Faster Convergence to Realized Volatility with Microstructure Noise

Yuh-Dauh Lyuu* Philip Sheih† Yu-cheng Tien‡ Yun-Cheng Tsai§

Abstract

The volatility of asset return plays a critical role for high-frequency trading. However, the microstructure noise could bias the estimated realized volatility. Zhang et al. (2005) propose a batch estimator for the realized volatility of high-frequency data in the presence of microstructure noise. It gives the estimates after all the data arrive. This paper proposes a recursive version of their estimator that outputs volatility estimates as the data arrive. Our estimator gives excellent estimates well before all the data arrive. Both real high-frequency futures data and simulation data confirm the performance of the recursive estimator.

Keywords: quadratic variation, realized volatility, high-frequency, microstructure noise.

*Professor, Department of Finance and Department of Computer Science & Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan. E-mail: lyuu@csie.ntu.edu.tw. The author was supported in part by the National Science Council of Taiwan, under grant 100-2221-E-002-111-MY3.

†Futures Proprietary Trading, Chinatrust Commercial Bank, No.3, Songshou Rd., Taipei 11051, Taiwan. E-mail: philip.sheih@chinatrust.com.tw.

‡Futures Proprietary Trading, Chinatrust Commercial Bank, No.3, Songshou Rd., Taipei 11051, Taiwan. E-mail: yu.cheng.tien@chinatrust.com.tw.

§Corresponding author. Ph.D. Candidate, Department of Computer Science & Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan. E-mail: d98922012@csie.ntu.edu.tw. The author was supported in part by the National Science Council of Taiwan, under grant 100-2221-E-002-111-MY3.

1 Introduction

Realized volatility is one of the most important statistics in financial markets. It is defined as the square root of the integrated instantaneous variance over a period of time (see Shreve (2004)). Realized volatility is central to risk management, pricing, and forecasting. The prices of most derivatives depend critically on it. In particular, the variance option's payoff depends explicitly on the realized volatility (see Schroder (1989), Bandi et al. (2008), Carr and Lee (2007)). Unfortunately, the realized volatility is not directly observable and must be estimated. Quadratic variation is often used to estimate the realized volatility (see Barndorff-Nielsen and Shephard (2002)), and the estimation error diminishes as the sampling frequency increasing (see Shreve (2004)). Estimating the realized volatility in the financial markets based on high-frequency intraday data has been much studied (see Andersen and Bollerslev (1998), Andersen et al. (2001), Maheu and McCurdy (2002), Areal and Taylor (2002)). According to Hendershott et al. (2011), high-frequency trading is responsible for 73% of the trading volume in the U.S. in 2009. Kozhan and Tham (2012) discuss the execution risk faced by high-frequency arbitrage trading strategies.

But as returns are sampled at finer intervals, the microstructure noise issue becomes more pronounced (see Hansen and Lunde (2006)). The idea of microstructure noise attempts to capture a variety of frictions inherent in the trading process: bid-ask spread, discreteness of price changes, differences in trade sizes or informational content of price changes, gradual response of prices to a block trade, strategic component of the order flow, inventory control effects, and so on. Ait-Sahalia and Yu (2009) and Barndorff-Nielsen (2002) show that the microstructure noise may cause quadratic variation to give a biased estimate of the realized volatility.

Hasbrouck (1993) and Zhou (1996) assume the microstructure noises are independent and identically distributed Gaussian with variance ($E\epsilon^2$) and mean zero and the every return is independent of microstructure noise. Following Zhou (1996), Zhang et al. (2005) propose a batch estimator that uses high-frequency data to estimate the realized volatility, even in the presence of the microstructure noise. They show that the estimator performs well for the stochastic-volatility model of Heston (1993).

The estimator of Zhang et al. (2005) runs in the batch mode and hence starts only after all the returns are received, such as at the end of a trading day. This paper proposes a recursive version which outputs the estimates in real-time as the data arrive, without having to wait until the end of the trading day. The recursive estimator gives good estimates well before all the data arrive. This is confirmed by both real-world high-frequency data and simulation data. With faster convergence, the user can fine-tune the trading strategies earlier.

The rest of the paper is organized as follows. Section 2 provides basic definitions and a summary of the batch estimator by Zhang et al. (2005). Section 3 describes the

recursive estimator. Section 4 compares the effectiveness of the batch and recursive estimator with both the secondly Taiwan Stock Exchange Capitalization Weighted Stock Index futures prices and simulation data for geometric Brownian motion and Heston's model. Section 5 concludes.

2 Preliminaries

2.1 Basic Definitions

To fix ideas, let $S(t)$ denote the price process of the security at time t. The log price is defined as $X(t) = \log S(t)$. Assume it follows the following Ito process:

$$dX(t) = u(t)\,dt + \sigma(t)\,dB(t), \tag{1}$$

where $B(t)$ is a Brownian motion, $\sigma(t)$ is volatility and $u(t)$ is mean rate of return. The realized volatility for the time interval from 0 to t_n is defined as

$$\sigma_{t_n}^{\text{real}} \equiv \sqrt{\int_0^{t_n} \sigma(s)^2\,ds}. \tag{2}$$

In reality, the realized volatility must be estimated by discrete sampling. Let $0 = t_0 < t_1 < \cdots < t_n$ be a set of time points when the process X is sampled. Assume identical sampling frequency $\Delta t = t_i - t_{i-1}$. For example, with $n = 18{,}000$ and $\Delta t =$ 1 second, it means 5 hours of secondly sampling. The realized volatility is usually estimated by the square root of the following quantity called the quadratic variation:

$$\sum_{i=1}^{n} (X(t_i) - X(t_{i-1}))^2. \tag{3}$$

For example, the realized volatility can be measured between 8:45AM and 13:45PM on a trading day by calculating the quadratic variation of secondly log returns. It is known that (Shreve (2004))

$$\sqrt{\operatorname*{plim}_{\Delta t \to 0} \sum_{i=1}^{n} \left(X\left(t_i\right) - X\left(t_{i-1}\right)\right)^2} = \sigma_{t_n}^{\text{real}}. \tag{4}$$

So the error of quadratic variation diminishes as the sampling frequency increases.

From equation (4), the quadratic variation computed from the highest-frequency data should provide the best possible estimation for the realized volatility. However, when prices are sampled at higher frequencies, microstructure noise will matter. Zhou (1996) incorporates the microstructure noise in the price process by postulating the observation log price process to be

$$Y(t) = X(t) + \epsilon(t), \tag{5}$$

where $\epsilon(t)$ is a zero-mean random noise independent of $X(t)$. Note that the realized volatility $\sigma_{t_n}^{\text{real}}$ is defined for X not Y, but Y is observed, not X.

2.2 The Two-Scales Estimator

Zhang et al. (2005) propose the estimator $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ below for $\left(\sigma_{t_n}^{\text{real}}\right)^2$, called the two-scales estimator, from samples of Y. Let $g = \{t_0, t_1, \ldots, t_n\}$ be the set of times the return process Y is observed. Let

$$[Y,Y]_{t_n} \equiv \sum_{i=1}^{n} (Y(t_i) - Y(t_{i-1}))^2,$$

be the quadratic variation of $Y(t_1), Y(t_2), \ldots, Y(t_n)$.

Divide the observation times into K nonoverlapping groups. The kth group $g(k)$ starts at t_{k-1} and then picks every Kth time point until t_n is reached. So $g(k) = \{t_{k-1}, t_{k-1+K}, t_{k-1+2K}, \ldots, t_{k-1+n_k K}\}$, for $k = 1, 2, \ldots, K$, where n_k is the integer making $t_{k-1+n_k K}$ the last point in $g(k)$. Note that $\bigcup_{k=1}^{K} g(k) = \{t_0, t_1, \ldots, t_n\}$. So $K \approx n/\bar{n}$, where $\bar{n} \equiv \lfloor \frac{n-K+1}{K} \rfloor$. We pick $K = O\left(n^{2/3}\right)$ according to Zhang et al. (2005). Define

$$[Y,Y]_{t_n}^{(k)} \equiv \sum_{t_{i-K}, t_i \in g(k)} (Y(t_i) - Y(t_{i-K}))^2,$$

the quadratic variation of the kth group. Note that

$$\sum_{k=1}^{K} [Y,Y]_{t_n}^{(k)} = \sum_{i=K}^{n} (Y(t_i) - Y(t_{i-K}))^2. \tag{6}$$

$[Y,Y]_{t_n}^{(\text{avg})}$ averages the above:

$$[Y,Y]_{t_n}^{(\text{avg})} \equiv \frac{1}{K} \sum_{k=1}^{K} [Y,Y]_{t_n}^{(k)}.$$

Finally, the two-scales estimator $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ combines $[Y,Y]_{t_n}$ and $[Y,Y]_{t_n}^{(\text{avg})}$ as follows:

$$\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})} \equiv \left(1 - \frac{\bar{n}}{n}\right)^{-1} \left([Y,Y]_{t_n}^{(\text{avg})} - \frac{\bar{n}}{n} [Y,Y]_{t_n} \right). \tag{7}$$

Zhang et al. (2005) choose $K = cn^{2/3}$ and show that

$$\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})} - \left(\sigma_{t_n}^{\text{real}}\right)^2 \approx \frac{1}{n^{1/6}} \left[\frac{4}{c^2} (E\epsilon^4) + c \frac{4T}{3} \int_0^{t_n} \sigma_s^4 ds \right]^{1/2} Z_{\text{total}}, \tag{8}$$

where Z_{total} is an asymptotically standard normal random variable independent of the X process.

3 Methodology

3.1 The Recursive Two-Scales Estimator

The two-scales estimator must wait until all the return data $Y(t_1), Y(t_2), \ldots, Y(t_n)$ are received before the calculation can commence. We now revise this two-scales estimator so that it can be updated on-the-fly as the data arrive and gives real-time estimates. The following is our recursive two-scales estimator.

Definition 3.1

$$\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})} \equiv \frac{1}{K-1} \sum_{j=K}^{i} \left[(Y(t_j) - Y(t_{j-K}))^2 - (Y(t_j) - Y(t_{j-1}))^2 \right], i = K, K+1, \ldots, n.$$

Hence $\langle \widehat{X,X} \rangle_{t_K}^{(\text{rec})}, \langle \widehat{X,X} \rangle_{t_{K+1}}^{(\text{rec})}, \ldots, \langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ are the estimates given at times $t_K, t_{K+1}, \ldots, t_n$. Note that $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ is calculated the moment $Y(t_i)$ is received. The final one $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ will be close to $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ (see Theorem 3.4). The quantity $\Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ defined below will help us calculate $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ from $\langle \widehat{X,X} \rangle_{t_{i-1}}^{(\text{rec})}$ efficiently instead of from scratch.

Definition 3.2

$$\begin{aligned}
&\Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})} = 0, \text{ where } i = 0, 1, \ldots, K-1, \\
&\Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})} = \frac{1}{K-1} \left[(Y(t_i) - Y(t_{i-K}))^2 - (Y(t_i) - Y(t_{i-1}))^2 \right], i = K, K+1, \ldots, n.
\end{aligned}$$

Note that $\Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ requires only $O(1)$ arithmetic operations. The next theorem shows how to update the recursive two-scales estimator as each $Y(t_i)$ arrives.

Theorem 3.3

$$\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})} = \langle \widehat{X,X} \rangle_{t_{i-1}}^{(\text{rec})} + \Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}, i = K, K+1, \ldots, n.$$

Proof:

$$\begin{aligned}
\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})} = & \quad \frac{1}{K-1} \sum_{j=K}^{i} \left[(Y(t_j) - Y(t_{j-K}))^2 - (Y(t_j) - Y(t_{j-1}))^2 \right] \\
= & \quad \frac{1}{K-1} \sum_{j=K}^{i-1} \left[(Y(t_j) - Y(t_{j-K}))^2 - (Y(t_j) - Y(t_{j-1}))^2 \right] + \\
& \quad \frac{1}{K-1} \left[(Y(t_i) - Y(t_{i-K}))^2 - (Y(t_i) - Y(t_{i-1}))^2 \right] \\
= & \quad \langle \widehat{X,X} \rangle_{t_{i-1}}^{(\text{rec})} + \Delta \langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}.
\end{aligned}$$

■

Next, we show that $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ at the end of day (time t_n) are practically identical.

Theorem 3.4

$$\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})} - \langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})} \approx \frac{1}{K-1} \left[\sum_{i=1}^{K-1} (Y(t_i) - Y(t_{i-1}))^2 \right].$$

Proof:

$$
\begin{aligned}
& \langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})} - \langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})} \\
= \quad & \frac{1}{K-1} \sum_{i=K}^{n} \left[(Y(t_i) - Y(t_{i-K}))^2 - (Y(t_i) - Y(t_{i-1}))^2 \right] - \left(1 - \frac{\bar{n}}{n}\right)^{-1} \left([Y,Y]_{t_n}^{(\text{avg})} - \frac{\bar{n}}{n} [Y,Y]_{t_n} \right) \\
\approx \quad & \frac{1}{K-1} \sum_{i=K}^{n} \left[(Y(t_i) - Y(t_{i-K}))^2 - (Y(t_i) - Y(t_{i-1}))^2 \right] - \left(\frac{1}{K-1} \sum_{k=1}^{K} [Y,Y]_{t_n}^{(k)} - \frac{1}{K-1} [Y,Y]_{t_n} \right) \\
= \quad & \frac{1}{K-1} \left[\sum_{i=K}^{n} \left[(Y(t_i) - Y(t_{i-K}))^2 - (Y(t_i) - Y(t_{i-1}))^2 \right] - \sum_{k=1}^{K} [Y,Y]_{t_n}^{(k)} + [Y,Y]_{t_n} \right] \\
= \quad & \frac{1}{K-1} \left[\sum_{i=K}^{n} (Y(t_i) - Y(t_{i-K}))^2 - \sum_{i=K}^{n} (Y(t_i) - Y(t_{i-1}))^2 - \sum_{k=1}^{K} [Y,Y]_{t_n}^{(k)} + [Y,Y]_{t_n} \right] \\
= \quad & \frac{1}{K-1} \left[\sum_{i=1}^{n} (Y(t_i) - Y(t_{i-1}))^2 - \sum_{i=K}^{n} (Y(t_i) - Y(t_{i-1}))^2 \right] \qquad (9) \\
= \quad & \frac{1}{K-1} \left[\sum_{i=1}^{K-1} (Y(t_i) - Y(t_{i-1}))^2 \right] \\
= \quad & \frac{1}{K-1} [Y,Y]_{t_{K-1}},
\end{aligned}
$$

where equation (9) is due to equation (6). ∎

So the difference between $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ is $\frac{1}{K-1}[Y,Y]_{t_{K-1}}$, which is insignificant because K is small relative to n.

3.2 The Computational Costs of Calculating $\langle \widehat{X,X} \rangle_{t_n}^{(\textbf{batch})}$ and $\langle \widehat{X,X} \rangle_{t_i}^{(\textbf{rec})}$

Computational efficiency is measured in terms of the number of arithmetic operations. Below, we derive the total number of arithmetic operations for $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ for $i = K, K+1, \ldots, n$. Clearly, $[Y,Y]_{t_n}$ takes $O(n)$ operations to calculate and so does $[Y,Y]_{t_n}^{(\text{avg})}$. Hence the number of operations to calculate $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ is $O(n)$. It takes $O(1)$ operations to obtain $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ from $\langle \widehat{X,X} \rangle_{t_{i-1}}^{(\text{rec})}$ by Theorem 3.3. The total number of operations to calculate $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ is thus $O(n)$, asymptotically the same as the original two-scales estimator $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$. So our estimators do not lose efficiency even though they provide real-time information on volatility.

4 Experimental Results

This section compares the convergence of the recursive two-scales estimator $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ to $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$, where $i = K, K+1, \ldots, n$. Following Zhang et al. (2005), we choose $K = cn^{2/3}$. Subsection 4.1 deals with simulation results using geometric Brownian motion and Heston's stochastic-volatility model. Subsection 4.2 uses the Taiwan Stock Exchange Capitalization Weighted Stock Index futures prices (ticker symbol TX).

4.1 Simulation Results

Assume the microstructure noise ϵ is Gaussian and set $(E\epsilon^2)^{1/2} = 0.0005$, same as Zhang et al. (2005) (so the volatility of the noise is 0.05% per sample). That roughly corresponds to a daily volatility of 4%. For each model, the initial price is $X_0 = 100$, a day consists of 6.5 trading hours, $n = 23{,}400$ and $K = 300$ as Zhang et al. (2005) (so $c = 0.367$). Each sample path lasts for $T = 1$ day (equivalently, 1/252 year). The sampling frequency is $\Delta t = 1$ second (equivalently, $1/(252 \times 23{,}400)$ year). A total of $M = 25{,}000$ sample paths are generated for each model. For each estimator, there are three validation methods: bias, RMSE and relative RMSE. The "bias" row reports the average of $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})} - \left(\sigma_{t_n}^{\text{real}}\right)^2$ and that of $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})} - \left(\sigma_{t_n}^{\text{real}}\right)^2$. The "RMSE" row reports the root-mean-square errors of $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ with respect to $\left(\sigma_{t_n}^{\text{real}}\right)^2$. The "relative RMSE" row reports the RMSE divided by $\left(\sigma_{t_n}^{\text{real}}\right)^2$.

4.1.1 Geometric Brownian Motion

We use the geometric Brownian motion (GBM) as our data-generating process:

$$dX_t = \left(r - \frac{1}{2}\sigma^2\right) dt + \sigma \, dB_t, \tag{10}$$

where B_t is the Brownian motion. Assume annual variance $\sigma^2 = 0.036288$ (daily volatility $\sigma\sqrt{T} = 0.012$, same as Lux (1996)) and risk-free rate $r = 0$.

We will estimate the theoretical realized volatility σ from the GBM process in the presence of microstructure noise ϵ. The theoretical realized volatility from 0 to t_i is simply $\sigma_{t_i}^{\text{real}} = \sigma\sqrt{t_i}$, where $i = K, K+1, \ldots, n$. On each sample path, we compare $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ with $\left(\sigma_{t_n}^{\text{real}}\right)^2 = \sigma^2 t_n$ at t_n and also compare $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ with $\left(\sigma_{t_i}^{\text{real}}\right)^2 = \sigma^2 t_i$, where $i = K, K+1, \ldots, n$.

The "bias" row of Table 1 shows that in spite of the microstructure noise, $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ are able to estimate $\left(\sigma_{t_n}^{\text{real}}\right)^2$ at t_n without bias. It also shows $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ are essentially the same at time point t_n. Any slight difference between them is due to $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ dropping the first K data (recall Lemma 3.4). The small

RMSEs in the "RMSE" row of Table 1 shows that both $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ estimate $\left(\sigma_{t_n}^{\text{real}}\right)^2$ at t_n accurately. The "relative RMSE" row of Table 1 again confirms the accuracy of $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$. Figure 1 plots the $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}}$ of a particular sample path for $i = K, K+1, \ldots, n$. Note that $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}}$ quickly converges to the theoretical realized volatility $\sigma_{t_n}^{\text{real}}$. Figure 2 plots the relative RMSE of $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ at every 1,800th second based on a particular sample path. Clearly, the relative RMSE trends lower with increasing i. Table 2 shows that the relative RMSEs of $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ at every 1,800th second trend lower with increasing i; hence $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ tends to be more accurate with increasing i. As a result, $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ converges to $\left(\sigma_{t_i}^{\text{real}}\right)^2$ well before t_n.

4.1.2 Heston's Model

Heston (1993) proposes the following stochastic-volatility model:

$$dX_t = (\mu - v_t/2)\,dt + \sigma_t\,dB_t, \tag{11}$$

$$dv_t = \kappa(\alpha - v_t/2)\,dt + \gamma\sqrt{v_t}\,dW_t, \tag{12}$$

where $v_t = \sigma_t^2$, B_t and W_t are Brownian motions, and ρ is the correlation between them. We will use it as our data-generating process. Assume Feller's condition $2\kappa\alpha \geq \gamma^2$ to make the zero boundary unattainable by the process v_t. Following Zhang et al. (2005), we set:

$$v_0 = 0, \mu = 0.05, \kappa = 5, \alpha = 0.04, \gamma = 0.5, \rho = -0.5.$$

As the theoretical realized volatility $\sigma_{t_i}^{\text{real}}$ has no closed-form solutions, it is replaced by the square root of the quadratic variation in equation (3) in the experiments below. But, for brevity, we continue to use the symbol $\sigma_{t_i}^{\text{real}}$.

We will estimate $\sigma_{t_n}^{\text{real}}$ from the X_t in the presence of microstructure noise ϵ using both batch and recursive two-scales estimator. On each sample path, we compare $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ with $\left(\sigma_{t_n}^{\text{real}}\right)^2$ at t_n and also compare $\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}$ with $\left(\sigma_{t_i}^{\text{real}}\right)^2$, where $i = K, K+1, \ldots, n$.

The "bias" row of Table 3 shows that in spite of the microstructure noise, $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ are able to estimate $\left(\sigma_{t_n}^{\text{real}}\right)^2$ at t_n without bias. It also shows $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ are essentially the same at time point t_n. Any slight difference between them is due to $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ dropping the first K data (recall Lemma 3.4). The small RMSEs in the "RMSE" row of Table 3 show that both $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$ estimate $\left(\sigma_{t_n}^{\text{real}}\right)^2$ at t_n accurately. The "relative RMSE" row of Table 3 again confirms the accuracy of $\langle \widehat{X,X} \rangle_{t_n}^{(\text{batch})}$ and $\langle \widehat{X,X} \rangle_{t_n}^{(\text{rec})}$. Figure 3 plots the $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}}$ of a particular sample path for $i = K, K+1, \ldots, n$. Note that $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\text{rec})}}$ quickly converges to $\sigma_{t_i}^{\text{real}}$.

Figure 4 plots the relative RMSE of $\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}$ at every 1,800th second. Clearly, the relative RMSE trends lower with increasing i based on a particular sample path. Table 4 shows that the relative RMSEs of $\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}$ at every 1,800th second trend lower with increasing i; hence $\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}$ tends to be more accurate with increasing i. As a result, $\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}$ converges to $\left(\sigma_{t_i}^{\mathrm{real}}\right)^2$ well before t_n.

4.2 Taiwan Stock Exchange Capitalization Weighted Stock Index Futures Contracts (TX)

This subsection compares $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$ with $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ using the TX futures prices for the period August 26, 2011 08:45:00 am–13:45:00 pm. This data set is tick-by-tick and contains 93,382 TX futures transactions. The total observation time is 5 hours, so $n = 18{,}000$ seconds. We choose the same $c = 0.367$ as in Subsection 4.1 (so $K = 252$). Figure 5 shows that $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{rec})}}$ and $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ are essentially equal at the end, as expected by Lemma 3.4, and $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$ converges to $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ fast. In order to prove our estimator $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$ converges fast in general, we use the secondly TX futures prices from November 23, 2011 to December 21, 2011. Table 5 shows the average relative difference between $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$ and $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$,

$$\frac{\left| \sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}} - \sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}} \right|}{\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}},$$

at every 2,000th second; so $i = 2{,}000, 4{,}000, \ldots, n$. Again, the last column is essentially zero. From that table, we can safely infer that the recursive two-scales estimator converges to $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ (which is expected to be close to $\sigma_{t_n}^{\mathrm{real}}$ from equation (8)) reasonably fast without waiting for all the data to arrive.

5 Conclusions

Zhang et al. (2005) propose a two-scales estimator that estimates the realized volatility in spite of the microstructure noise. However, it is a batch estimator and has to wait until all the data are received before any computation can commence. This paper proposes a recursive version of their two-scales estimator that runs in real-time and outputs the estimates for the realized volatility as the data arrive. Our findings are based on the simulation results and the secondly Taiwan Stock Exchange Capitalization Weighted Stock Index futures prices. They show that the recursive two-scales estimators gives essentially identical estimates as the batch-version two-scales estimators at the end of

each trading day. More importantly, the recursive two-scales estimator converges fast, well before the end of the trading day.

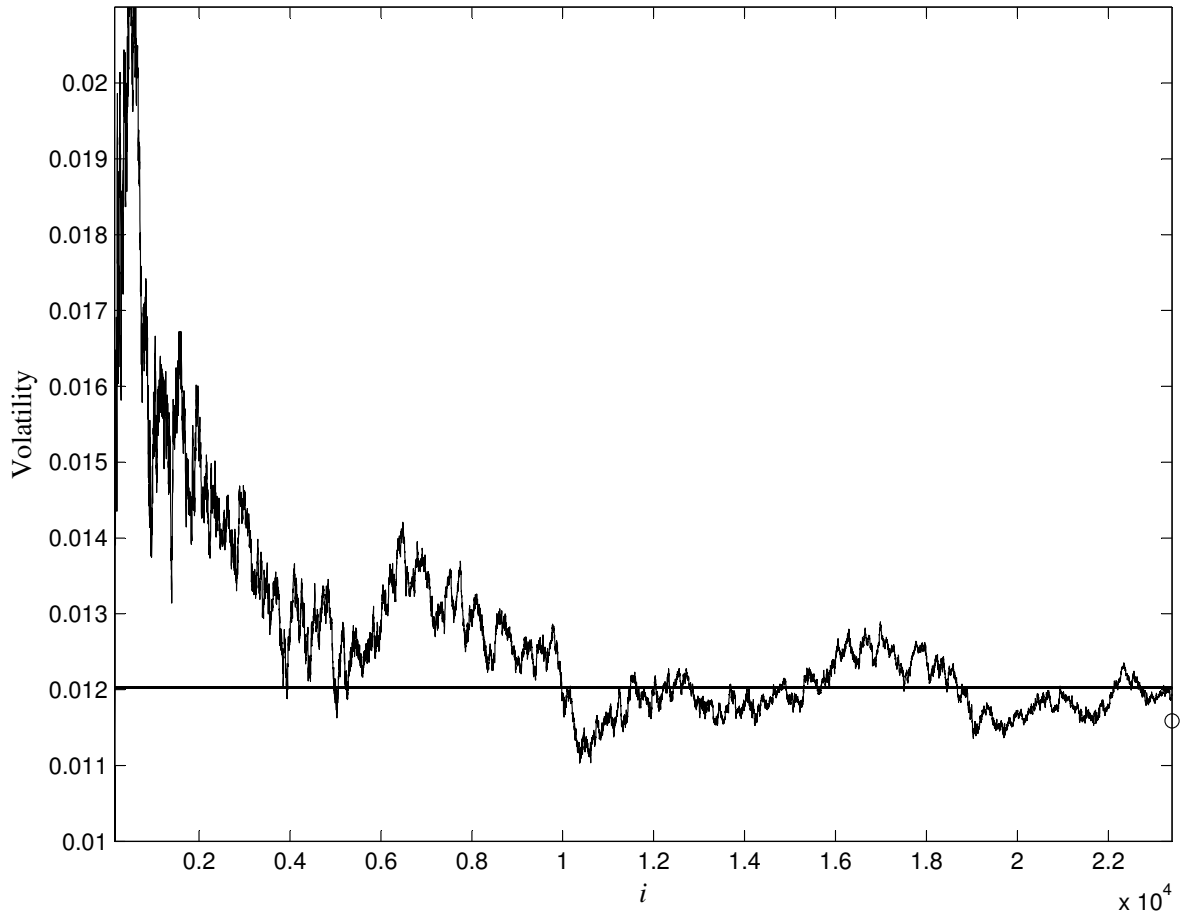

Figure 1: $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathbf{rec})}}$ **with geometric Brownian motion.** The horizontal line is $\sigma\sqrt{t_n} = 0.012$. The fluctuating line follows $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$, where $i = K, K+1, \ldots,$ 23,400 (second) and $K = 300$. The "o" marks $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ at $n =$ 23,400 (second) and $K = 300$.

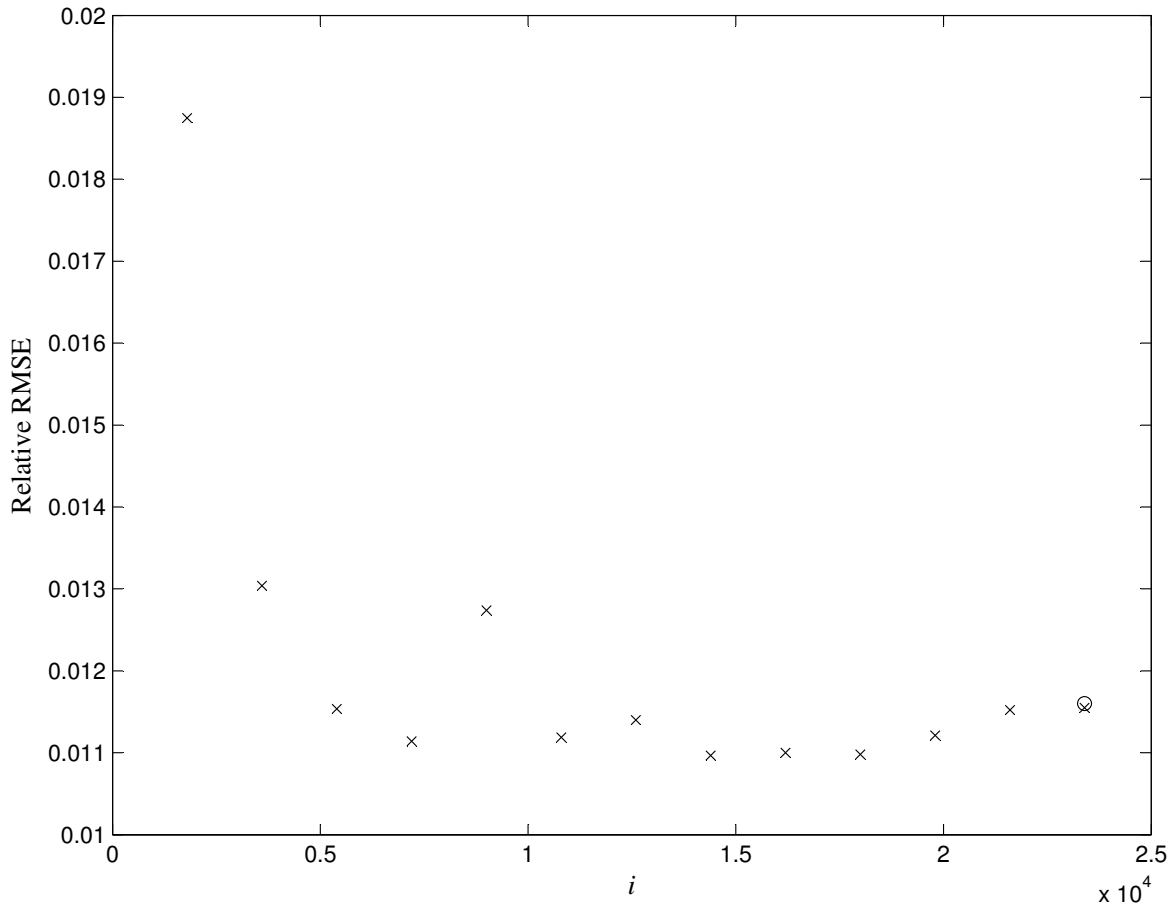

Figure 2: **Relative RMSE of $\langle \widehat{X, X} \rangle_{t_i}^{(\mathbf{rec})}$ with geometric Brownian motion.** An "x" marks the relative RMSE of $\langle \widehat{X, X} \rangle_{t_i}^{(\mathrm{rec})}$ at i = 1,800, 3,600, ..., 23,400 (second). The "o" marks the relative RMSE of $\langle \widehat{X, X} \rangle_{t_n}^{(\mathrm{batch})}$ at n = 23,400 (second).

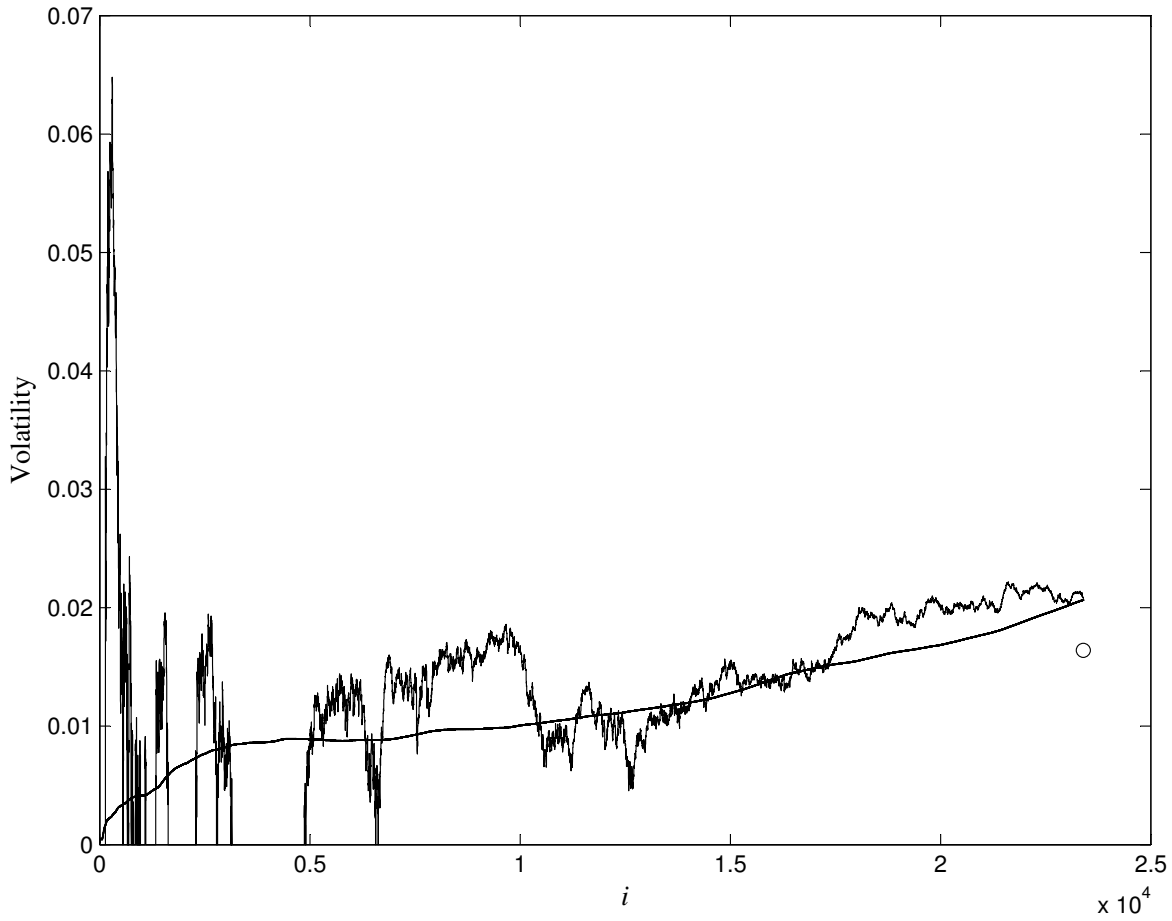

Figure 3: $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathbf{rec})}}$ **with Heston's model.** The smooth line is $\sigma_{t_i}^{\mathrm{real}}$ with Heston's model. The fluctuating line is $\sqrt{\langle \widehat{X,X} \rangle_{t_i}^{(\mathrm{rec})}}$, where $i = K, K+1, \ldots,$ 23,400 (second) and $K = 300$. The "o" marks $\sqrt{\langle \widehat{X,X} \rangle_{t_n}^{(\mathrm{batch})}}$ at $n =$ 23,400 (second) and $K = 300$.

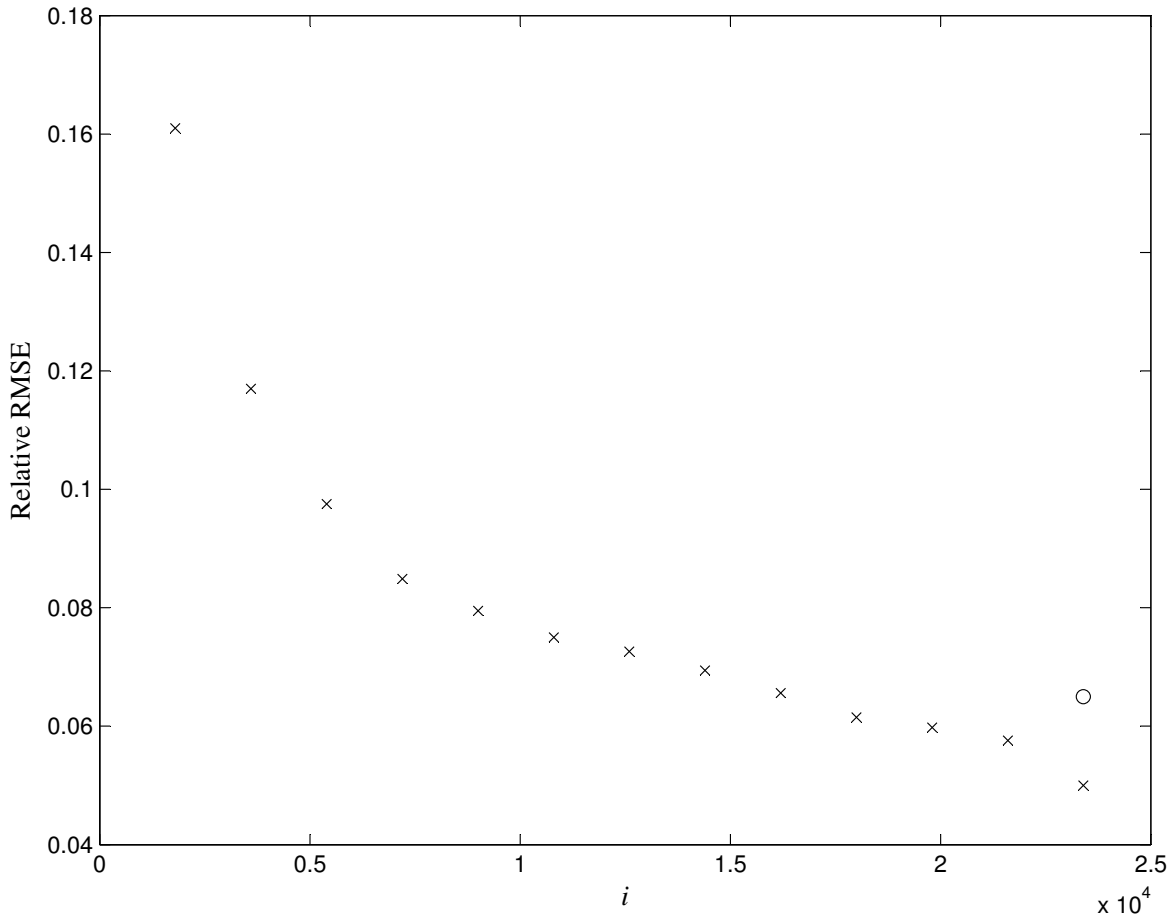

Figure 4: **Relative RMSE of $\langle \widehat{X, X} \rangle_{t_i}^{(\mathbf{rec})}$ with Heston's model.** An "x" marks the relative RMSE of $\langle \widehat{X, X} \rangle_{t_i}^{(\mathrm{rec})}$ at i = 1,800, 3,600, ..., 23,400 (second) and $K = 300$. The "o" marks the relative RMSE of $\langle \widehat{X, X} \rangle_{t_n}^{(\mathrm{batch})}$ at n = 23,400 (second) and $K = 300$.

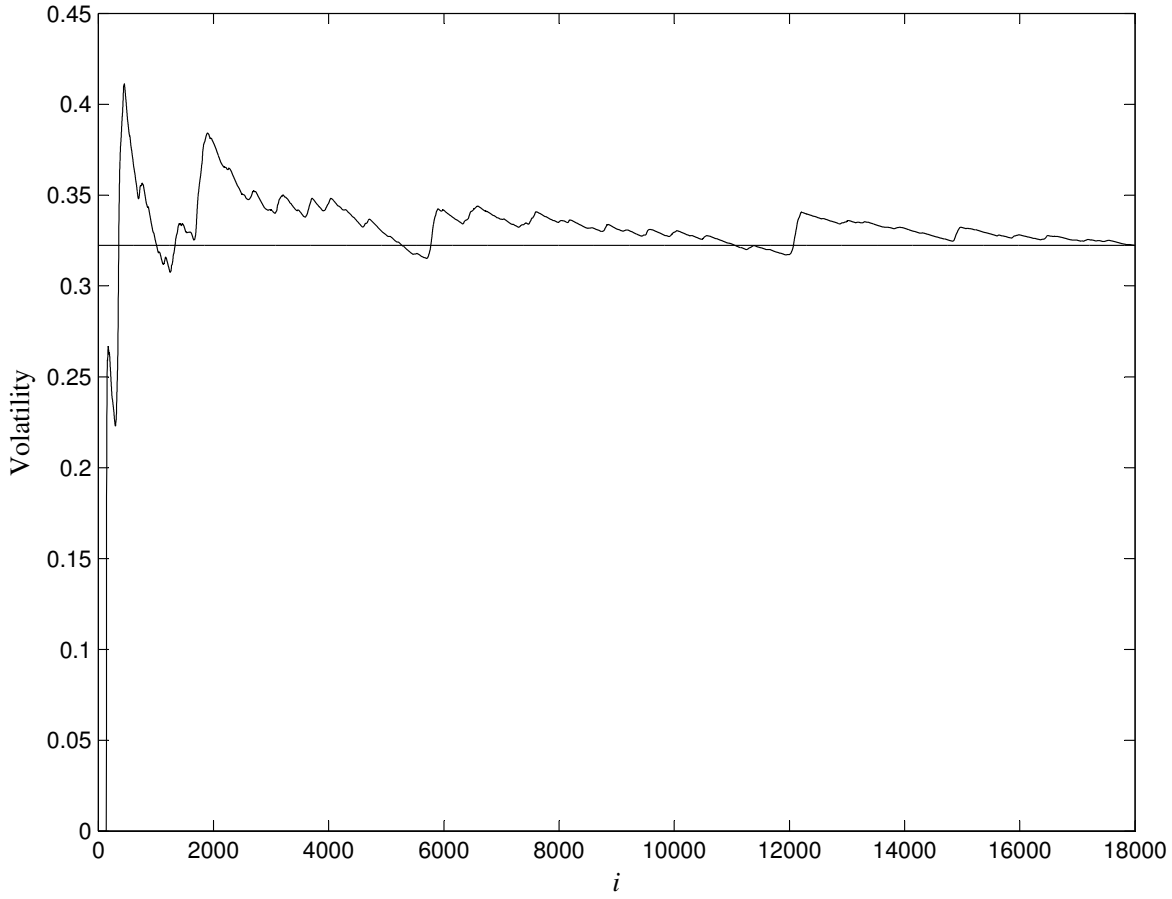

Figure 5: $\sqrt{\widehat{\langle X,X\rangle}_{t_i}^{(\mathbf{rec})}}$ **price with of TX futures for the period August 26, 2011 08:45:00 am–13:44:59 pm.** The horizontal line is $\sqrt{\widehat{\langle X,X\rangle}_{t_n}^{(\mathrm{batch})}}$ at $n = 18{,}000$ (second) and $K = 252$. The fluctuating line follows $\sqrt{\widehat{\langle X,X\rangle}_{t_i}^{(\mathrm{rec})}}$, where $i = K, K+1, \ldots, 18{,}000$ (second) and $K = 252$.

	$\widehat{\langle X, X\rangle}^{(\mathrm{batch})}_{t_n}$	$\widehat{\langle X, X\rangle}^{(\mathrm{rec})}_{t_n}$
Bias	1.40×10^{-8}	1.20×10^{-8}
RMSE	1.68×10^{-6}	1.60×10^{-6}
Relative RMSE	1.16×10^{-2}	1.15×10^{-2}

Table 1: **Monte Carlo simulation results for $n = 23{,}400$ with geometric Brownian motion.**

i	Relative RMSE of $\widehat{\langle X, X\rangle}^{(\mathrm{rec})}_{t_i}$
1800	1.87×10^{-2}
3600	1.30×10^{-2}
5400	1.15×10^{-2}
7200	1.11×10^{-2}
9000	1.27×10^{-2}
10800	1.12×10^{-2}
12600	1.14×10^{-2}
14400	1.09×10^{-2}
16200	1.10×10^{-2}
18000	1.09×10^{-2}
19800	1.12×10^{-2}
21600	1.15×10^{-2}
23400	1.15×10^{-2}

Table 2: **Relative RMSEs of $\widehat{\langle X, X\rangle}^{(\mathbf{rec})}_{t_i}$ with geometric Brownian motion.**

	$\widehat{\langle X, X\rangle}^{(\mathrm{batch})}_{t_n}$	$\widehat{\langle X, X\rangle}^{(\mathrm{rec})}_{t_n}$
bias	2.00×10^{-8}	1.99×10^{-8}
RMSE	9.40×10^{-6}	7.23×10^{-6}
relative RMSE	6.50×10^{-2}	5.00×10^{-2}

Table 3: **Monte Carlo simulation results for $n =$ 23,400 with Heston's model.**

i	Relative RMSE of $\widehat{\langle X, X\rangle}^{(\mathrm{rec})}_{t_i}$
1800	1.61×10^{-1}
3600	1.17×10^{-1}
5400	9.75×10^{-2}
7200	8.49×10^{-2}
9000	7.95×10^{-2}
10800	7.50×10^{-2}
12600	7.26×10^{-2}
14400	6.94×10^{-2}
16200	6.56×10^{-2}
18000	6.15×10^{-2}
19800	5.98×10^{-2}
21600	5.76×10^{-2}
23400	5.00×10^{-2}

Table 4: **Relative RMSEs of $\widehat{\langle X, X\rangle}^{(\mathbf{rec})}_{t_i}$ with Heston's model.**

i	2,000	4,000	6,000	8,000	10,000	12,000	14,000	16,000	18,000
2011/11/23	0.095	0.085	0.074	0.063	0.052	0.040	0.028	0.014	1.043×10^{-5}
2011/11/24	0.095	0.085	0.076	0.063	0.052	0.041	0.028	0.014	2.598×10^{-5}
2011/11/25	0.095	0.087	0.078	0.070	0.057	0.046	0.032	0.017	1.048×10^{-5}
2011/11/28	0.095	0.089	0.080	0.072	0.055	0.044	0.029	0.014	2.986×10^{-5}
2011/11/29	0.099	0.098	0.096	0.094	0.092	0.060	0.037	0.021	7.985×10^{-7}
2011/11/30	0.096	0.089	0.081	0.073	0.062	0.048	0.035	0.018	1.629×10^{-5}
2011/12/01	0.095	0.087	0.078	0.069	0.055	0.042	0.029	0.017	7.761×10^{-5}
2011/12/02	0.096	0.089	0.080	0.070	0.058	0.044	0.028	0.013	1.790×10^{-5}
2011/12/05	0.097	0.088	0.075	0.064	0.053	0.039	0.026	0.013	9.785×10^{-6}
2011/12/06	0.095	0.087	0.077	0.066	0.056	0.045	0.033	0.018	2.631×10^{-5}
2011/12/07	0.098	0.095	0.093	0.089	0.084	0.080	0.074	0.069	3.903×10^{-5}
2011/12/08	0.095	0.087	0.078	0.069	0.058	0.046	0.033	0.020	2.803×10^{-5}
2011/12/09	0.097	0.090	0.081	0.071	0.061	0.050	0.032	0.014	1.485×10^{-5}
2011/12/12	0.093	0.085	0.076	0.064	0.053	0.041	0.028	0.013	1.859×10^{-5}
2011/12/13	0.095	0.087	0.076	0.067	0.057	0.047	0.029	0.014	1.969×10^{-5}
2011/12/14	0.096	0.090	0.078	0.069	0.058	0.048	0.038	0.025	1.940×10^{-5}
2011/12/15	0.096	0.089	0.079	0.066	0.056	0.043	0.031	0.014	2.999×10^{-5}
2011/12/16	0.095	0.088	0.081	0.072	0.062	0.048	0.037	0.019	3.368×10^{-5}
2011/12/19	0.093	0.084	0.074	0.064	0.052	0.040	0.027	0.013	2.268×10^{-5}
2011/12/20	0.098	0.095	0.091	0.088	0.074	0.057	0.038	0.018	4.668×10^{-6}
2011/12/21	0.091	0.081	0.071	0.060	0.047	0.035	0.024	0.012	5.725×10^{-5}

Table 5: **The average relative differences between $\sqrt{\langle \widehat{X, X} \rangle_{t_i}^{(\mathbf{rec})}}$ and $\sqrt{\langle \widehat{X, X} \rangle_{t_n}^{(\mathbf{batch})}}$.** The data are based on the secondly TX futures prices from November 23, 2011 to December 21, 2011 at every 2,000th second.

References

Ait-Sahalia, Y. and J. Yu (2009). High frequency market microstructure noise estimates and liquidity measures. *The Annals of Applied Statistics 3*(1), 422–457.

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

Andersen, T., T. Bollerslev, F. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Areal, N. and S. Taylor (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets 22*(7), 627–648.

Bandi, F., J. Russell, and C. Yang (2008). Realized volatility forecasting and option pricing. *Journal of Econometrics 147*(1), 34–46.

Barndorff-Nielsen, O. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B 64*(2), 253–280.

Barndorff-Nielsen, O. and N. Shephard (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics 17*(5), 457–477.

Carr, P. and R. Lee (2007). Realized volatility and variance: options via swaps. *Risk 20*(5), 76–83.

Hansen, P. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics 24*(2), 127–161.

Hasbrouck, J. (1993). Assessing the quality of a security market: a new approach to transaction-cost measurement. *Review of Financial Studies 6*(1), 191–212.

Hendershott, T., C. Jones, and A. Menkveld (2011). Does algorithmic trading improve liquidity? *The Journal of Finance 66*(1), 1–33.

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

Kozhan, R. and W. Tham (2012). Execution risk in high-frequency arbitrage. *Management Science*, Online version at http://dx.doi.org/10.1287/mnsc.1120.1541.

Lux, T. (1996). The stable paretian hypothesis and the frequency of large returns: an examination of major german stocks. *Applied Financial Economics 6*(6), 463–475.

Maheu, J. and T. McCurdy (2002). Nonlinear features of realized FX volatility. *Review of Economics and Statistics 84*(4), 668–681.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *Journal of Finance 44*(1), 211–219.

Shreve, S. E. (2004). *Stochastic Calculus for Finance II: Continuous-Time Models*. New York: Springer.

Zhang, L., P. Mykland, and Y. Ait-Sahalia (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.

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