge (1-x)$, as in Jacod et al. (2009). Further, we choose the constants c_k and c_m following Figueroa-López and Wu (2022, Section 3).

The PaReMeDI estimator is defined as

$$\widetilde{\sigma}^2_{PAR}(t) = \frac{1}{d_{n(t)}h_n\phi_0^n}\sum_{i=n(t)}^{n(t)+l_n-h_n}\bar{P}^2_{t_i} - \frac{1}{d_{n(t)}h_n^2\phi_0^n}\sum_{i=n(t)+k_n}^{n(t)+l_n-h_n}\sum_{|l|\le l_n}\tilde{\phi}_l^n\left(\tilde{p}_{t_{i+l}} - \tilde{p}_{t_{i+l+k_n}}\right)\left(\tilde{p}_{t_i} - \tilde{p}_{t_{i-k_n}}\right)$$

where $\{h_n\}_{n\ge1}$, $\{l_n\}_{n\ge1}, \{k_n\}_{n\ge1}$ are sequences of integers, $n(t)$ denotes the number of observations up to time t, $d_{n(t)} = n(t+l_n) - n(t)$, $\bar{P}(t_i)$ is given in (2.92), the function g is the same as in the case of the PAK estimator and, finally,

$$\phi_0^n = \frac{1}{h_n}\sum_{i\in\mathbb{Z}} g\left(\frac{i}{h_n}\right)^2, \quad \tilde{\phi}_l^n = h_n\sum_{i\in\mathbb{Z}}\tilde{g}_i^n\tilde{g}_{i-l}^n, \quad \tilde{g}_i^n = \left(\frac{i+1}{h_n}\right) - g\left(\frac{i}{h_n}\right).$$

For the implementation of the PAR estimator, we choose again $g(x) = x \wedge (1-x)$, as in Li and

Linton (2022), and select h_n, k_n and l_n based on the discussion in Li and Linton (2022, Section 5).

The MISE values resulting from the application of the STS, PAK and PAR estimators on the 1-minute estimation grid are shown in Table 2.3. Based on the comparison of Table 2.3 with Table 2.2, the feasible application of the Fourier estimator yields the lowest MISE in every scenario and for every noise-to-signal ratios considered, with the only exception of scenario I, where, for $\zeta = 2, 3$, the PAK estimator shows a slightly better performance. These numerical evidences may be explained by the fact that the Fourier estimation method is based on the computation of integrated quantities (namely, the Fourier coefficients of the volatility process) rather than on a differentiation procedure, with benefits in terms of improved numerical stability, as already remarked in Malliavin and Mancino (2002) and Malliavin and Mancino (2009). It is also worth noticing that the Fourier method is relatively more parsimonious, compared to alternative methods considered, since it is based on the selection of only two parameters, N and M. In other words, by requiring no pre-averaging and no bias-correction, the efficient implementation of the Fourier method in the presence of noise entails only the careful reduction of N, without the need to select any additional parameters or auxiliary functions.

scenario	ζ	STS	PAK	PAR
I	1	$5.1 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$
	2	$5.0 \cdot 10^{-3}$	$2.2 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$
	3	$4.9 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$	$4.4 \cdot 10^{-3}$
II	1	$7.5 \cdot 10^{-2}$	$3.5 \cdot 10^{-3}$	$4.4 \cdot 10^{-3}$
	2	$8.2 \cdot 10^{-2}$	$4.1 \cdot 10^{-3}$	$4.7 \cdot 10^{-3}$
	3	$9.2 \cdot 10^{-2}$	$4.4 \cdot 10^{-3}$	$5.2 \cdot 10^{-3}$
III	1	$8.7 \cdot 10^{-3}$	$4.7 \cdot 10^{-3}$	$5.3 \cdot 10^{-3}$
	2	$8.7 \cdot 10^{-3}$	$7.9 \cdot 10^{-3}$	$8.2 \cdot 10^{-3}$
	3	$9.5 \cdot 10^{-3}$	$8.1 \cdot 10^{-3}$	$8.8 \cdot 10^{-3}$
IV	1	$8.9 \cdot 10^{-3}$	$5.2 \cdot 10^{-3}$	$6.5 \cdot 10^{-3}$
	2	$9.0 \cdot 10^{-3}$	$8.7 \cdot 10^{-3}$	$8.5 \cdot 10^{-3}$
	3	$9.3 \cdot 10^{-2}$	$9.0 \cdot 10^{-3}$	$8.6 \cdot 10^{-3}$

Table 2.3: MISE values produced by the STS, PAK and PAR estimators in the different scenarios and for the different values of the noise-to-signal ratio ζ.

2.4 Empirical study

The accuracy of spot volatility estimates can be investigated empirically by testing the distribution of high-frequency standardized log-returns, see Mancino and Recchioni (2015, Section 4). We briefly recall the testing procedure. Define the standardized log-return on $[t, t+h]$, $h > 0$, as

$$z_h(t) := \frac{r_h(t)}{\sqrt{\sigma^2(t)\, h}}, \tag{2.93}$$

where $r_h(t) := p(t+h) - p(t)$. Moreover, denote by $\hat{z}_h(t)$ the estimated standardized log-return, obtained by replacing $\sigma^2(t)$ in (2.93) with an estimate thereof. Under Assumption (A.I), if h is

sufficiently small, the sequence $\{z_h(t)\}_{t\in\tau}$, $\tau = \{0, h, 2h, ..., (\lfloor T/h \rfloor - 1)h\}$, approximates a sequence of i.i.d. standard Normal random variables. Thus, by testing the empirical distribution of the estimated standardized log-returns $\{\hat{z}_h(t)\}_{t\in\tau}$ for Gaussianity, one can obtain insight into the accuracy of spot volatility estimates.

We empirically implement the Fourier spot volatility estimator in (2.30) using high-frequency prices of three US stocks, Intel (INTC), Comcast (CMCSA) and Starbucks (SBUX), sampled over the period December 1, 2022 - December 30, 2022, which corresponds to 21 trading days. During this period, the average trading volume of INTC, CMCSA and SBUX is approximately equal to, respectively, 14, 7 and 2 million shares, reflecting three different liquidity levels.

Spot volatility paths are estimated on the daily horizon, that is, on the interval from 9:30 a.m. to 4:00 p.m. We measure time in trading days, hence we set $T = 1$. The Fourier estimator is implemented using tick-by-tick log-returns. As a consequence, the sample size n employed changes from day to day, reflecting the number of ticks available on each day. For the given n available, on each day we select N and M using the adaptive method described in Section 2.3.3. The rectangle S and the optimization parameters λ, ϑ, K are chosen as in, respectively, (2.88) and (2.89). Specifically, we set $c_\lambda = 1$, as in (2.90).

Price jumps are identified using the test by Lee and Mykland (2012) and removed from the sample employed for the estimation[4]. Overnight returns, that is, the difference between the opening log-price and the closing log-price of the previous day, are not employed in the estimation, as they often contain jumps due to the fact that market-related news are typically released when the market is closed.

To compute the sequence of the standardized log-returns, one needs to select a value for the interval h which is, at the same time, small enough to satisfactorily approximate their true distribution and large enough to make the impact of micro-structure noise negligible. Based on a common practice adopted in the literature, we select h equal to 5 minutes. Accordingly, Fourier estimates of the spot volatility are obtained on the equally-spaced grid with mesh size equal to 5 minutes. However, it is worth stressing that the Fourier method allows selecting an arbitrary grid for the estimation of the volatility path and that the selection of the 5-minute grid is only convenient for the purpose of the testing of the distribution of the reconstructed standardized log-return. The sequences of the estimated daily volatility paths of INTC, CMCSA and SBUX is displayed in Figure 2.5.

To evaluate the accuracy of estimates we perform the tests by Jarque and Bera (1987) and Anderson and Darling (1952) at the 95% confidence level on the 21 daily sub-samples of estimated standardized log-returns. Table 2.4 illustrates the number of rejections of the null hypothesis of Gaussianity and the average p-values.

The results in Table 2.4 suggest that the daily empirical distributions of the reconstructed standardized log-returns approximate a Gaussian distribution quite well for each of the three stocks considered, thereby suggesting that the Fourier spot volatility estimates obtained are accurate under the different liquidity levels considered.

[4]As in Lee and Mykland (2012, Section 6), we apply the jump test at the 99% confidence level on seven intra-day sub-intervals (namely, 9.30-10, 10-11, 11-12, 12-13, 13-14, 14-15 and 15-16) and, if the null hypothesis of absence of jumps is rejected on a given sub-interval, we replace the highest log-return in absolute value on that interval with a zero log-return. For each sub-interval the procedure is iterated until the test does not reject the null. Overall, we find evidence of the presence of jumps for 9 days (CMCSA), 9 days (INTC) and 12 days (SBUX).

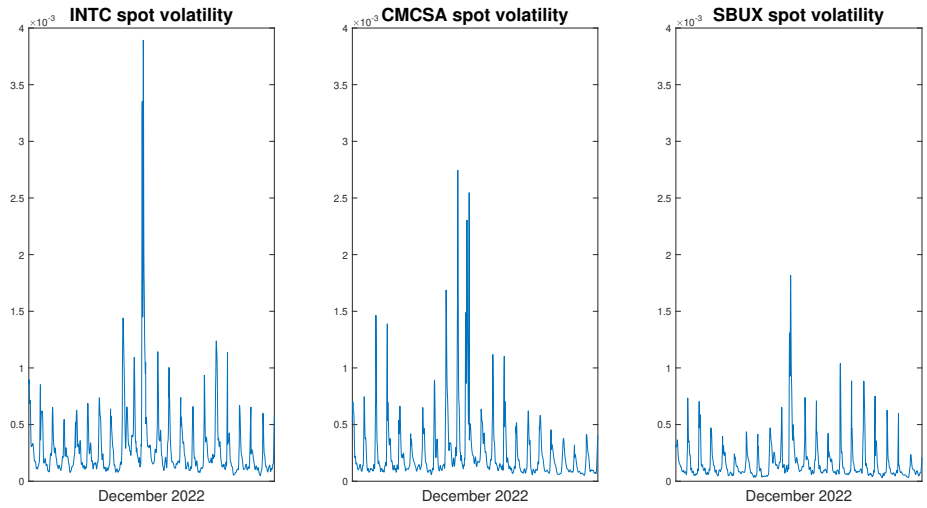

Figure 2.5: Sequences of the estimated daily volatility paths of CMCSA, INTC and SBUX in December 2022, on the 5-minute grid.

	JB test		AD test	
stock	rej. freq.	av. p-value	rej. freq.	av. p-value
INTC	2/21	0.520	1/21	0.584
CMCSA	1/21	0.498	1/21	0.488
SBUX	1/21	0.433	2/21	0.408

Table 2.4: Rejection frequencies and average p-values of the Jarque-Bera test (JB) and the Anderson-Darling (AD) test, performed at the 95% confidence level on the 21 daily samples of 5-minute reconstructed standardized returns referring to the period December 1, 2022 - December 30, 2022.

A similar analysis has also been carried on considering a stock index, and using in particular the S&P500 high-frequency trade prices, sampled over the period March, 2018 - April, 2018. The removal of intra-day jumps is carried out using the test by Lee and Mykland (2012), as describe above, and the sequence of the reconstructed daily volatility paths at the frequency of 5 minutes is displayed in Figure 2.6.

At the 95% confidence level, the null hypothesis of Gaussianity was never rejected over the 41 daily samples of reconstructed standardized log-returns. The average p-value was equal to 0.374. Moreover, Table 2.5 reports the averages and standard deviations, over the 41 daily samples analyzed, of the sample mean, variance, skewness and kurtosis of the reconstructed standardized log-returns. The averages are very close to the theoretical moments of a standard Normal, while the standard deviations are relatively small. Overall, the results of the Gaussanity tests, along with the empirical evidence on the sample moments, suggest that the daily empirical distributions of the reconstructed standardized log-returns approximate a standard normal quite well, thereby supporting the accuracy of the Fourier spot volatility estimates obtained also on this S&P500 data set.

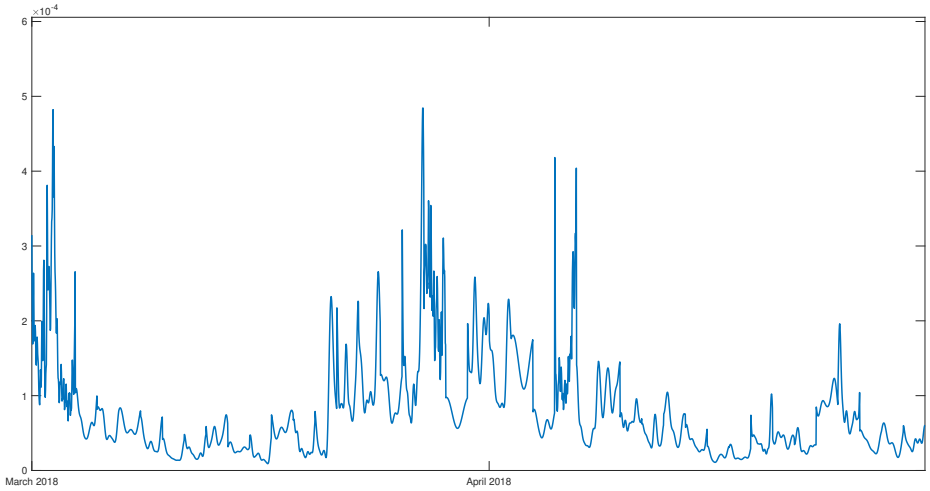

Figure 2.6: S&P500 (March, 2018 - April, 2018): sequence of the reconstructed daily volatility paths on the 5-minute grid.

	average	std. dev.
mean	-0.005	0.132
variance	1.008	0.182
skewness	0.006	0.222
kurtosis	2.940	0.268

Table 2.5: S&P500 (March, 2018 - April, 2018): averages and standard deviations, over the 41 daily samples analyzed, of the sample statistics of the reconstructed standardized log-returns.

Appendix

Appendix 2.A Properties of the Dirichlet and Féjèr kernels

This Appendix resumes some results about the rescaled Dirichlet kernel and the Féjèr kernel, respectively defined as in Equations 2.4 and 2.6. This results are needed for the proofs of Theorems 2.2.1 and 2.2.7.

Lemma 2.A.1. *i) For any M,*

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

Moreover, under the assumption $\lim_{n,M\to\infty} \frac{M^\tau}{n} = a$, for a constant $a > 0$ and for some $\tau > 1$, it holds:

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

and, in particular,

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

ii) For any $M \geq 1$,

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

iii) Under the assumption $\lim_{n,M\to\infty} \frac{M^\tau}{n} = a$, for a constant $a > 0$ and for some $\tau > 1$, it holds:

$$\lim_{M\to\infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(x)dx = \lim_{M,n\to\infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(\varphi_n(x))dx = \frac{4\pi}{3}. \tag{2.94}$$

iv) If $\lim_{n,M\to\infty} \frac{M^\tau}{n} = a$, for a constant $a > 0$ and for some $\tau > 1$, and f is Hölder continuous with parameter $\alpha \in (0, 1]$, then

$$\lim_{n,M\to\infty} \frac{1}{M} \int_{-\pi}^{\pi} F_M^2(x - \varphi_n(y))f(y)dy = \lim_{M\to\infty} \int_{-\pi}^{\pi} \frac{1}{M} F_M^2(x - y)f(y)dy = \frac{4\pi}{3} f(x). \tag{2.95}$$

Proof. See Cuchiero and Teichmann (2015, Lemma 5.1).

Remark 2.A.2. *As observed in Cuchiero and Teichmann (2015, Remark 5.2), all above results similarly hold true on $[0, T]$.*

Lemma 2.A.3. *Suppose that $\lim_{n,M\to\infty} \frac{M^\tau}{n} = a$, for some constant $a > 0$ and $\tau > 1$. Then, it holds that:*

$$\lim_{n,M\to\infty} \int_{-\pi}^{\pi} \frac{1}{M^3} |F'_M(\varphi_n(x))|^2 \, dx = \lim_{M\to\infty} \int_{-\pi}^{\pi} \frac{1}{M^3} |F'_M(x)|^2 \, dx = \frac{2}{15}\pi, \tag{2.96}$$

$$\lim_{n,M\to\infty}\int_{-\pi}^{\pi}\frac{1}{M^5}|F''_M(\varphi_n(x))|^2\,dx=\lim_{M\to\infty}\int_{-\pi}^{\pi}\frac{1}{M^5}|F''_M(x)|^2\,dx=\frac{4}{105}\pi. \tag{2.97}$$

Proof. See Toscano et al. (2022, Lemma 8.1).

Lemma 2.A.4. *i) If* $\lim_{n,N\to\infty}\frac{N}{n}=c>0$*, then for any* $p>1$ *there exists a constant* C_p *such that*

$$\lim_{n,N\to\infty} n\sup_{x\in[0,2\pi]}\int_0^{2\pi}D_N^p(\varphi_n(x)-\varphi_n(y))dy\le C_p.$$

ii) If $\lim_{n,N\to\infty}\frac{N}{n}=c>0$*, it holds that*

$$\lim_{n,N\to\infty} n\int_0^x D_N^2(\varphi_n(x)-\varphi_n(y))dy=\pi(1+2K(2c))$$

and, for any α*-Hölder continuous function* f *with* $\alpha\in(0,1]$*,*

$$\lim_{N,n\to\infty} n\int_0^x D_N^2(\varphi_n(x)-\varphi_n(y))f(y)dy=\pi(1+2K(2c))\,f(x),$$

where

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

being $r(c)=c-[c]$*, with* $[c]$ *denoting the integer part of* c*.*
iii) If $\lim_{n,N\to\infty}\frac{N}{n}=c>0$*, then for any* $s\in[0,2\pi)$ *and for any* $\varepsilon>0$*,*

$$\lim_{n,N\to\infty}\ n\int_0^{x-\varepsilon}D_N^2(\varphi_n(x)-\varphi_n(y))dy=0.$$

Proof. See Clement and Gloter (2011, Lemma 3).

Lemma 2.A.5. *If* $\lim_{N,n\to\infty}\frac{N^\tau}{n}=c$*, for a constant* $c>0$ *and some* $\tau>1$*, then*

$$\lim_{n,N\to\infty} N\int_0^x D_N^2(\varphi_n(x)-\varphi_n(y))dy=\lim_{N\to\infty} N\int_0^x D_N^2(x-y)dy=\frac{\pi}{2}.$$

Proof. See Cuchiero and Teichmann (2015), Lemma 5.1, using the fact that

$$D_N^2(x)=(2N+1)^{-1}F_{2N}(x).$$

Lemma 2.A.6. *Under the condition* $\lim_{N,n\to\infty}\frac{N^\tau}{n}=c$*, for a constant* $c>0$ *and some* $\tau>1$*, it holds:*

$$i)\quad \lim_{n,N\to\infty}\frac{1}{N}\int_0^{2\pi}|D'_N(\varphi_n(x))|^2dx=\lim_{N\to\infty}\frac{1}{N}\int_0^{2\pi}|D'_N(x)|^2dx=\frac{\pi}{3}.$$

$$ii)\quad \lim_{n,N\to\infty}\frac{1}{N^3}\int_0^{2\pi}|D''_N(\varphi_n(x))|^2dx=\lim_{N\to\infty}\frac{1}{N^3}\int_0^{2\pi}|D''_N(x)|^2dx=\frac{\pi}{5}.$$

Proof. For i) see Mancino and Toscano (2022). As for ii), it holds that:

$$\frac{1}{N^3}\int_0^{2\pi}|D''_N(x)|^2\,dx=\frac{1}{N^3}\frac{2\pi}{(2N+1)^2}\sum_{|k|\le N}k^4=\frac{2\pi}{15}\frac{N(N+1)(3N^2+3N-1)}{N^3(2N+1)}\to\frac{\pi}{5}.$$

Chapter 3

From Zero-Intelligence to Queue-Reactive: Limit-Order-Book modeling for high-frequency volatility estimation and optimal execution

As shown in the previous chapter high-frequency data provide, in principle, the possibility of obtaining very precise estimates of the volatility; moreover availability of efficient estimates of the volatility of financial assets may be crucial for a number of applications, such as model calibration, risk management, derivatives pricing, high-frequency trading and optimal execution. The (infill) asymptotic theory of volatility estimators is usually derived, as we have seen, under the assumption that the (efficient) asset price follows an Itô semimartingale (see also Chapter 3 of Aït-Sahalia and Jacod (2014)). The Itô semimartingale hypothesis ensures the absence of arbitrage opportunities (see Delbaen and Schachermayer (1994)) and, at the same time, is rather flexible, as it does not require to specify any parametric form for the dynamics of the asset price. To account for empirical evidences that indicate that the prices of financial assets do not conform to the semimartingale hypothesis at high frequencies, due to the presence of microstructure phenomena such as, e.g., bid-ask bounces or price rounding, we introduced in Section 2.2.3 an additive noise component, following the most established approach present in the recent literature (see Chapter 7 of Aït-Sahalia and Jacod (2014)). In its most basic form, the noise due to microstructure can be assumed to be i.i.d. and independent of the semimartingale driving the price dynamics (i.e., the so-called efficient price, see Section 2.2), however, as introduced in Section 2.3, more sophisticated forms have been studied, such as, for instance, an additive noise which is auto-correlated or correlated with the efficient price (see, e.g., Hansen and Lunde (2006)).

As mentioned, the literature on the estimation of the volatility in the presence of noise is very rich. In fact, there exists a number of alternative methodologies making an efficient use of high-frequency prices to reconstruct both the integrated and the spot volatility. In addition to the Fourier method these include for example, among others, the two-scale and multi-scale approach by, respectively, Zhang et al. (2005) and Zhang (2006), the kernel-based method, originally proposed in Barndorff-Nielsen et al. (2008), and the pre-averaging approach by Jacod et al. (2009). Given this variety of alternative methodologies, it is not straightforward to establish which specific noise-robust estimator should be preferred for high-frequency financial applications, being the finite sample performances not necessarily in line with the asymptotic properties of each estimator.

Gatheral and Oomen (2010) proposed to compare the finite-sample performance of different high-frequency estimators via simulations based on a market simulator which is able to reproduce the actual mechanism of price formation at high frequencies with sufficient realism. In this regard, the authors used simulations obtained via the zero-intelligence (ZI) limit order book model by Smith et al. (2003) to compare the performance of different integrated volatility estimators. However, the ZI model is based on several simplistic assumptions on the dynamics of the limit order book, and thus it may fail to replicate the actual behavior of high frequency financial data and microstructure noise with satisfactory accuracy. For example, under the ZI model, the order flow is described by independent Poisson processes, while it is well-known that the order flow is a long-memory process (see Lillo and Farmer (2004)) and the different components of the order flow are lead-lag cross-correlated (Eisler et al. (2012)). Moreover, as pointed out by Bouchaud et al. (2018), the ZI model leads to systematically profitable market making strategies. These properties are likely to have an effect on the dynamics of the volatility and market microstructure noise, and thus an analysis based on a more realistic limit order book model is needed.

In this chapter, instead, we use a more realistic limit order book model, namely the queue-reactive (QR) model by Huang et al. (2015). Under this model, the arrival rates of orders depend on the state of the limit order book. This implicitly introduces auto- and cross-correlations of the book components, thereby generating more realistic dynamics for the price process at high-frequencies. Secondly, we compare not only the performance of a number of estimators of the integrated volatility (expanding the collection of estimators considered in the study by Gatheral and Oomen (2010)), but also that of different estimators of the spot volatility.

For what concerns the integrated variance, we find that the pre-averaging estimator by Jacod et al. (2009) is favorable in terms of bias minimization. Instead, when looking at the optimization of the mean squared error, the situation appears to be more nuanced. Indeed, the Fourier estimator by Malliavin and Mancino (2009) obtains the best average ranking across the considered price series (mid-, micro- and trade- prices) without actually achieving the best ranking for any of these individual series. The best rankings are instead achieved by the unified volatility estimator by Li et al. (2018) (for mid- and micro-prices) and by the alternation estimator by Large (2011) (for trade-prices). Instead, for what concerns the spot variance, the Fourier estimator provides the relative best performance for the three prices series, both in terms of bias and mean-squared-error optimization.

Finally, we investigate, via simulations of the QR model, how the use of different volatility estimators affects the inference of the variance of the cost of the execution strategy. To do so, we consider the instance where the trader is set to execute a volume-weighted average price (VWAP) strategy and assumes that market impact is described by the Almgren and Chriss (Almgren and Chriss (2001)) model. We compare the empirical variance of the implementation shortfall of the simulated executions with the corresponding model-based prediction, evaluated with different spot volatility estimators. As a result, we find that the estimator that yields the optimal performance in terms of bias and mean-squared-error optimization, namely the Fourier estimator, also gives the optimal forecast of the cost variance. More generally, our results suggest that the choice of the spot estimator is not irrelevant, as it may lead to significantly different forecasts of the variance of the implementation shortfall.

The chapter is organized as follows. In Section 3.1 we recall the main characteristics of the ZI and QR limit-order-book models, discuss their calibration on empirical data and compare their ability

to reproduce realistic volatility and noise features. In Section 3.2 we illustrate the estimators of the integrated and spot variance, while in Section 3.3 we evaluate their finite-sample performance with simulated data from the QR model. Finally, Section 3.4 contains the study of the impact of efficient volatility estimates on optimal execution.

3.1 Limit-order-book models: zero-intelligence vs queue-reactive

Electronic financial markets are often based on a double auction mechanism, with a bid (buy) side and an ask (sell) side. The limit order book (LOB) is the collection of all the outstanding limit orders, which are orders of buying or selling a given quantity of the asset at a given price, expressed as a multiple of the tick size (i.e., the minimum price movement allowed) of the asset. Other two types of orders can be placed: a cancellation, that erases a limit order previously inserted by the same agent, thereby reducing the volume at a given price level, and a market order, that is, an order to immediately buy/sell the asset at the best possible price. The best bid is the highest price at which there is a limit order to buy, and the best ask is the lowest price at which there is a limit order to sell. The spread is the difference between the best ask and the best bid, and is typically expressed in tick size. For a detailed overview of the LOB see Abergel et al. (2016).

In the following, we will be interested in three price series that can be retrieved from LOB data: the mid-price, the micro-price and the trade price.

Definition 3.1.1. *We define the mid-price p_{mid} and the micro-price p_{micro} of an asset at time t as, respectively, the arithmetic average and the volume-weighted average of the best bid and best ask quotes at time t, i.e.,*

$$p_{mid}(t) := \frac{p^b(t) + p^a(t)}{2}, \qquad p_{micro}(t) := \frac{p^b(t)v^a(t) + p^a(t)v^b(t)}{v^b(t) + v^a(t)},$$

where p^b, p^a, v^b and v^a denote, respectively, the best bid, the best ask, the volume (i.e., the number of outstanding limit orders) at the best bid and the volume at the best ask. Finally, the trade price p_{trade} series is defined as the series of prices arising from the execution of market orders.

An example of the different behavior of the three series considered is shown in Figure 3.1, where the trajectories are obtained simulated the evolution of the LOB with the QR model by Huang et al. (2015).

In our study, we will consider two models for the simulation of the LOB. The simplistic ZI model by Smith et al. (2003) and the more sophisticated QR model by Huang et al. (2015). In the next paragraphs we briefly recall the main characteristics of the two models. Please refer to the original papers for a more thorough description.

3.1.1 Model descriptions

The zero-intelligence model

The ZI model, originally proposed by Smith et al. (2003), is a statistical representation of the double action mechanism used in most stock markets. Despite its simplicity, the model is able to generate a relatively complex dynamic for the order book. It is based on three parameters: the intensity of

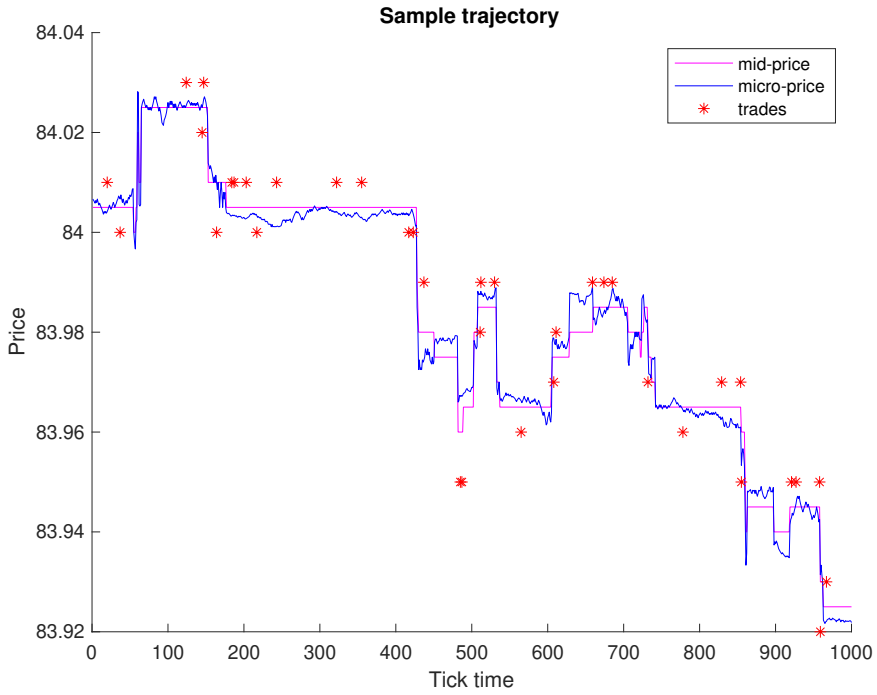

Figure 3.1: Sample trajectory for a simulated LOB: the mid-, micro- and trade-price. Simulations obtained using the QR model.

limit orders, λ^L, the intensity of cancel orders, λ^C, and the intensity of market orders, λ^M. The three components of the order flow follow independent Poisson processes, thus the type of order extracted at each time is independent of the previous orders and the current state of the LOB, and orders may arrive at every price level with the same probability. Each order is assumed to have unitary size. For a detailed discussion about the flexibility of this model, see Gatheral and Oomen (2010).

As mentioned, the ZI model may be deemed as too simplistic. Indeed, the assumptions that the intensities of order arrival are independent of the state of the book and that the intensities are equal for each price level are highly unrealistic. Moreover, this model produces purely endogenous order-book dynamics, without considering the effect of exogenous information. Further, as shown in Bouchaud et al. (2018), under the ZI model the market impact of new orders is such that profitable market-making opportunities can be created, even if they are usually absent in real markets. Some of the weakness of the ZI model are overcome by the QR model.

The queue-reactive model

The QR model (Huang et al. (2015)) is a LOB model suitable to describe large tick assets, i.e., assets whose bid-ask spread is almost always equal to one tick. This model is able to reproduce a richer and more realistic behavior of the LOB, compared to the ZI model. In other words, the QR model attempts to fix some of the flaws of the ZI model. This is achieved, in the first place, by assuming different intensities for each level of the LOB. Moreover, the degree of realism is increased by introducing a correlation not only between order-arrival intensities and the corresponding queue

size at each level, but also between intensities and the queue size at the corresponding level at the opposite side of the book. Further, a dependence between the volume at the best level and order arrivals at the other levels is assumed. Finally, differently form the ZI model, the QR model allows for exogenous dynamics by taking into account the flow of exogenous information that hits the market.

Following Huang et al. (2015), under the QR model the LOB is described by a $2K-$dimensional vector, with K denoting the number of available price levels at the bid and ask sides of the book. At the level $Q_{\pm i}$, $i = 1, ..., K$, the corresponding price is equal to $p \pm i(tick)$, where p denotes the center of the $2K$ dimensional vector. Precisely, Q_{-i} denotes a level order at the bid side and Q_i denotes a level order at the ask side. Moreover, $q_{\pm i}$ denotes the volume at the level $Q_{\pm i}$.

The process $X(t) = (q_{-K}(t), ..., q_{-1}(t), q_1(t), ..., q_K(t))$ is a continuous-time Markov process with the following infinitesimal generator matrix $\mathcal{Q}$:

$$\begin{aligned}
\mathcal{Q}_{q,q+e_i} &= f_i(q), \\
\mathcal{Q}_{q,q-e_i} &= g_i(q), \\
\mathcal{Q}_{q,q} &= -\sum_{p \neq q} \mathcal{Q}_{q,p}, \\
\mathcal{Q}_{q,q} &= 0 \quad \text{otherwise},
\end{aligned}$$

with $q = (q_{-K}, ..., q_{-1}, q_1, ..., q_K)$ and e_i denoting the $i - th$ vector of the standard base of $\mathbb{R}^{2K}$.

Thus, with different intensities at each queue, we have that

$$\begin{aligned}
f_i(q) &= \lambda^L(q_i, S_{m,l}(q_{-i})), \\
g_i(q) &= \lambda^C(q_i, S_{m,l}(q_{-i})) + \lambda^M_{\text{buy}} \mathbb{1}_{\text{bestask}(q)=i}, \quad \text{if} \quad i > 0, \\
g_i(q) &= \lambda^C(q_i, S_{m,l}(q_{-i})) + \lambda^M_{\text{sell}} \mathbb{1}_{\text{bestbid}(q)=i}, \quad \text{if} \quad i < 0,
\end{aligned}$$

where the function $S_{m,l}(q_{-i})$ is responsible for the interaction between the bid and the ask side of the book, that is,

$$\begin{aligned}
S_{m,l}(q_{-i}) &= Q^0 \quad \text{if} \quad q_{-i} = 0, \\
S_{m,l}(q_{-i}) &= Q^- \quad \text{if} \quad 0 < q_{-i} \leq m, \\
S_{m,l}(q_{-i}) &= \bar{Q} \quad \text{if} \quad m < q_{-i} \leq l, \\
S_{m,l}(q_{-i}) &= Q^+ \quad \text{if} \quad q_{-i} > l,
\end{aligned}$$

being m and l two fixed thresholds. For example, given a certain volume at the bid side, a new bid limit order has different intensities depending on whether the volume at the ask is, e.g., $q_i = 0$, $0 < q_i \leq 5$, $5 < q_i \leq 10$ or $q_i > 10$. Market orders may arrive only at the best quote. Moreover, we assume that λ^L_i and λ^C_i are also functions of $\mathbf{1}_{q_{\pm 1} \geq 0}$, for $i \neq \pm 1$, to allow for interactions between the best level and the dynamics far from the best level.

Conditionally on the LOB state, the arrival of different orders at a given limit is assumed to be independent, and follows a Poisson distribution, with intensity equal to λ. However, since the queue sizes depend on the order flow, the model reproduces some auto- and cross-correlations between the components of the order flow, as observed in empirical data.

Contrary to the ZI model, the QR gives a dynamic that is not entirely driven by the orders arriving on the LOB. To achieve this, two additional parameters are introduced, θ^{ref} and θ^{reinit}. Whenever the mid-price changes, an auxiliary (not observed) price, called the reference price and denoted by p_{ref}, changes in the same direction with probability θ^{ref} and by an amount equal to the tick size of the asset. Moreover, when p_{ref} changes, the LOB state is redrawn from its invariant distribution around the new value of p_{ref}, with probability θ^{reinit}. In fact, as the authors explain, the parameter θ^{reinit} captures the percentage of price changes due to exogenous information.

In the following we address the procedures that we followed to estimate the two LOB models and discuss the differences between them in terms of ability of reproducing realistic features of empirical high-frequency data.

3.1.2 Calibration procedures

For the calibration of the ZI and QR models, we used order-book data of the stock Microsoft (MSFT) over the period April 1, 2018 - April 30, 2018. Data were retrieved from the LOBSTER database. Microsoft is a very liquid stock with an average spread approximately equal to 1.25 ticks and thus can be considered a large tick asset, suitable to be modeled by the queue-reactive design.

Before the calibration, for each day we removed from the sample the first hour of trading activity after the market opening and the last 30 minutes before the market closure; this is a standard procedure adopted when working with high-frequency data, since during these two moments of the day the trading activity is known to be more intense and volatile, thereby possibly leading to a violation of the large tick asset hypothesis, even for a liquid stock like Microsoft. Given the average spread observed, and being the activity almost fully concentrated at the best limits, we implemented the ZI and QR models using two limits, $Q_{\pm 1}$ and $Q_{\pm 2}$. This is in line with Huang et al. (2015).

To estimate the intensities of order arrivals under the QR model, the following inputs are needed:

- the type of each event, i.e., limit order, cancel order or market order;
- the time between events that happen at Q_1 and Q_2, along with the queue sizes q_1 and q_2 before each event;
- the size of each event, as q_i is expressed as a multiple of the median event size.

The estimation of the intensities is performed via maximum likelihood, as in Huang et al. (2015). The parameters m and l that capture the bid-ask dependence are set equal to the the 33% lower and upper quantiles of the q_{-i}'s (conditional on positive values). Given the symmetry property of the LOB, intensities are computed for just one side.

The parameters θ^{ref} and θ^{reinit} are calibrated using the mean-reversion ratio $\bar{\zeta}$ of the mid-price, which is defined as

$$\bar{\zeta} := \frac{n_c}{2n_a},$$

where n_c is the number of continuations (i.e., the number of consecutive price moves in the same directions) and n_a is the number of alternations (i.e., the number of consecutive price moves in opposite directions). For more details about the relation between the mean-reversion ratio and the microstructure of large tick assets, see Robert and Rosenbaum (2011).

We carried out the calibration using a two-step generalized method of moments (GMM), which is more robust than the heuristic approach proposed by Huang et al. (2015). Denote by σ^{emp} and

$\bar{\zeta}^{emp}$ the empirical estimates of the standard deviation and mean-reversion ratio of the mid-price returns, computed at the 1-second frequency using the last tick rule. Further, denote by $\sigma_t(\bar{\theta})$ and $\bar{\zeta}_t(\bar{\theta})$ the quantities estimated in simulation t, with $t = 1, ..., T$, and $\bar{\theta} = (\theta^{\text{ref}}, \theta^{\text{reinit}})$.

Two-step GMM-based procedure for the calibration of $\bar{\theta}$.

Step 1. Obtain a consistent estimate of $\bar{\theta}$ via the estimator

$$\bar{\theta}_1 := \underset{\bar{\theta}\in[0,1]\times[0,1]}{\arg\min} \frac{1}{T^2}\left[\left(\sum_t g_t^{\sigma}(\bar{\theta})\right)^2 + \left(\sum_t g_t^{\bar{\zeta}}(\bar{\theta})\right)^2\right],$$

where $g_t^{\sigma}(\bar{\theta}) = \dfrac{\sigma_t(\bar{\theta})}{\sigma^{emp}} - 1$, $g_t^{\bar{\zeta}}(\bar{\theta}) = \dfrac{\bar{\zeta}_t(\bar{\theta})}{\bar{\zeta}^{emp}} - 1$.

Step 2. Obtain the GMM-estimate of $\bar{\theta}$

$$\bar{\theta}_{GMM} := \underset{\bar{\theta}\in[0,1]\times[0,1]}{\arg\min} \left(\frac{1}{T}\sum_t \bar{g}_t(\bar{\theta})\right)' W \left(\frac{1}{T}\sum_t \bar{g}_t(\bar{\theta})\right),$$

where $\bar{g}_t(\bar{\theta}) = (g_t^{\sigma}(\bar{\theta}), g_t^{\bar{\zeta}}(\bar{\theta}))$ and $W^{-1} = \dfrac{1}{T}\sum_t (\bar{g}_t(\bar{\theta}_1))'(\bar{g}_t(\bar{\theta}_1))$.

The estimator is asymptotically efficient in the GMM class. For the implementation we used $T = 100$ simulations with an horizon of one trading day. The tick size was set equal to the minimum bid-ask spread recorded in the data, namely 1 cent. As a final estimate, we obtained $\theta^{\text{ref}} = 0.6$ and $\theta^{\text{reinit}} = 0.85$.

The asymptotic distribution of the queue size, needed for the re-initialization of the LOB state, was obtained following the approach proposed by Huang et al. (2015). For each simulated path, the starting LOB state was randomly chosen using the asymptotic distribution of the queue size.

For the calibration of the ZI model, one only needs to reconstruct the intensities of order arrivals. To do that, the only information needed involves the type of order, the order arrival time and the order size. The ensuing estimators read

$$\lambda^L = \frac{\#L}{\#O\overline{\Delta}_t}, \quad \lambda^C = \frac{\#C}{\#O\overline{\Delta}_t}, \quad \lambda^M = \frac{\#M}{\#O\overline{\Delta}_t}$$

where $\#O = \#L + \#C + \#M$ and $\#L$, $\#C$, $\#M$ denote, respectively, the total number of limit, cancel and market orders arriving at the best quotes or between the spread, while $\overline{\Delta}_t$ is the average elapsed time between two consecutive orders.

3.1.3 Comparison of volatility and noise features

As pointed out by Gatheral and Oomen (2010), neither the efficient price nor its volatility are well-defined under the ZI model. The same holds under the QR model. However, if one assumes constant model parameters, thanks to the ergodicity of the processes (see Huang et al. (2015)), the variance

of the efficient price can be defined, following Gatheral and Oomen (2010), as

$$\sigma^2 := \lim_{m\to\infty} \frac{1}{m} E\left[(p(m) - p(0))^2\right], \tag{3.1}$$

where p may denote any price process among those considered, that is, the mid-price, the micro-price and the trade price.

Based on (3.1), given the (calibrated) values of the LOB parameters, it is possible to estimate the true value of σ^2 via simulations. Specifically, in our study, given the LOB parameters calibrated on the Microsoft sample data, numerical results over 2500 simulations show that for $m > 18000$ the volatility of mid-prices, micro-prices and trade-prices stabilizes around the value $\sigma^2 = 1.039 \cdot 10^{-8}$ in the case of the QR model.

Since the knowledge of the value of σ^2 is crucial to compare the finite-sample performance of volatility estimators, we wish to verify whether the two order-book models give similar results. The estimates of σ^2 for the two models and for the three price series considered are compared in Table 3.1. The levels of variance obtained with the ZI model and the QR are significantly different, even when considering the error in the estimation procedure, with the ZI producing a level about 50% higher than the one observed for the QR model. This highlights a first fact to be considered: the choice of the LOB model is not irrelevant for applications involving the volatility parameter. In fact, it is known that the ZI model calibrated on real data only partially reproduces the actual empirical variance, with a bias which depends on the relative magnitude of the intensities (see, e.g., Bouchaud et al. (2018)).

Order-book model	statistics	mid-price	micro-price	trade price
ZI model	Variance	$1.539 \cdot 10^{-8}$	$1.539 \cdot 10^{-8}$	$1.539 \cdot 10^{-8}$
	Std. error	$1.259 \cdot 10^{-9}$	$2.434 \cdot 10^{-9}$	$2.587 \cdot 10^{-9}$
QR model	Variance	$1.039 \cdot 10^{-8}$	$1.039 \cdot 10^{-8}$	$1.039 \cdot 10^{-8}$
	Std. error	$3.336 \cdot 10^{-10}$	$3.330 \cdot 10^{-10}$	$3.319 \cdot 10^{-10}$

Table 3.1: Comparison of the estimated variance reconstructed via Eq. (3.1) using simulations of the ZI and the QR models.

Moreover, since we are interested in comparing the performance of noise-robust volatility estimators, we wish to discriminate between the two LOB models on the basis of how well they mimic the noise accumulation observable in empirical data at different frequencies. To do so, we use the Hausman test for the null hypothesis of the absence of noise by Aït-Sahalia and Xiu (2019a). In particular, we use the formulation of the test in Equation (16) of Aït-Sahalia and Xiu (2019a), which is coherent with the use of LOB models with a constant variance parameter.

Tables 3.2 illustrates the frequencies (in seconds) at which the Hausman test rejects the null hypothesis of the absence of noise with a significance level of 5% (★) and the frequencies at which the null is instead not rejected (†) for, respectively, the MSFT sample and the simulated samples from the ZI and QR models. The results of Hausman test suggest that the noise accumulation mechanism at different frequencies under the QR model is more realistic than the one observed under the ZI model, based on the comparison with the noise-detection pattern in the MSFT sample. This aspect is clearly relevant when analyzing the finite-sample performance of noise-robust volatility

estimators, and adds empirical support to the use of the QR model for that purpose.

	1 sec.	2 sec.	5 sec.	10 sec.	15 sec.	30 sec.	60 sec.
Microsoft (April 2018)							
Mid-price	★	★	★	★	★	†	†
Micro-price	★	★	★	★	★	†	†
Trade-price	★	★	★	★	★	★	†
ZI model							
Mid-price	★	★	★	★	†	†	†
Micro-price	★	★	★	†	†	†	†
Trade-price	★	★	★	†	†	†	†
QR model							
Mid-price	★	★	★	★	★	†	†
Micro-price	★	★	★	★	★	†	†
Trade-price	★	★	★	★	★	†	†

Table 3.2: Hausman test results at different frequencies (in seconds) for Microsoft (April 2018), the ZI model and the QR model. The symbol ★ (†) indicates that the null of absence of noise is rejected (not rejected) with a significance level of 5%.

Lastly, we look at the average spread that the two order-book models are able to generate. This aspect is of paramount importance, being the spread a crucial characteristic of LOB and one of the main sources of market microstructure noise. Table 3.3 suggests that, even though both the ZI and the QR models generate an average spread lower than the one empirically observed, the underestimation is less severe in the case of the QR. The underestimation of spread under the ZI model has been documented in Bouchaud et al. (2018).

Data	Average spread
MSFT sample	1.25
ZI simulations	1.03
QR simulations	1.14

Table 3.3: Comparison of the average spread (expressed in ticks) computed from the empirical MSFT sample with the corresponding values obtained from simulations of the ZI and QR samples.

3.2 Volatility estimators

In this section, we briefly describe the noise-robust integrated- and spot-volatility estimators whose performance will be studied in the next section. The formulae for the tuning parameters involved in the computation of the estimators that optimize the mean squared error (MSE) are also reported.

Preliminary notation

We consider the estimation horizon $[t, t+h]$, $t, h > 0$, and assume that the price p is sampled on the equally-spaced grid with mesh h/n, where n denotes the number of price observations. The quantity

p_i denotes the log-price of the asset at time $t_i := t + ih/n$, $i = 0, 1, ..., n$. Note that p may refer indifferently to the trade-, mid- or micro-price.

The spot volatility at time t is denoted by $\sigma^2(t)$ and

$$IV_{t,u} := \int_t^{t+u} \sigma^2(s)ds$$

denotes the integrated volatility on $[t, t+u]$, $u \leq h$. Clearly, in a setting with constant volatility $\sigma^2(t) = \sigma^2$ for all t, the latter simplifies to $\sigma^2 u$.

As an auxiliary quantity for the implementation of some estimators, we will need to estimate the integrated quarticity, that is,

$$IQ_{t,u} := \int_t^{t+u} \sigma^4(s)ds.$$

In the rest of the chapter, we drop the subscript of both IV and IQ as we always refer to the interval $[t, t+h]$.

Furthermore, we recall that the asymptotic properties of high-frequency volatility estimators are typically derived under a setting analogous to the one presented in Section 2.2, that we briefly recall. For all t, the observable price $\tilde{p}$ is decomposed as

$$\tilde{p}(t) = p(t) + \eta(t),$$

where p denotes the efficient price, whose dynamics follow an Itô semimartingale, while η is an i.i.d. zero-mean noise due to the market microstructure. As additional auxiliary quantities, we need estimates of the second moment of η, i.e.,

$$\xi = E[\eta^2].$$

Finally, we denote the floor function as $\lfloor \cdot \rfloor$ and the rounding to the nearest integer as $[\cdot]$.

3.2.1 Integrated volatility

Bias-corrected realized variance

The realized variance, that is, the sum of squared log-returns over a given time horizon, represents the most natural rate-efficient estimator of the integrated volatility in the absence of noise. However, in the presence of noise, as it is typically the case for high-frequency settings, the realized variance is biased. The bias-corrected realized variance by Zhou (1996) corrects for the bias due to noise by taking into account the first order auto-covariance of the log-returns. The estimator reads:

$$IV_{BC} := c \sum_{j=0}^{q-1} \left(\sum_{i=1}^{c_2} (\tilde{p}_{iq+j} - \tilde{p}_{(i-1)q+j} \right)^2 + 2 \sum_{i=1}^{c_2-1} (\tilde{p}_{iq+j} - \tilde{p}_{(i-1)q+j})(\tilde{p}_{(i+1)q+j} - \tilde{p}_{iq+j}),$$

where $c = \frac{n}{n-q+1} \frac{1}{q}$ and $c_2 = \lfloor (n-j+1)/q \rfloor$.

The MSE-optimal value of q is attained as

$$q^* = \max \left(1, \left[\frac{2n\xi}{IV\sqrt{3}} \right] \right).$$

Fourier estimator

Introduced by Malliavin and Mancino (2002) and Malliavin and Mancino (2009), the Fourier estimator of the integrated volatility relies on the computation of the zero-th Fourier coefficient of the volatility, given the Fourier coefficients of the log-returns. The noise is filtered out by suitably selecting the cutting frequency N. If one uses the Fejér (respectively, Dirichlet) kernel to weight the convolution product, the estimator is defined as

$$IV_F^{Fej} := \frac{(2\pi)^2}{N+1} \sum_{|k|\leq N} \left(1 - \frac{|k|}{N+1}\right) c_k(d\tilde{p}_n) c_{-k}(d\tilde{p}_n),$$

and

$$IV_F^{Dir} := \frac{(2\pi)^2}{2N+1} \sum_{|k|\leq N} c_k(d\tilde{p}_n) c_{-k}(d\tilde{p}_n),$$

where $c_k(d\tilde{p}_n)$, i.e. the k-th discrete Fourier coefficient of the log-return, is defined as in Equation 2.1.

The optimal value of the integer N in the presence of noise can be selected by performing a feasible minimization of the MSE, see Mancino and Sanfelici (2008).

In this analysis, we implemented IV_F^{Fej}, as unreported simulations suggest that it performs better than IV_F^{Dir}.

Maximum likelihood estimator

The maximum-likelihood estimator by Aït-Sahalia et al. (2005) is based on the assumption that noisy log-returns follow an MA(1) model, consistently with the seminal microstructure model for the bid-ask spread by Roll (1984). under the MA(1) assumption, it holds that

$$\delta_i(p) = w_i + \phi w_{i-1}, \tag{3.2}$$

and the maximum-likelihood estimator reads

$$IV_{ML} := n\hat{\sigma}_w^2 (1+\hat{\phi})^2,$$

where the pair $(\hat{\phi}, \hat{\sigma}_w^2)$ is the result of the standard maximum-likelihood estimation of the MA(1) model.

Two-scale estimator

The two-scale realized variance by Zhang et al. (2005) eliminates the noise-induced bias of the realized variance by combining two different realized variance values, one computed at a higher frequency and one computed at a lower frequency. The estimator reads:

$$IV_{TS} := c \left(\frac{1}{q} \sum_{j=0}^{q-1} \sum_{i=1}^{\lfloor (n-j+1)/q \rfloor} (\tilde{p}_{iq+j} - \tilde{p}_{(i-1)q+j})^2 - \frac{n-q+1}{nq} \sum_{i=0}^{n-1} \delta_i(\tilde{p})^2 \right),$$

where $c = \left(1 - \frac{n-q+1}{nq}\right)^{-1}$. The MSE-optimal value of q is equal to

$$q^* = n^{2/3} \left(\frac{12\xi^2}{IQ} \right)^{1/3}.$$

Multi-scale estimator

Zhang (2006) also proposed a more sophisticated combination of realized variances at various frequencies that smooths out the effect of microstructure noise. The multi-scale estimator reads:

$$IV_{MS} := \sum_{j=1}^{q} \frac{a_j}{j} \sum_{k=0}^{j-1} \sum_{i=1}^{\lfloor (n-k+1)/j \rfloor} (\tilde{p}_{ij+k} - \tilde{p}_{(i-1)j+k})^2,$$

where

$$a_j = j(1 - 1/q^2)^{-1} \left(\frac{12(1/q - 1/2)}{q^2} - \frac{6}{q^3} \right).$$

The MSE-optimal q is given by

$$q^* = \sqrt{n} \left(\beta + \sqrt{\beta^2 + \frac{144\xi^2}{104/35 IQ}} \right)^{1/2},$$

where $\beta = \dfrac{48/5\xi(IV + \xi^2/2)}{208/35 IQ}$.

Kernel estimator

Kernel-based estimators, originally introduced by Barndorff-Nielsen et al. (2008), correct for the bias due to noise of the realized variance by taking into account the autocorrelation of returns at different lags, suitably weighted by means of a kernel function $k(\cdot)$. The estimator reads:

$$IV_K := \sum_{i=0}^{n-1} \delta_i(p)^2 + 2 \sum_{j=1}^{q} k\left(\frac{j-1}{q} \right) \sum_{i=0}^{n-1-j} \delta_i(\tilde{p}) \delta_{i+j}(\tilde{p}).$$

Moreover, the MSE-optimal value of q is equal to

$$q^* = \sqrt{n} \left(\beta + \sqrt{\beta^2 + \frac{c\xi^2}{aIQ}} \right)^{1/2},$$

where $\beta = \dfrac{b\xi(IV + \xi/2)}{2aIQ}$, $a = 4\displaystyle\int_0^1 k(x)^2 dx$, $b = -8\displaystyle\int_0^1 k(x) k''(x) dx$, $c = 12(k'''(0) + \displaystyle\int_0^1 k(x) k'''(x) dx$.

In this exercise we implemented this estimator by using the Tukey Hanning 2 kernel, i.e., we set $k(x) = \sin^2(\pi/2(1-x)^2)$. This kernel was shown to perform satisfactorily, compared to other kernels (see Barndorff-Nielsen et al. (2008)).

Pre-averaging estimator

The pre-averaging estimator, proposed by Jacod et al. (2009), relies on the averaging of the price values over a window h to compute the realized variance, together with a bias-correction term. The estimator is as follows:

$$IV_{PA} := \frac{6}{h} \sum_{s=0}^{n-h+1} \left(\frac{1}{h/2} \sum_{s'=0}^{h/2-1} \tilde{p}_{s+s'+h/2} - \frac{1}{h/2} \sum_{s'=0}^{h/2-1} \tilde{p}_{s+s'} \right)^2 - \frac{6}{h^2} \sum_{i=0}^{n-1} \delta_i(\tilde{p})^2.$$

Jacod et al. (2009) suggest that the estimator is robust to the choice of h; in our simulation study we set h to obtain a window of approximately 4 minutes, in line with Li et al. (2021).

Alternation estimator

Proposed by Large (2011), the alternation estimator corrects the realized variance with a factor dependent on the number of alternations and continuations in the sample, that is, the estimator reads

$$IV_{Alt} := \frac{n_c}{n_a} \sum_{i=0}^{n-1} \delta_i(\tilde{p})^2,$$

where n_c is the number of consecutive price movements in the same direction, while n_a the number of consecutive price movement in the opposite direction.

MinRV and MedRV estimators

Andersen et al. (2012) introduced two jump-robust estimators of the integrated variance which consist, respectively, in the (scaled) sum of the minimum or the median between consecutive returns, that is,

$$IV_{Min} = \frac{\pi}{\pi - 2} \left(\frac{n}{n-1} \right) \sum_{i=0}^{n-2} \min(|\delta_i(\tilde{p})|, |\delta_{i+1}(\tilde{p})|)^2;$$

$$IV_{Med} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{n}{n-2} \right) \sum_{i=1}^{n-2} \text{med}(|\delta_{i-1}(\tilde{p})|, |\delta_i(\tilde{p})|, |\delta_{i+1}(\tilde{p})|)^2.$$

To make the estimators robust to the presence of microstructure noise, pre-averaging may be applied to price observations, as shown in Andersen et al. (2012), Appendix B. Accordingly, in this comparison we used the noise-robust version of IV_{Min} and IV_{Med} with price pre-averaging.

Range estimator

The main idea behind the range estimator by Vortelinos (2014) is to substitute the simple returns with the difference between the maximum and minimum observed price over a given window, to obtain the estimator

$$IV_{RG} := \frac{1}{4\log(2)} \sum_{i=1}^{n/q} (max_{qi} - min_{qi})^2,$$

where $max_{qi} = \max(\tilde{p}_{(i-1)q}, ..., \tilde{p}_{iq})$ and $min_{qi} = \min(\tilde{p}_{(i-1)q}, ..., \tilde{p}_{iq})$. The MSE-optimal frequency is

$$q^* = \left(\frac{IQ}{\xi^2} \right)^{1/3}.$$

Unified estimator

Li et al. (2018) proposed a unified approach to volatility estimation, obtaining an estimator which is consistent not only in the presence of the typical i.i.d. noise, but also when the noise comes from price rounding. The estimators is defined as follows:

$$IV_U := \sum_{l=1}^{m} \left(\frac{1}{m} - \bar{n} \frac{n_l - \bar{n}}{\sum_{j=1}^{L} n_j^2 - h\bar{n}^2} \right) \frac{1}{q_l} \sum_{k=0}^{q_l - 1} \sum_{i=1}^{n_l} (\tilde{p}_{(k+i)q_l} - \tilde{p}_{k+(i-1)q_l})^2,$$

where $\bar{n} = (\sum_{l=1}^{m} n_l)/h$, $n_l = n/q_l$ and $q_{l+1} = q_1 + l$, $l = 1, ..., m-1$. The optimal q_1 and h can be selected via the data-driven procedure detailed in Li et al. (2018).

3.2.2 Spot volatility

Fourier estimator

As outined in Chapter 2 the Fourier method allows reconstructing the trajectory of the volatility as a function of time on a discrete grid, and a consistent and asymptotically unbiased estimator in presence of noise is the one defined in Section 2.1, provided that the cutting frequencies N and M are properly chosen.

Note that, differently from the other estimators detailed below, the Fourier estimator is *global*, in the sense that it estimates the entire volatility function on a discrete grid over the interval $[0, 2\pi]$[1], instead of a local value at a specific time t.

The selection of N and M for this exercise is performed based on the numerical results given Mancino and Recchioni (2015).

Regularized estimator

Proposed by Ogawa (2008), the regularized estimator is based on a regularization procedure that involves data around the estimation point. The estimator reads

$$\sigma^2_{REG}(t) := \frac{3s^2}{q\sqrt{[n/q]}(3sq - q^2 + 1)} \sum_{i=0}^{[n/q]-1} \delta_{t+i}(\bar{p})^2,$$

where here $\bar{p}$ is given by

$$\bar{p}_t = \frac{1}{s} \sum_{j=1}^{s} \tilde{p}_{tq-j+1},$$

Following Ogawa and Sanfelici (2011), for the implementation we set $q = [n/n_t]$ and $s = 2q$, where n_t is the number of points on which the spot variance trajectory is reconstructed.

Kernel estimator

Fan and Wang (2008) (see also Kristensen (2010)) proposed an estimator of the spot variance based on the localization of the kernel-weighted realized variance over a window of length q, which reads

[1] By suitably re-scaling the unit of time, the estimator can be applied to any arbitrary interval.

$$\sigma_K^2(t) := \frac{1}{q} \sum_{t_i=t-q}^{t+q} k\left(\frac{t_i - t}{q}\right) \delta_{t_i}(\tilde{p})^2,$$

with

$$q = [\sqrt{n}/\log n].$$

For the implementation, we used the Fejér kernel, following Mancini et al. (2015).

Pre-averaging estimator

The pre-averaging estimator of the spot volatility by Jing et al. (2014) and Li et al. (2021) relies on the localization of the pre-averaging integrated estimator on a window of length h and reads

$$\sigma_{PA}^2(t) := \frac{n}{q}\frac{12}{s} \sum_{j=0}^{q-s+1} \frac{1}{2} \left(\frac{1}{s/2} \sum_{j'=0}^{s/2-1} \tilde{p}_{t+j+j'+s/2} - \frac{1}{s/2} \sum_{j'=0}^{s/2-1} \tilde{p}_{t+j+j'} \right)^2 - \frac{n}{q}\frac{6}{s} \sum_{j=0}^{q} \delta_{t+j}(\tilde{p})^2,$$

with $q = [c_1 n^{3/4}]$ and $s = [c_2 n^{1/2}]$.

The constants c_1 and c_2 can be chosen based on the numerical results in Li et al. (2021).

Two-scale estimator

The localized two-scale estimator, proposed by Zu and Boswijk (2014), reads

$$\sigma_{TS}^2(t) := \frac{1}{s} \sum_{j=t-s}^{t} \frac{(\tilde{p}_j - \tilde{p}_{j-q})^2}{q} - \frac{\bar{n}}{n}\frac{1}{s} \sum_{j=t-s}^{t} \delta_j(\tilde{p})^2,$$

where $\bar{n} = \frac{ns-q+1}{qs}$. The MSE-optimal values of s and q are given by

$$q^* = q^0 n^{2/3} \qquad s^* = s^0 n^{-1/6},$$

where

$$q^0 = \left(\frac{12\xi^2}{IQ}\right)^{1/3}$$

and

$$s^0 = n \left(\frac{8\xi^2/(q^0)^2 + \frac{4}{3} q^0 IQ}{\frac{1}{3}\gamma} \right)^{1/2},$$

with

$$\gamma = \sum_{i=0}^{n-1} (\delta_i(\sigma^2))^2, \qquad \sigma^2(t_i) = \frac{1}{s} \sum_{j=0}^{s-1} \delta_{t+j}(\tilde{p})^2.$$

Optimal candlestick estimator

The optimal candlestick estimator proposed by Li et al. (2022) relies on candlestick data, i.e., the opening, closing, highest and lowest prices within a given interval. Formally, denote by $I_{n,i} := \{(i-1)\rho(n), i\rho(n)\}$ the interval associated with the ith candlestick, where $\rho(n) \to 0$ as $n \to \infty$.

Moreover, let O_i, C_i, H_i and L_i denote, respectively, the opening, closing, highest and lowest prices on $I_{n,i}$. The estimator reads

$$\sigma^2_{OK}(t) = \frac{(\lambda_1|O_i - C_i| + \lambda_2(H_i - L_i))^2}{\rho(n)}, \quad t \in I_{n,i}$$

where λ_1 and λ_2 are constants. For the implementation, we set $\rho(n) = 1$ minute (see Section 3.3) and selected (λ_1, λ_2) based on the optimality conditions detailed in Li et al. (2022), Section 2.2.

Pre-averaging kernel estimator

Figueroa-López and Wu (2022) proposed an estimator of the spot variance which consists in a localization of the realized kernel estimator with pre-averaging, i.e., the pre-averaging kernel estimator already defined and discussed in Section 2.3.4, that again we chose to implement and apply following the optimal criterion as outlined in Figueroa-López and Wu (2022).

3.2.3 Feasible selection of tuning parameters

The feasible implementation of the optimization formulae for the tuning parameters which appear in the previous section may require the estimation of IV, IQ and ξ. In this regard, we use the following estimators, as in Gatheral and Oomen (2010):

$$\widehat{IV} = \frac{n}{n-q+1}\frac{1}{q}\sum_{s=0}^{q-1}\sum_{i=1}^{\lfloor (n-s+1)/q \rfloor}(\tilde{p}_{iq+s} - \tilde{p}_{(i-1)q+s})^2$$

$$\widehat{IQ} = \left(\frac{n}{n-q+1}\right)^2 \frac{26}{q}\sum_{s=0}^{q-1}\sum_{i=1}^{\lfloor (n-s+1)/q \rfloor}(\tilde{p}_{iq+s} - \tilde{p}_{(i-1)q+s})^4$$

$$\hat{\xi} = -\hat{\phi}\hat{\sigma}^2_w, \quad \hat{\xi}^2 = (\hat{\xi})^2,$$

where the pair $(\hat{\sigma}^2_w, \hat{\phi})$ is obtained as in (3.2).

Note that q is selected in correspondence of the 5-minute (noise-free) sampling frequency.

3.3 Comparative performance study of volatility estimators

In this section we present the results of a study of the finite-sample performance of the volatility estimators described in the previous section, which is based on simulations of the QR model. In the study, we considered two alternative scenarios. In the first scenario, we assumed constant values of the parameters θ^{ref} and θ^{reinit}, which translate into a constant volatility parameter. Instead, in the second scenario we allowed θ^{ref} and θ^{reinit} to change, so that the volatility parameter is no longer constant. We emphasize the fact that this second scenario introduces a novelty compared to the study by Gatheral and Oomen (2010), where the volatility parameter is constant. In fact, a scenario with time-varying volatility offers a more realistic framework to assess the performance of estimators.

For each scenario, we simulated 2500 daily paths. For each couple $(\theta^{\text{ref}}, \theta^{\text{reinit}})$ considered, the corresponding true value of the volatility parameter was obtained via additional simulations

exploiting Equation 3.1, as described in Subsection 3.1.3. Estimators were computed using 1-second price observations. The integrated volatility was estimated on daily intervals, while for the spot volatility we reconstructed daily trajectories on the 1-minute grid. The selection of the tuning parameters involved in the computation of the different estimators was performed based on the feasible formulae and the suggestions reported in the previous section.

3.3.1 Constant $\theta^{\mathbf{ref}}$ and $\theta^{\mathbf{reinit}}$

In the first scenario, simulations were performed with θ^{ref} and θ^{reinit} constant and equal to, respectively, 0.6 and 0.85, that is, the parameter values calibrated on the MSFT sample (see Section 3.1). We recall that the resulting reference value of the spot variance parameter is $\sigma^2 = 1.0387 \cdot 10^{-8}$ (see Subsection 3.1.3).

Tables 3.4-3.7 display the ranking of the estimators for the three series of mid-price, micro-price and trade-price in terms of their finite-sample performance. The latter is evaluated by means of, resp., the relative bias and MSE for the integrated volatility and the relative integrated bias and MSE for the spot volatility. In each table, the estimators are ordered based on the average ranking, obtained as the arithmetic mean of the rankings for the three prices series.

For what concerns integrated estimators, the pre-averaging estimator and the Fourier estimator provide the relative best performance in terms, resp., of bias and MSE minimization, based on the average ranking. However, note that the Fourier estimator achieves the best MSE average ranking without resulting the first in the ranking for any price series. Indeed, the unified volatility estimator, which is robust to i.i.d. noise and rounding, yields the best performance for mid- and micro-prices, while the best result for trade-prices is achieved by the alternation estimator, which is robust to price discreteness and rounding (see Large (2011)). Overall, this may suggest that robustness to rounding may be crucial to optimize the mean squared error. Instead, the pre-averaging estimator is more clearly favorable in terms of bias reduction, given the fact that it results the first in terms of bias minimization for micro- and trade-price series and the second for mid-price series. Moreover, it is worth noticing that the Min RV and Med RV, which are also computed from pre-averaged data, occupy the second and third places in the average ranking for the bias.

As for spot estimators, simulations indicate that the Fourier estimator outperforms the other estimators considered, both in terms of bias and MSE optimization, for all the price series considered. Only the regularized estimator is capable of obtaining comparable performances in terms of bias. The Fourier estimator differs from the other spot estimators considered in that it relies on the integration of the Fourier coefficients of the volatility rather than on the differentiation of the (estimated) integrated variance, and this appears to represent a solid numerical advantage. Figure 3.2 shows sample trajectories of the spot estimators computed from the mid-price series, along with the true volatility value, to help better understand the difference in performance among the estimators. Appendix 3.B contains the analogous figures for the micro-price and the trade-price.

Finally, in the case of both integrated and spot estimators we investigated whether the volatility values obtained using different methods were statistically different. To this end, for each pair of estimators, we performed a t-test of the null hypothesis that the mean value of the estimated volatility is the same. As a result, we found that all average estimations are pairwise significantly different at the 1% confidence level. Moreover, to better assess the differences in the performances of estimators, we also applied the model confidence set (MCS) procedure by Hansen et al. (2011), with

a significance level of 1%. Following Patton (2011), for the MCS we used the qlike loss function. As a result, the MCS procedure always chooses as the optimal model set the one containing only the estimator with the best ranking in terms of MSE (see Tables 3.5 and 3.7), thus supporting the soundness of our results. Overall, the results of the t-tests and the MCS offer additional support to the fact that the careful selection of the estimation method is not irrelevant.

Integrated variance estimators - relative bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Pre-averaging	-0.01106	2	-0.00643	1	-0.00534	1	1.33
Min RV	0.01071	1	0.01050	2	0.01181	3	2
Med RV	0.01367	3	0.01348	3	0.01456	4	3.33
Unified	0.05518	4	0.04679	4	0.11442	7	5
Fourier	0.09154	6	0.08290	5	0.09234	6	5.66
Range	-0.09196	7	-0.08302	6	-0.06997	5	6
Alternation	0.06799	5	0.11612	12	-0.06226	2	6.33
Kernel	0.12679	8	0.10075	8	0.24836	8	8
Maximum likelihood	0.13990	12	0.10011	7	0.27243	9	9.33
Two-scale RV	0.12827	9	0.10465	9	0.32015	10	9.66
Multi-scale RV	0.12841	10	0.10478	10	0.32030	11	10.33
Bias-corrected RV	0.12852	11	0.10489	11	0.32042	12	11.33

Table 3.4: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price, according to the relative bias. The average ranking is reported in the last column.

Integrated variance estimators - relative MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	0.00981	3	0.00658	2	0.01039	2	2.33
Unified	0.00433	1	0.00346	1	0.01446	6	2.66
Alternation	0.00593	2	0.01145	5	0.00485	1	2.66
Med RV	0.01075	4	0.01174	6	0.01181	3	4.33
Min RV	0.01275	5	0.012754	7	0.01282	4	5.33
Maximum likelihood	0.01647	7	0.01112	3	0.07564	9	6.33
Kernel	0.01726	9	0.01128	4	0.06314	8	7
Pre-averaging	0.01566	6	0.015653	11	0.01568	7	8
Multi-scale RV	0.01648	8	0.012070	8	0.10401	11	9
Range	0.01794	12	0.016448	12	0.01304	5	9.66
Two-scale RV	0.01757	10	0.012901	10	0.10391	10	10
Bias-corrected RV	0.01734	11	0.012094	9	0.10409	12	10.66

Table 3.5: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price, according to the relative mean squared error. The average ranking is reported in the last column.

Spot variance estimators - relative integrated bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	0.00362	1	0.00299	1	0.00713	1	1
Regularized	0.00676	2	0.00663	2	0.00802	2	2
Optimal candlestick	-0.16165	3	-0.12525	3	-0.07430	3	3
Pre-averaging kernel	-0.14755	4	-0.13462	4	-0.08459	4	4
Pre-averaging	0.19939	5	0.19923	5	0.19936	6	5.33
Kernel	0.22205	6	0.23801	6	0.73200	7	6.33
Two-scale	-0.24781	7	-0.26271	7	-0.11798	5	6.33

Table 3.6: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price, according to the relative integrated bias. The average ranking is reported in the last column.

Spot variance estimators - relative integrated MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	0.03095	1	0.03060	1	0.03128	1	1
Pre-averaging kernel	0.13392	2	0.13745	2	0.09148	2	2
Regularized	0.20632	3	0.20628	3	0.20635	3	3
Optimal candlestick	0.31790	4	0.31436	4	0.33429	4	4
Kernel	0.46111	5	0.47095	5	0.84365	5	5
Two-scale	1.62190	6	1.59843	6	2.24599	6	6
Pre-averaging	2.59134	7	2.59064	7	2.59247	7	7

Table 3.7: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price, according to the relative integrated mean squared error. The average ranking is reported in the last column.

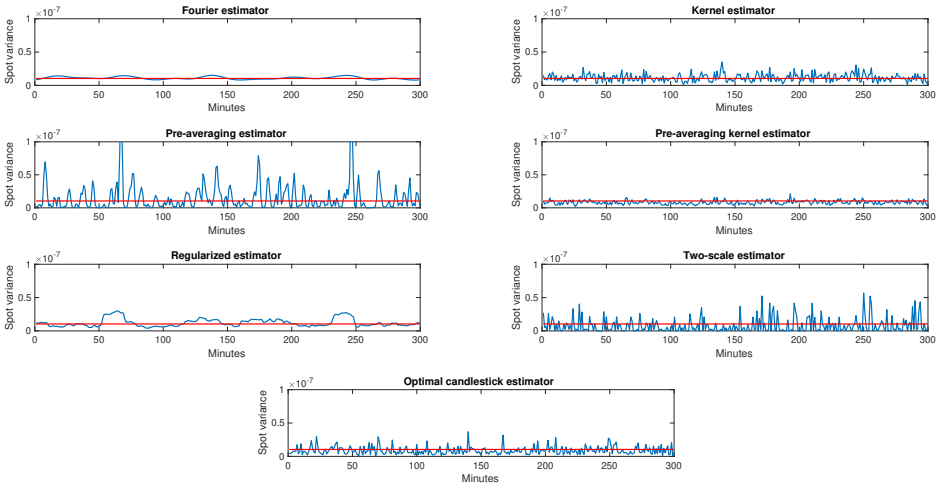

Figure 3.2: Constant θ^{ref} and θ^{reinit}: sample trajectories of spot variance estimators computed from mid-prices (in blue) and true volatility (in red).

3.3.2 Variable θ^{ref} and θ^{reinit}

A nice feature of the QR model, compared to the ZI model, is the flexibility introduced by the parameters θ^{ref} and θ^{reinit}. In the second scenario we assessed the effect of time-varying values of θ^{ref} and θ^{reinit} on the accuracy of volatility estimators. Specifically, we allowed for piece-wise constant volatility dynamics, which might describe a regime-shifting scenario driven, for example, by the flow of information hitting the market.

We considered two sub-scenarios, with increasing variability of the volatility parameter. In the first one, the LOB follows five regimes (each with length equal to 1/5 of a day), which translate into a double u-shaped pattern for the volatility parameter, as illustrated in Table 3.8.

Period	θ^{ref}	θ^{reinit}	σ^2
I	0.7	0.6	$1.473 \cdot 10^{-8}$
II	0.4	0.6	$8.553 \cdot 10^{-9}$
III	0.6	0.85	$1.039 \cdot 10^{-8}$
IV	0.4	0.9	$7.195 \cdot 10^{-9}$
V	0.8	0.9	$1.366 \cdot 10^{-8}$

Table 3.8: θ^{ref} and θ^{reinit} values used for the simulations with five regimes and the corresponding values of σ^2.

In the second sub-scenario, we allowed for ten regimes (each with length equal to 1/10 of a day), which recreate a u-shape pattern for the volatility parameter, see Table 3.9.

Period	θ^{ref}	θ^{reinit}	σ^2	Period	θ^{ref}	θ^{reinit}	σ^2
I	0.7	0.6	$1.473 \cdot 10^{-8}$	VI	0.3	0.5	$4.813 \cdot 10^{-9}$
II	0.8	0.9	$1.366 \cdot 10^{-8}$	VII	0.4	0.9	$7.195 \cdot 10^{-9}$
III	0.8	0.7	$1.281 \cdot 10^{-8}$	VIII	0.4	0.6	$8.553 \cdot 10^{-9}$
IV	0.5	0.5	$7.753 \cdot 10^{-9}$	IX	0.6	0.85	$1.039 \cdot 10^{-8}$
V	0.2	0.8	$3.068 \cdot 10^{-9}$	X	0.9	0.4	$1.469 \cdot 10^{-8}$

Table 3.9: θ^{ref} and θ^{reinit} values used for the simulations with ten regimes and the corresponding values of σ^2.

Tables 3.10 and 3.11 illustrate the average performance rankings in the second scenario[2]. Specifically, average rankings in correspondence of five and ten intra-day regimes are compared with the case of a unique regime (i.e., the case with constant θ^{ref} and θ^{reinit} illustrated in the previous subsection). Full performance results in terms of bias and MSE are detailed in Appendix 3.A.

Overall, it appears that the introduction of time-varying parameters does not significantly affect the performance rankings previously obtained with constant parameters. In other words, most of the rankings are quite stable (with a few exceptions, see, e.g., the average ranking for the MSE of the Range and Pre-averaging estimators), compared to the first scenario. The stability is more evident for spot variance estimators.

[2]In Tables 3.10 and 3.11 the estimators are presented in the same order in which they appear in Section 3.2, i.e., in the chronological order of publication, since the average ranking may change when moving from the bias to the MSE and from one simulating scenario to the other.

As a further investigation, it might be interesting to study whether the performance ranking remains similar also when the volatility path has infinite regimes, for example when the evolution of θ^{ref} and θ^{reinit} is driven by stochastic differential equations. This would require to build a precise mapping between the volatility and θ^{ref} and θ^{reinit} in order to simulate the LOB for each value of the simulated volatility. This interesting investigation is beyond the scope of this work and is left for future research.

As in the previous subsection, to help better understand the difference in performance among the estimators in the second scenario, Figures 3.3 and 3.4 contain sample trajectories of the spot variance estimators computed from mid-price observations, together with the path of the true variance parameter; the analogous figures for micro- and trade-prices are in Appendix 3.B.

Average rankings for different volatility regimes (integrated variance)						
	relative bias			relative MSE		
Estimator	1 regime	5 regimes	10 regimes	1 regime	5 regimes	10 regimes
Pre-averaging	1.33	1	2	8	4	4.33
Fourier	5.66	6.33	6.33	2.33	3	3.33
Med RV	3.33	2	1	4.33	5.33	6.66
Min RV	2	3	3.33	5.33	7	8.66
Unified	5	6	5.66	2.66	3	3.33
Range	6	4	5	9.66	3	3.33
Alternation	6.33	7.66	6.33	2.66	5	4.33
Maximum likelihood	9.33	8.66	8	6.33	7.66	6.33
Kernel	8	8.33	9	7	8	8
Two-scale RV	9.66	9.33	9.33	10	9.66	9.66
Multi-scale RV	10.33	10.33	10.66	9	10.66	9
Bias-corrected RV	11.33	11.66	11.66	10.66	11.66	11

Table 3.10: Average rankings of integrated variance estimators for an increasing number of intra-day volatility regimes.

Average rankings for different volatility regimes (spot variance)						
	relative int. bias			relative int. MSE		
Estimator	1 regime	5 regimes	10 regimes	1 regime	5 regimes	10 regimes
Fourier	1	1	1	1	1	1
Regularized	2	2	2	3	3	3
Pre-averaging kernel	4	3	3	2	2	2
Optimal candlestick	3	4	4.33	4	4	4
Kernel	6.33	6	5.66	5	5	5
Pre-averaging	5.33	5	5.66	7	7	6.66
Two-scale	6.33	7	7	6	6	6.33

Table 3.11: Average rankings of spot variance estimators for an increasing number of intra-day volatility regimes.

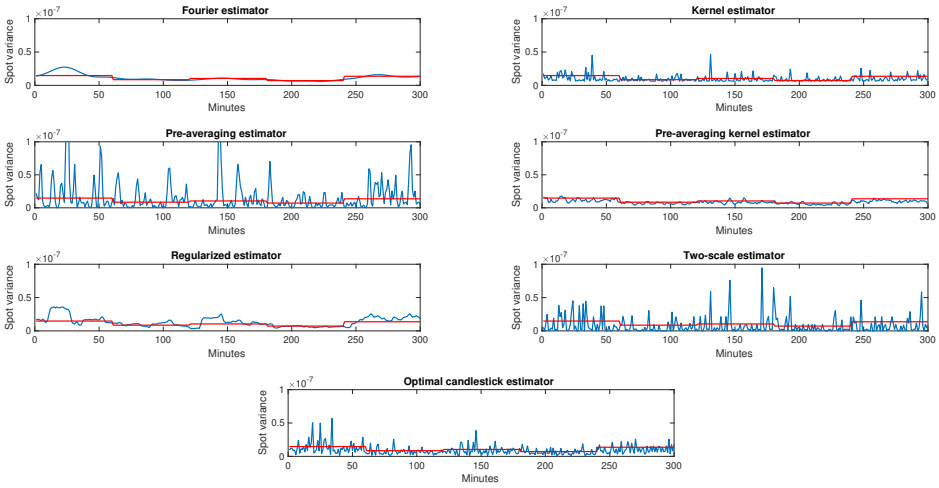

Figure 3.3: Variable θ^{ref} and θ^{reinit} (5 regimes): sample trajectories of spot variance estimators computed from mid-prices (in blue) and true volatility (in red).

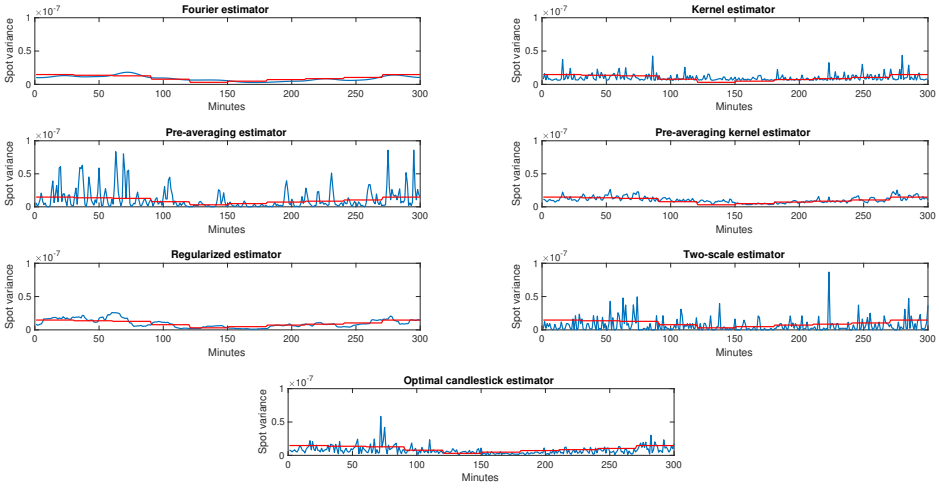

Figure 3.4: Variable θ^{ref} and θ^{reinit} (10 regimes): sample trajectories of spot variance estimators computed from mid-prices (in blue) and true volatility (in red).

3.4 The impact of efficient volatility estimates on optimal execution

In this section, we discuss the results of a study which aims at providing insights into the impact of the use of efficient volatility estimates on the prediction of the variance of the cost of a VWAP execution.

Consider the following problem. A trader has S shares to buy within the interval $[0, T]$. The interval is divided into N_τ time periods of length $\tau = T/N_\tau$ and v_k, $k = 1, ..., N_\tau$, denotes the (signed) number of shares to be traded in interval k. Clearly, $\sum_{k=1}^{N_\tau} v_k = S$. Moreover, let $\tilde{p}_k$ be the price at which the investor trades in interval k (in general different from the average price in the interval, p_k) and p_0 the price before the start of the execution. The objective function is the Implementation Shortfall (IS), defined as

$$C(\mathbf{v}) \equiv \sum_{k=1}^{N_\tau} v_k \tilde{p}_k - S p_0 \tag{3.3}$$

i.e., as the difference between the cost and the cost in an infinitely liquid market. The IS is in general a stochastic variable, therefore one often wishes to minimize

$$\mathbb{E}[C(\mathbf{v})] + \lambda Var[C(\mathbf{v})]$$

where λ measures the risk aversion of the trader.

Let us consider a traders which models markt impact according to the Almgren and Chriss model (Almgren and Chriss (2001)). In this model, the price of the stock at step k is equal to the previous price plus a linear permanent market impact term and a random shock, that is,

$$p_k = p_{k-1} + \vartheta v_k + u_k \qquad u_k \sim \text{IID}(0, \sigma^2 \tau), \tag{3.4}$$

where σ^2 is instantaneous volatility of the unaffected price. Moreover, the actual price paid $\tilde{p}_k$ is different from the average price p_k in the interval and reads

$$\tilde{p}_k = p_k + \rho v_k, \tag{3.5}$$

where ρv_k represents a linear temporary impact.

The expected cost of an execution is then given by

$$E[C(\mathbf{v})] = (\vartheta + \rho) \sum_{k=1}^{N_\tau} v_k^2 + \vartheta \sum_{i>j} v_i v_j$$

and its variance is equal to

$$Var[C(\mathbf{v})] = \sigma^2 \tau \mathbf{v}^T B \mathbf{v},$$

where

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & ... \\ 1 & 2 & 2 & 2 & ... \\ 1 & 2 & 3 & 3 & ... \\ 1 & 2 & 3 & 4 & ... \\ ... & ... & ... & ... & ... \end{pmatrix}.$$

Note that the variance does not depend on the impact parameters ρ and ϑ, but only on the volatility. Further, note that the above expression is more general, as it remains valid also when the temporary impact is nonlinear, see Guéant (2016) for more details.

For the sake of simplicity we assume that the trader performs a VWAP execution (i.e., $\mathbf{v} = \frac{S}{N_\tau}(1, 1, ..., 1)^T$), so that the variance of the cost is equal to

$$Var[C_{VWAP}] = \frac{\sigma^2 S^2}{N_\tau^2}\tau \sum_{k=0}^{N_\tau - 1}(2k+1)(N_\tau - k) = S^2\sigma^2\tau\frac{1}{6N_\tau}(2N_\tau^2 + 3N_\tau + 1) \tag{3.6}$$

This expression shows that, in order to estimate the variance of the execution cost, the traders must have a reliable estimation of σ. We investigate whether the availability of an efficient estimate of the latent volatility parameter could allow the trader to reliably infer the variance of the cost of the strategy. More specifically, we are interested in assessing whether the use of a specific spot volatility estimator, among those studied in Section 3.3, leads to a gain in the accuracy of such inference.

To this aim, we use Monte Carlo scenarios of the QR model to simulate a VWAP execution and we compare the variance of the cost of the simulated executions with the corresponding value predicted by the Almgren and Chriss model (see Equation 3.6), evaluated with the (average) value of σ^2 obtained through a specific estimator. Since it is not obvious that the Almgren-Chriss model faithfully describes the market impact in the QR model, we opted for a more robust comparison that considers the ratio of the aforementioned quantities in correspondence of two different values of the couple $(\theta^{\text{ref}}, \theta^{\text{reinit}})$, i.e. of the volatility. In this way, the effect of the strategy parameters S, τ and N_τ in Equation 3.6, that depend on the specific market-impact model, disappear.

For the simulation of the execution strategy, we set $T = 3$ hours and 20 minutes, $\tau = 10$ minutes and $S = 60$, so that $N_\tau = 20$ and $\mathbf{v} = (3, ..., 3)^T$. The strategy was simulated on top of QR dynamics simulated in two different scenarios with parameters $(\theta^{\text{ref}}, \theta^{\text{reinit}})$ equal to $(0.6, 0.85)$ and $(0.4, 0.6)$. We considered 100 VWAP executions and we computed the empirical variance of their execution cost. The average spot variance values were retrieved from the study of Section 3.3.

Table 3.12 compares the ratios between the empirical variance cost when $(\theta^{\text{ref}}, \theta^{\text{reinit}}) = (0.6, 0.85)$ and when $(\theta^{\text{ref}}, \theta^{\text{reinit}}) = (0.4, 0.6)$) for all estimators. As stated above, we employ a ratio for the purpose of eliminating the constants appearing in Equation 3.6, which depend on the setting of the specific execution.

Results in Table 3.12 suggest that the Fourier estimator and the regularized estimator produce the relative best forecasts of the variance of the strategy costs, as they are associated to a ratio approximately equal to 1.23, which is the closest to the benchmark value of 1.397. As these two estimators provide also the relative best performance in terms of bias and MSE (see Section 3.3), our

Ratios of variances of the VWAP execution cost	
Empirical variance	1.397
Fourier	1.235
Regularized	1.234
Pre-averaging kernel	1.563
Optimal candlestick	1.566
Pre-averaging	1.208
Kernel	1.186
Two-scale	1.030

Table 3.12: Ratios of variances of the VWAP implementation shortfall in correspondence of, respectively, $(\theta^{ref}, \theta^{reinit}) = (0.6, 0.85)$ *(numerator) and* $(\theta^{ref}, \theta^{reinit}) = (0.4, 0.6)$ *(denominator).*

study suggests that efficient volatility estimates may be linked to a better forecast of the variance of the execution cost. Furthermore, note that the range of variation of the ratios in Table 3.12 suggests that the the choice of the estimator is not irrelevant and may lead to significant differences in the forecast of the execution strategy. It seems however that, in general, the use of the formula in Equation. 3.6 leads to a certain underestimation of the the variance of the implementation shortfall of the considered strategy.

Appendix

Appendix 3.A Additional results for variable $\theta^{\mathbf{ref}}$ and $\theta^{\mathbf{reinit}}$

Integrated variance estimators - relative bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Pre-averaging	0.02365	1	0.02519	1	0.02364	1	1
Med RV	-0.03787	2	-0.03778	2	-0.03738	2	2
Min RV	-0.03828	3	-0.03821	3	-0.03782	3	3
Range	-0.07194	4	-0.05015	4	-0.06590	4	4
Unified	0.092375	5	0.08478	5	0.16468	7	6
Fourier	0.12680	7	0.10451	6	0.12580	6	6.33
Alternation	0.09803	6	0.19524	12	0.09607	5	7.66
Kernel	0.17164	9	0.14899	8	0.31278	8	8.33
Maximum likelihood	0.16986	8	0.14813	7	0.41152	11	8.66
Two-scale RV	0.17577	10	0.15552	9	0.41140	9	9.33
Multi-scale RV	0.17592	11	0.15569	10	0.41152	10	10.33
Bias-corrected RV	0.17606	12	0.15583	11	0.41164	12	11.66

Table 3.13: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price with 5 regimes of θ^{ref} and θ^{reinit}, according to the relative bias. The average ranking is reported in the last column.

Integrated variance estimators - relative MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Unified	0.01002	1	0.00864	1	0.02881	7	3
Fourier	0.01525	3	0.01251	2	0.01823	4	3
Range	0.01769	4	0.01253	3	0.01029	2	3
Pre-averaging	0.01888	5	0.01691	4	0.01685	3	4
Alternation	0.01020	2	0.03927	12	0.00332	1	5
Med RV	0.02022	6	0.02022	5	0.02010	5	5.33
Min RV	0.02511	7	0.02510	8	0.02492	6	7
Maximum likelihood	0.03008	8	0.02311	6	0.12163	9	7.66
Kernel	0.03075	9	0.02334	7	0.10014	8	8
Two-scale RV	0.03211	10	0.02534	9	0.17089	10	9.66
Multi-scale RV	0.032161	11	0.02539	10	0.17098	11	10.66
Bias-corrected RV	0.03222	12	0.02543	11	0.17102	12	11.66

Table 3.14: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price with 5 regimes of θ^{ref} and θ^{reinit}, according to the relative mean square error. The average ranking is reported in the last column.

Spot variance estimators - relative integrated bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	-0.00099	1	-0.00083	1	-0.00597	1	1
Regularized	-0.00904	2	-0.00912	2	-0.00877	2	2
Pre-averaging kernel	-0.06324	3	-0.04842	3	-0.05288	3	3
Optimal candlestick	-0.16978	4	-0.12938	4	-0.07699	4	4
Pre-averaging	0.17632	5	0.17629	5	-0.11810	5	5
Kernel	0.20633	6	0.26122	6	0.27361	6	6
Two-scale	-0.24749	7	0.26543	7	0.81305	7	7

Table 3.15: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price with 5 regimes of θ^{ref} and θ^{reinit}, according to the relative integrated bias. The average ranking is reported in the last column.

Spot variance estimators - relative integrated MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	0.035801	1	0.03585	1	0.03616	1	1
Pre-averaging kernel	0.19077	2	0.19206	2	0.14159	2	2
Regularized	0.21590	3	0.21592	3	0.21582	3	3
Optimal candlestick	0.31354	4	0.31110	4	0.32879	4	4
Kernel	0.42691	5	0.49152	5	0.82784	5	5
Two-scale	1.62405	6	1.61549	6	2.38642	6	6
Pre-averaging	2.52237	7	2.52255	7	2.52422	7	7

Table 3.16: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price with 5 regimes of θ^{ref} and θ^{reinit}, according to the relative integrated mean square error. The average ranking is reported in the last column.

Integrated variance estimators - relative bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Med RV	0.06614	1	0.06609	1	0.06660	1	1
Pre-averaging	0.07184	2	0.07182	2	0.07237	2	2
Min RV	0.07669	3	0.07667	3	0.07680	4	3.33
Range	-0.12325	5	-0.11637	5	0.13146	5	5
Unified	0.12626	6	0.11419	4	0.19885	7	5.66
Fourier	0.16911	7	0.14236	6	0.17286	6	6.33
Alternation	0.09319	4	0.16445	12	-0.0741	3	6.33
Maximum likelihood	0.18492	8	0.15720	7	0.37057	9	8
Kernel	0.18758	10	0.15928	8	0.34518	8	9
Two-scale RV	0.18740	9	0.16030	9	0.43099	10	9.33
Multi-scale RV	0.18759	11	0.16048	10	0.43118	11	10.66
Bias-corrected RV	0.18773	12	0.16061	11	0.43131	12	11.66

Table 3.17: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price with 10 regimes of θ^{ref} and θ^{reinit}, according to the relative bias. The average ranking is reported in the last column.

Integrated variance estimators - relative MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Unified	0.01742	2	0.01447	1	0.04113	7	3.33
Fourier	0.02189	3	0.02270	2	0.03208	5	3.33
Range	0.02660	4	0.02659	4	0.02070	2	3.33
Alternation	0.01003	1	0.02796	11	0.00111	1	4.33
Pre-averaging	0.02732	5	0.02632	5	0.02734	3	4.33
Maximum likelihood	0.03541	7	0.02586	3	0.13900	9	6.33
Med RV	0.02881	6	0.02712	10	0.02987	4	6.66
Kernel	0.03641	10	0.02655	6	0.12112	8	8
Min RV	0.03572	8	0.03674	12	0.03674	6	8.66
Multi-scale RV	0.03639	9	0.02689	8	0.18767	10	9
Two-scale RV	0.03642	11	0.02684	7	0.18770	11	9.66
Bias-corrected RV	0.03645	12	0.02694	9	0.18779	12	11

Table 3.18: Performance and ranking of the integrated variance estimators for the series of mid-price, micro-price and trade-price with 10 regimes of θ^{ref} and θ^{reinit}, according to the relative mean square error. The average ranking is reported in the last column.

Spot variance estimators - relative integrated bias							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	-0.02342	1	-0.02412	1	-0.01745	1	1
Regularized	-0.02472	2	-0.02471	2	-0.02438	2	2
Pre-averaging kernel	-0.08816	3	-0.09235	3	-0.07796	3	3
Optimal candlestick	-0.20641	5	-0.15948	4	-0.0838	4	4.33
Pre-averaging	0.24979	6	0.24976	5	0.24958	5	5.33
Kernel	0.20568	4	0.25983	6	0.79978	6	5.66
Two-scale	-0.27408	7	-0.28400	7	-0.28864	7	7

Table 3.19: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price with 10 regimes of θ^{ref} and θ^{reinit}, according to the relative integrated bias. The average ranking is reported in the last column.

Spot variance estimators - relative integrated MSE							
Estimator	mid-price	rank	micro-price	rank	trade-price	rank	av. rank
Fourier	0.05523	1	0.05527	1	0.05618	1	1
Pre-averaging kernel	0.19695	2	0.19833	2	0.15153	2	2
Regularized	0.23490	3	0.23502	3	0.23458	3	3
Optimal candlestick	0.31762	4	0.31529	4	0.34982	4	4
Kernel	0.57819	5	0.68831	5	0.79963	5	5
Two-scale	1.61446	6	1.65113	6	2.78745	7	6.33
Pre-averaging	2.49265	7	2.49293	7	2.49131	6	6.66

Table 3.20: Performance and ranking of the spot variance estimators for the series of mid-price, micro-price and trade-price with 10 regimes of θ^{ref} and θ^{reinit}, according to the relative integrated mean square error. The average ranking is reported in the last column.

Appendix 3.B Sample daily volatility trajectories of spot variance estimates for micro-price and trade-price

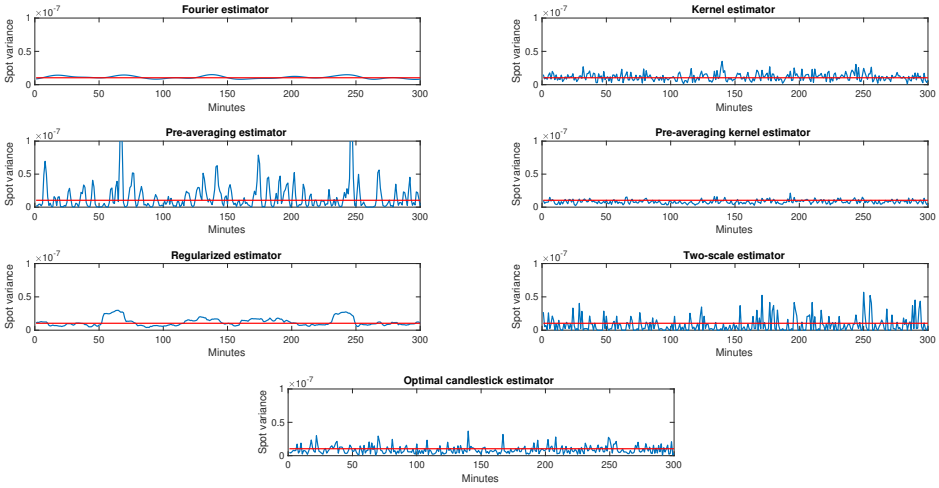

Figure 3.5: Constant θ^{ref} and θ^{reinit}: sample trajectories of spot variance estimators computed from micro-prices (in blue) and true volatility (in red).

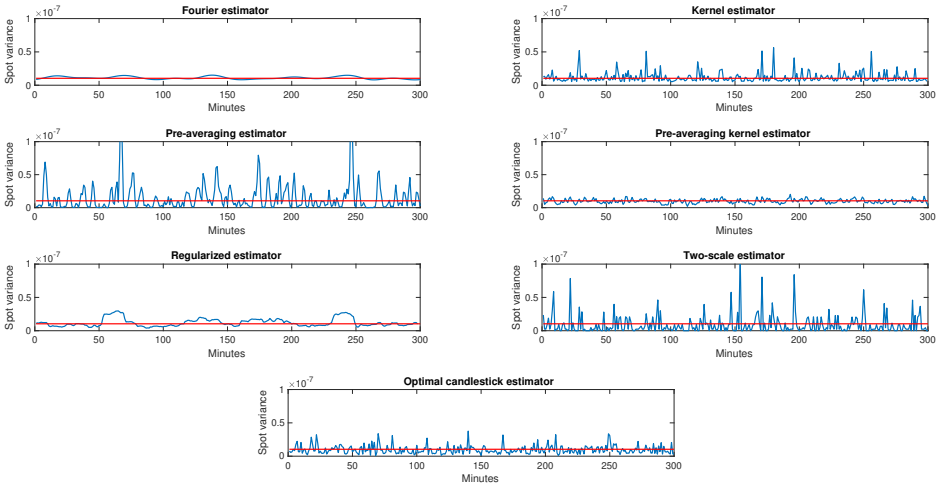

Figure 3.6: Constant θ^{ref} and θ^{reinit}: sample trajectories of spot variance estimators computed from trade-prices (in blue) and true volatility (in red).

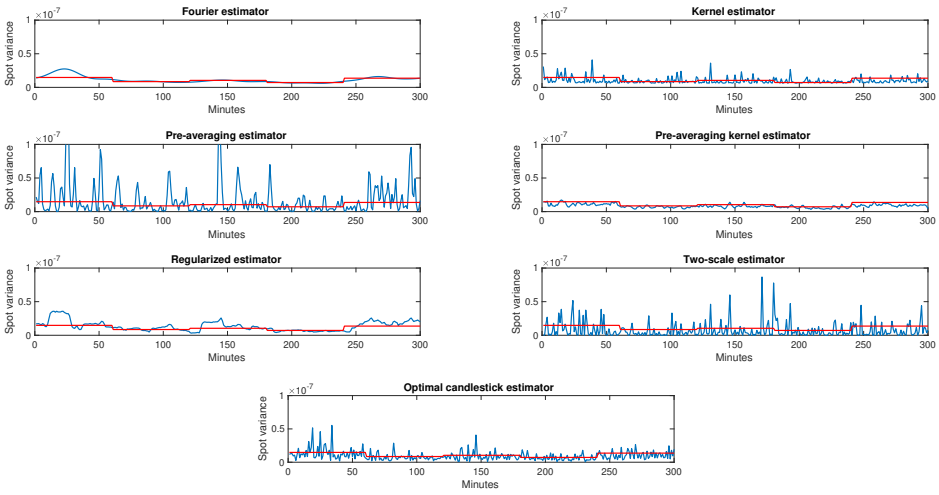

Figure 3.7: Variable θ^{ref} and θ^{reinit} (5 regimes): sample trajectories of spot variance estimators computed from micro-prices (in blue) and true volatility (in red).

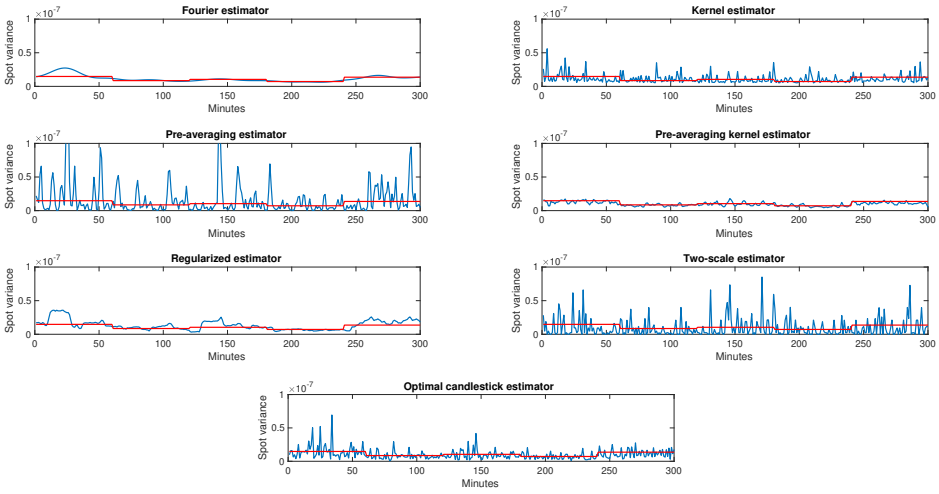

Figure 3.8: Variable θ^{ref} and θ^{reinit} (5 regimes): sample trajectories of spot variance estimators computed from trade-prices (in blue) and true volatility (in red).

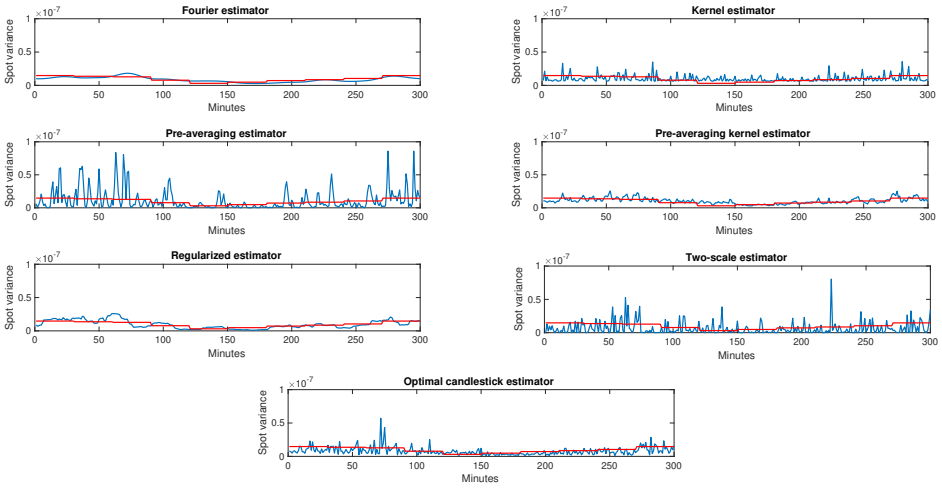

Figure 3.9: Variable θ^{ref} and θ^{reinit} (10 regimes): sample trajectories of spot variance estimators computed from micro-prices (in blue) and true volatility (in red).

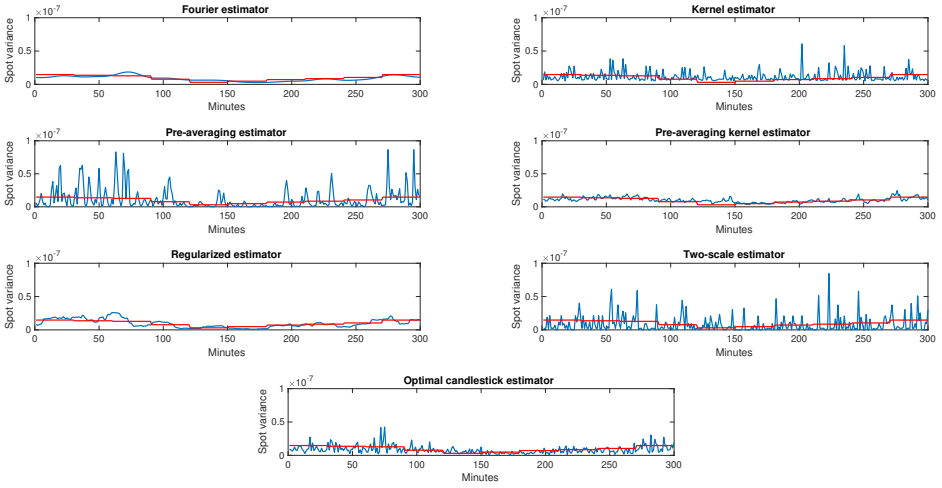

Figure 3.10: Variable θ^{ref} and θ^{reinit} (10 regimes): sample trajectories of spot variance estimators computed from trade-prices (in blue) and true volatility (in red).

Chapter 4

Symmetric positive semi-definite Fourier estimator of instantaneous variance-covariance matrix

While in the previous chapters the focus was on estimation of volatility of one asset, empirical studies have pointed out the importance of considering distinct time variations in correlations between asset prices, thus considering a multivariate setting. In particular in the last years, several studies have addressed the issue of efficiently estimating covariances using high frequency data asynchronously sampled across different assets. While the literature is becoming rich as it concerns the estimation of integrated covariances, it is still sparse for the spot covariances estimation. In contrast with the other estimators which rely on a pre-processing of data in order to make them synchronous, such as linear interpolation, piecewise constant (previous-tick) interpolation or the refresh-time procedure proposed by Barndorff-Nielsen et al. (2011), the Fourier estimator, as clear from Equation 2.5, uses all the available data, being based on an integration procedure. The possibility of using all data, avoiding any preliminary manipulation of them (such as pre-averaging), translates into the direct use of unevenly sampled returns and even asynchronous data in the multivariate case. The estimator proposed by Hayashi and Yoshida (2005) proposes to circumvent the drawbacks caused by asynchronicity by considering the contribution of the product of the price increments only if the corresponding observation intervals are overlapping. The Pre-averaged All-Overlapping estimator by Christensen et al. (2010) builds on the same idea, with the aim to render the estimator by Hayashi and Yoshida (2005) robust to microstructure noise.

Apart from the asymptotic properties such as the rate of convergence to their asymptotic distributions and the finite-sample accuracy, a third factor must now be taken into consideration: the fact that the estimated covariance matrix preserves its symmetry and positive semi-definiteness is a primary issue. This property has important consequences in several contexts, such as the recently developed field of principal component analysis with high-frequency data (Liu and Ngo (2017), Aït-Sahalia and Xiu (2019b), Chen et al. (2020)) or the asset allocation framework (see, e.g., Engle and Colacito (2006)). While this point has been addressed by some authors for the integrated covariances estimators (see, e.g, Barndorff-Nielsen et al. (2011), Mancino and Sanfelici (2011) Park et al. (2016), Cui et al. (2019)), at the best of our knowledge the estimator of spot covariances proposed in the present work is the first to guarantee positive semi-definiteness of the estimation itself, a problem

that so far has not been addressed in the literature. For example, when dealing with spot volatility, Chen et al. (2020) integrate the estimations before computing the eigenvalues of the covariance matrix, while in Bu et al. (2022) positive semi-definiteness is imposed applying suitable shrinkage techniques to the estimation, thus introducing a manipulation of the estimated matrix.

Also the Fourier spot estimator, however, due to lack of symmetry in the Fejèr kernel, may fail to provide positive semi-definite estimations when the asset prices are observed on asynchronous grids. To guarantee that the estimations are symmetric and positive semi-definite, in this chapter we introduce a modified version of the Fourier estimator, which we call Positive-Definite Fourier (hereafter PDF) estimator, and we prove that indeed fulfill the desired property.

The proposed estimator relies on two parameters: the cutting frequency N, and the localization M. The question of how to optimally choose them in order to minimize the error is assessed via a simulation study, where a grid of possible values is tested against several different model specification. We find that the optimal cutting frequency N is stable between different models for the volatility process, and in absence of noise $N = \frac{1}{2}n^{3/4}$ should be chosen, while in presence of noise a value of $N = \frac{1}{2}n^{2/3}$ is more suitable, in line with the other Fourier-type estimators. In both cases, choosing $N < n/2$, i.e. smaller than the Nyquist frequency, leads to a reduced error. The parameter M seems not to be significantly affected by microstructure noise, and the optimal value appears to be $M = N^{4/9}$ or $M = N^{1/2}$ depending on the model for volatility, with small differences between the two choices for the cases analyzed in our study.

In addition to that, the sensitivity of N is further studied in a dedicated simulation exercise, where we specifically analyze the impact of introducing asynchronicity, in comparison to a scenario with synchronous sampling. The above-mentioned choice of N is confirmed as preferable to the Nyquist frequency, since when the correlation between the two assets is smaller than 1 both the absolute bias and the mean square error increase for large values of N in presence of asynchronicity, showing however that exists an interval of values of N that emerges substantially unaffected.

Thereafter, to evaluate the finite-sample performance of the proposed PDF estimator, we compare its accuracy and the percentage of positive semi-definite estimations that it is able to produce with the ones obtained employing alternative estimators able to manage asynchronous observations. First of all we consider a case in absence of market microstructure noise; secondly we focus on the level of asynchronicity, considering different intensities of the Poisson processes that drives the observation frequency; last we analyze the presence of noise, considering noise coming from rounding, i.i.d noise, autocorrelated noise, noise correlated with the efficient price process and heteroskedastic noise. It is shown that, in this exercise, the PDF estimator is the only one to consistently produce positive semi-definite estimations in 100% of the cases, while maintaining an hedge with respect to the competitors in terms of mean square error. Moreover, while the issue of dimensionality in the literature has been mainly considered in relation to PCA, through the comparison we consider whether increasing dimension may change the empirical distribution of the eigenvalues to the point of invalidate the ability of estimators to produce positive semi-definite matrices. As a result, the competing estimators are shown to produce a lower proportion of positive semi-definite matrices when the number of assets considered increases.

The robustness of all the simulation results are confirmed changing the model for the efficient price process; in particular we consider: an Heston Stochastic Volatility model, a One Factor Volatiltiy model, a Two Factor Volatility model (Chernov et al. (2003)) and a Rough Heston model (El Euch

and Rosenbaum (2019)), getting in each case comparable results.

Finally, a technique to reduce the computational cost of the estimator is presented, and the trade-off between accuracy and speed of execution that arises is analyzed with respect to the parameters of the estimator and the dimension of the considered price series.

The remainder of this chapter is organized as follows. In Section 4.1 the positive semi-definite Fourier estimator of spot covariance is introduced, and its positivity is proved. Section 4.2 contains the simulation study, and the comparison between the proposed estimator and alternative estimators, in which accuracy and ability to produce positive semi-definite matrices are considered. Section 4.3 discuss an alternative version of the estimator, in which a randomization technique is used to reduce the execution time of the estimation algorithm.

4.1 The positive semi-definite estimator

Contrary to what done in Chapter 2 and 3, in the following we consider a multivariate setting with multiple assets. In particular, let the logarithmic price process $p = (p^1, ..., p^d)$ be a d-dimensional Itô semimartingale of the form

$$p^j(t) = p^j(0) + \sum_{k=1}^{d} \int_0^t \sigma_{jk}(s) dW_s^k + \int_0^t b_j(s) ds, \quad j = 1, ..., d$$

where $W = (W^1, ..., W^d)$ is a d-dimensional Brownian motion on the filtered probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and b_j and σ_{jk} are adapted processes.

In this work, we are interested in the non-parametric estimation of the $d \times d$ instantaneous (spot) covariance matrix $V(t)$, with entries

$$V^{j,j'}(t) \coloneqq \sum_{k=1}^{d} \sigma_{jk}(t)\sigma_{j'k}(t), \quad \text{for} \quad j, j' = 1, ..., d \quad \text{and} \quad t \in [0, T].$$

We assume, similarly to what was done in Chapter 2, that the prices are observed on discrete, possibly irregular and asynchronous time grids

$$0 = t_0^j \leq t_1^j \leq, ..., \leq t_{n_j}^j = 2\pi \quad \text{for } j = 1, ..., d,$$

where we remark that the assumption $T = 2\pi$ is done for simplicity of notation and is not restrictive. In line with the one dimension case, $\delta_i^j(p)$ denotes the discrete return $p^j(t_{i+1}) - p^j(t_i)$ for $j = 1, ..., d$ and $i = 0, ..., n_j - 1$.

In this setting, we propose the following estimator of spot covariance.

Definition 4.1.1. *Let $\mathcal{K}$ be a finite subset of $\mathbb{Z}$, $\mathcal{S} \coloneqq \{\mathcal{S}(k) \subset_{finite} \mathbb{Z}^2 \colon k \in \mathcal{K}, (s, s') \in \mathcal{S}(k) \implies s + s' = k\}$, and c be a complex function on $\mathcal{K}$; we define the estimator for $V^{j,j'}(t)$ as:*

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t) = \sum_{i=0}^{n_j-1} \sum_{i'=0}^{n_{j'}-1} \sum_{k\in\mathcal{K}} c(k) e^{\mathrm{i}kt} \sum_{(s,s')\in\mathcal{S}} e^{-\mathrm{i}st_i^j} e^{is't_{i'}^{j'}} \delta_i^j(p)\delta_{i'}^{j'}(p). \tag{4.1}$$

Remark 4.1.2. *If we take $\mathcal{K} = \{0, \pm1, \pm2, ..., \pm M\}$ for some positive integer M and $\mathcal{S}(k) =$*

$\{(s,s')|s+s'=k, s=0,\pm1,\pm2,...,\pm N\}$ *for some positive integer* N*, and*

$$c(k)=\frac{1}{2\pi}\left(1-\frac{|k|}{M+1}\right)\frac{1}{2N+1},$$

we obtain:

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t)=\sum_{k=-M}^{M}\left(1-\frac{|k|}{M+1}\right)e^{\mathrm{i}kt}\sum_{s=-N}^{N}\sum_{i=0}^{n_j-1}\sum_{i'=0}^{n_{j'}-1}e^{-\mathrm{i}st_i^j}e^{\mathrm{i}(k-s)t_{i'}^{j'}}\delta_i^j(p)\delta_{i'}^{j'}(p)$$

Using the rescaled Dirichlet and the Féjèr kernels defined respectively in Equation 2.4 and Equation 2.6, the estimator can be written as follows:

$$\hat{V}^{j,j'}_{\mathcal{K},\mathcal{S}}(t)=\frac{1}{2\pi}\sum_{i=0}^{n_j-1}\sum_{i'=0}^{n_{j'}-1}F_M(t-t_i^j)D_N(t_i^j-t_{i'}^{j'})\delta_i^j(p)\delta_{i'}^{j'}(p),$$

and the estimator (4.1) coincides with the Fourier spot covariance estimator introduced by Malliavin and Mancino (2009), whose asymptotic properties (in th case of $d=1$*) have been studied in Mancino and Recchioni (2015) (in the absence of noise) and, both in the presence and in the absence of noise, in Chapter 2. However, while the positivity of the corresponding estimator of the integrated covariance matrix is proved in Mancino and Sanfelici (2011), the spot covariance estimator may fail in producing symmetric positive semi-definite estimations, being* $F_M(t-t_i^j)D_N(t_i^j-t_{i'}^{j'})$ *not symmetric in* j,j'*, leading to complex eigenvalues in* $\hat{V}_{\mathcal{K},\mathcal{S}}(t)$*.*

The main theoretical result of this work concerns the positive semi-definiteness of the proposed estimator, and is stated in the following theorem.

Theorem 4.1.3. *Suppose that* $\mathcal{K}=\{0,\pm1,\pm2,...,\pm2N\}$ *for some positive integer* N*,* $c(k)$ *is a positive semi-definite function on* $\mathcal{K}$ *and*

$$\mathcal{S}(k)=\begin{cases}\{(-N+k+v,N-v):v=0,...,2N-k\} & 0\le k\le 2N\\ \{(N+k-v,-N+v):v=0,...,2N+k\} & -2N\le k<0.\end{cases}$$

Then, $\hat{V}_{\mathcal{K},\mathcal{S}}(t)$ *defined in (4.1) is symmetric and positive semi-definite.*

The proof of Theorem 4.1.3 is reported in Section 4.1.1.

Moreover, it emerges that $\hat{V}_{\mathcal{K},\mathcal{S}}(t)$ can be rewritten as:

$$\hat{V}_N^{j,j'}(t)=\frac{1}{2\pi}\sum_{i=0}^{n_j-1}\sum_{i'=0}^{n_{j'}-1}\sum_{u=-N}^{N}\sum_{u'=-N}^{N}c(u-u')e^{\mathrm{i}u(t-t_i^j)}e^{-\mathrm{i}u'(t-t_{i'}^{j'})}\delta_i^j(p)\delta_{i'}^{j'}(p),\tag{4.2}$$

for two asset j and j' and $t\in(0,2\pi)$, where $c(k)$ is still a positive semi-definite function.

Remark 4.1.4. *The positive semi-definite Fourier (PDF) estimator defined above is a consistent estimator for the spot covariance, as proved in Akahori et al. (2023), if the function* $c(k)$ *is properly chosen. In particular, a Gaussian kernel is sufficient to ensure consistency, and our choice for* $t\in(0,2\pi)$ *is:*

$$c(k)=\frac{1}{2N+1}e^{-\frac{2\pi^2k^2}{M}},\quad \textit{with}\quad M\to\infty\quad \textit{as}\quad N\to\infty.$$

To underline the presence of the localization parameter M, the left side of Equation (4.2) can be written as $\hat{V}^{j,j'}_{N,M}(t)$; insights about the sensitivity of the proposed estimator to the choice of the parameters N and M are reported in Sections 4.2.2 and 4.2.3.

Remark 4.1.5. *By Bochner's theorem, we know that for each positive semi-definite function c, there exists a bounded measure μ on $\mathbf{R}$ such that*

$$c(x) = \int_{\mathbf{R}} e^{2\pi \mathrm{i} y x} \mu(dy).$$

Therefore, we may also rewrite the PDF estimator (4.2) using the measure μ instead of the positive semi-definite function $c(k)$, and obtain

$$\hat{V}^{j,j'}_{N,M}(t) = \frac{2N+1}{2\pi} \sum_{i=0}^{n_j-1} \sum_{i'=0}^{n_{j'}-1} \int_{\mathbf{R}} D_N(t - t^j_i + y) D_N(t - t^{j'}_{i'} + y) \mu_M(dy) \delta^j_i(p) \delta^{j'}_{i'}(p). \tag{4.3}$$

The choice of density should aligned with the choice of $c(k)$ discussed in Remark 4.1.4, for a more detailed description of the assumptions on μ, see Akahori et al. (2023).

4.1.1 Proof of Theorem 4.1.3

Let a_j for $j = 1, 2, 3$ be arbitrary functions on $\mathbb{Z}$, from the definitions of $\mathcal{K}$ and $\mathcal{S}(k)$ we notice that:

$$\begin{aligned}
&\sum_{k \in \mathcal{K}} \sum_{(s,s') \in \mathcal{S}(k)} a_1(k) a_2(s) a_3(s') \\
&= \sum_{k=0}^{2N} \sum_{v=0}^{2N-k} a_1(k) a_2(-N+k+v) a_3(N-v) + \sum_{k=-2N}^{-1} \sum_{v=0}^{2N+k} a_1(k) a_2(N+k-v) a_3(-N+v) =: A + B.
\end{aligned}$$

For the first term we have:

$$\begin{aligned}
A &= \sum_{k=0}^{2N} \sum_{u=k-N}^{N} a_1(k) a_2(k-u) a_3(u) \\
&= \sum_{u=-N}^{N} \sum_{k=0}^{u+N} a_1(k) a_2(k-u) a_3(u) \\
&= \sum_{u=-N}^{N} \sum_{u'=-N}^{N} a_1(u+u') a_2(u') a_3(u),
\end{aligned}$$

where we set $u = N - v$ in the first line, changed the order of the summations in the second line, and put $u' = k - u$. Similarly, using the convention that $\sum_{u=0}^{-1} = 0$, for the second term we have:

$$
\begin{aligned}
B &= \sum_{k=-2N}^{-1} \sum_{u=-N}^{N+k} a_1(k)a_2(k-u)a_3(u) \\
&= \sum_{u=-N}^{N} \sum_{k=u-N}^{-1} a_1(k)a_2(k-u)a_3(u) \\
&= \sum_{u=-N}^{N} \sum_{u'=-N}^{-u-1} a_1(u+u')a_2(u')a_3(u).
\end{aligned}
$$

Thus we see that

$$
\sum_{k\in\mathcal{K}} \sum_{(s,s')\in\mathcal{S}(k)} a_1(k)a_2(s)a_3(s') = \sum_{u=-N}^{N} \sum_{u'=-N}^{N} a_1(u+u')a_2(u')a_3(u).
$$

When $a_1(k) = c(k)e^{\mathrm{i}kt}$, $a_2(s) = e^{-\mathrm{i}st_{i'}^{j'}}$ and $a_3(s') = e^{-ist_i^j}$, using the change of variable $u- \to -u'$, we obtain:

$$
\hat{V}_N^{j,j'}(t) = \sum_{i=0}^{n_j-1} \sum_{i'=0}^{n_{j'}-1} \sum_{u=-N}^{N} \sum_{u'=-N}^{N} c(u-u')e^{\mathrm{i}u(t-t_i^j)}e^{-\mathrm{i}u'(t-t_{i'}^{j'})}\delta_i^j(p)\delta_{i'}^{j'}(p).
$$

Then, for $x \in \mathbf{C}^d$

$$
\begin{aligned}
&\sum_{j,j'} \hat{V}_N^{j,j'}(t)x_j\overline{x}_{j'} \\
&= \sum_{u=-N}^{N} \sum_{u=-N}^{N} c(u-u') \left(\sum_{j=1}^{d} x_j \sum_{i=o}^{n_j-1} e^{\mathrm{i}u(t-t_i^j)}\delta_i^j(p) \right) \left(\sum_{j'=1}^{d} x_j \sum_{i'=0}^{n_{j'}-1} e^{-\mathrm{i}u(t-t_{i'}^{j'})}\delta_{i'}^{j'}(p) \right) \\
&= \sum_{u=-N}^{N} \sum_{u'=-N}^{N} c(u-u')f(u)\overline{f(u')} \geq 0.
\end{aligned}
$$

with $f(u) \coloneqq \sum_{j=1}^{d} x_j \sum_{i=0}^{n_j-1} e^{\mathrm{i}u(t-t_i^j)}\delta_i^j(p)$. The proof is complete.

4.2 Simulation study

4.2.1 Simulation settings

In this section we present an extensive numerical simulated study. The aim of this study is twofold: first in Section 4.2.2 we analyze the sensitivity of the estimator to the choice of the parameters N and M and, with an unfeasible optimization, we find out how much the optimal choice changes in different scenarios. Secondly, in Section 4.2.4 we evaluate the accuracy of the proposed PDF estimator and its ability to produce symmetric and positive semi-definite estimations in a comparison with two alternative estimators that are present in the literature.

To give robustness to the results of our study, in the above-mentioned exercises we consider many

different simulation scenarios, focusing on both the two components of high-frequency financial data: the efficient price and the additive noise component given by market microstructure, so that the observed price $\tilde{p}$ is the multivariate analogous of Equation 2.31, i.e., :

$$\tilde{p}^j_{t_i} = p^j_{t_i} + \eta^j_{t_i}, \qquad j = 1, ..., d \tag{4.4}$$

with η being the noise component.

In particular, for the efficient price process we consider the following specifications:

- Heston stochastic volatility model by Heston (1993);
- the One Factor stochastic volatility model (SVF1) used, e.g., in Huang and Tauchen (2005);
- Two Factor stochastic volatility model (SVF2) by Chernov et al. (2003);
- the Rough Heston model (RH) by El Euch and Rosenbaum (2019);

while for the additive microstructure noise we take into account the following possibilities:

- no noise;
- noise coming from rounding;
- i.i.d. noise;
- autocorrelated noise;
- noise correlated with the efficient price process;
- heteroskedastic noise.

Since in the different cases when noise is present we analyze respectively 2, 4, 3, 3, and 3 different levels of rounding, noise intensity, noise autocorrelation, noise correlation and maximum noise intensity, in our simulated analysis we study a total of 64 different scenarios.

For simplicity of the computations, through Sections 4.2.2 - 4.2.4, all the simulated analysis is conducted on the interval $[0, 1]$; for that reason, and in light of Remark 4.1.4, the PDF estimator is given by:

$$\hat{V}^{j,j'}_{N,M}(t) = \frac{1}{2N+1} \sum_{i=0}^{n_j-1} \sum_{i'=0}^{n_{j'}-1} \sum_{u=-N}^{N} \sum_{u'=-N}^{N} e^{-\frac{2\pi^2(u-u')^2}{M}} e^{\mathrm{i}u(t-t^j_i)} e^{-\mathrm{i}u'(t-t^{j'}_{i'})} \delta^j_i(\tilde{p}) \delta^{j'}_{i'}(\tilde{p}), \tag{4.5}$$

Where not stated otherwise, the simulations consist of $K = 500$ daily trajectories, considering a trading day of length 6.5 hours, and are carried out on an equally spaced grid of width 2 seconds. To introduce asynchronicity in the data, observations are drawn from a Poisson process with intensity $\bar{\Delta}t = 10$, i.e. a process that produces on average one observation every 10 seconds. Moreover, where not explicitly stated, the correlation between Brownian motions driving the efficient processes of different assets, following Bibinger et al. (2019), is fixed to mimic the median estimated realized correlation of the Nasdaq components, i.e:

$$\left\langle W^j, W^i \right\rangle = 0.312, \qquad j, i = 1, ..., d, \quad j \neq i.$$

In Sections 4.2.1 and 4.2.1 we define the models used for the efficient price process and the microstructure noise.

Efficient price process

Heston model

The Heston stochastic volatility model by Heston (1993). The multivariate version of Equation 2.81 takes the form:

$$\begin{cases} dp^j(t) & = (\mu - \frac{1}{2}(\sigma^j(t))^2)dt + \sigma^j(t)dW_t^j \\ d(\sigma^j(t))^2 & = \theta(\alpha - (\sigma^j(t))^2)dt + \gamma\sigma^j(t)dZ_t^j, \end{cases}$$

with $\left\langle W^j, Z^j \right\rangle = \lambda$ to account for the leverage effect, and $\left\langle W^j, Z^i \right\rangle = 0$ for $i \neq j$.

The parameters are set to be:

$$(\mu, \theta, \alpha, \gamma, \lambda) = (0.05/252, 5/252, 0.1, 0.5/252, -0.5),$$

that is the same choice made by Zu and Boswijk (2014), Mancino and Recchioni (2015) and Figueroa-López and Wu (2022).

Factor volatility models

Factor volatility models have been long used in the literature, see for example Huang and Tauchen (2005). First we consider the One Factor Stochastic Volatility model (SV1F) of the form:

$$\begin{cases} dp^j(t) & = bdt + \sigma^j(t)dW_t^j \\ \sigma^j(t) & = \mathrm{e}^{\beta_0 + \beta_1 \tau^j(t)} \\ d\tau^j(t) & = \alpha\tau^j(t) + dZ_t^j \end{cases}$$

for $j = 1, ..., d$, with $\left\langle W^j, Z^j \right\rangle = \lambda$, and $\left\langle Z^j, Z^{j'} \right\rangle = 0$ for $j \neq j'$. The simulation is conducted using as parameters:

$$(\mu, \beta_1, \alpha, \beta_0, \lambda) = (0.03, 0.125, -0.025, \beta_1/(2\alpha), -0.3)$$

that are the parameters used also in Zu and Boswijk (2014), Mancino and Recchioni (2015) and Figueroa-López and Wu (2022).

Second, we consider the Two Factors Stochastic Volatility model (SV2F), proposed by Chernov et al. (2003), and able to reproduce higher values of volatility of volatility. It has the form:

$$\begin{cases} dp^j(t) & = bdt + \text{s-exp}[\beta_0 + \beta_1 \tau^{j,1}(t) + \beta_2 \tau^{j,2}(t)]dW_t^j \\ d\tau^{j,1}(t) & = \alpha_1 \tau^{j,1}(t) + dZ_t^{j,1} \\ d\tau^{j,2}(t) & = \alpha_2 \tau^{j,2}(t) + (1 + \beta_v \tau^{j,2}(t))dZ_t^{j,2} \end{cases}$$

with

$$\text{s-exp(x)} = \begin{cases} \exp(x), & \text{if} \quad x \leq x_0 = \log(1.5) \\ \frac{e^{x_0}}{\sqrt{x_0}}\sqrt{x_0 - x_0^2 + x^2}, & \text{otherwise} \end{cases}$$

for $j = 1, ..., d$, with $\langle W^j, Z^{j,1} \rangle = \langle W^j, Z^{j,2} \rangle = \lambda$, and $\left\langle Z^{j,i}, Z^{j',i'} \right\rangle = 0$ for $j \neq j'$ or $i \neq i'$, $i, i' = 1, 2$. For the parameters our choice is to use:

$$(\mu, \beta_0, \beta_1, \beta_2, \beta_v, \alpha_1, \alpha_2, \lambda) = (0.03, -1.1, 0.04, 0.3, -0.003, -0.6, 0.25).$$

Rough volatility

A new strand of financial econometric literature has grown considering dynamics of the volatility process that are not driven by a standard Brownian motion, but instead are driven by a fractional Brownian motion, with Hurst index $H < 0.5$, see Alòs et al. (2007), Gatheral et al. (2018) and El Euch and Rosenbaum (2018). The proposed PDF estimator is consistent even in presence of rough volatility, see Akahori et al. (2023).

Rough volatility may also be modeled through stochastic Volterra equation, as in the Rough Heston model studied by El Euch and Rosenbaum (2019) and that we intend to use in this analysis:

$$\begin{cases} p^j(t) & = p^j(0) + \int_0^t p^j(s)\sigma(s)^j dW_s^j \\ (\sigma^j(t))^2 & = (\sigma^j(0))^2 + \int_0^t K(t-s)\left((\theta - \gamma(\sigma^j(s))^2)ds + \nu\sigma^j(s)dZ_s^j\right) \end{cases}$$

with $\langle W^j, Z^j \rangle = \lambda$ and $K(t) = Ct^{H-\frac{1}{2}}$ for a Hurst index $H \in \left(0, \frac{1}{2}\right)$ and a constant C. In order to simulate the rough Heston model we apply the discrete-time Euler-type scheme studied in Richard et al. (2023), referring in particular to Equation (8) thereof. The parameters of the model are set to ensure that in the exercise the simulated volatility process does not exhibits negative values, and in particular they take values:

$$(\theta, \gamma, \nu, \lambda, H) = (0.2, 0.3, 0.2, -0.7, 0.1),$$

where the choice for the Hurst parameters is driven by the empirical evidence present in the literature; see, e.g., Gatheral et al. (2018).

Market microstructure noise specifications

In the following analysis we consider some of the most used specification on market microstructure noise that have been proposed in the literature so far. While the basic ideas underlying each model for noise (e.g. presence or absence of correlation with the price process) may be common to the four scenarios presented in the simulated analysis in Section 2.3, the specific forms chosen for this exercise may differ, thus we introduce again each of them separately.

Noise coming form rounding

In presence of rounding the observed price process has the following form:

$$\tilde{p}_t^j = \log([\exp(p_t^j)/r]r),$$

where X denotes the efficient price process. We consider two levels of rounding, corresponding to $r = 1$ or 5 cents, which are the most used in financial markets.

Noise i.i.d.

The most widely used characterization of noise is to consider it the i.i.d. additive component as in Section 2.2.3, with mean equal to zero and a given variance:

$$\tilde{p}_t^j = p_t^j + \eta_t^j, \tag{4.6}$$
$$\eta^j \sim i.i.d. \quad E[\eta^j] = 0, \quad E[(\eta^j)^2] = \text{std}(\delta_{10sec}^j(p))\zeta,$$

where p_{10sec}^j denotes the regularly-spaced series obtained sampling the simulated series every 10 seconds.

Here we specify a Gaussian distribution for the noise, and we set its variance proportional to the average time between two consecutive observations $\bar{\Delta}t$, to avoid that the noise is wiped out by subsampling effect. We consider four values for the noise-to-signal ratio: $\zeta = 1, 1.5, 2, 2.5$.

Autocorrelated noise.

In this case, autocorrelation is introduced in the noise component and the noise, while maintaining the additive form of Equation (4.6), is modeled through an Ornstein-Uhlenbeck process defined as:

$$d\eta^j(t) = -\theta_\eta \eta^j(t)dt + \sigma_\eta dE_t^j,$$

where E is a standard Brownian motion independent of W. Three different levels of autocorrelation are considered, using $\theta_\eta = 0.2, 0.3, 0.4$. σ_η^2 is set to obtain the same level of variance as the second case in the previous scenario.

Noise correlated with the efficient price process

A natural generalization of the previous specification of the noise is to allow for correlation between the Brownian motion that drives the noise component and the one behind the efficient price dynamic:

$$d\eta^j(t) = -\theta_\eta \eta^j(t)dt + \sigma_\eta dE_t^j, \qquad \left\langle W^j, E^j \right\rangle = \rho_\eta$$

where ρ_η is a constant. In light of the empirical evidence present in the literature (see, e.g. Hansen and Lunde (2006)), we impose ρ_η to be negative, with three considered values: -0.1, -0.3, -0.5, and we consider a level of noise equal to the one obtained in the i.i.d. case with noise to signal ratio equal to 2, and absence of autocorrelation in the noise process is assumed.

Heteroskedastic noise

In the last formulation for MMN we allow for the variance of noise to be time changing over the interval. In particular, to reproduce the stylized fact observed by Kalnina and Linton (2008) that the noise is higher at the beginning and at the end of the trading day, following Bu et al. (2022), we use the function:

$$\sigma_\eta(t) = \bar{\sigma}_\eta \left(\frac{1}{2}(\cos(2\pi t) + 1)0.9 + 0.1 \right)$$

choosing $\bar{\sigma}_\eta = 3, 3.5, 4$. For the noise we maintain for this case the Ornstein-Uhlenbeck formulation stated above, considering also the presence of both autocorrelation and dependence between noise and efficient price process, with $\theta = 0.3$ and $\rho_n = -0.3$.

This last formulation allow for a quite general form of microstructure noise, that includes the stylized facts observed in the literature. In a last, unreported, exercise, we add also a rounding of 1 cent to this simulation scheme, without significant changes in the results.

4.2.2 Selection of frequencies N and M

In this section we want to evaluate the sensitivity of the estimator to the choice of the frequencies N and M appearing in the definition of the PDF estimator. The performance of the estimator for each couple of parameters is evaluated in order to search for the optimal parameters over the entire time interval, across the K simulated independent trajectories considered in our unfeasible optimization exercise. In all the following analysis the spot variance trajectory is reconstructed on a regular grid of width 20 minutes.

In this optimization study we let the values of the cutting frequency N to be dependent on n, while the localization parameter M takes values that depend on N itself, building a grid of couples (N, M), with values:

$$N = \frac{n^\alpha}{2}, \qquad \alpha = 1, \frac{5}{6}, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}, \frac{1}{3},$$
$$M = N^\beta, \qquad \beta = \frac{5}{6}, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}, \frac{4}{9},$$

where the choice for the possible values of N relies on the known fact that, as for all the other Fourier-type estimators, in order to avoid aliasing effects, the cutting frequency must be smaller than the Nyquist frequency, i.e., $N \leq n/2$.

For each scenario, in a setting with $d = 2$, on the grid of values described above, we look at the error made in estimating the variance $\hat{V}^{1,1}$ and the covariance $\hat{V}^{1,2}$, using in particular the integrated error:

$$MISE_j = (K)^{-1} \sum_{k=1}^{K} \int_0^1 (\hat{V}_k^{1,j}(t) - V_k^{1,j}(t))^2 dt, \qquad j = 1, 2 \tag{4.7}$$

and choosing as optimal the pair that minimize

$$0.1 \cdot MISE_1 + 0.9 \cdot MISE_2,$$

where a higher weight is given to the estimation of the covariance, being the dominant component of a generic variance-covariance matrix.

Table 4.1 shows the optimal couple of N and M for each scenario. The first thing to notice is that the optimal value for N seems to be pretty stable across the different model for the efficient price process, and it is always smaller than the Nyquist frequency $n/2$, even in absence of noise. When the data are more heavily affected by noise, the optimal N is smaller, coherently with what is

known for the other Fourier-type estimators of both integrated and spot volatility, see, e.g., Mancino et al. (2017). Moreover, the choice of N, while being sensitive to the intensity of noise, does not seems to be dependent of its correlation structure. For what concerns the parameter M, it seems to have almost the same optimum in all the scenarios considered in this exercise. Moreover, for the four models for the efficient price, the difference between $M = N^{4/9}$ and $M = N^{1/2}$ seems to be negligible, as shown in Table 4.2, in which $MISE_2$ on the defined grid is reported for selected scenarios. From Table 4.2 it is also clear that, overall, the estimator is much more sensible to the choice of N than to the choice of M, and changes in the former cause higher shifts in the estimation error; even if, for sake of simplicity, Table 4.2 shows the figures only for the Heston and the SVF2 models, the behavior for the remaining models is analogous. It is important stress that, in line with what observed also for the original spot Fourier estimator in Chapter 2, the removal of the noise is carried out with an appropriate choice of the cutting frequency N. In fact, N controls the number of Fourier coefficients of the asset returns to be considered in the convolution to obtain the Fourier coefficient of the covariance process. It is well know that by reducing the number of frequencies N, the noise is filtered out (see Mancino and Recchioni (2015).

	Heston	SVF1	SVF2	RH
/	No noise			
/	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
r	Noise from rounding			
0.01	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
0.05	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
ζ	I.i.d. noise			
1	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
1.5	3/4, 4/9	3/4, 4/9	3/4, 1/2	3/4, 1/2
2	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
2.5	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
θ	Autocorrelated noise			
0.2	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
0.3	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
0.4	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
ρ_η	Noise correlated with p			
-0.1	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
-0.3	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
-05	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
$\bar{\sigma}_\eta$	Heteroskedastic noise			
3	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
3.5	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2
4	2/3, 4/9	2/3, 4/9	2/3, 1/2	2/3, 1/2

Table 4.1: Optimal couple of α, β in the considered grid across the different models for volatility and microstructure noise.

Remark 4.2.1. *While in this exercise and in the following analysis, focusing on obtaining the best level of performance for the whole variance-covariance matrix for a single couple of parameters N*

Heston - No noise					
α/β	5/6	3/4	2/3	1/2	4/9
1	1.0386e-03	1.0204e-03	1.0069e-03	9.9026e-04	9.8691e-04
5/6	2.3748e-04	2.0522e-04	1.7974e-04	1.4534e-04	1.3932e-04
3/4	2.3898e-04	1.9593e-04	1.6083e-04	1.1666e-04	1.1096e-04
2/3	3.3382e-04	2.8090e-04	2.3916e-04	1.9747e-04	1.9399e-04
1/2	7.4551e-04	7.0530e-04	6.8504e-04	6.7519e-04	6.7487e-04
1/3	2.0294e-03	2.0282e-03	2.0278e-03	2.0277e-03	2.0277e-03

SVF2 - No noise					
α/β	5/6	3/4	2/3	1/2	4/9
1	1.4352e-03	1.4189e-03	1.4101e-03	1.4065e-03	1.4081e-03
5/6	4.0236e-04	3.7224e-04	3.5338e-04	3.3913e-04	3.4125e-04
3/4	3.6879e-04	3.2276e-04	2.9026e-04	2.6256e-04	2.6423e-04
2/3	4.2664e-04	3.7789e-04	3.4306e-04	3.2132e-04	3.2378e-04
1/2	8.8461e-04	8.5025e-04	8.3790e-04	8.4012e-04	8.4159e-04
1/3	2.9639e-03	2.9649e-03	2.9659e-03	2.9668e-03	2.9669e-03

Heston - I.i.d. noise, $\sigma_\eta = 2.5$					
α/β	5/6	3/4	2/3	1/2	4/9
1	6.6533e-03	5.2534e-03	4.2323e-03	2.9124e-03	2.6191e-03
5/6	1.7481e-03	1.3619e-03	1.0611e-03	6.6181e-04	5.9198e-04
3/4	7.3901e-04	6.0224e-04	4.9336e-04	3.5747e-04	3.3973e-04
2/3	4.7485e-04	3.9982e-04	3.4147e-04	2.8342e-04	2.7856e-04
1/2	7.7739e-04	7.3628e-04	7.1565e-04	7.0578e-04	7.0548e-04
1/3	2.0670e-03	2.0658e-03	2.0654e-03	2.0653e-03	2.0653e-03

SVF2 - I.i.d. noise, $\sigma_\eta = 2.5$					
α/β	5/6	3/4	2/3	1/2	4/9
1	9.3093e-03	7.1797e-03	5.5577e-03	3.4186e-03	2.9621e-03
5/6	2.8428e-03	2.2971e-03	1.8831e-03	1.3697e-03	1.2876e-03
3/4	1.2835e-03	1.0975e-03	9.5587e-04	7.9035e-04	7.7287e-04
2/3	6.1650e-04	5.3221e-04	4.6916e-04	4.2024e-04	4.2082e-04
1/2	9.1933e-04	8.8559e-04	8.7380e-04	8.7665e-04	8.7821e-04
1/3	3.0359e-03	3.0369e-03	3.0380e-03	3.0390e-03	3.0391e-03

Table 4.2: Error in estimating covariance over the considered grid, for selected scenarios.

and M, a higher weight is given to covariance estimation, it is worth noting that, when dealing with estimating variance alone (i.e. in an univariate setting) the optimal choice is usually different from the one obtained for covariance, for the differences in the fluctuations of the two quantities. In particular, as Table 4.3 shows for the case of the Rough Heston model, in absence of noise the best choice is obtained for higher values of N and M (in particular setting $\alpha = 1$), while in presence of noise the differences are reduced, leading to a best choice of $\alpha = 2/3$ and $\beta = 5/6$. The values for N are in line with the one observed in similar experiments for the traditional Fourier estimator of spot volatility, see, e.g., Mancino and Recchioni (2015), while the impact of M is confirmed to be less pronounced also in this case. More insights on the differences in optimizing variance and covariance in presence of asynchronicity are given in the next section.

Rough Heston - No noise					
α/β	5/6	3/4	2/3	1/2	4/9
1	6.4138e-03	6.7283e-03	7.1840e-03	8.0398e-03	8.8810e-03
5/6	7.5422e-03	7.8732e-0.3	8.2934e-03	9.3735e-03	9.8061e-03
3/4	8.4182e-03	8.7230e-03	9.1001e-03	1.0076e-02	1.0431e-02
2/3	9.5777e-03	9.7728e-03	1.0058e-03	1.0887e-02	1.1144e-02
1/2	1.3381e-02	1.3525e-02	1.3756e-02	1.4132e-02	1.4190e-02
1/3	2.5303e-02	2.5353e-02	2.5384e-02	2.5407e-02	2.5409e-02

Rough Heston - noise					
α/β	5/6	3/4	2/3	1/2	4/9
1	4.6703e+00	4.6520e+00	4.6384e+00	4.6204e+00	4.6161e+00
5/6	4.1665e-01	4.1473e-01	4.1331e-01	0.41143-01	4.1117e-01
3/4	6.2516e-02	6.2070e-02	6.1793e-02	6.1843e-02	6.2042e02
2/3	1.3675e-02	1.3813e-02	1.4038e-02	1.4408e-02	1.4464e-02
1/2	1.4980e-02	1.4994e-02	1.5127e-02	1.5778e-02	1.6012e-02
1/3	2.6021e-02	2.6071e-02	2.6102e-02	2.6125e-02	2.6127e-02

Table 4.3: Error in estimating variance over the considered grid, for the Rough Heston model.

4.2.3 The impact of introducing asynchronicity in covariance estimation

In this section we study more in detail how much the effect of asynchronicity in the data affects the optimal choice of the cutting frequency N, since introduction of asynchronicity might influence the efficiency of spot variance estimation, as shown in Park et al. (2016), and may lead to a different optimal value for the parameters of a Fourier-type estimator, as shown in Mancino et al. (2017) for the case of integrated variance. While in this section we compare the synchronous and the asynchronous case to evaluate the impact on the choice of N, an additional analysis on the influence of sampling schemes on the error of estimation is presented in Section 4.2.4.

We work here in a simplified environment along the lines of Mancino et al. (2017), and we maintain

in this section a fixed $M = N^{4/9}$, in light of the results of the previous section. In particular we consider two asset, whose prices p^1 and p^2 are equal to two standard Brownian Motions W^1 and W^2, with $\langle W^1, W^2 \rangle = \rho$. To induce asynchronicity in the observations here we simply introduce two distinct vectors of observation times:

$$t_i^1 = i/n \quad \text{for} \quad i = 0, ..., n \quad \text{and}$$

$$t_i^2 = i/n + 0.5/n \quad \text{for} \quad i = 1, ..., n-1,$$

with $t_0^2 = 0$ and $t_n^2 = 1$. We set $n = 500$ and we consider a set of 10000 simulated price paths, on which we estimate the covariance between W^1 and W^2 considering two scenarios for the sampling scheme: the first one in which the two assets are observed synchronously on to the time grid t^1, and the second in which the the two assets are observed respectively on t^1 and t^2. $\hat{V}_N^{1,2}(t)$ is computed for $N = 0, ..., n$ and for $t = 0.5$ (i.e. in the middle of the time window), to evaluate the performance of the estimators while varying the parameter N. The analysis is carried out considering multiple values of correlation between W^1 and W^2, namely $\rho = \{0.2, \pm 0.3, \pm 0.5, \pm 0.7, \pm 1\}$, so that, under this point of view, we expand the range possibilities with respect to the previous section, where a single value for the correlation between assets was taken into account.

Figure 4.1 - 4.5 show the results of this analysis. In particular, each figure show the plot of the relative mean square error and relative bias of the estimation against the number of Fourier coefficients N, for one value of correlation ρ, in the two cases of synchronous and asynchronous observations. The results appear to be consistent across the different values of correlation, with the exception of the case in which two perfectly correlated assets are considered.

It is interesting to notice that while in presence of perfect correlation the introduction of asynchronicity seems to have barely any effect, this is not the case when correlation between the two assets decreases. In particular we can notice that setting $\rho \neq 1$ leads to an increased mean square error for large values of N (in particular for values above the Nyquist frequency, as expected, but also in an interval on the left of $n/2$), while the observed increase in th mean square error for the very small values of N seems to be due to a finite sample effect of the change in the correlation. There is, however, a region for which the mean square error appears to be stable when asynchronicity is introduced, and the starting point of this region seems to be close to the value $N = \frac{1}{2}n^{3/4}$, that we found to be optimal in absence of noise in the previous analysis. Moreover, the introduction of asynchronicity seems to induce a significant bias for large values of N, and therefore the choice of a smaller N also improves the quality of estimation under this point of view. This exercise shows that choice previously adopted seems to be influenced by the presence of asynchronicity in the data, and is robust to different choices of ρ.

While a central limit theorem for the estimator is left for future work, the similar behavior that emerges for different values of correlation appears to be coherent with an asymptotic distribution that does not depend on the estimated quantity itself, as usually happens for high frequency estimators, see, e.g., Aït-Sahalia and Jacod (2014).

4.2.4 Comparison

After having assessed the sensibility of our estimator to the choice of its parameters, in this section we replicate the extensive simulation study adopted in Section 4.2.2 to evaluate the accuracy and

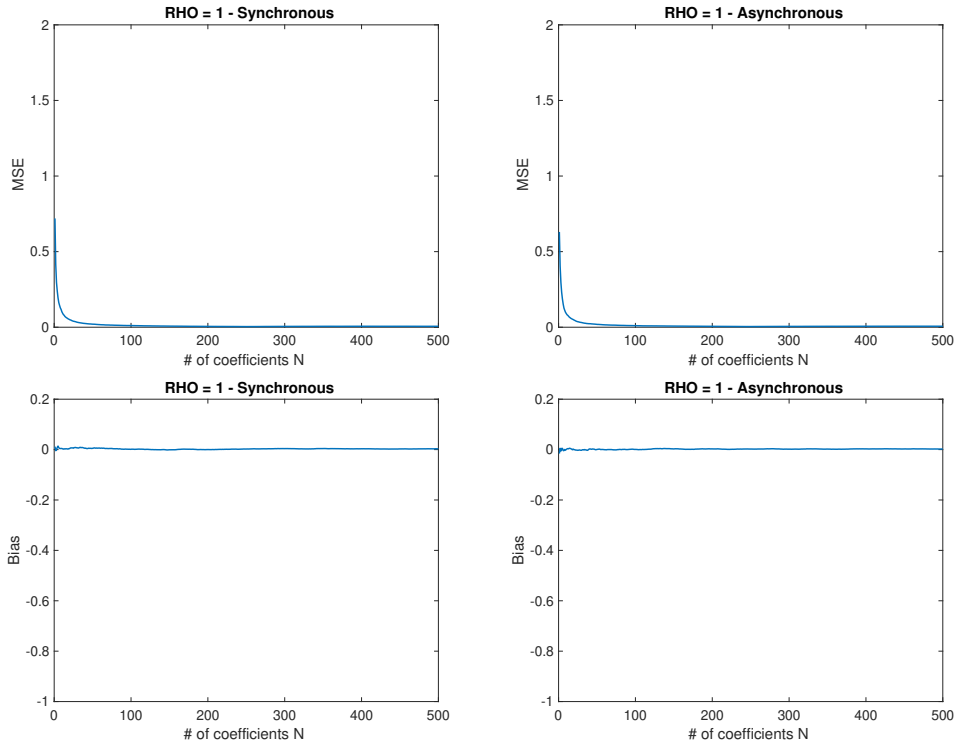

Figure 4.1: Mean square error (first row) and bias (second row) for $\rho = 1$.

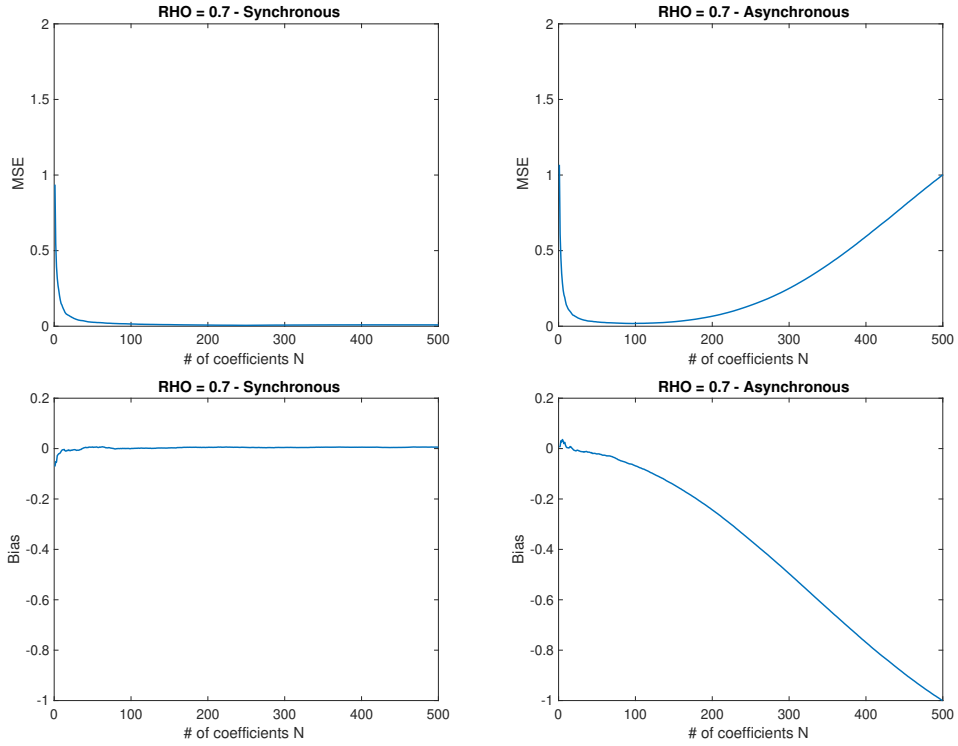

Figure 4.2: Mean square error (first row) and bias (second row) for $\rho = 0.7$.

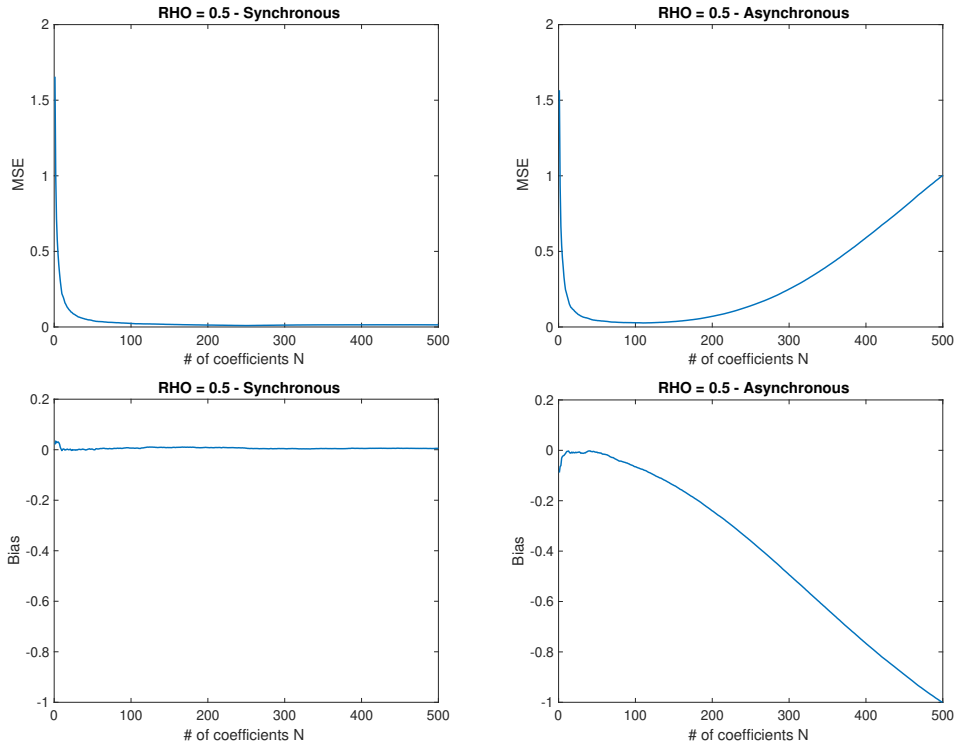

Figure 4.3: Mean square error (first row) and bias (second row) for $\rho = 0.5$.

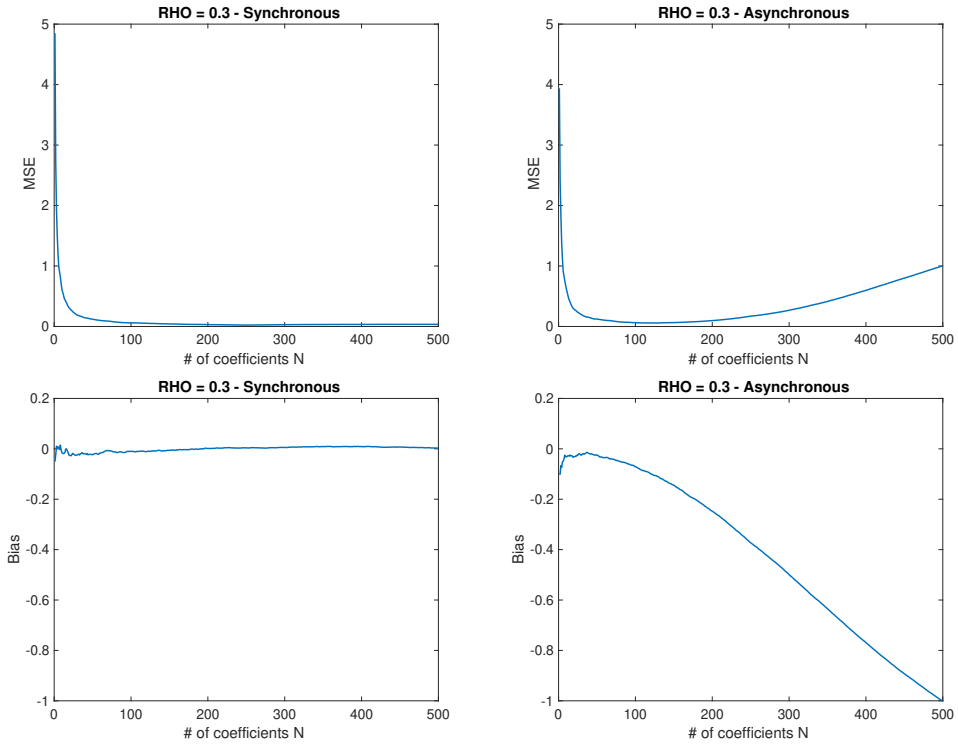

Figure 4.4: Mean square error (first row) and bias (second row) for $\rho = 0.3$.

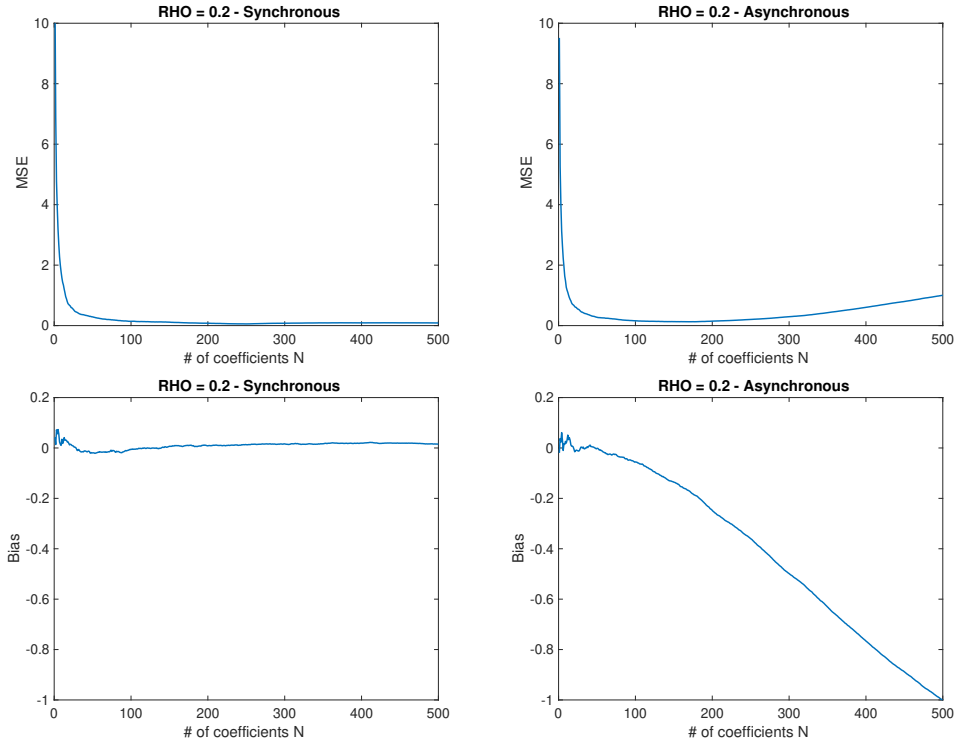

Figure 4.5: Mean square error (first row) and bias (second row) for $\rho = 0.2$.

the ability of the proposed PDF estimator in comparison with the ones of two competing estimators that are present in the literature. Focusing on the estimators able to manage irregularly spaced and asynchronous observations, here we consider:

- the modified Fourier estimator proposed in this chapter (PDF);
- the smoothed two-scale spot estimator, by Mykland et al. (2019) (STS), that is the multivariate analogous of the one defined in Section 2.3.4;
- the local method of moments spot estimators, by Bibinger et al. (2019) (LMM), defined as:

$$\text{vec}\left(\hat{V}(t)\right) = (U_t - L_t + 1)^{-1} \sum_{k=L_t}^{U_t} \sum_{z=1}^{J_n} \hat{W}_z \times \text{vec}\left(S_{zk} S'_{zk} - \pi^2 z^2 h^{-2} \hat{H}_k\right)$$

with the optimal weights W proportional to the local Fisher information matrix

$$W_z = I_k^{-1} I_{zk},$$

where I_{zk} is the Fisher information matrix associated with the k-th block and the spectral frequency i, and $I_k = \sum_{z=1}^{J_n} I_{zk}$, computed an an auxiliary pre-estimation with sub-optimal weights

$$\text{vec}\left(\hat{V}(t)^{pre}\right) = (U_t - L_t + 1)^{-1} \sum_{k=L_t}^{U_t} J^{-1} \sum_{z=1}^{J} \text{vec}\left(S_{zk} S'_{zk} - \pi^2 z^2 h^{-2} \hat{H}_k\right).$$

$\hat{H}_k$ denotes the block-wise constant diagonal matrix of the intensity of noise, with elements obtained as, given an estimate of the nose variance $\hat{\xi}^j$:

$$\hat{H}_k^{jj} = \frac{\hat{\xi}^j}{h} \sum_{kh \le t_i \le (k+1)h} \delta_i (t^j)^2,$$

while to obtain the block-wise covariance matrix the spectral statistics S_{zk} is used

$$S_{zk} = \pi z h_n^{-1} \left(\sum_{i=0}^{n_j - 1} \delta_i^j(\tilde{p})) \Phi_{zk} \left(\frac{t_i^j + t_{i+1}^j}{2} \right) \right)_{1 \le j \le d},$$

with Φ the orthogonal sine function with frequency z:

$$\Phi_{zk}(x) = \frac{\sqrt{2h}}{z\pi} \sin \left(z\pi h^{-1} (x - kh) \right) \mathbb{1}_{[kh,(k+1h))}(x)$$

Finally, we have $J_n = O(\log(n))$ and $h_n = \mathcal{O}(\log(n) n^{1/2})$. For more details on the construction of the estimator and on how L_t, U_t, and J should be set, see Bibinger et al. (2019).

Remark 4.2.2. *The kernel-based estimators proposed by Bu et al. (2022), while it may be extended to manage irregularly spaced and asynchronous observations, relies on specifically tuned shrinkage techniques to impose positive semi-definiteness of the estimation, and is therefore not considered in this analysis.*

The analysis is carried out considering a maximum of $d = 40$ assets, and the performances of each estimator are evaluated according to the mean integrated square error (MISE) and its relative counterpart (RMISE), defined for the volatility matrix as

$$MISE = (Kd^2)^{-1} \sum_{k=1}^{K} \int_0^1 \sum_{j,i=1}^{d} (\hat{V}_k^{ij}(t) - V_k^{ij}(t))^2 dt,$$

$$RMISE = (Kd^2)^{-1} \sum_{k=1}^{K} \int_0^1 \sum_{j,i=1}^{d} (\hat{V}_k^{ij}(t) - V_k^{ij}(t))^2 / V_k^{ij}(t)^2 dt.$$

Unreported results show that using a different loss function, in particular the Frobenius norm, the Euclidean norm or the 1-norm of the difference between the estimated and the real spot volatility matrix, does not affect the rankings that emerge from Tables 4.4-4.10. Most importantly, in this analysis we give particular attention to the percentage of symmetric positive semi-definite (psd) variance-covariance matrix that each estimator is able to produce in the different scenarios.

For the LMM estimator, we use the parameters that where found to be optimal in the numerical analysis by Bibinger et al. (2019), while for the STS estimator we choose the parameters in a neighborhood of the one used by Chen et al. (2020), minimizing the mean square error obtained on auxiliary simulations.

Absence of noise

In Sections 4.2.4 and 4.2.4 we report the results of the comparison, in terms of MISE and % of psd estimations produced by the three competing estimators, when the efficient price process follows the Heston model. In this settings, the results in terms of RMISE are analogous. We begin considering noise-free data, and focusing on two major features of high-frequency covariance matrix estimation: the dimensionality of the matrix and the frequency of observations.

It is known that the spectral distribution of a sample covariance matrix depends on the dimension of the matrix itself (or, better, on the ratio between the number of observations and the dimension of the observed random vector); see, e.g, Bai and Silverstein (2010) for theory of spectral distribution of random matrices. The impact of matrix dimension is relevant also on applications, for example in principal component analysis (the so-called "curse of dimensionality" and its impact on estimation of eigenvalues, see, e.g., Chen et al. (2020)). However, for what concerns the spot variance estimators considered in this work, nothing is known about the spectral density of the estimated matrices, and the estimated matrices (for the alternative estimators) are not even guaranteed to be positive semi-definite. While addressing the issue of studying in detail the spectral distribution of estimated matrices (that may be useful to build, for example, a finite-sample PCA framework that relies on spot covariance matrices), goes beyond the aim of this work, it is of interest here to assess whether changing the dimension of the problem influence the proportion of positive semi-definite estimations obtained with the considered estimators, hence considering a spectral distribution with possibly negative support. This issue is addressed in all the following exercises considering varying levels of d, i.e., an increasing number of assets.

Table 4.4 shows the results for data not considered by noise, looking at matrices of dimension $d = 2, 5, 10, 15, 20, 25, 30, 40$. In this exercise, our modified Fourier estimator is the top performer in terms of mean square error, for any dimension of the volatility matrix, when the efficient price process is observed. The effectiveness of the smoothed two-scale estimator to produce positive semi-definite estimations seems to decrease while the number of assets increase, in particular with $d > 20$, while the other two estimators both produce 100% of psd matrices. For the important role that plays in estimating variance-covariance matrices, and for the influence that it has on the positivity of the estimation, dimensionality will always be taken into consideration in the remaining analysis. For sake of simplicity, the dimension considered in the following are limited to $d = 5, 10, 15, 20$.

In a setting of absence of noise we take into consideration the presence of different sampling schemes for the data. To do so, we study the changes in the performances when the average time between two consecutive observations increases; in particular we extract the observation from the simulated trajectories according to Poisson processes that produce on average one observation every 15, 30 and 40 seconds. Table 4.5 shows that, still maintaining an edge in terms of MISE with respect to the competitors, the PDF estimator is the only one that is able to produce psd estimations in 100% of the cases, while both the STS and the LMM estimator may fails with increased frequency when $\bar{\Delta}t$ increase, even though the impact of this kind of changes seems to be quite small in terms of percentage of positive estimation obtained. It also seems that the accuracy of all the estimators decreases with higher values of $\bar{\Delta}t$; this is of course in line with the fact that the consistency of these estimators is an asymptotic property.

Estimator	MISE	% PSD	MISE	% PSD
	d=2		d=20	
PDF	1.8453e-04	100%	6.5495e-04	100%
LMM	5.2131e-04	100%	1.6008e-03	100%
STS	8.2259e-04	100%	2.5759e-03	100%
	d=5		d=25	
PDF	5.7567e-04	100%	6.1793e-04	100%
LMM	1.5514e-03	100%	1.5090e-03	100%
STS	2.3805e-03	100%	2.4401e-03	99.20%
	d=10		d=30	
PDF	8.3119e-04	100%	6.8819e-04	100%
LMM	2.0574e-03	100%	1.6779e-03	100%
STS	3.3500e-03	100%	2.7116e-03	98.56%
	d=15		d=40	
PDF	6.9991e-04	100%	6.3139e-04	100%
LMM	1.7120e-03	100%	1.5390e-03	100%
STS	2.7634e-03	100%	2.4887e-03	33.10%

Table 4.4: Accuracy (MISE) and % of psd matrix produced by each estimator, when the dimension d of V increase.

Estimator	MISE	% PSD	MISE	% PSD	MISE	% PSD
	d=5, $\bar{\Delta}t = 15$		d=5, $\bar{\Delta}t = 30$		d=5, $\bar{\Delta}t = 40$	
PDF	7.7661e-04	100%	1.2284e-03	100%	1.5871e-03	100%
LMM	1.5937e-03	100%	2.0398e-03	100%	2.3025e-03	99.83%
STS	2.5185e-03	100%	3.4587e-03	100%	4.3867e-03	100%
	d=10, $\bar{\Delta}t = 15$		d=10, $\bar{\Delta}t = 30$		d=10, $\bar{\Delta}t = 40$	
PDF	1.1008e-03	100%	1.7661e-03	100%	2.2329e-03	100%
LMM	2.2549e-03	100%	2.7865e-03	100%	3.0841e-03	99.66%
STS	3.4760e-03	100%	4.8142e-03	100%	6.4107e-03	100%
	d=15, $\bar{\Delta}t = 15$		d=15, $\bar{\Delta}t = 30$		d=15, $\bar{\Delta}t = 40$	
PDF	9.0624e-04	100%	1.4832e-03	100%	1.8593e-03	100%
LMM	1.8481e-03	100%	2.2859e-03	100%	2.5646e-03	98.62%
STS	2.8926e-03	100%	4.0185e-03	99.98%	5.3583e-03	99.75 %
	d=20, $\bar{\Delta}t = 15$		d=20, $\bar{\Delta}t = 30$		d=20, $\bar{\Delta}t = 40$	
PDF	8.5236e-04	100%	1.3852e-03	100%	1.8593e-03	100%
LMM	1.7170e-03	100%	2.1356e-03	99.96%	2.5646e-03	98.62%
STS	2.6941e-03	99.98%	3.7669e-03	98.83%	5.3583e-03	99.08 %

Table 4.5: Accuracy and % of psd matrix produced by each estimator, when the average time between consecutive observations $\bar{\Delta}t$ changes.

Data contaminated by MMN

In this section we run our comparison considering the noise specification described in Section 4.2.1. It is useful to remark that, while the LMM estimator entails an explicit noise correction and the STS estimator relies on pre-averaging of the data, for the proposed PDF estimator, in line with the original Fourier estimator of spot volatility, there is no need to manipulate data or correct the estimator to manage the presence of noise, but it is sufficient to cut the frequency N, as shown in Section 4.2.2.

Table 4.6 shows that the presence of rounding seems not to affect significantly the accuracy of the estimators and the positive semi-definiteness of the estimations. This effect may be due to the scheme adopted to simulate irregularly sampled data, that implies a subsampling with respect to the rounded simulated series, reducing the intensity of this source of noise.

Estimator	MISE	% PSD	MISE	% PSD
	d=5, r=0.01		d=5, r=0.05	
PDF	5.6886e-04	100%	5.8434e-04	100%
LMM	1.4588e-03	100%	1.4744e-03	100%
STS	2.4015e-03	100%	2.4342e-03	100%
	d=10, r=0.01		d=10, r=0.05	
PDF	8.2010e-04	100%	8.2171e-04	100%
LMM	2.0337e-03	100%	2.0445e-03	100%
STS	3.3250e-03	100%	3.3612e-03	100%
	d=15, r=0.01		d=15, r=0.05	
PDF	6.7300e-04	100%	6.7597e-04	100%
LMM	1.6870e-03	100%	1.6806e-03	100%
STS	2.7879e-03	100%	2.7996e-03	100%
	d=20, r=0.01		d=20, r=0.05	
PDF	6.3197e-04	100%	6.3562e-04	100%
LMM	1.5806e-03	100%	1.5791e-03	100%
STS	2.6104e-03	100%	2.6063e-03	100%

Table 4.6: Accuracy and % of psd matrix produced by each estimator, when a rounding of 1 or 5 cents is present.

Table 4.7 shows that i.i.d. noise, especially with high ζ, is able to negatively affect the ability of the LMM and STS estimators to produce psd estimations, with a stronger impact as the dimension of the estimated matrix grows. Also, the accuracy of all the estimators deteriorates with higher noise, with the PDF confirmed as the top performer also in this scenario. Since i.i.d. noise, form a market microstructure perspective, is usually linked to the presence bid-ask spread as modelled, e.g., as in Roll (1984), and being the bid-ask spread usually related to the liquidity of an asset, the ability of managing this kind of noise may be regarded as the ability to estimate correctly covariance also for illiquid assets.

Table 4.8 shows that autocorrelation is able to significantly effect the ability of the STS estimators to produce psd matrices, in particular with low values of θ_η; also the LMM estimator is affected, even if more slightly, Low values of θ_η also reduce the accuracy of all the three estimators, while

Estimator	MISE	% PSD	MISE	% PSD	MISE	% PSD	MISE	% PSD
	d=5, $\zeta = 1$		d=5, $\zeta = 1.5$		d=5, $\zeta = 2$		d=5, $\zeta = 2.5$	
PDF	5.8957e-04	100%	1.2411e-03	100%	2.1739e-03	100%	3.5475e-03	100%
LMM	1.4774e-03	100%	2.0615e-03	100%	3.2557e-03	99.91%	5.5029e-03	98.18%
STS	2.4258e-03	100%	3.0744e-03	100%	3.9064e-03	99.95%	5.1760e-03	97.85%
	d=10, $\zeta = 1$		d=10, $\zeta = 1.5$		d=10, $\zeta = 2$		d=10, $\zeta = 2.5$	
PDF	8.3152e-04	100%	1.7158e-03	100%	2.4388e-03	100%	3.5411e-03	100%
LMM	2.0447e-03	100%	2.6918e-03	100%	3.6557e-03	99.56%	5.5120e-03	91.14%
STS	3.3806e-03	100%	4.3057e-03	99.83%	5.4130e-03	90.31%	7.0991e-03	52.19 %
	d=15, $\zeta = 1$		d=15, $\zeta = 1.5$		d=15, $\zeta = 2$		d=15, $\zeta = 2.5$	
PDF	6.8032e-04	100%	1.4012e-03	100%	1.8502e-03	100%	2.5615e-03	100%
LMM	1.6911e-03	100%	2.1709e-03	99.98%	2.8183e-03	99.20%	3.9896e-03	80.66%
STS	2.8036e-03	99.87%	3.5990e-03	91.42%	5.8970e-03	34.15 %	6.5234e-03	26.37%
	d=20, $\zeta = 1$		d=20, $\zeta = 1.5$		d=20, $\zeta = 2$		d=20, $\zeta = 2.5$	
PDF	6.3538e-04	100%	1.3026e-03	100%	1.6698e-03	100%	2.2492e-03	100%
LMM	1.5694e-03	100%	2.0025e-03	99.96%	2.5535e-03	97.44%	3.5224e-03	64.45%
STS	2.6111e-03	97.51%	3.3587e-03	42.48%	4.1644e-03	11.71 %	5.4918e-03	0.50%

Table 4.7: Accuracy and % of psd matrix produced by each estimator, when the data is contaminated by i.i.d. noise.

maintaining their ranking unchanged.

Estimator	MISE	% PSD	MISE	% PSD	MISE	% PSD
	d=5, $\theta_\eta = 0.2$		d=5, $\theta_\eta = 0.3$		d=5, $\theta_\eta = 0.4$	
PDF	5.6169e-03	100%	2.8806e-03	100%	2.1414e-03	100%
LMM	1.0856e-02	100%	4.8943e-03	100%	3.3544e-03	100%
STS	8.5059e-03	97.31%	5.1553e-03	99.08%	4.1302e-03	99.92%
	d=10, $\theta_\eta = 0.2$		d=10, $\theta_\eta = 0.3$		d=10, $\theta_\eta = 0.4$	
PDF	5.1481e-03	100%	3.0019e-03	100%	2.3519e-03	100%
LMM	9.7818e-03	100%	5.0765e-03	99.77%	3.7482e-03	100%
STS	1.0814e-02	60.36%	6.8148e-03	90.08%	5.6242e-03	91.37%
	d=15, $\theta_\eta = 0.2$		d=15, $\theta_\eta = 0.3$		d=15, $\theta_\eta = 0.4$	
PDF	3.5769e-03	100%	2.2202e-03	100%	1.8000e-03	100%
LMM	6.6905e-03	92.90%	3.7483e-03	95.98%	2.8798e-03	96.80%
STS	8.7234e-03	3.70%	5.6586e-03	29.92%	4.6585e-03	35.03%
	d=20, $\theta_\eta = 0.2$		d=20, $\theta_\eta = 0.3$		d=20, $\theta_\eta = 0.4$	
PDF	3.0654e-03	100%	1.9632e-03	100%	1.6249e-03	100%
LMM	5.6948e-03	93.90%	3.3182e-03	92.98%	2.6127e-03	95.48%
STS	8.0435e-03	4.87%	5.2368e-03	9.65%	4.3666e-03	8.23%

Table 4.8: Accuracy and % of psd matrix produced by each estimator, when the data is contaminated by autocorrelated noise.

Table 4.9 shows that introducing correlation between E and W, i.e. between the Brownian motions that drives the noise and the one that drives the dynamic of the efficient price, does not seem to influence particularly the positive semi-definiteness of the estimations. Moreover, it seems

that the accuracy on slightly increase for all the estimators, with respect to the baseline i.i.d. case; this is due to the fact that a negative autocorrelation between efficient price and noise may induce a bias with opposite sign whit respect to the bias induced by uncorrelated noise; this is usually observed using volatility signature plots, as done in Hansen and Lunde (2006).

Estimator	MISE	% PSD	MISE	% PSD	MISE	% PSD
	d=5, $\rho_\eta = -0.1$		d=5, $\rho_\eta = -0.3$		d=5, $\rho_\eta = -0.5$	
PDF	2.0352e-03	100%	1.9671e-03	100%	1.9375e-03	100%
LMM	3.1654e-03	100%	3.0462e-03	100%	2.9484e-03	100%
STS	3.9369e-03	99.81%	3.8525e-03	99.96%	3.8145e-03	99.92%
	d=10, $\rho_\eta = -0.1$		d=10, $\rho_\eta = -0.3$		d=10, $\rho_\eta = -0.5$	
PDF	2.2896e-03	100%	2.3440e-03	100%	2.2369e-03	100%
LMM	3.5607e-03	99.94%	3.4779e-03	99.92%	3.4024e-03	99.98%
STS	5.3662e-03	91.56%	5.2421e-03	94.05%	5.1648e-03	95.49%
	d=15, $\rho_\eta = -0.1$		d=15, $\rho_\eta = -0.3$		d=15, $\rho_\eta = -0.5$	
PDF	1.7608e-03	100%	1.7866e-03	100%	1.7255e-03	100%
LMM	2.7502e-03	98.62%	2.7090e-03	99.44%	2.6483e-03	99.79%
STS	4.4396e-03	37.17%	4.3377e-03	44.13%	4.2908e-03	53.57%
	d=20, $\rho_\eta = -0.1$		d=20, $\rho_\eta = -0.3$		d=20, $\rho_\eta = -0.5$	
PDF	1.6065e-03	100%	1.6253e-03	100%	1.5582e-03	100%
LMM	2.5180e-03	98.35%	2.4725e-03	98.89%	2.3988e-03	95.84%
STS	4.1353e-03	0.82%	4.0389e-03	2.92%	3.9787e-03	5.43%

Table 4.9: Accuracy and % of psd matrix produced by each estimator, when the data is contaminated by noise correlated with the efficient price process.

Table 4.10 shows the results for the last specification of noise that we consider, with noise autocovariance, correlation with the efficient price and time-varying noise variance. Our results show that, in this setting, both the LMM and the STS estimators may have difficulties in reaching satisfactory percentages of psd estimations, depending on the intensity of the microstructure component. We confirm once again the ability of the PDF estimator to produce variance-covariance matrix with the desired property, and with relatively low estimation error.

4.2.5 Alternative volatility models

In the previous sections we have exposed the results of our comparison for what concerns simulation of the efficient price process made with an Heston model. Even though the error produced by the three estimators may be different changing the simulation model behind our analysis, and in particular the differences in MISE between the PDF and the LMM estimators are reduced when using the SVF2 or the Rough Heston model, the results are substantially confirmed: the PDF estimator remain the only one able to consistently produce positive semi-definite estimations, and is the best performer in terms of mean square error in almost any scenario. Table 4.11 shows the percentage of psd estimations obtained under the alternative volatility models, in absence of microstructure noise. We can see that, in this exercise, it seems that, moving to the SV1F, to the SV2F or Rough Heston, does not influence significantly the ability of the estimators of producing positive matrices, and the LMM

Estimator	MISE	% PSD	MISE	% PSD	MISE	% PSD
	d=5, $\bar{\sigma}_\eta = 3$		d=5, $\bar{\sigma}_\eta = 3.5$		d=5, $\bar{\sigma}_\eta = 4$	
PDF	4.7998e-03	100%	7.8667e-03	100%	1.3212e-02	100%
LMM	6.8314e-03	100%	1.1376e-02	98.81%	1.8293e-02	93.86%
STS	8.6164e-03	84.19%	1.3869e-02	71.84%	2.0190e-02	61.96%
	d=10, $\bar{\sigma}_\eta = 3$		d=10, $\bar{\sigma}_\eta = 3.5$		d=10, $\bar{\sigma}_\eta = 4$	
PDF	4.5443e-03	100%	6.9249e-03	100%	1.1111e-02	100%
LMM	6.7069e-03	99.89%	1.0364e-02	95.89%	1.6497e-02	83.48%
STS	1.1635e-02	52.86%	1.7614e-02	36.65%	2.6044e-02	30.51%
	d=15, $\bar{\sigma}_\eta = 3$		d=15, $\bar{\sigma}_\eta = 3.5$		d=15, $\bar{\sigma}_\eta = 4$	
PDF	3.2065e-03	100%	4.7542e-03	100%	7.2763e-03	100%
LMM	4.8670e-03	99.37%	7.2946e-03	89.26%	1.1172e-02	68.90%
STS	9.4853e-03	35.25%	1.4449e-02	24.56%	2.1268e-02	19.98%
	d=20, $\bar{\sigma}_\eta = 3$		d=20, $\bar{\sigma}_\eta = 3.5$		d=20, $\bar{\sigma}_\eta = 4$	
PDF	2.7842e-03	100%	4.0233e-03	100%	6.0704e-03	100%
LMM	4.2737e-03	98.14%	6.2574e-03	80.82%	9.4823e-03	54.03%
STS	8.7501e-03	25.69%	1.3279e-02	16.33%	1.9663e-02	11.38%

Table 4.10: Accuracy and % of psd matrix produced by each estimator, when the data is contaminated by heteroskedastic noise process.

estimator shows the same ability of producing psd matrices than the PDF estimator. More extensive results about the alternative models in the presence of microstructure noise, showing in particular the percentage of psd estimations in the cases with i.i.d., autocorrelated and heteroskedastic noise, are reported Tables 4.12 - 4.14, whose results broadly confirms the previous findings under those alternative settings. It is worth underlying, as final consideration, that the results in terms or RMISE always see the PDF estimator as the top performer, and also the ranking in terms of MISE remains unchanged in most of the cases, when moving to the alternative models.

Estimator	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=2			d=20		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	99.86%	99.86%	100%
	d=5			d=25		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	98.45%	97.20%	96.54%
	d=10			d=30		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	94.25%	92.80	29.72%
	d=15			d=40		
PDF	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%
STS	100%	100%	100%	34.04%	32.12%	27.71%

Table 4.11: % of psd matrix produced by each estimator, when the efficient price process is produced by alternative models.

Estimator	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\zeta = 1$			d=5, $\zeta = 1.5$			d=5, $\zeta = 2$			d=5, $\zeta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	100%	100%	100%	100%	99.85%	99.68%	97.28%	93.98%	94.98%
STS	100%	100%	100%	100%	100%	100%	99.77%	99.77%	99.79%	97.95%	97.33%	97.70%
	d=10, $\zeta = 1$			d=10, $\zeta = 1.5$			d=10, $\zeta = 2$			d=10, $\zeta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	100%	100%	99.70%	99.25%	95.26	90.03%	85.48%	87.11%
STS	100%	100%	100%	99.60%	99.69%	99.54%	90.91%	89.16%	91.10	52.28%	49.62%	54.59%
	d=15, $\zeta = 1$			d=15, $\zeta = 1.5$			d=15, $\zeta = 2$			d=15, $\zeta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	99.80%	99.80%	99.25%	97.85%	98.02%	79.28%	75.16%	77.98%
STS	100%	99.90%	99.96%	90.52%	88.66%	90.81%	37.21%	29.88%	39.24%	23.72%	15.64%	18.25%
	d=20, $\zeta = 1$			d=20, $\zeta = 1.5$			d=20, $\zeta = 2$			d=20, $\zeta = 2.5$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	100%	100%	99.96%	98.27%	98.82%	97.95%	95.23%	95.88%	65.91%	63.84%	65.95%
STS	97.37%	96.68%	97.14%	42.42%	49.39%	44.44%	28.56%	22.64%	25.84%	1.80%	0.28%	0.02%

Table 4.12: % of psd matrix produced by each estimator, when the efficient price process is produced by alternative models, in presence of i.i.d. noise.

Estimator	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\theta_\eta = 0.2$			d=5, $\theta_\eta = 0.3$			d=5, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	99.96%	100%	100%	100%	100%	100%	100%	100%
STS	97.39%	97.20%	97.83%	99.21%	99.14%	99.83%	99.77%	99.25%	99.73%
	d=10, $\theta_\eta = 0.2$			d=10, $\theta_\eta = 0.3$			d=10, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.85%	98.20%	100%	98.16%	96.90%	100%	100%	100%	100%
STS	49.23%	47.87%	53.55%	73.25%	71.10%	73.25%	90.39%	88.34%	89.65%
	d=15, $\theta_\eta = 0.2$			d=15, $\theta_\eta = 0.3$			d=15, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.12%	97.90%	96.55%	99.85%	99.73%	98.16%	99.88%	95.36%	97.22%
STS	10.22%	10.63%	9.36%	11.32%	15.38%	14.64%	21.32%	20.44%	20.89%
	d=20, $\theta_\eta = 0.2$			d=20, $\theta_\eta = 0.3$			d=20, $\theta_\eta = 0.4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	98.52%	97.50%	96.55%	98.63%	99.20%	98.25%	99.85%	96.17%	96.82%
STS	1.83%	1.32%	3.74%	4.57%	3.22%	4.72%	8.94%	8.55%	9.33%

Table 4.13: % of psd matrix produced by each estimator, when the efficient price process is produced by alternative models, in presence of autocorrelated noise.

Estimator	SV1F	SV2F	RH	SV1F	SV2F	RH	SV1F	SV2F	RH
	d=5, $\bar{\sigma}_\eta = 3$			d=5, $\bar{\sigma}_\eta = 3.5$			d=5, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	100%	99.96%	99.96%	99.25%	98.95%	99.18%	92.79%	91.65%	91.31%
STS	89.60%	82.73%	84.52%	72.26%	72.97%	74.06%	61.17%	60.53%	63.01%
	d=10, $\bar{\sigma}_\eta = 3$			d=10, $\bar{\sigma}_\eta = 3.5$			d=10, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.85%	99.73%	99.77%	95.78%	95.74%	95.61%	79.97%	80.03%	81.27%
STS	45.15%	44.82%	45.95%	36.34%	36.11%	36.38%	30.33%	29.74%	30.45%
	d=15, $\bar{\sigma}_\eta = 3$			d=15, $\bar{\sigma}_\eta = 3.5$			d=15, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	99.12%	98.02%	99.25%	90.68%	89.58%	91.22%	67.04%	66.54%	68.96%
STS	30.58%	29.88%	30.81%	23.92%	23.39%	23.66%	19.40%	19.32%	18.94%
	d=20, $\bar{\sigma}_\eta = 3$			d=20, $\bar{\sigma}_\eta = 3.5$			d=20, $\bar{\sigma}_\eta = 4$		
PDF	100%	100%	100%	100%	100%	100%	100%	100%	100%
LMM	97.58%	97.09%	97.58%	83.06%	80.64%	83.44%	52.92%	51.88%	55.24%
STS	21.37%	21.37%	21.30%	15.77%	15.66%	15.31%	11.09%	11.05%	10.99%

Table 4.14: % of psd matrix produced by each estimator, when the efficient price process is produced by alternative models, in presence of heteroskedastic noise.

Remark 4.2.3. *After having analyzed how much the different specifications for noise may affect the ability of some spot volatility estimator to produce positive semi-definite estimations in finite samples, we would like to point out, as a final remark on the performance comparison, that, even applying some manipulations to the estimated matrices useful to overcome this shortcoming, the gap in the estimation error is not reduced. To this extent we follow the seminal work by Higham (1988), in which it is shown how to obtain the nearest symmetric positive semi-definite matrix to an arbitrary real-valued matrix, where the distance is measured in the Frobenius norm. In particular, the unique*

positive approximant of a matrix $\hat{V}$ is:

$$(\hat{V}_1 + \hat{V}_2)/2 \tag{4.8}$$

where $\hat{V}_1 = (\hat{V} + \hat{V}')/2$ and $\hat{V}_2$ is the symmetric positive semi-definite matrix coming form the polar decomposition $\hat{V}_1 = A\hat{V}_2$ with $A'A = I$.

Table 4.15 show, as an example in which we use the SVF2 model to simulate the efficient price process, that, even though applying this technique to the estimated matrix obtained with the STS estimator allows to recover a matrix with the desired positivity, this seems to have a negligible impact on the MISE across the different scenarios for the noise, thus not being able to produce a final results that attains the same level of accuracy of the PDF estimator which, in turn, does not need any post manipulation of the output, whose impact on the asymptotic properties of the estimators is, in principle, unknown.

	PDF	STS	STSPSD
ζ	I.i.d. noise		
1	6.5931e-04	1.2245e-03	1.3075e-03
1.5	7.7088e-04	1.2458e-03	1.2445e-03
2	9.0900e-04	1.5532e-03	1.5397e-03
2.5	1.0126e-03	2.1353e-03	2.0746e-03
θ_η	Autocorrelated noise		
0.2	1.4802e-03	2.7254e-03	2.6464e-03
0.3	9.0024e-04	1.6698e-03	1.6373e-03
0.4	8.5164e-04%	1.5723e-03	1.5587e-03
ρ_η	Noise correlated with p		
-0.1	8.9725e-04	1.5249 e-03	1.5105e-03
-0.3	8.7500e-04	1.5096e-03	1.4992e-03
-0.5	8.5526e-04	1.5041e-03	1.4554e-03
$\bar{\sigma}_\eta$	Heteroskedastic noise		
3	1.3781e-03	2.9249e-03	2.7322e-03
3.5	1.8631e-03	4.0452e-03	4.0321e-03
4	2.5679e-03	5.8471e-03	5.1773e-03

Table 4.15: MISE for the PDF estimator, the STS estimator and its nearest psd matrix (TSPSD) computed according to Equation 4.8, under different noise scenarios, for the SVF2 model and $d = 20$.

4.3 Randomization

One additional aspect to consider when dealing with an estimator for high-frequency data, apart from the asymptotic properties such as the speed of convergence and the asymptotic error (analyzed for example for the Fourier spot volatility estimator in Chapter 2), the finite sample performance in terms, e.g., of mean square error (as done in greater detail in Chapter 3 for both integrated and spot volatility estimators) and the ability to produce some desirable characteristic of the output (that was is the main topic of this chapter), is the computational time needed to perform the estimation itself.

This issue is particularly relevant when high frequency data are employed, being the number of operation required by the implementation algorithm of each estimator a non-decreasing function of the number of observations n. The problem may be even more relevant when the estimation of spot volatility is involved since, on the contrary of what happens for integrated volatility, many different estimations are needed for each trading day, depending on the grid on which the volatility process is reconstructed. Finally, in the multivariate setting, when at each point in time we need to estimate not only one value, but a full volatility matrix, controlling the speed of execution of the estimation algorithms acquires the utmost importance, with an execution time that will increase also with the dimension of the matrix d. The aim of this section is to propose an alternative formulation of the PDF estimator that, while maintaining the property of producing positive semi-definite estimations, allows to increase the speed of execution of the estimation routine.

Form Remark 4.1.5 we know that the PDF estimator can be rewritten as in Equation 4.3. The idea of the alternative version, called Randomized PDF or RPDF, is to approximate the integral over the (gaussian) measure μ that appear in Equation 4.3 with an average over K random shifts Y sampled form a gaussian distribution, so that the RPDF estimator, for $t \in (0, 2\pi)$, reads as:

$$\hat{V}^{j,j'}_{N,M,K}(t) = \frac{2N+1}{2\pi K} \sum_{k=1}^{K} \sum_{i=0}^{n_j-1} \sum_{i'=0}^{n_{j'}-1} D_N(t - t^j_i + Y_k) D_N(t - t^{j'}_{i'} + Y_k) \delta^j_i(p) \delta^{j'}_{i'}(p), \tag{4.9}$$

with $K \geq 0$ and $Y_k \sim \mathcal{N}(0, 1/M)$. The estimator defined in this way is still symmetric and positive semi-definite, while the impact of the randomization on the execution time and the accuracy of the estimator is analyzed in the next section.

Remark 4.3.1. *The computational effort needed for the PDF estimator, similarly to what happens to the Fourier spot volatility estimator, is mainly due to the computation of all the discrete Fourier coefficients that appear in the convolution. As noted in Mancino et al. (2017), the burden of this operation may be reduced using Fast Fourier Transform algorithm or,in presence of irregularly spaced observations, Non-uniform Fast Fourier Transform, as done by Chang et al. (2020). The proposed randomization is an alternative, simple scheme to reduce the computational burden.*

Another aspect to consider, as noted in Chang et al. (2020), is the possibility of vectorization of the code, but a full vectorization of the the cross products appearing in the definition of the estimators may lead, however, to filling the memory of the machine for high values of n and, for the seek of the following analysis, we avoided full vectorization of the code.

4.3.1 Accuracy and speed of execution

We use the following simple stochastic volatility model to generate the trajectories of $d = 2$ logarithmic stock prices and their volatility:

$$dp^j(t) = \sigma^j(t) dW^j_t \qquad j = 1, 2$$

$$d\sigma^j(t) = \alpha \sigma^j(t) dZ^j_t \qquad j = 1, 2$$

with $\langle W^j, Z^j \rangle = \lambda^j$ for $j = 1, 2$ and $\langle W^1, W^2 \rangle = \rho$.

The simulations are performed using Euler scheme on an equally-spaced grid with step corresponding to $\frac{1}{2}$ a second, for on horizon of 23400 seconds, corresponding to 6.5 hours of market activity.

Irregularly spaced and asynchronous observations are obtained as before by drawing random times obtained by uniform Poisson process with intensity set to obtain an average of $\bar{n}$ observations for each trajectory. Sampling is performed according to the previous tick method. Moreover, simulations are carried out considering the following set of parameters:

$$(\alpha^1, \alpha^2, \lambda^1, \lambda^2, \rho) = (0.2, 0.25, -0.3, -0.7, 0.5)$$

while the initial price is fixed to $X_0 = \log(100)$ for both assets.

The main aim of our numerical analysis is to assess if, how much, and under which conditions the RPDF estimator may give an advantage, in terms of time required to compute the estimation,with respect to the implementation of the standard PDF estimator[1].

The performance are compared considering the minimum execution time over $K = 100$ replication, while looking at accuracy, we consider the relative mean integrated square error that, analogously to Equation 4.7 is, defined as in Section 4.2.4.

We start by focusing on the frequency N and the parameters of the design of the exercise that may influence the speed of execution, considering in particular:

- the parameter $\alpha = 1, 3/4, 2/3$ for the PDF and RPDF estimator, recalling that α is related to the choice of N via $N = \frac{n^\alpha}{2}$;
- the average number of observations $\bar{n} = 2340, 11700, 23400, 46800$, corresponding to an average of respectively one observation every 10, 2 and 1 and 0.5 seconds;
- the number of points $\#\tau = 78,390$ on which the trajectory of the spot volatility matrix is reconstructed, corresponding to an equally-spaced grid of width 5 and 1 minutes respectively.

The results are presented in Table 4.16, where we fixed $K = 250$, and show that indeed the RPDF is able to reduce the execution time in all the scenarios, and in particular for high values of n or when we need to chose a high value of α (see Section 4.2.2), being the speed of the RPDF not dependent on level of α.

Table 4.17 show how the situation changes when varying the value of K, considering in this case also the accuracy of the estimators in terms of RMISE. It appears that there is a clear trade-off for the RPDF estimator between speed and accuracy, with the latter converging to the one of the PDF as K increase, as expected, and the former linearly decreasing in K. The parameter K should then be chosen according to the needs of the statistician, depending on whether of the two aspects is considered more valuable.

Finally the parameters M, which is present in both the two estimators and is set as in the previous sections as $M = N^\beta$, as we can see from Equations 4.5 and 4.9, should no real influence in the execution speed the two estimators. This is confirmed by Table 4.18, that shows execution time and RMISE for the PDF and RPDF for a range of values of β. The results in terms of RMISE, for a small dimension $d = 2$, confirms the behavior underlined in Remark 4.2.1.

[1]Benchmarking is done using 2.30GHz clock speed Intel i7-11800H with 16GB RAM on Arcolinux kernel 6.2.7-arch1-1 with Python 3.9.12.

Execution Time												
	$\bar{n} = 2340$			$\bar{n} = 11700$			$\bar{n} = 23400$			$\bar{n} = 46800$		
α	1	3/4	2/3	1	3/4	2/3	1	3/4	2/3	1	3/4	2/3
Estimator	$\#\tau = 78$											
PDF	49.05	2.16	0.90	1060	14.45	4.30	3600+	35.46	9.01	3600+	95.99	20.42
RPDF	0.68	0.69	0.68	3.57	3.56	3.58	6.49	6.48	6.43	13.50	13.85	13.52
Estimator	$\#\tau = 390$											
PDF	229.6	10.36	4.44	3600+	67.85	19.99	3600+	167.11	40.48	3600+	426.03	90.85
RPDF	3.76	3.77	3.77	17.50	17.64	17.87	32.18	32.28	32.62	66.89	66.53	67.86

Table 4.16: Execution time (in seconds) of PDF and RPDF estimators in different simulation settings, with $K = 250$.

Accuracy and speed over K for RPDF												
K	50	100	150	200	250	300	350	400	450	500	750	1000
Time	1.35	2.72	3.94	5.30	6.48	7.82	9.25	10.70	12.06	13.28	19.08	23.57
RMISE	8.877	4.465	3.820	3.396	2.521	2.453	2.329	2.307	2.266	1.750	1.789	1.644
Reference for PDF												
Time	35.46											
RMISE	1.475											

Table 4.17: RMISE and execution time of RPDF for varying values of K. Values for RMISE are of order 10^{-2}.

Accuracy and speed over β for RPDF				
	PDF		RPDF	
β	Time	RMISE	Time	RMISE
1/2	35.25	2.414	6.46	3.516
2/3	35.39	1.888	6.48	3.216
3/4	35.35	1.663	6.49	3.108
3/4	35.35	1.316	6.50	2.654
5/6	35.30	1.475	6.48	2.521

Table 4.18: RMISE and execution time of PDF and RPDF for varying values of β. Values for RMISE are of order 10^{-2}.

Chapter 5

Conclusions

In this work we analyze the behavior of the Fourier spot volatility estimator introduced by Malliavin and Mancino (2002), focusing in particular on two of the main features of high-frequency data: microstructure noise and trade asynchronicity.

While realized-type estimators of the spot volatility typically require the manipulation of price observations and a bias-correction to become efficient and robust in the presence of microstructure noise, in this work we show that the Fourier estimator of the spot volatility can be efficient and robust at high-frequencies without the need for any manipulation of the original data or bias-correction. In fact, for this purpose it is sufficient to carefully reduce the value of the cutting frequency N, which controls the number of log-return coefficients to be used in the convolution formula that yields the Fourier coefficients of the volatility.

Specifically, we address both asymptotic and finite-sample efficiency. Firstly, we prove two rate-efficient Central Limit Theorems in the absence and in the presence of noise. Secondly, we derive a feasible procedure for selecting estimation parameters with high-frequency samples, based on the minimization of the integrated asymptotic-error variance. Further, we numerically show the robustness of this feasible procedure in more realistic scenarios, in which the hypotheses of the Central Limit Theorem are partially violated. Moreover we test this feasible procedure with tick-by-tick prices of three US stocks, characterized by different degrees of liquidity, and one US stock index, obtaining empirical evidence that supports its accuracy.

Additionally, we perform a comparative simulation study of the finite-sample performance of the Fourier estimator with that of alternative noise-robust spot volatility estimators and obtain numerical evidences that suggest that the use of the Fourier estimator may lead to a gain in finite-sample efficiency. This gain appears to be larger in scenarios characterized by more realistic microstructure noise patterns.

To extend this comparative analysis we focus our attention also on statistical models for the Limit Order Book, able to generate price trajectories in a way that closely mimic the functioning of electronic financial markets. Using this simulation setting to address the performance of a comprehensive selection fo volatility estimator we extend the work by Gatheral and Oomen (2010) in two directions. First, we use a more sophisticated LOB simulator compared to the ZI model, namely the QR model, which, by introducing correlations between the current state of the LOB and the intensities of order arrival, and thanks to a more complex re-initialization mechanism, is able to produce more realistic market microstructure dynamics, and we show with test statistics the superior ability

of the QR model to reproduce the noise pattern observed in empirical data.

For what concerns the comparison, we address not only integrated volatility estimators, but also spot volatility estimators. For what concerns integrated estimators, we find that the pre-averaging estimator by Jacod et al. (2009) appears to be favorable in terms of bias optimization. Instead, when looking at the minimization of the MSE, the situation is more nuanced, with the Fourier estimator obtaining the best average ranking across the three different price series considered without actually achieving the best ranking for any of the individual series. Specifically, the MSE is optimized by the unified estimator by Li et al. (2018) (in the case of mid- and micro-prices) and the alternation estimator by Large (2011) (in the case of trade-prices). As for the spot volatility, the Fourier estimator appears to yield the optimal accuracy both in terms of bias and MSE, outperforming the other estimators considered, confirming our previous findings in a more challenging simulation environment.

Moreover, as a financial application, we show that the careful choice of the spot volatility estimator may be relevant for optimal execution. Specifically, we investigate the impact of different spot volatility estimators on the prediction of the variance of the cost of a VWAP strategy and find that the use of the Fourier estimator, which gives the relative most accurate volatility estimates, leads also to a significant gain in predicting the cost variance.

After having analyzed the asymptotic and finite-sample properties of the Fourier spot volatility estimator in the univariate setting, we focus on the multivariate case, considering in particular the relevant case of asynchronous observation, being the synchronous case analogous to the univariate one. In particular, being the original estimator cursed by possible asymmetry in the estimation, we propose a modified version of the Fourier estimator, called PDF estimator, to overcame the difficulty of obtaining symmetric and positive semi-definite estimation of the spot variance-covariance matrix.

The proposed estimator is know to be consistent (Akahori et al. (2023)), and here we prove its positive semi-definiteness. A numerical study is carried out to evaluate the optimal choice of the frequencies N and M in a variety of settings. The optimal couple seems to be quite stable, and, as for the original spot Fourier estimator, if the data is heavily contaminated by noise N should be reduced. Based on those findings, a thorough simulation study is conducted to evaluate the accuracy of the PDF estimator and its actual ability in producing psd estimations. Comparing the results with the ones of two alternative estimators present in the literature, that are not proven to be positive semi-definite, we found out that the proposed PDF estimator usually outperforms the competitors in terms of mean square error, showing that it is possible to obtain desirable properties of the estimated spot volatility matrix without reducing the accuracy of the estimation. Moreover, the proposed estimator is found to be robust to a variety of specification of market microstructure noise, via an appropriate choice of the cutting frequency N. The robustness of our results are finally confirmed using alternative data generating processes.

As a last issue we also consider the possibility of reducing the computational time required by the PDF estimator, introducing an alternative randomization algorithm, that may significantly improve the execution time, at the cost of reducing the accuracy of the estimation, while maintaining the desired positivity of te volatility matrix. The trade-off between execution speed and accuracy is analyzed in detail, considering the impact of the different parameters of the two estimators in relation to the number of available observations.

References

Abergel, Frédéric, Marouane Anane, Anirban Chakraborti, Aymen Jedidi, and Ioane Muni Toke (2016). *Limit order books*. Cambridge University Press.

Aït-Sahalia, Yacine and Jean Jacod (2014). *High-frequency financial econometrics*. Princeton University Press.

Aït-Sahalia, Yacine, Per A Mykland, and Lan Zhang (2005). "How often to sample a continuous-time process in the presence of market microstructure noise". In: *The review of financial studies* 18.2, pp. 351–416.

Aït-Sahalia, Yacine and Dacheng Xiu (2019a). "A Hausman test for the presence of market microstructure noise in high frequency data". In: *Journal of Econometrics* 211.1, pp. 176–205.

— (2019b). "Principal component analysis of high-frequency data". In: *Journal of the American Statistical Association* 114.525, pp. 287–303.

Akahori, Jirô, Reika Kambara, and Gô Yûki (2023). "Consistency of PDF estimator". In: *arXiv preprint forthcoming*.

Almgren, Robert and Neil Chriss (2001). "Optimal execution of portfolio transactions". In: *Journal of Risk* 3, pp. 5–40.

Alòs, Elisa, Jorge A León, and Josep Vives (2007). "On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility". In: *Finance and Stochastics* 4.11, pp. 571–589.

Andersen, Torben G and Tim Bollerslev (1997). "Intraday periodicity and volatility persistence in financial markets". In: *Journal of empirical finance* 4.2-3, pp. 115–158.

Andersen, Torben G, Tim Bollerslev, and Francis X Diebold (2010). "Parametric and nonparametric volatility measurement". In: *Handbook of financial econometrics: Tools and techniques*. Elsevier, pp. 67–137.

Andersen, Torben G, Tim Bollerslev, Francis X Diebold, and Paul Labys (2003). "Modeling and forecasting realized volatility". In: *Econometrica* 71.2, pp. 579–625.

Andersen, Torben G, Dobrislav Dobrev, and Ernst Schaumburg (2012). "Jump-robust volatility estimation using nearest neighbor truncation". In: *Journal of Econometrics* 169.1, pp. 75–93.

Anderson, Theodore W and Donald A Darling (1952). "Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes". In: *The Annals of Mathematical Statistics* 23.2, pp. 193–212.

Bai, Zhidong and Jack W Silverstein (2010). *Spectral analysis of large dimensional random matrices*. Vol. 20. Springer.

Bandi, Federico M and Jeffrey R Russell (2006). "Separating microstructure noise from volatility". In: *Journal of Financial Economics* 79.3, pp. 655–692.

Bandi, Federico M and Jeffrey R Russell (2008). "Microstructure noise, realized variance, and optimal sampling". In: *The Review of Economic Studies* 75.2, pp. 339–369.

Barndorff-Nielsen, Ole E, Peter Reinhard Hansen, Asger Lunde, and Neil Shephard (2008). "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise". In: *Econometrica* 76.6, pp. 1481–1536.

— (2011). "Multivariate realised kernels: consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading". In: *Journal of Econometrics* 162.2, pp. 149–169.

Barucci, Emilio and Roberto Reno (2002). "On measuring volatility and the GARCH forecasting performance". In: *Journal of International Financial Markets, Institutions and Money* 12.3, pp. 183–200.

Bibinger, Markus, Nikolaus Hautsch, Peter Malec, and Markus Reiss (2019). "Estimating the spot covariation of asset prices - statistical theory and empirical evidence". In: *Journal of Business & Economic Statistics* 37.3, pp. 419–435.

Bouchaud, Jean-Philippe, Julius Bonart, Jonathan Donier, and Martin Gould (2018). *Trades, quotes and prices: financial markets under the microscope*. Cambridge University Press.

Bu, Ruijun, Degui Li, Oliver Linton, and Hanchao Wang (2022). "Nonparametric Estimation of Large Spot Volatility Matrices for High-Frequency Financial Data". In: *arXiv:2307.01348*.

Chang, Patrick, Etienne Pienaar, and Tim Gebbie (2020). "Malliavin–Mancino Estimators Implemented with Nonuniform Fast Fourier Transforms". In: *SIAM Journal on Scientific Computing* 42.6, pp. 1378–1403.

Chen, Dachuan, Per A Mykland, and Lan Zhang (2020). "The five trolls under the bridge: principal component analysis with asynchronous and noisy high-frequency data". In: *Journal of the American Statistical Association* 115.532, pp. 1960–1977.

Chernov, Mikhail, A Ronald Gallant, Eric Ghysels, and George Tauchen (2003). "Alternative models for stock price dynamics". In: *Journal of Econometrics* 116.1-2, pp. 225–257.

Christensen, Kim, Silja Kinnebrock, and Mark Podolskij (2010). "Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data". In: *Journal of econometrics* 159.1, pp. 116–133.

Christensen, Kim, Roel CA Oomen, and Mark Podolskij (2014). "Fact or friction: Jumps at ultra high frequency". In: *Journal of Financial Economics* 114.3, pp. 576–599.

Clement, Emmanuelle and Arnaud Gloter (2011). "Limit theorems in the Fourier transform method for the estimation of multivariate volatility". In: *Stochastic Processes and their Applications* 121.5, pp. 1097–1124.

Cuchiero, Christa and Josef Teichmann (2015). "Fourier transform methods for pathwise covariance estimation in the presence of jumps". In: *Stochastic Processes and their Applications* 125.1, pp. 116–160.

Cui, Liyuan, Yongmiao Hong, Yingxing Li, and Junhui Wang (2019). "Large-Dimensional Positive Definite Covariance Estimation for High Frequency Data via Low-rank and Sparse Matrix Decomposition". In: *SSRN 3414910*.

Delbaen, Freddy and Walter Schachermayer (1994). "A general version of the fundamental theorem of asset pricing". In: *Mathematische annalen* 300.1, pp. 463–520.

Eisler, Zoltán., Jean-Philippe Bouchaud, and Julien Kockelkoren (2012). "The price impact of order book events: Market orders, limit orders and cancellations". In: *Quantitative Finance* 12, pp. 1395–1419.

El Euch, Omar and Mathieu Rosenbaum (2019). "The characteristic function of rough Heston models". In: *Mathematical Finance* 29.1, pp. 3–38.

El Euch, Omar El and Mathieu Rosenbaum (2018). "Perfect hedging in rough Heston models". In: *The Annals of Applied Probability* 28.6, pp. 3813–3856.

Engle, Robert and Riccardo Colacito (2006). "Testing and valuing dynamic correlations for asset allocation". In: *Journal of Business & Economic Statistics* 24.2, pp. 238–253.

Fan, Jianqing and Yazhen Wang (2008). "Spot volatility estimation for high-frequency data". In: *Statistics and its Interface* 1.2, pp. 279–288.

Figueroa-López, José E and Cheng Li (2020). "Optimal kernel estimation of spot volatility of stochastic differential equations". In: *Stochastic Processes and their Applications* 130.8, pp. 4693–4720.

Figueroa-López, José E and Bei Wu (2022). "Kernel estimation of spot volatility with microstructure noise using pre-averaging". In: *Econometric Theory*, pp. 1–50.

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum (2018). "Volatility is rough". In: *Quantitative finance* 18.6, pp. 933–949.

Gatheral, Jim and Roel CA Oomen (2010). "Zero-intelligence realized variance estimation". In: *Finance and Stochastics* 14.2, pp. 249–283.

Guéant, Olivier (2016). *The Financial Mathematics of Market Liquidity: From optimal execution to market making*. Vol. 33. CRC Press.

Hansen, Peter R and Asger Lunde (2006). "Realized variance and market microstructure noise". In: *Journal of Business & Economic Statistics* 24.2, pp. 127–161.

Hansen, Peter R, Asger Lunde, and James M Nason (2011). "The model confidence set". In: *Econometrica* 79.2, pp. 453–497.

Hasbrouck, Joel (2007). *Empirical market microstructure: the institutions, economics, and econometrics of securities trading*. Oxford University Press.

Hayashi, Takaki and Nakahiro Yoshida (2005). "On covariance estimation of non-synchronously observed diffusion processes". In: *Bernoulli* 11.2, pp. 359–379.

Heston, Steven L (1993). "A closed-form solution for options with stochastic volatility with applications to bond and currency options". In: *The Review of Financial Studies* 6.2, pp. 327–343.

Higham, Nicholas J (1988). "Computing a nearest symmetric positive semidefinite matrix". In: *Linear algebra and its applications* 103, pp. 103–118.

Huang, Weibing, Charles-Albert Lehalle, and Mathieu Rosenbaum (2015). "Simulating and analyzing order book data: the queue-reactive model". In: *Journal of the American Statistical Association* 110.509, pp. 107–122.

Huang, Xin and George Tauchen (2005). "The relative contribution of jumps to total price variance". In: *Journal of financial econometrics* 3.4, pp. 456–499.

Jacod, Jean (1997). "On continuous conditional Gaussian martingales and stable convergence in law". In: *Séminaire de Probabilités XXXI* 31, pp. 232–246.

— (2019). "Estimation of volatility in a high-frequency setting: a short review". In: *Decisions in Economics and Finance* 42, pp. 351–385.

Jacod, Jean, Yingying Li, Per A Mykland, Mark Podolskij, and Mathias Vetter (2009). "Microstructure noise in the continuous case: the pre-averaging approach". In: *Stochastic Processes and their Applications* 119.7, pp. 2249–2276.

Jacod, Jean, Yingying Li, and Xinghua Zheng (2017). "Statistical properties of microstructure noise". In: *Econometrica* 85.4, pp. 1133–1174.

Jacod, Jean and Philip Protter (1998). "Asymptotic error distributions for the Euler method for stochastic differential equations". In: *Annals of Probability*, pp. 267–307.

Jarque, Carlos M and Anil K Bera (1987). "A test for normality of observations and regression residuals". In: *International Statistical Review/Revue Internationale de Statistique*, pp. 163–172.

Jing, Bing-Yi, Zhi Liu, and Xin-Bing Kong (2014). "On the estimation of integrated volatility with jumps and microstructure noise". In: *Journal of Business & Economic Statistics* 32.3, pp. 457–467.

Kalnina, Ilze and Oliver Linton (2008). "Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error". In: *Journal of econometrics* 147.1, pp. 47–59.

Kalnina, Ilze and Dacheng Xiu (2017). "Nonparametric estimation of the leverage effect: a trade-off between robustness and efficiency". In: *Journal of the American Statistical Association* 112.517, pp. 384–396.

Kristensen, Dennis (2010). "Nonparametric filtering of the realized spot volatility: a kernel-based approach". In: *Econometric Theory* 26.1, pp. 60–93.

Large, Jeremy (2011). "Estimating quadratic variation when quoted prices change by a constant increment". In: *Journal of Econometrics* 160.1, pp. 2–11.

Lee, Suzanne S and Per A Mykland (2012). "Jumps in equilibrium prices and market microstructure noise". In: *Journal of Econometrics* 168.2, pp. 396–406.

Li, Jia, Dishen Wang, and Qiushi Zhang (2022). "Reading the Candlesticks: An OK Estimator for Volatility". In: *The Review of Economics and Statistics*, pp. 1–45.

Li, Yingying, Guangying Liu, and Zhiyuan Zhang (2021). "Volatility of volatility: Estimation and tests based on noisy high frequency data with jumps". In: *Journal of Econometrics*.

Li, Yingying, Zhiyuan Zhang, and Yichu Li (2018). "A unified approach to volatility estimation in the presence of both rounding and random market microstructure noise". In: *Journal of Econometrics* 203.2, pp. 187–222.

Li, Z Merrick and Oliver B Linton (2022). "Robust Estimation of Integrated and Spot Volatility". In: *SSRN 4184933*.

Lillo, Fabrizio and J Doynne Farmer (2004). "The long memory of the efficient market". In: *Studies in nonlinear dynamics & econometrics* 8, p. 3.

Liu, Nien-Lin and Hoang-Long Ngo (2017). "Approximation of eigenvalues of spot cross volatility matrix with a view toward principal component analysis". In: *Japan Journal of Industrial and Applied Mathematics* 34.3, pp. 747–761.

Malliavin, Paul and Maria Elvira Mancino (2002). "Fourier series method for measurement of multivariate volatilities". In: *Finance and Stochastics* 6.1, pp. 49–61.

— (2009). "A Fourier transform method for nonparametric estimation of multivariate volatility". In: *The Annals of Statistics* 37.4, pp. 1983–2010.

Mancini, Cecilia, Vanessa Mattiussi, and Roberto Renò (2015). "Spot volatility estimation using delta sequences". In: *Finance and Stochastics* 19, pp. 261–293.

Mancino, Maria Elvira and Maria Cristina Recchioni (2015). "Fourier spot volatility estimator: asymptotic normality and efficiency with liquid and illiquid high-frequency data". In: *PLOS ONE* 10.9, e0139041.

Mancino, Maria Elvira, Maria Cristina Recchioni, and Simona Sanfelici (2017). *Fourier-Malliavin volatility estimation: Theory and practice*. Springer.

Mancino, Maria Elvira and Simona Sanfelici (2008). "Robustness of Fourier estimator of integrated volatility in the presence of microstructure noise". In: *Computational Statistics & Data Analysis* 52.6, pp. 2966–2989.

— (2011). "Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise". In: *Journal of financial econometrics* 9.2, pp. 367–408.

— (2012). "Estimation of quarticity with high-frequency data". In: *Quantitative Finance* 12.4, pp. 607–622.

Mancino, Maria Elvira and Giacomo Toscano (2022). "Rate-efficient asymptotic normality for the Fourier estimator of the leverage process". In: *Statistics and Its Interface* 15.1, pp. 73–89.

Mariotti, Tommaso, Fabrizio Lillo, and Giacomo Toscano (2023). "From zero-intelligence to queue-reactive: limit-order-book modeling for high-frequency volatility estimation and optimal execution". In: *Quantitative Finance* 23.3, pp. 367–388.

Mykland, Per A, Lan Zhang, and Dachuan Chen (2019). "The algebra of two scales estimation, and the S-TSRV: high-frequency estimation that is robust to sampling times". In: *Journal of Econometrics* 208.1, pp. 101–119.

Nielsen, Morten Ørregaard and Per Frederiksen (2008). "Finite-sample accuracy and choice of sampling frequency in integrated volatility estimation". In: *Journal of Empirical Finance* 15.2, pp. 265–286.

Ogawa, Shigeyoshi (2008). "Real-time scheme for the volatility estimation in the presence of microstructure noise". In: *Monte Carlo Methods and Applications* 14.4, pp. 331–342.

Ogawa, Shigeyoshi and Simona Sanfelici (2011). "An improved two-step regularization scheme for spot volatility estimation". In: *Economic Notes* 40.3, pp. 107–134.

Park, Sujin, Seok Young Hong, and Oliver Linton (2016). "Estimating the quadratic covariation matrix for asynchronously observed high-frequency stock returns corrupted by additive measurement error". In: *Journal of Econometrics* 191.2, pp. 325–347.

Patton, Andrew J (2011). "Volatility forecast comparison using imperfect volatility proxies". In: *Journal of Econometrics* 160.1, pp. 246–256.

Richard, Alexandre, Xiaolu Tan, and Fan Yang (2023). "On the discrete-time simulation of the rough Heston model". In: *SIAM Journal on Financial Mathematics* 14.1, pp. 223–249.

Robert, Christian Y and Mathieu Rosenbaum (2011). "A new approach for the dynamics of ultra-high-frequency data: The model with uncertainty zones". In: *Journal of Financial Econometrics* 9.2, pp. 344–366.

Roll, Richard (1984). "A simple implicit measure of the effective bid-ask spread in an efficient market". In: *The Journal of finance* 39.4, pp. 1127–1139.

Ruder, Sebastian (2017). "An overview of gradient descent optimization algorithms". In: *arXiv:1609.04747*.

Sanfelici, Simona, Imma Valentina Curato, and Maria Elvira Mancino (2015). "High-frequency volatility of volatility estimation free from spot volatility estimates". In: *Quantitative Finance* 15.8, pp. 1331–1345.

Smith, Eric, J Doyne Farmer, László Gillemot, and Supriya Krishnamurthy (2003). "Statistical theory of the continuous double auction". In: *Quantitative finance* 3.6, p. 481.

Toscano, Giacomo, Giulia Livieri, Maria Elvira Mancino, and Stefano Marmi (2022). "Volatility of Volatility Estimation: central Limit Theorems for the Fourier Transform Estimator and Empirical Study of the Daily Time Series Stylized Facts". In: *Journal of Financial Econometrics*.

Veraart, Almut ED and Luitgard AM Veraart (2012). "Stochastic volatility and stochastic leverage". In: *Annals of Finance* 8, pp. 205–233.

Vortelinos, Dimitrios I (2014). "Optimally sampled realized range-based volatility estimators". In: *Research in International Business and Finance* 30, pp. 34–50.

Zhang, Lan (2006). "Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach". In: *Bernoulli* 12.6, pp. 1019–1043.

Zhang, Lan, Per A Mykland, and Yacine Aït-Sahalia (2005). "A tale of two time scales: Determining integrated volatility with noisy high-frequency data". In: *Journal of the American Statistical Association* 100.472, pp. 1394–1411.

Zhou, Bin (1996). "High-frequency data and volatility in foreign exchange rates". In: *Journal of Business & Economic Statistics* 14.1, pp. 45–52.

Zu, Yang and H Peter Boswijk (2014). "Estimating spot volatility with high-frequency financial data". In: *Journal of Econometrics* 181.2, pp. 117–135.