Dipartimento di Economia
Facoltà di Economia
Università degli Studi di Parma
Via Kennedy, 6
43100 PARMA

An improved two-step regularization scheme for spot volatility estimation

Shigeyoshi Ogawa
Dept. Mathematics, Ritsumeikan University, Japan
ogawa-s@se.ritsumei.ac.jp

Simona Sanfelici
Dept. Economics, University of Parma, Italy
simona.sanfelici@unipr.it

WP 02/2008

Serie: Matematica

Dicembre 2008

Abstract

We are concerned with the problem of parameter estimation in Finance, namely the estimation of the spot volatility in the presence of the so-called microstructure noise. In [16] a scheme based on the technique of multi-step regularization was presented. It was shown that this scheme can work in a real-time manner. However, the main drawback of this scheme is that it needs a lot of observation data. The aim of the present paper is to introduce an improvement of the scheme such that the modified estimator can work more efficiently and with a data set of smaller size. The technical aspects of implementation of the scheme and its performance on simulated data are analyzed. The proposed scheme is tested against other estimators, namely a realized volatility type estimator, the Fourier estimator and two kernel estimators.

JEL: G10, C14, C22.
Keywords: Spot volatility; Nonparametric estimation; Multi-step regularization; Microstructure.

1 Introduction

Let $p(t),\ (0 \le t \le T)$ be a stochastic process representing the evolution in time of the log-price of an asset, which in the theory of finance is supposed to be a real or an $\mathbb{R}^d$ valued Itô process generated by the following mechanism;

$$dp(t) = a(t,\omega)dt + b(t,\omega)dW(t), \quad t \in [0,T] \tag{1}$$

where $W(t,\omega),\ (t \ge 0, \omega \in \Omega)$ is the Brownian motion process defined on a probability space $(\Omega, \mathcal{F}, P)$ and $a(\cdot), b(\cdot)$ are measurable functions, adapted to the natural filtration $\{\mathcal{F}_t,\ t \ge 0\}$ with $\mathcal{F}_t = \sigma\{W_s; 0 \le s \le t\} \vee \sigma\{p(0)\}$. Here we suppose that the random variable $p(0)$ is independent of the $W(\cdot)$. For simplicity we suppose that the following conditions hold

$$A := \sup_t \sqrt{E[|a(t)|^2]} < \infty \quad \text{and} \quad B := \sup_{t\in[0,T]} E[b^4(t)] < \infty. \tag{2}$$

This is a standard model for the price process but the coefficients $a(\cdot), b(\cdot)$ are not known in real situation. We add one more hypothesis on the regularity of these coefficients as follows

Hypothesis (H) The functions $a(t,\omega), b(t,\omega)$ are Hölder continuous in $t \in T$ of order $\alpha \in (0,1)$, in $L^2(\Omega)$ sense; namely, for some constants L_A, L_B the following inequalities hold

$$E[|b(t)-b(s)|^2] \le L_B^2|t-s|^{2\alpha}, \quad E[|a(t)-a(s)|^2] \le L_A^2|t-s|^{2\alpha}, \quad \forall t,s \ge 0.$$

We are to study the problem of estimating the value $b^2(t,\omega_0)$ at any t, called the *spot volatility* in the theory of Finance, starting from a finite set of observations $\{X(t_k,\omega_0)\}$ of a single trajectory $\omega_0 \in \Omega$.

The recent availability of high frequency financial data has motivated a growing literature in the field of nonparametric estimation of volatility, see e.g [2], [19] for reviews. In particular, the so-called *Realized Volatility* measure has received considerable attention and is widely used in the empirical Finance literature. More general and alternative approaches to volatility estimation are pursued in [7, 3, 11, 8, 12].

However, when high frequency data are used, market microstructure effects may spoil volatility estimates at a very large extent. Indeed, the efficiency of all the methodologies proposed for accurately estimating the volatility builds on the observability of the true price process, while observed asset prices are contaminated by market microstructure effects, such as price discreteness, separate trading prices for buyers and sellers and other contaminations; as a consequence observed asset prices diverge from their efficient values. The study of the implications of market microstructure noise on volatility measurement has largely focused on the so-called integrated (e.g. daily) volatility and several methods have been proposed to correct the realized volatility estimator for the effect of market microstructure noise, in order to obtain unbiased estimators of the true integrated volatility [20, 1, 4, 5, 21, 6, 9, 14].

A very common way to model microstructure effects is assuming that the true value of the process $p(t)$ can not be observed but with an additive noise $Z(t)$, namely what we can observe is the values of the following contaminated process

$$X(t) = p(t) + Z(t). \tag{3}$$

The noise $Z(t)$ is supposed to be of very high frequency as featured in the following condition (Z) and is called the *microstructure noise*.

Hypothesis (Z) Z has the following properties.

(a) $E[Z] = 0, \quad \sup_t E[Z_t^2] = C_z^2 < \infty \quad (C_z = \text{unknown positive constant})$

(b) For an arbitrary point set $\{t_k \in [0,T],\ 1 \le k \le n\}$ of arbitrary size n, the random variables $\{Z(t_k)\}$ are independent.

A common assumption in the literature supposes that the noise process Z is independent of the price process $p(t)$. Nevertheless, we remark that in our discussion we do not need this assumption.

In this context, we are concerned with the problem of estimating the spot volatility value $b^2(t_k)$ at any prefixed time $\{t_k\} \in [0,T]$. Our aim is to design an estimator that can work in a real-time manner, namely an estimator that can estimate the spot volatility at each time almost immediately on receiving the observed data around that time. A preliminary form for such an estimator was proposed in [16] and was called *real-time estimator*. Given the observed log-price process $X(t)$, a first regularization is performed to obtain a smoothed process $\bar{X}(t)$ which is then used to recover spot volatility. The author showed that already in its preliminary form the estimator works very fast in computation time with sufficient precision. The basic idea of this scheme is in the adoption of multi-step regularizations which make use of non-overlapping windows of data and by this reason it requires numerous observations around each point t_k where the estimation is computed. This yields an unmotivated limitation on the number of estimation points when only a finite sample of data is available. In this paper we propose a modification of this scheme which allows to use overlapping blocks of data for the regularization. This feature makes the scheme much simpler to implement since it allows to use the same observed data

repeatedly to estimate volatility at different estimation times, thus reducing the size of necessary observation data.

We conclude this section with an important comment on our real-time scheme. The literature on nonparametric estimation of volatility is nowadays well established. A recently proposed method called *kernel estimator* ([11])

$$\hat{\sigma}^2(t) = \lim_{h\to 0} \sum_{i=1}^{N_T} \frac{1}{h} K\left(\frac{t_{i-1}-t}{h}\right) [X(t_i) - X(t_{i-1})]^2$$

is believed to cover a wide class of estimators for the spot volatility. However, this is not the case for our estimator, as it can be seen by the following arguments. First we should notice that the essential idea of our estimator is that we apply the regularization procedure twice at two different stages, while the kernel type estimator given above is performed on the data process $X(t)$ and not on the smoothed one $\bar{X}(t)$. Therefore, looking at the definition of the real-time estimator

$$\frac{G(M)^{-1}}{L} \sum_{l=1}^{L} \frac{(\Delta_{k+l-1}\bar{X})^2}{\Delta},$$

we can say that our scheme is a kernel-type estimator over a regularized set of data and this is of crucial importance to deal with microstructure effects. In particular, microstructure noise is smoothed out by the inclusion in our estimator of a weighted sum of products of lagged intra-day returns. Moreover, the specified kernel was found by an optimization argument and allowed us to construct an easily implementable estimator having a very simple form and working in a real-time manner. Finally we notice that, generally speaking, the validity of the kernel method is assured by some kind of continuity of the function to be estimated. Our theoretical analysis make this point much clear, since we clearly specify the conditions under which our scheme provides good estimates of $b^2(t)$. More precisely, this is assured by the hypothesis that $b(t)$ is Hölder continuous in $L^2(\Omega)$-sense.

The remaining part of the paper is structured as follows. Section 2 briefly describes the first version of the real-time estimator and states its error estimate. Section 3 is divided in three subsections which introduce notations, present the new form of the real-time estimator and prove the new bias estimate. In Section 4, we discuss the implementation of the proposed estimator and of other well known spot volatility estimators and test their performance on simulated data. Section 5 concludes.

2 Old scheme based on multi-step regularization

In this paragraph we briefly resume the old scheme based on the idea of *multi-step regularization*. We will denote by $\{t_k,\ 0 \le k \le N\}$ the set of N points in $[0,T]$ such that $t_k = k\Delta,\ \Delta = T/N$. At each point t_k we want to estimate the value of the spot volatility $b^2(t_k,\omega)$. To smear out the influence of the noise Z we need a regularization procedure and for that purpose we introduce around each estimation point t_k a set of finer partition points $\{t_k^i\},\ -M \le i \le M$ given by

$$t_k^i := t_k + \frac{i}{2M}\Delta. \tag{4}$$

To a data process $Y(= X, p, \text{and } Z)$ the regularization procedure is applied in the following manner to obtain the regularized process $\bar{Y}(t)$

$$\bar{Y}(t_k) := \frac{1}{2M+1} \sum_{i=-M}^{M} Y(t_k^i), \quad \text{where } Y = X, p, Z. \tag{5}$$

Remark 2.1 *We may also think about regularizations of different forms as follows*

$$\bar{Y}^{(-)}(t_k) := \frac{1}{M} \sum_{i=-M+1}^{0} Y(t_k^i)$$

or

$$\bar{Y}^{(+)}(t_k) := \frac{1}{M} \sum_{i=0}^{M-1} Y(t_k^i).$$

The former variant $\bar{Y}^{(-)}(t_k)$ is important since the estimator of this form uses only trades available up to time t_k, hence it is adapted to the observed process. But for the sake of simplicity we will mainly discuss the regularization given in the definition (5) above.

Thus, after the first regularization procedure we have

$$\bar{X}(t_k) = \bar{p}(t_k) + \bar{Z}(t_k) \quad 1 \le k \le N-1. \tag{6}$$

The aim of the procedure is to calm down the influence from the noise process as we see in the following

$$E[\bar{Z}(t_k)] = 0, \quad Var(\bar{Z}(t_k)) = \frac{1}{(2M+1)^2} \sum_{|i| \le M} Var(Z(t_k^i)) \le \frac{C_z^2}{2M+1}.$$

Due to the fact that the noise process $Z(t)$ exhibits a very high frequency movement compared to the signal process $p(t)$, the simple scheme introduced in [17] and based on the coupling of two techniques of quadratic variation and a further regularization procedure is expected to apply to the data process $\bar{X}$ given in (6). This consideration is the basis for the construction of the estimator $\hat{b}^2(t_k)$ defined in [16] as

Definition 2.2 (Estimator)

$$\hat{b}^2(t_k) = \frac{G(M)^{-1}}{2L+1} \sum_{l=0}^{2L} \frac{(\Delta_{k+l-L}\bar{X})^2}{\Delta} \tag{7}$$

where

$$G(M) = \frac{1}{(2M+1)^2} \sum_{i,j}^{2M} \left(1 - \frac{|i-j|}{2M}\right) = \frac{8M^2+6M+1}{3(2M+1)^2}$$

and $\Delta_k \bar{X} = \bar{X}(t_k) - \bar{X}(t_{k-1})$.

Remark 2.3 *The estimator which is adapted to the observed process should be given in the following form*

$$\hat{b}^2(t_k) = \frac{G(M)^{-1}}{2L+1} \sum_{i=0}^{2L} \frac{(\Delta_{k+i-2L}\bar{X}^{(-)})^2}{\Delta}. \tag{8}$$

As for the efficiency of this estimator (7) we have the following result

Theorem 2.4 ([16]) *For some positive constants C_1, C_2, C_3, C_4, independent of the parameters L, M, N the following estimate holds at every point $t_k = k\Delta$*

$$E\left[|\hat{b}^2(t_k) - b^2(t_k)|\right] \leq C_1 \left(\frac{L}{N}\right)^{\alpha} + C_2 \frac{1}{N^{\alpha \wedge 1/2}} + C_3 \sqrt{\frac{N}{M}} + C_4 \frac{1}{\sqrt{L}}.$$

By checking the proof given in [16] we see the validity of the inequality with appropriately chosen constants as follows

$$\begin{aligned} C_1 &= 2B^{1/4} L_B^{1/2} T^{\alpha} \\ C_2 &= 3C_\delta = 6\{\sqrt{B^{1/2}(\tfrac{L_B^2}{2\alpha+1} + 4C_z^2)} + 1\} \\ C_3 &= 3(\sqrt{C_p C_z} + C_z) \geq 3\tfrac{A^2}{T}(\sqrt{B} + 1/2) \\ C_4 &= 2\sqrt{3B}. \end{aligned} \tag{9}$$

where C_p is a constant such that $E[(\Delta_k \bar{p})^2/\Delta] \leq C_p$.

Already in this preliminary form the estimator works very fast in computation time with sufficient precision. However, the first regularization makes use of non-overlapping windows of data and by this reason it requires numerous observations around each point t_k where the estimation is computed. In the next section, we propose a modification of this scheme which allows to use overlapping blocks of data for the regularization. This feature makes the scheme much simpler to implement since it allows to use the same observed data repeatedly to estimate volatility at different estimation times.

3 The new scheme

By Theorem 2.4, we see that an appropriate tuning of the parameters L, M, N can assure a good behavior of our estimator. However, we also notice as a weak point of our estimator that it needs a lot of data to compute the estimation. For the computation of estimates at all N points $\{t_k\}$ we need $2MN$ observation data, or MN for the case of causal estimator (8). Looking at the error terms that appear on the right hand side of the estimation inequality given in Theorem 2.4, we know that the three parameters L, M, N should be chosen in the following way

$$1 << L << N << M.$$

Thus the total number of data $2MN$ should become very large, compared to the number of estimation points N. Hence we have great interest in finding some technique such that the total number of observation points can be smaller and not depending on N and on the bandwidth M according to the relations above. This can be done by allowing, in the first regularization procedure, the multiple use of the data in overlapping intervals. We are going to explain this idea and how the new scheme would work. In what follows we still suppose that the assumptions (H) and (Z) hold.

3.1 Point sets and parameters

For the convenience of expressing the formula of the estimator we need to introduce one more parameter, N_T standing for the total number of the observed data, while N stands for the total number of estimating points which we denote by $\{t'_k\}$. According to this change in the system of parameters, we will resume their meaning in the following

- N_T: we will denote by $\{t_k, 0 \le k \le N_T\}$ the set of points where the process $X(t)$ is observed. For simplicity we suppose that the points t_k are equally spaced; therefore $t_k = k\frac{T}{N_T} = k\Delta$, with the convention $\Delta = \frac{T}{N_T}$.
- D and N: the estimation of the value $b^2(t)$ is computed only at the points $\{t'_k\} \subset \{t_k\}$, where $t'_k = t_{kD}$ for some $D \in \mathbb{N}$ and $k = 1, 2, \ldots, N$. Clearly, $ND = N_T$.
- M : at each t'_k the first regularization $\bar{X}(t'_k)$ is formed by taking M data at points $\{t_{kD-i+1},\ 1 \le i \le M\}$, that is the first regularization is computed by the causal form.
- ρ : the ratio between two parameters M, D is denoted by the symbol $\rho = \dfrac{M}{D}$.
- L: the number of points $\{t'_{k+l},\ 1 \le l \le L\}$ used for the second regularization at time t'_k (see Section 3.3).

3.2 A new scheme for estimation

We introduce the first regularization procedure in the following way

$$\bar{X}(t'_k) = \frac{1}{M}\sum_{i=1}^{M} X(t_{kD-i+1}), \quad 1 \le k \le N, \tag{10}$$

where $N = N_T/D$.

The regularization of other processes $p(t), Z(t)$ at $t = t'_k = t_{kD}$ is defined in the same way. Thus we have

$$\bar{X}(t'_k) = \bar{p}(t'_k) + \bar{Z}(t'_k).$$

When $M = D$ the procedure is just the same as the old one. Hence we are interested in this article only in the case $D < M$.

Notice that $E[(\bar{Z})^2] = \dfrac{C_z}{M}$ by hypothesis (Z). Now taking the increment over $[t'_{(k-1)}, t'_k]$ of each member in the relation above, we have

$$\Delta'_k\bar{X} = \Delta'_k\bar{p} + \Delta'_k\bar{Z}$$

where,

$$\Delta'_k\bar{Y} = \bar{Y}(t'_k) - \bar{Y}(t'_{k-1}) = \bar{Y}(t_{kD}) - \bar{Y}(t_{(k-1)D}), \quad (Y = X, p, Z).$$

Given this, we are to present the new scheme of our estimator in the following

Definition 3.1 *For each t'_k $(1 \le k \le N)$ we set*

$$\hat{b}^2(t'_k) = \frac{G^{-1}(M)}{L}\sum_{l=1}^{L}\frac{(\Delta'_{k+l}\bar{X})^2}{\Delta'}, \tag{11}$$

where, $\Delta' = T/N = D\Delta$ and $G(M) = \dfrac{3MD - D^2 + 1}{3M^2}$.

We remark that the new normalizing constant $G(M)$ is derived from the following equality

$$G(M) = \frac{1}{DM^2}\sum_{ij=1}^{D}\{M - |i-j|\}. \tag{12}$$

As for the precision level of this estimator, we have the following result

Theorem 3.2 *Under the condition that $D < M$, the following error estimate holds*

$$E|b^2(t'_k) - \hat{b}^2(t'_k)| \leq$$
$$C_1\max\{\rho, |L-\rho|\}\sqrt[\alpha]{\frac{M}{N_T}} + C_2\frac{1}{3\rho\sqrt{\rho}}\sqrt{\frac{1}{L}} + C_3\frac{3}{3\rho^2-1}(\frac{D}{N_T})^{\alpha\wedge 1/2} + C_4\sqrt{\frac{N_T}{M^2}},$$

where the C_1, C_2, C_3, C_4 are positive constants independent of all control parameters N_T, N, D, M and L.

3.3 Proof of Theorem 3.2

For the verification of Theorem we need some lemmas. First notice that $t'_k = t_{kD}$ and that the increment over the interval $[t'_{k-1}, t'_k]$ of the regularized processes $\bar{Y}(= \bar{X}, \bar{p}, \bar{Z})$ can be expressed in the following way

$$\begin{aligned}\Delta'_k\bar{Y} : \ &= \bar{Y}(t'_k) - \bar{Y}(t'_{k-1}) = \frac{1}{M}\sum_{i=1}^{M}\{Y(t_{kD-i+1}) - Y(t_{kD-D-i+1})\} \\ &= \frac{1}{M}\sum_{i=1}^{D}\{Y(t_{kD-i+1}) - Y(t_{kD-i+1-M})\}.\end{aligned} \tag{13}$$

Thus by Hypothesis (Z) we have

$$E[(\Delta'_k\bar{Z})^2] = \frac{2DC_z}{M^2}.$$

From equation (13) we see $\Delta'_k\bar{p} = \frac{1}{M}\sum_{i=1}^{D}\Delta^M_{k,i}p$, where

$$\Delta^M_{k,i}p := p(t_{kD-i+1}) - p(t_{kD-i+1-M}) = \int_{t_{kD-i+1-M}}^{t_{kD-i+1}}\{a(s)ds + b(s)dW_s\}. \tag{14}$$

Set

$$\epsilon_{k,i} = \int_{t_{kD-i+1-M}}^{t_{kD-i+1}}[\{a(s) - a(t_{kd-i+1-M})\}ds + \{b(s) - b(t_{kD-i+1-M})\}dW_s],$$

then we see from equation (14) that

$$\Delta^M_{k,i}\bar{p} = a(t_{kD-i+1-M})M\Delta + b(t_{kD-i+1-M})\Delta^M_{k,i}W + \epsilon_{k,i} \tag{15}$$

where $\Delta^M_{k,i}W = W(t_{kD-i+1}) - W(t_{kD-i+1-M})$ is a random variable following the normal law $N(0, M\Delta)$. Moreover following the same argument given in the previous article ([16]) we readily obtain the following estimate

Lemma 3.3 *We have*

$$E[|\epsilon_{k,i}|^2] \leq \frac{(M\Delta)^{2\alpha+1}}{2\alpha+1}\{M\Delta L_A^2 + L_B^2\},$$

which yields the following estimate when $M\Delta < 1$

$$E[|\epsilon_{k,i}|^2] \leq \frac{L_A^2 + L_B^2}{2\alpha+1}(\frac{M}{D})^{2\alpha+1}(\Delta')^{2\alpha+1}.$$

Proof Since ϵ is just the one-step error of the Euler-Maruyama scheme for the discrete approximation of the SDE, taking the hypothesis (H) into account and following a usual discussion in the theory of numerical solution to SDE, we get the conclusion (see for example [18]). *q.e.d.*

By definition of the process $X(t)$ we have

$$\begin{aligned}(\Delta'_k\bar{X})^2 &= \frac{1}{M^2}\sum_{i,j=1}^{D}\Delta_{k,i}^M X\Delta_{k,j}^M X \\ &= \frac{1}{M^2}\sum_{j,j=1}^{D}\Delta_{k,i}^M p\Delta_{k.j}^M p + (\Delta'_k\bar{Z})^2 + 2(\Delta'_k\bar{p})(\Delta'_k\bar{Z}).\end{aligned} \tag{16}$$

To evaluate the difference between $(\Delta'_k\bar{X})^2$ and $b(t'_k)^2\Delta'$ we first analyse the quantity $\Delta_{k,i}^M p\Delta_{k,j}^M p$. From equation (15) we see

$$\begin{aligned}\Delta_{ki}^M p\Delta_{kj}^M p &= \{a(t_{kD-i+1-M})(M\Delta) + b(t_{kD-i+1-M})\Delta_{k,i}^M W + \epsilon_{k,i}\} \\ &\times\{a(t_{kD-j+1-M})(M\Delta) + b(t_{kD-j+1-M})\Delta_{k,j}^M W + \epsilon_{k,j}\}, \quad (1 \leq i,j \leq D).\end{aligned}$$

Set

$$\delta_k^{ij} = \Delta_{ki}^M p\Delta_{kj}^M p - b(t_{kD-i+1-M})b(t_{kD-j+1-M})\Delta_{k,i}^M W\Delta_{k,j}^M W, \tag{17}$$

then by Lemma 3.3 we easily get the following estimate

Lemma 3.4 *For any* $1 \leq i,j \leq D$ *and any* $1 \leq k \leq N$ *the following inequality holds*

$$E|\delta_M^{ij}| = O((M\Delta)^{(\alpha\wedge 1/2)+1} = O((\Delta')^{(\alpha\wedge 1/2)+1}.$$

A simple calculus yields the following equality

$$\Delta_{ki}^M W\Delta_{kj}^M W = \{M - |i-j|\}\Delta + \theta_k^{ij} \quad (i \geq j), \tag{18}$$

where

$$\begin{aligned}\theta_k^{ij} &= 2\int_{t_{kD-j+1-M}}^{t_{kD-i+1}}\{W(s) - W(t_{kD-j+1-M})\}dW_s \\ &+\{W(t_{kD-j+1}) - W(t_{kD-i+1})\}\{W(t_{kD-i+1}) - W(t_{kD-j+1-M})\} \\ &+\{W(t_{kD-i+1}) - W(t_{kD-j+1-M})\}\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}.\end{aligned}$$

We notice that θ_k^{ij} has following properties.

Lemma 3.5 *For any* $1 \leq i,j,i',j' \leq D$ *and any* $k,l(k \neq l)$ *it holds*

(a) $E[\theta_k^{ij}] = 0,\ E[\theta_k^{ij}\theta_l^{i'j'}] = 0$

(b) $E[(\theta_k^{ij})^2] = 2M(M-|i-j|)\Delta^2 = 2\frac{M}{D^2}(\Delta')^2$

(c) $E[(\theta_k^{ij})^4] \le 3078(M\Delta)^4$.

Proof Properties (a),(b) being almost immediate by simple calculus, we will show the property (c). We treat the case $i \ge j$.
Set $\theta_k^{ij} = \Theta_1 + \Theta_2 + \Theta_3$, where

$$\begin{aligned}
\Theta_1 &= 2\int_{t_{kD-j+1-M}}^{t_{kD-i+1}} \{W(s) - W(t_{kD-j+1-M})\}dW_s \\
\Theta_2 &= \{W(t_{kD-j+1}) - W(t_{kD-i+1})\}\{W(t_{kD-i+1}) - W(t_{kD-j+1-M})\} \\
\Theta_3 &= \{W(t_{kD-i+1}) - W(t_{kD-j+1-M})\}\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}.
\end{aligned}$$

For each of these 3 terms we have the following estimates

- $E[\Theta_1^4] = 96E\{\int_{t_{kD-j+1-M}}^{t_{kD-i+1}}(W_s - W(t_{kD-j+1-M}))^2 ds\}^2$
  $\le 96(M-|i-j|)\Delta E\{\int_{t_{kD-j+1-M}}^{t_{kD-i+1}}(W_s - W(t_{kD-j+1-M}))^4 ds\}$
  $= 96\{(M-|i-j|)\Delta\}^4$
- $E[\Theta_2^4] = 9(i-j)^2(M-|i-j|)^2\Delta^4$
- $E[\Theta_3^4] = 9(i-j)^2(M-|i-j|)^2\Delta^4$

Substituting these estimates into the inequality

$$E[(\theta_k^{ij})^4] \le 27E[\Theta_1^4 + \Theta_2^4 + \Theta_3^4]$$

and noting that $(i-j)^2, (M-|i-j|)^2 \le M^2$ we obtain the result (c). *q.e.d.*

Now we set

$$\eta_k^{ij} := b(t_{kD-j+1-M})b(t_{kD-i+1-M})\theta_k^{ij}$$

and find the following estimate

Lemma 3.6 *For any indexes $1 \le i, j, i', j' \le D$ we have the following equalities and inequality*

$$\begin{aligned}
E[\eta_k^{ij}] &= 0, \\
E[\eta_k^{ij}\eta_l^{i'j'}] &= 0 \quad \textit{for } |k-l| \ge M/D+1 \\
E[(\eta_k^{ij})^2] &\le 2B(M\Delta)^2 = 2B(\tfrac{M}{D})^2(\Delta')^2,
\end{aligned}$$

where $B = \sup_t E[b^4(t)]$.

Proof The first two equalities being evident, we are going to show the last inequality. Set

$$H_i := b(t_{kD-j+1-M})b(t_{kD-i+1-M})\Theta_i \quad (i = 1, 2, 3).$$

Since $E(H_iH_j) = 0$ $i \ne j$, we have $E[(\eta_k^{ij})^2] = E[H_1^2 + H_2^2 + H_3^2]$. For the terms H_1, H_2 we have

$$\begin{aligned}
E[H_1^2] &= E[b^2(t_{kD-j+1-M})b^2(t_{kD-i+1-M})E\{\Theta_1^2|\mathcal{F}_{kD-i+1-M}\}] \\
&\le 2B[(M-|i-j|)\Delta]^2, \\
E[H_2^2] &\le B|i-j|(M-|i-j|)\Delta^2 \le B\tfrac{M}{D}(\Delta')^2.
\end{aligned} \tag{19}$$

The estimation of the term H_3 is a little bit complicated. We begin by noting the following

$$\begin{aligned} E[H_3^2] &= E[\{W(t_{kD-i+1}) - W(t_{kD-j+1-M})\}^2] \\ &\times E[b^2(t_{kD-j+1-M})b^2(T_{kD-i+1-M})\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}^2] \\ &= (M - |i-j|)\Delta \\ &\times E[b^2(t_{kD-i+1-M})E[b^2(t_{kD-j+1-M})\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}^2|\mathcal{F}_{kD-i+1-M}]]. \end{aligned} \tag{20}$$

On the other hand

$$\begin{aligned} &E[b^2(t_{kD-j+1-M})\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}^2|\mathcal{F}_{kD-i+1-M}] \\ &\sqrt{E[b^4(t_{kD-j+1-M})|\mathcal{F}_{kD-i+1-M}]E[\{W(t_{kD-j+1-M}) - W(t_{kD-i+1-M})\}^4|\mathcal{F}_{kD-i+1-M}]} \\ &= \sqrt{6}|i-j|\Delta\sqrt{E[b^4(t_{kD-j+1-M})|\mathcal{F}_{kD-i+1-M}]}. \end{aligned}$$

Combining this with (20), we obtain

$$E[H_3^2] \leq \sqrt{6}B(M - |i-j|)|i-j|\Delta^2. \tag{21}$$

Hence by (19) and (21) we get the desired inequality

$$\begin{aligned} E[(\eta_k^{ij})^2] &\leq B(M-|i-j|)\Delta^2\{2(M-|i-j|) + (\sqrt{6}+1)|i-j|\} \\ &\leq 2B(M^2 - |i-j|^2)\Delta^2 \\ &\leq 2B(M\Delta)^2 = 2B(\frac{M}{D})^2(\Delta')^2. \end{aligned}$$

q.e.d.

Let us give a look on each component of our estimator. By definition of the estimator and by equations (16)-(18), we have the following equality

$$\begin{aligned} \hat{b}^2(t'_k) &= \frac{1}{G(M)L}\sum_{l=1}^{L}\frac{1}{DM^2}\sum_{ij=1}^{D} b(t_{(k+l)D-i+1-M})b(t_{(k+l)D-j+1-M})(M-|i-j|) \\ &+\frac{1}{G(M)M^2}\sum_{ij=1}^{D}\frac{1}{L}\sum_{l=1}^{L}\frac{\eta_{k+l}^{ij}}{\Delta'} \\ &+\frac{1}{G(M)L}\sum_{l=1}^{L}\{\frac{(\Delta'_{k+l}\bar{Z})^2}{\Delta'} + 2\frac{(\Delta'_{k+l}\bar{Z})(\Delta'_{k+l}\bar{p})}{\Delta'}\} \\ &+\frac{1}{G(M)L}\sum_{l=1}^{L}\frac{1}{M^2}\sum_{ij=1}^{D}\frac{\delta_{k+l}^{ij}}{\Delta'}. \end{aligned}$$

Knowing the definition (12) of $G(M)$, from the above expression we obtain the following

estimate

$$
\begin{aligned}
E|b^2(t'_k) - \hat{b}^2(t'_k)| \;&\le \frac{1}{G(M)L}\sum_{l=1}^{L}\frac{1}{DM^2}\sum_{ij=1}^{D}(M-|i-j|)\\
&\quad\times E[|b(t_{(k+l)D-i+1-M})b(t_{(k+l)D-j+1-M}) - b^2(t_{kD})|]\\
&+\frac{1}{G(M)M^2}\sum_{ij=1}^{D}E[|\frac{1}{L}\sum_{l=1}^{L}\frac{\eta^{ij}_{k+l}}{\Delta'}|]\\
&+\frac{1}{G(M)L}\sum_{l=1}^{L}\frac{1}{M^2}\sum_{ij=1}^{D}\frac{E|\delta^{ij}_{k+l}|}{\Delta'} \qquad (22)\\
&+\frac{1}{G(M)L}\sum_{l=1}^{L}\{E|\frac{(\Delta'_{k+l}\bar{Z})^2}{\Delta'}| + 2E|\frac{\Delta'_{k+l}\bar{p}\Delta'_{k+l}\bar{Z}}{\Delta'}|\},\\
&\text{we denote these 4 terms by the following symbols,}\\
&=: T_1 + T_2 + T_3 + T_4.
\end{aligned}
$$

To evaluate the term T_1 we notice that, when $LD \le 2M$, the quantity

$$E[|b(t_{(k+l)D-i+1-M})b(t_{(k+l)D-j+1-M}) - b^2(t_{kD})|]$$

can be evaluated in the following way

$$
\begin{aligned}
&E[|b(t_{(k+l)D-i+1-M})b(t_{(k+l)D-j+1-M}) - b^2(t_{kD})|]\\
&\le \{\sqrt{B}E|b(t_{(k+l)D-j+1-M}) - b(t_{kD})|^2\}^{1/2}\\
&+\{\sqrt{B}E|b(t_{(k+l)D-i+1-M}) - b(t_{kD})|^2\}^{1/2}\\
&\le 2B^{1/4}L_B(\max_{l,j}|M - lD + j - 1|\Delta)^{\alpha} = 2B^{1/4}L_B(\frac{M}{D})^{\alpha}(\Delta')^{\alpha},
\end{aligned}
$$

since we have $\max_{l.j}\{|lD - j + 1 - M|\} = \max\{LD - M, M - 1\} \le M$ when $LD \le 2M$.

Thus

$$
T_1 \le \begin{cases} 2B^{1/4}L_B|L - \dfrac{M}{D}|^{\alpha}(\Delta')^{\alpha} & \text{for } LD > 2M\\ 2B^{1/4}L_B(\dfrac{M}{D})^{\alpha}(\Delta')^{\alpha} & \text{for } LD \le 2M. \end{cases}
$$

In other words we confirm the next

Lemma 3.7 *Set* $C(L,M) = \max\{|L - \frac{M}{D}|, |\frac{M-1}{D}|\}$, *then we have*

$$T_1 \le 2B^{1/4}L_B C(L,M)^{\alpha}(\Delta')^{\alpha} = C_1 \max\{\rho, |L-\rho|\}\sqrt[\alpha]{\frac{M}{N_T}}),$$

where $C_1 = 2B^{1/4}L_B\sqrt[\alpha]{T}$.

For the estimation of the term T_2, we notice that

$$E\left|\frac{1}{L}\sum_{l=1}^{L}\frac{\eta^{ij}_{k+l}}{\Delta'}\right| = O(\sqrt{\frac{M}{DL}}).$$

In fact, by Lemma 3.6 we see that

$$\begin{aligned} E|\frac{1}{L}\sum_{l=1}^{L}\frac{\eta^{ij}_{k+l}}{\Delta'}|^2 &= \frac{1}{L^2}\sum_{|k-s|\le M/D+1} E|\frac{\eta^{ij}_{k+l}\eta^{ij}_{s+l}}{(\Delta')^2}| \\ &\le \frac{1}{L^2}2B(\frac{M}{D})^2\sum_{|k-s|\le M/D+1} \\ &= 2B\frac{M^2}{D^2}\frac{1}{L^2}\{L^2-(L-|\frac{M}{D}-1|)^2\} \\ &\le 2(2B+1)\frac{M^3}{D^3}\frac{1}{L}. \end{aligned}$$

Hence

$$T_2 \le \sqrt{2(2B+1)}\sqrt{\frac{D}{M}}\frac{1}{G(M)}\frac{1}{\sqrt{L}}.$$

Moreover, from

$$G(M) = \frac{(3\rho-1)D^2+1}{3(\rho D)^2} \le \frac{1}{3\rho},$$

we obtain the following estimate

Lemma 3.8 $T_2 \le C_2\frac{1}{3\rho\sqrt{\rho}}\frac{1}{\sqrt{L}}$, *where* $C_2 = \sqrt{2(2B+1)}$.

For the terms T_3, T_4, it is not difficult to show the following results by using Lemma 3.4 and Lemma 3.3

Lemma 3.9 *(1)* $T_3 \le \frac{3}{3\rho^2-1}(\Delta')^{\alpha\wedge 1/2} = C_3\frac{3}{3\rho^2-1}(\frac{D}{N_T})^{\alpha\wedge 1/2}$, *with* $C_3 = \sqrt[\alpha\wedge 1/2]{T}$

(2) $T_4 \le C_4\sqrt{\frac{N_T}{M^2}}$, *with* $C_4 = \sqrt{\frac{2C_z^2}{T}}$.

We conclude this section giving the proof of the main theorem.

Proof of Theorem 3.2 Substituting the results of Lemmas 3.7, 3.8 and 3.9 into the inequality (22) we obtain the desired estimate. The forms of the constants C_1, C_2, C_3 and C_4 are also clear from the corresponding lemmas. *q.e.d.*

4 Numerical aspects of the estimators

In this section we analyze the technical aspects of implementation of the real-time scheme (11) and its performance as spot volatility estimator. The proposed scheme is tested against other estimators, namely a realized volatility type estimator, the Fourier estimator [12] and two kernel estimators [3, 7].

For the sake of completeness, we briefly recall the spot volatility estimators employed in the comparison. Given a discrete realization of the process X_t, namely given N_T+1 observations $X(t_0), X(t_1), \ldots, X(t_{N_T})$ in the interval $[0,T]$, we want to estimate the spot volatility $b^2(t'_k)$ at time t'_k. The simplest estimator is given by

$$\hat{\sigma}^2_0(t'_k) := \frac{(X(t'_k)-X(t'_{k-1}))^2}{\Delta'}.$$

Unfortunately, this estimator is very noisy. A more robust estimator of the *Realized Volatility* type is

$$\hat{\sigma}^2_R(t'_k) := \frac{1}{m}\sum_{j=1}^{m} \frac{(X(t'_{k-m/2+j}) - X(t'_{k-m/2+j-1}))^2}{\Delta'}.$$

$\hat{\sigma}^2_R$ is an average of m squared returns around t_k. The larger m, the smoother the estimates. In the presence of microstructure effects due to tick-by-tick quotes, data must be sampled at a lower frequency in order to filter out the high frequency noise component.

A general class of nonparametric spot volatility filters is based on a *kernel weighted measure* of the integrated volatility [11] and can be seen as a continuous-time weighted moving average of the instantaneous volatility. Given a kernel $K : \mathbb{R} \to \mathbb{R}$ normalized to $\int_{\mathbb{R}} K(z)dz = 1$ and a bandwidth $h > 0$, define $K_h(z) := K(z/h)/h$. Then, based on standard results for kernel estimators, a natural spot volatility estimator is

$$\hat{\sigma}^2(t) = \lim_{h\to 0} \sum_{i=1}^{N} K_h(t'_{i-1} - t)[X(t'_i) - X(t'_{i-1})]^2.$$

This is Nadaraya-Watson type kernel estimator and it is simply the limit for shrinking bandwidth sequences of a kernel weighted average of the squared increments of data or, in other words, a kernel regression estimator in the time domain. By an appropriate choice of the kernel, it can potentially deal with the presence of market microstructure effects. This broad class of estimators includes as a special case the standard realized volatility type estimators and the *rolling window estimators* proposed by [7, 3]. In [3] the *k-day spot volatility estimator* is defined as a one-sided moving k-day average of one-day spot volatility given by a rescaled sum of squared intraday returns. This estimator looks very close to the one proposed in the present paper; however, our real-time scheme allows for more freedom in the choice of the windows in the two regularization processes and by introducing suitable weights allows to deal with the presence of market microstructure effects. See also the discussion in Section 1 on this issue.

In particular, in our analysis we consider the following schemes from the class of kernel estimators: the *one-sided exponential filter*

$$\hat{\sigma}^2_E(t'_k) = (1-\lambda)\sum_{i=1}^{k} \lambda^i [X(t'_{k-i+1}) - X(t'_{k-i})]^2,$$

where λ is a smoothing parameter, which can be set equal to 0.94 for daily data according to J.P. Morgan standard; the *one-sided rolling daily window volatility*, defined as

$$\hat{\sigma}^2_W(t'_k) = \sum_{j=1}^{n_L} w_j [X(t'_{k-j+1}) - X(t'_{k-j})]^2,$$

where n_L is the lag length of the rolling window and $w_j = exp(-\alpha j)$. The parameter α can be optimized to minimize the asymptotic measurement error variance as explained in [7]. These two volatility measures are then normalized by the sampling interval in order to obtain a measure of volatility per unit time.

The Fourier estimator of the spot volatility reconstructs the volatility in the frequency domain and derives an estimator of it in terms of Fourier transforms [12]. To conform notation to the existing literature, in the following we will denote the diffusion coefficient by σ instead of b. Assume that the process p_t is observed on a fixed time window, which

can be always reduced to $[0, 2\pi]$ by change of origin and rescaling, and a and σ are adapted random processes satisfying hypothesis

$$E[\int_0^{2\pi} (a(t))^2 dt] < \infty, \quad E[\int_0^{2\pi} (\sigma(t))^4 dt] < \infty.$$

The Fourier method reconstructs $\sigma^2(t)$ on $[0, 2\pi]$ using the Fourier transform of $dp(t)$. Define the Fourier coefficients of dp and σ^2 as follows

$$a_0(dp) = \frac{1}{2\pi}\int_0^{2\pi} dp_t, \qquad a_0(\sigma^2) = \frac{1}{2\pi}\int_0^{2\pi} \sigma^2(t)dt,$$

$$a_k(dp) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt)dp_t, \qquad a_k(\sigma^2) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt)\sigma^2(t)dt,$$

$$b_k(dp) = \frac{1}{\pi}\int_0^{2\pi} \sin(kt)dp_t, \qquad b_k(\sigma^2) = \frac{1}{\pi}\int_0^{2\pi} \sin(kt)\sigma^2(t)dt.$$

In [12] it is shown that, given an integer $n_0 > 0$, we have

$$a_0(\sigma^2) = \lim_{N_F\to\infty} \frac{\pi}{N_F + 1 - n_0} \sum_{k=n_0}^{N_F} \frac{1}{2}(a_k^2(dp) + b_k^2(dp)) \tag{23}$$

$$a_q(\sigma^2) = \lim_{N_F\to\infty} \frac{\pi}{N_F + 1 - n_0} \sum_{k=n_0}^{N_F} \frac{1}{2}(a_k(dp)a_{k+q}(dp) + b_k(dp)b_{k+q}(dp)) \tag{24}$$

$$b_q(\sigma^2) = \lim_{N_F\to\infty} \frac{\pi}{N_F + 1 - n_0} \sum_{k=n_0}^{N_F} \frac{1}{2}(a_k(dp)b_{k+q}(dp) + b_k(dp)a_{k+q}(dp)) \tag{25}$$

almost surely. Then, σ_t^2 can be computed using its Fourier coefficients as

$$\hat{\sigma}^2(t) = \lim_{M\to\infty} \sum_{k=0}^{M} [a_k(\sigma^2)\cos(kt) + b_k(\sigma^2)\sin(kt)] \tag{26}$$

where convergence is in $L^2([0, 2\pi])$ norm.

From this convergence result, a suitable estimator for the spot volatility can be derived by truncating the expansion (26) and approximating the coefficients of dp_t by means of

$$a_k(dp) \cong \frac{p(2\pi) - p(0)}{\pi} - \frac{1}{\pi}\sum_{i=0}^{N_T-1} p(t_i)[\cos(kt_{i+1}) - \cos(kt_i)]$$

$$b_k(dp) \cong -\frac{1}{\pi}\sum_{i=0}^{N_T-1} p(t_i)[\sin(kt_{i+1}) - \sin(kt_i)].$$

In implementing the Fourier estimator $\hat{\sigma}_F^2$ on a finite sample, the limits in (23)-(25) cannot be achieved, since the smallest wavelength that can be evaluated in order to avoid aliasing effects is twice the smallest distance between two consecutive prices, which yields $N_F \leq N_T/2$ (*Nyquist frequency*). However, as shown both theoretically and empirically in [14] in a study of the integrated volatility, the Fourier estimator can be made robust to microstructure effects by considering a smaller number of Fourier coefficients in the expansions.

When actually implementing the estimator, we add a linear trend on the observed process X_t such that we get $X(2\pi) = X(0)$, which does not affect the volatility estimate. Moreover, we set $n_0 = 1$ and we filter progressively high modes by using the following approximation of the volatility

$$\hat{\sigma}^2_F(t) = \sum_{k=0}^{M} \varphi_F(\delta k)[a_k(\sigma^2)\cos(kt) + b_k(\sigma^2)\sin(kt)],$$

where $\varphi_F(x) = \sin^2 x/x^2$ is the modified Fejer kernel [13] and δ is a parameter adapted to the scale which is expected to give an appropriate resolution of the volatility and to filter out high frequency noise modes. It is advisable to take $\delta \geq 1/M$; the larger δ, the smoother and less detailed the estimated volatility. After choosing a suitable value of the cutting frequency N_F, it is convenient to use the maximum M that can be computed, namely $M = N_F - n_0$, and then tune δ to filter high frequencies. Finally, we remark that according to well-known diffraction phenomena of Fourier series near the boundary of the time window this estimator performs badly at the boundaries.

We simulate discrete data from Heston continuous time stochastic volatility model [10], with microstructure contaminations. From the simulated data, volatility estimates can be compared to the value of the true volatility path. The infinitesimal variation of the true log-price process and spot volatility is given by the following SDE's

$$\begin{aligned} dp(t) &= \mu(t)\ dt + \sigma(t)\ dW_1(t) \\ d\sigma^2(t) &= \alpha(\beta - \sigma^2(t))dt + \nu\sigma(t)\ dW_2(t), \end{aligned} \tag{27}$$

where W_1, W_2 are possibly correlated Brownian motions with $< dW_1, dW_2 >_t = \rho dt$ and α is the speed of reversion of the variance σ^2 to its long term mean β. Moreover, we assume that the logarithmic noises $Z(t_k)$ are i.i.d. Gaussian and independent from p; this is typical of bid-ask bounce effects in the case of exchange rates and, to a lesser extent, in the case of equities. Alternative, possibly discrete, distributions can be used to describe microstructure noise. In [15], for instance, a bid-ask bounce effect is described by an order-driven indicator discrete variable.

In order to avoid other data manipulations such as interpolation or imputation which might affect the numerical results, we generate (through simple Euler-Maruyama discretization scheme) second-by-second evenly sampled true and observed returns and variance paths over a daily trading period of $T = 6$ hours, for a total of 21600 observations per day. In the case of microstructure contaminations, observed returns must be sampled at a lower frequency in order to have consistent realized volatility type estimates.

Figure 1 shows the true and estimated variance paths for 100 days of trading and sampling frequency of 1 second in the presence of microstructure effects. The parameter values are taken from the original paper [10]: $\alpha = 2.0$, $\beta = 0.01$, $\nu = 0.1$, $\mu = 0.0$, $\rho = 0.0$, $p_0 = \log 100$, $\sigma_0 = \beta$; moreover, $var(Z_t) = 0.0142$. The other parameters involved in the real-time estimator design are $N_T = 2160000$, $N = 21600$, $D = 100$, while M and L are let to vary. In Panel A, we set $L = [2\sqrt{N}] = 294$ and $M = 50, 100, 200, 1500$ in the magenta, green, blue and yellow trajectories respectively. The corresponding Integrated Squared Error $E[\int_0^T (\hat{\sigma}^2_t - \sigma^2_t)^2 dt]$ between the generated and reconstructed trajectories changes from 3.0428e-5 to 6.8862e-6 to 4.0706e-6 to 6.7678e-6. In Panel B, we set $M = 200$ and $L = 100, 294, 500$ in the magenta, blue and green trajectories respectively. The corresponding Integrated Squared Error changes from 3.0946e-6 to 4.0706e-6 to 5.0613e-6. We can notice that the parameter M has the greatest effect on the real-time estimator: very small or very large values of M yield large peaks and oscillations in the estimated

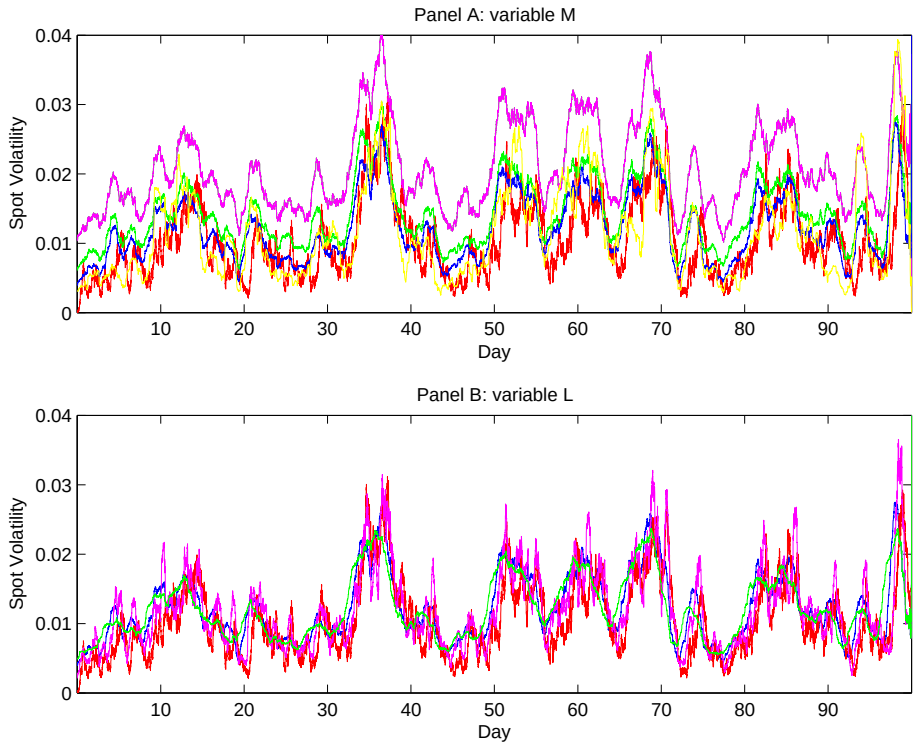

Figure 1: True and estimated variance paths obtained by the real-time scheme. Parameter values: $N_T = 2160000$, $N = 21600$, $D = 100$. Panel A: $L = 294$ and $M = 50, 100, 200, 1500$ in the magenta, green, blue and yellow trajectories respectively. Panel B: $M = 200$ and $L = 100, 294, 500$ in the magenta, blue and green trajectories respectively.

volatility path, while a larger L value yields a slight smoothing and anticipating effect. For this reason, M can be used as tuning parameter to design optimized estimators, as explained below. Figure 2 shows the Integrated Squared Error as a function of M and L. Its minimum is achieved at $M = 250$, $L = 115$ and is equal to 2.9782e-6.

Figure 3 shows the true and estimated variance paths for 100 days of trading and sampling frequency of 1 second in the presence of microstructure effects. When microstructure effects are included in the model, the realized volatility type estimators are spoiled unless we resort to lower frequency data sampling. Therefore, to compute $\hat{\sigma}^2_R$ we sample quotes each hour, while for the real-time scheme and the Fourier estimator we use all the data. In the first panel, the parameters for each estimator are arbitrarily set to $M = 100$, $L = 294$, $m = 30$, $\lambda = 0.75$, $\alpha = 0.665$ and $n_L = 26$, while in the second panel they are chosen in such a way to optimize the different estimators. More precisely, the optimized estimators are obtained by minimizing the Integrated Squared Error $E[\int_0^T (\hat{\sigma}^2_t - \sigma^2_t)^2 dt]$ between the generated and reconstructed trajectories with respect to M (and keeping L arbitrarily fixed to the suboptimal value $[2\sqrt{N}]$) for the real-time scheme, m for the realized volatility estimator $\hat{\sigma}^2_R$, λ for the exponential filter and α for the rolling window. The Fourier parameters are chosen arbitrarily as $N_F = 50$ and $\delta = 2$ in order to filter out high frequency noise modes, i.e. they are not optimized. Moreover, this estimator has been computed by reconstructing the Fourier expansion on each day separately, because otherwise it would provide a too smooth trajectory but nevertheless capable of capturing the peaks of the volatility path. The best volatility estimates are provided by the real-time scheme and the Fourier estimator, with an Integrated Squared Error equal to 1.45e-5 and 2.70e-6 respectively. The (suboptimal) Fourier estimator provides a very good recon-

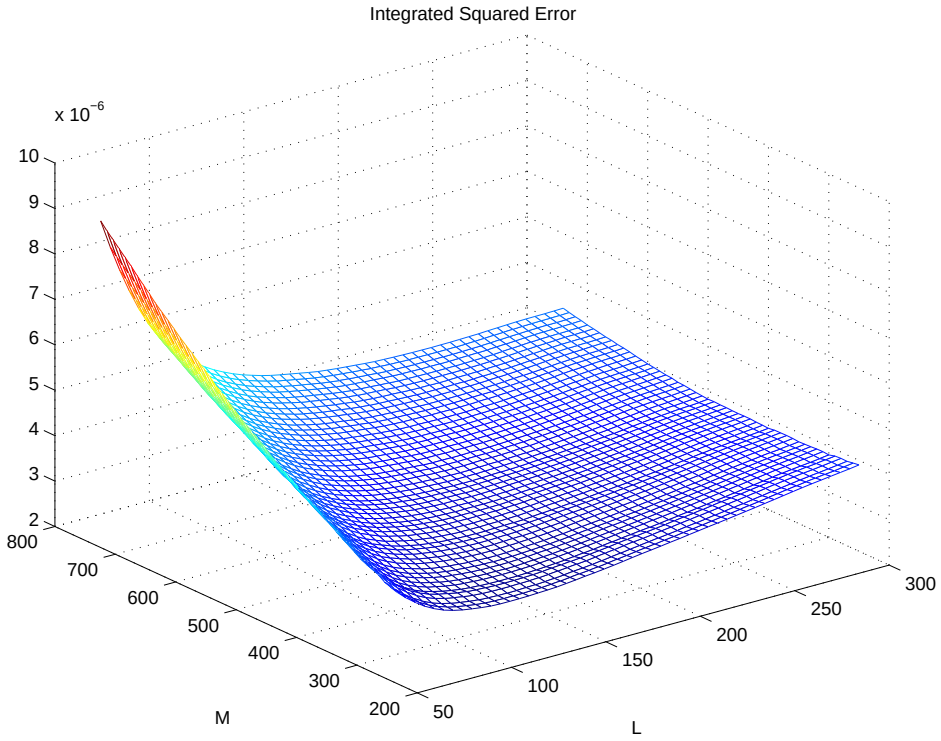

Figure 2: Integrated Squared Error as a function of M and L. Parameter values: $N_T = 2160000$, $N = 21600$, $D = 100$.

struction of the dynamics of the original process (27) as well, with a good representation of both the abrupt changes and the more regular sections in the volatility path. On the contrary, both the rolling window and the exponential filter are not much accurate.

One advantageous peculiarity of the real-time estimator consists in its ease of implementation and therefore in the fact that it works faster than others for a given accuracy level, namely in its real-time property. This is an issue of the outmost importance from a practitioner point of view. In order to better compare the computational speed of the different estimators, we fix a suitable precision threshold ($\epsilon = 2.0e-3$) and tune the estimator parameters in such a way that the Integrated Squared Error is below this level. The considered parameters are again M for the real-time scheme, m for the realized volatility estimator, λ for the exponential filter and α for the rolling window. The tuning is achieved by letting the specific parameters of each method vary in a range of 100 values; of course, the computational time may be influenced by how far the initial value of each parameter is from the threshold value and therefore the total number of iterations K is displayed for the reader's convenience. The results are shown in Table 1 and refer to 100 days of trading and a sampling frequency of 1 second. The CPU time refers to the computational time in seconds needed to identify the estimators having the Integrated Squared Error below ϵ. The corresponding Integrated Squared Error is displayed. Moreover, as a further information, the minimum achievable Integrated Squared Error is listed in the 5th column with the corresponding parameter value in the 6th one. The Fourier estimator has been excluded from the comparison because it is not competitive from a computational cost point of view. However, for the sake of completeness, the Integrated Squared Error achievable by the Fourier estimator and the corresponding parameter values used δ and N_F are listed as well. It is evident that in the presence of microstructure effects the real-time estimator $\hat{b}^2$ provides the highest accuracy at the expense of a slightly higher computational

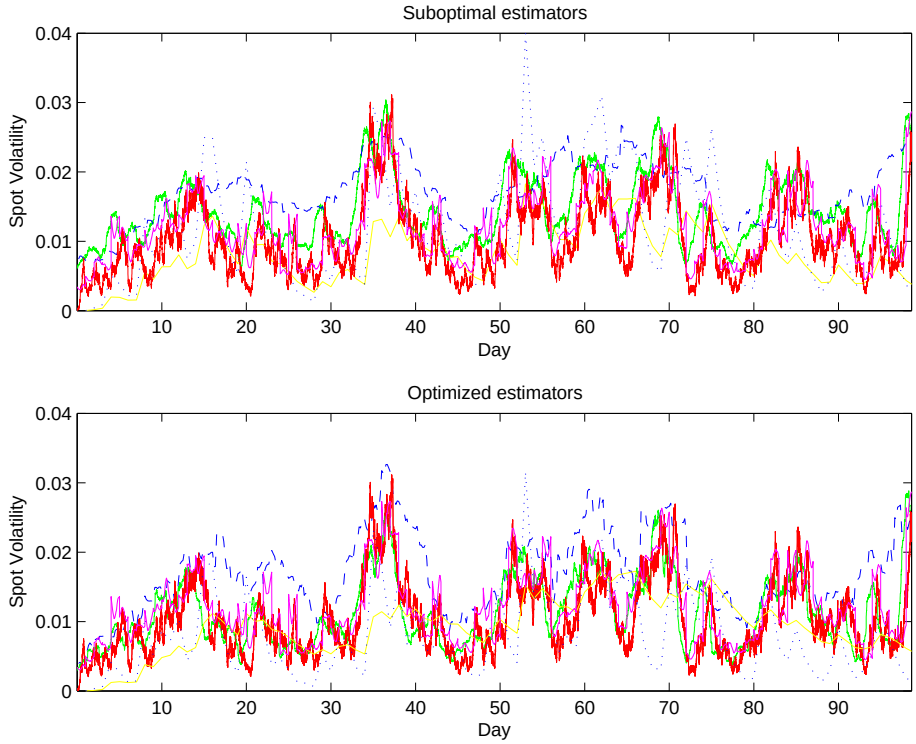

Figure 3: True and estimated variance paths. Red line: true spot volatility; green line: real-time scheme; magenta line: Fourier estimator; dashed line: Realized Volatility type estimators $\hat{\sigma}^2_R$; dotted line: rolling window; yellow line: exponential filter.

cost. We remark that, using daily data, the exponential filter and the rolling window do not provide accurate volatility estimates.

Finally, in Table 2 we consider the situation where $\rho = -0.35$ in order to investigate how robust our estimator is to leverage effects. We can see again that the real-time scheme provides the most accurate estimate at the expenses of a slightly higher computational cost. In particular, in order to avoid the end effects which would affect Fourier expansion on each trading day, the Fourier estimator has been estimated on the whole 100 days trading period, with $N_F = 200$ and $\delta = 0.1$. Finally, we remark that in all the simulations related to Tables 1-2 the optimal value for the parameter M in the real-time estimator is such that the windows of data involved in the first regularization overlap. In fact, $M_{opt} > D = 100$. This is a further support of the need for the more general form (11) of

Method	CPU time	Int. Sq. Err.	Param.	K	Min. Int. Sq. Err.	Param.
$\hat{b}^2$	4.14e+0	1.41e-5	60	1	3.62e-6	400
$\hat{\sigma}^2_R$	0.00e+0	1.82e-5	20	1	1.57e-5	32
$\hat{\sigma}^2_E$	0.00e+0	1.35e-3	0.51	1	8.35e-4	0.87
$\hat{\sigma}^2_W$	1.60e–2	1.95e-3	0.63	13	1.33e-3	0.87
$\hat{\sigma}^2_F$	-	2.70e-6	$\delta = 2.0$	$N_F = 50$	-	-

Table 1: CPU time in seconds needed to identify estimators having Integrated Squared Error below $\epsilon = 2e-3$. The other entries from left to right are: the corresponding Integrated Squared Error and parameter value, the number of iterations performed, the minimum achievable Integrated Squared Error and the corresponding parameter value.

Method	CPU time	Int. Sq. Err.	Param.	K	Min. Int. Sq. Err.	Param.
$\hat{b}^2$	4.25e+0	1.40e-5	60	1	3.39e-6	290
$\hat{\sigma}^2_R$	0.00e+0	2.50e-5	20	1	1.70e-5	208
$\hat{\sigma}^2_E$	0.00e+0	1.23e-3	0.51	1	7.70e-4	0.94
$\hat{\sigma}^2_W$	3.10e-2	2.00e-3	0.96	46	1.56e-3	1.31
$\hat{\sigma}^2_F$	-	4.32e-6	$\delta = 0.1$	$N_F = 200$	-	-

Table 2: CPU time in seconds needed to identify estimators having Integrated Squared Error below $\epsilon = 2e-3$. The other entries from left to right are: the corresponding Integrated Squared Error and parameter value, the number of iterations performed, the minimum achievable Integrated Squared Error and the corresponding parameter value.

the real-time estimator with respect to the orginal one (7).

5 Concluding Remarks

We have introduced a two-step regularization scheme for spot volatility estimation in the presence of microstructure effects. The basic idea of this estimator was already contained in [16], but the newly proposed modification of the scheme is much simpler to implement. In fact, in the original estimator the first regularization of the observed process is performed on non-overlapping windows of data and by this reason it requires numerous observations around each point t_k where the estimation is computed. This yields an unmotivated limitation on the number of estimation points when only a finite sample of data is available. The modified scheme allows to use overlapping blocks of data for the regularization. This feature allows to use the same observed data repeatedly to estimate volatility at different estimation times, thus reducing the size of necessary observation data. Our scheme can be seen as a kernel-type estimator over a regularized set of data and the choice of suitable weights is of crucial importance to deal with microstructure effects. In particular, microstructure noise is smoothed out by the inclusion in our estimator of a weighted sum of products of lagged intra-day returns. The proposed scheme is tested against other estimators, namely the realized volatility estimator, the Fourier estimator and two kernel estimators, in several Monte Carlo experiments which show the great efficiency and precision of the two-step estimator at low computational cost.

References

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

[2] T. Andersen, T. Bollerslev, and F. Diebold, "Parametric and Nonparametric Volatility Measurement", In: Ait-Sahalia, Y., Hansen, L.P. (Eds.), *Handbook of Financial Econometrics*, North Holland, Amsterdam, 2005.

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

[4] F.M. Bandi, and J.R. Russell, "Microstructure noise, realized variance and optimal sampling". *Review of Economic Studies*, 75(2), 339-369, 2008.

[5] F.M. Bandi, and J.R. Russell, "Separating market microstructure noise from volatility". *Journal of Financial Economics*, 79(3), 655-692, 2006.

[6] O.E. Barndorff-Nielsen, P.R. Hansen, A. Lunde, A. and N. Shephard, "Designing realised kernels to measure ex-post variation of equity prices in the presence of noise". *OFRC Working paper series*, 2006-FE-05, 2006.

[7] D.P. Foster, and D.B. Nelson, "Continuous record asymptotics for rolling sample variance estimators", *Econometrica*, 64, 139–174, 1996.

[8] V. Genon-Catalot, C. Laredo, and D. Picard, "Non-parametric estimation of the diffusion coefficient by wavelets methods", *Scandinavian Journal of Statistics*, 19, 317–335, 1992.

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

[10] S.L. Heston, "A closed-form solution for option with stochastic volatility with applications to bond and currency options", *The Review of Financial Studies*, 6(2), 327–343, 1993.

[11] D. Kristensen, "Nonparametric filtering of the realised spot volatility: a kernel-based apporach", *Working Paper*, 2006. Forthcoming in Econometric Theory.

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

[13] P. Malliavin, and A. Thalmaier, "Stochastic calculus of variations in Mathematical Finance", *Springer Finance*, 2005.

[14] M.E. Mancino and S. Sanfelici, "Robustness of Fourier Estimator of Integrated Volatility in the Presence of Microstructure Noise", *Comp. Stat. & Data Analysis*, 52, 2966–2989, 2008.

[15] Nielsen, M.O. and Frederiksen, P.H., "Finite sample accuracy and choice of sampling frequency in integrated volatility estimation", *Journal of Empirical Finance,* 15(2), 265–286, 2008.

[16] S. Ogawa, "Real time scheme for the volatility estimation in the presence of the microstructure noise", *Monte Carlo Methods and Applications*, 14 (4), 331–342, 2008.

[17] S. Ogawa and K. Wakayama, "Real time scheme for the volatility estimation" *Monte Carlo Methods and Applications*, 10(2), 2007.

[18] S. Ogawa, "Introduction to the discrete approximation of stochastic differential equations", *SUGAKU Expositions*, 18(1), 191–122, American Math. Society, 2005.

[19] N. Shephard, "Stochastic volatility: selected readings", *Oxford University Press*, 2005.

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

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

Stampato in proprio dal Dipartimento di Economia dell'Università degli Studi di Parma in Via J.F. Kennedy 6, 43100 Parma, adempiuti gli obblighi di cui all'articolo 1 del D.L. 31 agosto 1945 n. 660.