

QUANTITATIVE FINANCE  
RESEARCH CENTRE



UNIVERSITY OF  
TECHNOLOGY SYDNEY



## QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 299

January 2012

---

### Particle Filters for Markov Switching Stochastic Volatility Models

Yun Bao, Carl Chiarella and Boda Kang

---

ISSN 1441-8010

[www.qfrc.uts.edu.au](http://www.qfrc.uts.edu.au)

# PARTICLE FILTERS FOR MARKOV SWITCHING STOCHASTIC VOLATILITY MODELS

YUN BAO<sup>\*</sup>, CARL CHIARELLA<sup>#</sup> AND BODA KANG<sup>†</sup>

**ABSTRACT.** This paper proposes an auxiliary particle filter algorithm for inference in regime switching stochastic volatility models in which the regime state is governed by a first-order Markov chain. We propose an ongoing updated Dirichlet distribution to estimate the transition probabilities of the Markov chain in the auxiliary particle filter. A simulation-based algorithm is presented for the method which demonstrated that we are able to estimate a class of models in which the probability that the system state transits from one regime to a different regime is relatively high. The methodology is implemented to analyze a real time series: the foreign exchange rate of Australian dollars vs South Korean won.

**JEL Classification:** C61, D11.

**Keywords:** Particle filters, Markov switching stochastic volatility models, Sequential Monte Carlo simulation.

## 1. INTRODUCTION

Time-varying volatilities are broadly recognized as the nature of most financial time series data. Stochastic volatility (SV) models have been considered as a practical means to capture the time-varying variance, in particular the mean and log volatility have separate error terms.

Both autoregressive conditional heteroscedasticity (ARCH) models and stochastic volatility models are formulated under the belief of persistence of volatility to some

---

*Date:* Current Version January 30, 2012.

<sup>\*</sup> yunbao@gmail.com, Toyota Financial Services Australia;

<sup>#</sup> carl.chiarella@uts.edu.au; Finance Discipline, UTS Business School, University of Technology, Sydney, Australia;

<sup>†</sup> Corresponding author: boda.kang@uts.edu.au; Finance Discipline, UTS Business School, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia.

extent. Examples of empirical studies, which documented the evidence of volatility persistence, include Chou (1988), French, Schwert and Stambaugh (1987), Poon and Taylor (1992) and So, Lam and Li (1998). However as the economic environment changes, the magnitude of the volatility may shifts accordingly. Lamoureux and Lastrapes (1990) apply the generalized autoregressive conditional heteroscedasticity (GARCH) model to examining the persistence in volatility, while Kalimipalli and Susmel (2004) show that regime-switching SV model performs better than single-state SV model and GARCH family of models for short-term interest rates. So, Lam and Li (1998) advocate a Markov switching stochastic volatility (MSSV) model to measure the fluctuations in volatility according to economic forces.

Many methods have been developed to estimate Markov switching models. Examples of expectation maximization methods include Chib (1996), James et al. (1996), Elliott et al. (1998) and Elliott and Malcolm (2008). Examples of Bayesian Markov Chain Monte Carlo (MCMC) methods include Frühwirth-Schnatter (2001), Hahn et al. (2010), and Kalimipalli and Susmel (2004) among others. More recently, Fearnhead and Clifford (2003) as well as Carvalho and Lopes (2006) utilize particle filters to estimate Markov switching models.

The transition probabilities associated with MSSV model are the crucial parameters to estimate. They do not only determine the ergodic probability, but also determine how long the regime stays. Carvalho and Lopes (2006) combine a kernel smoothing, which is proposed by Liu and West (2001), and auxiliary particle filters (Pitt and Shephard (1999)) to estimate the parameters of the MSSV model. However this method is quite sensitive to the knowledge of prior distributions.

The modification that we made for the method in Carvalho and Lopes (2006) is to use an updated Dirichlet distribution to search reliable transition probabilities rather than applying a multi-normal kernel smoothing. The Dirichlet distribution has been used with MCMC in Chib (1996) and Frühwirth-Schnatter (2006). The combination of auxiliary particle filters and Dirichlet distribution of transition probabilities allows for an updating path of transition probabilities over time.

The rest of this paper is organized as follows. Section 2 presents a MSSV model. In section 3, the proposed method of auxiliary particle filter is described in detail.

Successively, the simulation results have been shown in Section 4, while the methodology is applied to the real data, namely the exchange rate of Australian dollars vs South Korean won, in Section 5. Conclusions are given in Section 6.

## 2. THE MARKOV SWITCHING STOCHASTIC VOLATILITY MODEL

Let  $y_t$  be a financial time series with a time-varying log volatility  $x_t$ . The observations  $y_1, \dots, y_t$  are conditionally independent with the latent variable  $x_t$ , and are normally distributed,

$$y_t = \exp\left(\frac{x_t}{2}\right) V_t,$$

and the log-volatility is assumed to be a linear autoregressive process

$$x_t = \alpha_{s_t} + \phi x_{t-1} + \sigma W_t,$$

where  $V_t$  and  $W_t$  are independent and identically distributed random variables of standard normal distribution. The drift parameter,  $\alpha = (\alpha_1, \dots, \alpha_k)$ , indicates the regime shifting effect. The elements in the set of regime switching are the labels for states, i.e.,  $s_t = \{1, 2, \dots, k\}$ , where  $k$  is the number of states. The transition probabilities are defined as

$$p_{ij} = \Pr(s_t = j | s_{t-1} = i) \quad \text{for } i, j = 1, 2, \dots, k,$$

where  $\sum_{j=1}^k p_{ij} = 1$ . In order to avoid the problem of identification, we assume

$$\alpha_{s_t} = \gamma_1 + \sum_{j=2}^k \gamma_j I_{jt},$$

where  $\gamma_1 \in R$ ,  $\gamma_i > 0$  for  $i > 1$ , and  $I_{jt}$  is the indicator function

$$I_{jt} = \begin{cases} 1 & \text{if } s_t = j, \\ 0 & \text{otherwise.} \end{cases}$$

In the MSSV model, the conditional probability distributions for observations  $y_t$ , and state variables  $x_t$  are given by

$$p(y_t | x_t) = (2\pi e^{x_t})^{-1/2} \exp\left(-\frac{y_t^2}{2e^{x_t}}\right),$$

$$p(x_t | x_{t-1}, \theta, s_t) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{(x_t - \alpha_{s_t} - \phi x_{t-1})^2}{2\sigma^2}\right),$$

where  $\Theta$  is the parameter vector, i.e.,  $\Theta = \{\alpha_1, \alpha_2, \sigma^2, \phi\}$ .

In this paper, we start with a simple MSSV model where there exist only two states, namely high- and low-volatility states, i.e.,  $k = 2$ . We also assume that only the mean of volatility shifts depending on the state, i.e.,  $\alpha_1 = \gamma_1$  and  $\alpha_2 = \gamma_1 + \gamma_2$ .

### 3. AUXILIARY PARTICLE FILTER

Let  $D_t$  denote a set of observations, i.e.,  $D_t = \{y_1, y_2, \dots, y_t\}$ . According to Bayes rule, the conditional probability density function of  $x_{t+1}$  is given by

$$(1) \quad p(x_{t+1}|D_{t+1}) = \frac{p(y_{t+1}|x_{t+1})p(x_{t+1}|D_t)}{p(y_{t+1}|D_t)}.$$

As shown in Eq.(1), the posterior density  $p(x_{t+1}|D_{t+1})$  consists of three components. They are the likelihood function  $p(y_{t+1}|x_{t+1})$ , the prior  $p(x_{t+1}|D_t)$ , and the denominator  $p(y_{t+1}|D_t)$ . The prior distribution for  $x_{t+1}$  is given by

$$p(x_{t+1}|D_t) = \int p(x_{t+1}|x_t) p(x_t|D_t) dx_t,$$

and the denominator is an integral

$$p(y_{t+1}|D_t) = \int p(y_{t+1}|x_t) p(x_t|D_t) dx_t.$$

Thus, the posterior distribution for  $x_{t+1}$  is a proportion of numerator of the right hand side of Eq.(1), i.e.,

$$p(x_{t+1}|D_{t+1}) \propto p(y_{t+1}|x_{t+1}) \int p(x_{t+1}|x_t) p(x_t|D_t) dx_t.$$

Suppose there is a set of particles  $\{x_t^1, \dots, x_t^N\}$  with discrete probabilities  $\{\omega_t^1, \dots, \omega_t^N\}$ , and  $\{x_t^j, \omega_t^j\}_{j=1}^N \sim p(x_t|D_t)$ . Therefore the posterior expectation is given by

$$(2) \quad \hat{p}(x_{t+1}|D_t) = \sum_{j=1}^N p(x_{t+1}|x_t^j) \omega_t^j.$$

Then at the time  $t + 1$  the posterior distribution is approximated by

$$(3) \quad \hat{p}(x_{t+1}|D_{t+1}) = p(y_{t+1}|x_{t+1}) \sum_{j=1}^N p(x_{t+1}|x_t^j) \omega_t^j.$$

Following Pitt and Shephard (1999), Eqs.(2) and (3) are called as empirical prediction density and the empirical filtering density, respectively. The auxiliary particle filter, which is also known as auxiliary sequential importance resampling (ASIR), adds an indicator in Eq.(3) to do the resampling. The indicator can be mean or mode, which

depends upon researchers' favor. As Pitt and Shephard (1999) claimed that if the measure of state variable does not vary over the particles, the ASIR is more efficient than general SIR. Since  $p(x_{t+1}|x_t)$  is more condense than  $p(x_{t+1}|D_t)$  compared to their conditional likelihood, using ASIR for MSSV is an good alternative to SIR.

In addition to tracking the unobserved state variables, we adopt a kernel smoothing in Liu and West (2001) to estimate the parameters, except for transition probabilities. The parameters using the kernel smoothing estimation are the volatility levels  $\alpha_1$  and  $\alpha_2$ , the volatility variance  $\sigma^2$ , and the volatility persistence  $\phi$ .

For the case of the kernel smoothing, the smooth kernel density form from West (1993) is given by

$$p(\Theta|D_t) \approx \sum_{j=1}^N \omega_t^j N(\Theta|\mathbf{m}_t^j, h^2 \mathbf{V}_t),$$

where  $\Theta$  is the parameter vector,  $h > 0$  is the smoothing parameter,  $\mathbf{m}$  and  $h^2 \mathbf{V}$  are the mean and variance of the multivariate normal density. Based on it, Liu and West (2001) proposed the conditional evolution density for  $\Theta$

$$p(\Theta_{t+1}|\Theta_t) \sim N(\Theta_{t+1}|a\Theta_t + (1-a)\bar{\Theta}_t, h^2 \mathbf{V}_t),$$

where  $a = \frac{3\delta-1}{2\delta}$ , and  $h^2 = 1 - a$ . The discount factor  $\delta$  is in  $(0, 1]$ ,  $\bar{\Theta}_t$  and  $\mathbf{V}_t$  are the mean and variance of the Monte Carlo approximation to  $p(\Theta|D_t)$ . Straightforwardly,  $\bar{\Theta}_t = \sum_{j=1}^N \omega_t^j \Theta_t^j$  and  $\mathbf{V}_t = \sum_{j=1}^N \omega_t^j (\Theta_t^j - \bar{\Theta}_t) (\Theta_t^j - \bar{\Theta}_t)'$ .

For the case of the transition probabilities, the parameters are updated by the Dirichlet distribution. Suppose the matrix of transition probabilities  $P$  is  $k \times k$ , and sum of each row is equal to 1. Then the  $i$ th row of  $P$  is denoted by  $p_{i.} = \{p_{i1}, \dots, p_{ik}\}$ , and let  $p_{i.}$  be the random variables of a Dirichlet distribution

$$p_{i.} \sim \mathcal{D}(\lambda_{i1}, \dots, \lambda_{ik}).$$

Each prior distribution of  $p_{i.}$  is independent upon  $p_{j.}$ ,  $i \neq j$ . According to Chib (1996), the updated distribution of  $P|S_t$  is also a Dirichlet distribution, where  $S_t = \{s_1, s_2, \dots, s_t\}$

$$p_{i.}|S_t \sim \mathcal{D}(\lambda_{i1} + n_{i1}, \dots, \lambda_{ik} + n_{ik}),$$

where  $n_{ik}$  is the number of one-step transition from state  $i$  to state  $k$  in sample  $S_t$ . In this case, we assume a two-state problem, i.e.,  $k = 2$ .

Initially, the starting-point parameters for each particle are draw from their prior distributions. Afterwards, in the case of Markov switching stochastic volatility, the starting-point state variable  $s_0^j$  is determined by ergodic probability. The ergodic probability for two states is  $\Pr(s_0^j = 1) = \frac{1-p_{22}^j}{2-p_{11}^j-p_{22}^j}$ , and  $\Pr(s_0^j = 2) = 1 - \Pr(s_0^j = 1)$ . If a random number from uniform distribution (0,1) is less than  $\Pr(s_0^j = 1)$ ,  $s_0^j = 1$ ; otherwise  $s_0^j = 2$ . Then given the state, the starting-point log-volatility is able to be drawn from a normal distribution,  $x_0^j \sim N\left(\frac{\alpha_{s_0^j}^j}{1-\phi_0^j}, (\sigma_0^j)^2\right)$ . Below is the algorithm for the ASIR with updated Dirichlet distribution afterwards.

While  $t \leq T$

Step 1: Determine the mean (our guessing)

For  $j = 1$  to  $N$ ,

$$\begin{aligned} s_{t+1}^{\mu^j} &= \arg \max_{i=1,2} \Pr(s_{t+1} = i | s_t^j) \\ \mu_{t+1}^j &= \alpha_{s_{t+1}^{\mu^j}} + \phi_t^j x_t^j \\ \omega_{t+1}^{\mu^j} &\propto p(y_{t+1} | \mu_{t+1}^j) \widetilde{\omega_t^j} \end{aligned}$$

End for

$$\text{Normalized importance weights } \widetilde{\omega_{t+1}^{\mu^j}} = \frac{\omega_{t+1}^{\mu^j}}{\sum_{j=1}^N \omega_{t+1}^{\mu^j}}.$$

Step 2: Resampling

For  $j = 1$  to  $N$ ,

$$\{\Theta_t^{j^l}, \mu_{t+1}^{j^l}, x_t^{j^l}, s_t^{j^l}\} = \text{resample} \left\{ \Theta_t^j, \mu_{t+1}^j, x_t^j, s_t^j, \widetilde{\omega_{t+1}^{\mu^j}} \right\},^1$$

$$\Theta_{t+1}^{j^l} \sim N(a\Theta_t^{j^l} + (1-a)\bar{\Theta}_t, h^2\mathbf{V}_t^l),$$

$$\text{updating } n_{t+1,ij}, \text{ and } p_{t+1,i} \sim \mathcal{D}(\lambda_{i1} + n_{t+1,i1}, \lambda_{i2} + n_{t+1,i2}), \quad i = 1, 2.$$

End for

Step 3: Sample the hidden variables  $\{s_{t+1}^j, x_{t+1}^j\}$

For  $j = 1$  to  $N$ ,

Filtering the conditional probability  $\Pr(s_{t+1}^j = k | \mathbf{y}^{t+1}, \Theta_{t+1}^j)$ ,  $k = 1$  or  $2$ .

(a) One-step ahead prediction probabilities

$$\Pr(s_{t+1}^j = k | \mathbf{y}^t, \Theta_t^j) = \sum_{i=1}^K \Pr(s_{t+1}^j = k | s_t^j = i) \Pr(s_t^j = k | \mathbf{y}^t, \Theta_t^j).$$

---

<sup>1</sup>Please refer to the Appendix for more details.

(b) Filtering for  $s_t$

$$\Pr(s_{t+1}^j = k | \mathbf{y}^{t+1}, \Theta_{t+1}^j) = \frac{p(y_{t+1} | s_{t+1}^j = k, \mathbf{y}^t, \Theta_{t+1}^j) \Pr(s_{t+1}^j = k | \mathbf{y}^t, \Theta_t^j)}{\sum_{i=1}^K p(y_{t+1} | s_{t+1}^j = i, \mathbf{y}^t, \Theta_{t+1}^j) \Pr(s_{t+1}^j = i | \mathbf{y}^t, \Theta_t^j)}.$$

(c) Draw  $\widetilde{p_{t+1}^j} \sim \text{uniform}(0, 1)$

if  $\widetilde{p_{t+1}^j} \leq \Pr(s_{t+1}^j = k | \mathbf{y}^{t+1}, \Theta_{t+1}^j)$ , then  $s_{t+1}^j = k$

otherwise,  $s_t$  would be another state.

Sample  $x_{t+1}^j \sim p(x_{t+1}^j | x_t^j, s_{t+1}^j, \Theta_{t+1}^j)$ ,

$$\omega_{t+1}^j \propto \frac{p(y_{t+1} | x_{t+1}^j)}{p(y_{t+1} | \mu_{t+1}^j)}$$

End for

Normalized importance weights  $\widetilde{\omega_{t+1}^j} = \frac{\omega_{t+1}^j}{\sum_{j=1}^N \omega_{t+1}^j}$

Step 4: Summarize  $\Theta_{t+1} = \widetilde{\omega_{t+1}^j} \Theta_{t+1}^j$ ,  $x_{t+1} = \widetilde{\omega_{t+1}^j} (x_{t+1}^j)'$ .

Step 5: Redo from Step 1 ( $t = t + 1$ ).

End while

#### 4. SIMULATION STUDY

In this section, we use four datasets to illustrate this method. Four datasets have been generated from the MSSV model with two states. The parameters of these four datasets are shown in Table 1, and the log-volatility of these datasets are shown in Figure 1. Parameter  $\Theta$  would be updated by a multinormal distribution, so we transfer some parameters as  $\log(\gamma_2)$ ,  $\log(\sigma^2)$ , and  $\log\left(\frac{\phi}{1-\phi}\right)$ . The first and the fourth sample let the transition probability matrix concentrate in the diagonal, but the persistence parameter varies. Thus, the unconditional mean for the volatility are different. The second and third samples have relative lower diagonal transition probabilities, which means the volatility regime changes frequently.

The starting-point values of the estimation are determined by their prior distribution with the central tendency close to their true values. Since  $k = 2$ , the transition probability matrix is  $\begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix}$ , where  $p_{11}$  is the probability of state of one given the previous state is one, and  $p_{11}$  is the probability state of two given the previous state is two. The discount rate  $\delta$  is set at 0.75 which implies  $a = 0.8333$  and  $h = 0.4082$ .



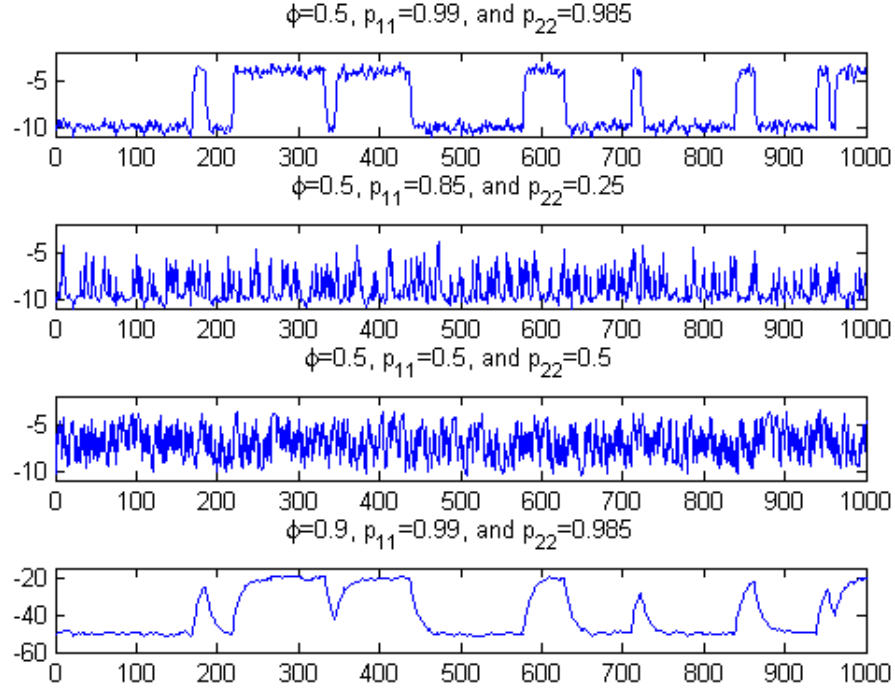


FIGURE 1. Log volatility for four dataset

TABLE 1. Parameter values for the simulated dataset

Parameter	Test 1	Test 2	Test 3	Test 4
$\gamma_1$	-5.0	-5.0	-5.0	-5.0
$\gamma_2$	3.0	3.0	3.0	3.0
$\sigma^2$	0.1	0.1	0.1	0.1
$\phi$	0.5	0.5	0.5	0.9
$p_{11}$	0.99	0.85	0.5	0.99
$p_{22}$	0.985	0.25	0.5	0.985

Figures 2-5 shows the simulation results of four datasets, and each figure has four graphs. The first/top graph shows the simulated time series data, and the second graph contains the simulated Markov Chain (the shifting states). The third graph compares the simulated log-volatility and estimated log-volatility. The last/bottom graph shows

the estimated probability that state is in high volatility regime given by the previous information.

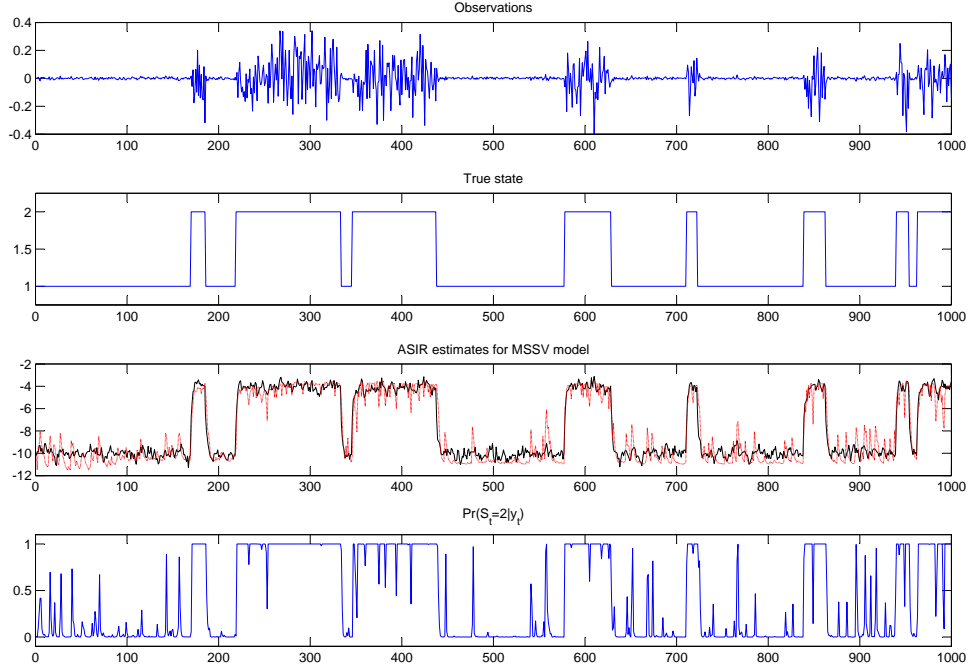


FIGURE 2. Simulated data 1: top graph—simulated time series  $y_t$ , second graph—true regime variables  $s_t$ , third graph—true (blank line) and estimated log-volatilities (red line), and bottom graph—estimated probability  $Pr(s_t = 2|D_t)$ .

The sequential estimation of the parameters of these four simulated datasets are shown in Figures 6-9, separately. The black lines denote the modes of the parameters, the green lines represent the 10% and 90% quantiles, and the red lines represent the true values of the parameters. In addition, the modes of the parameters are summarized in Table 2.

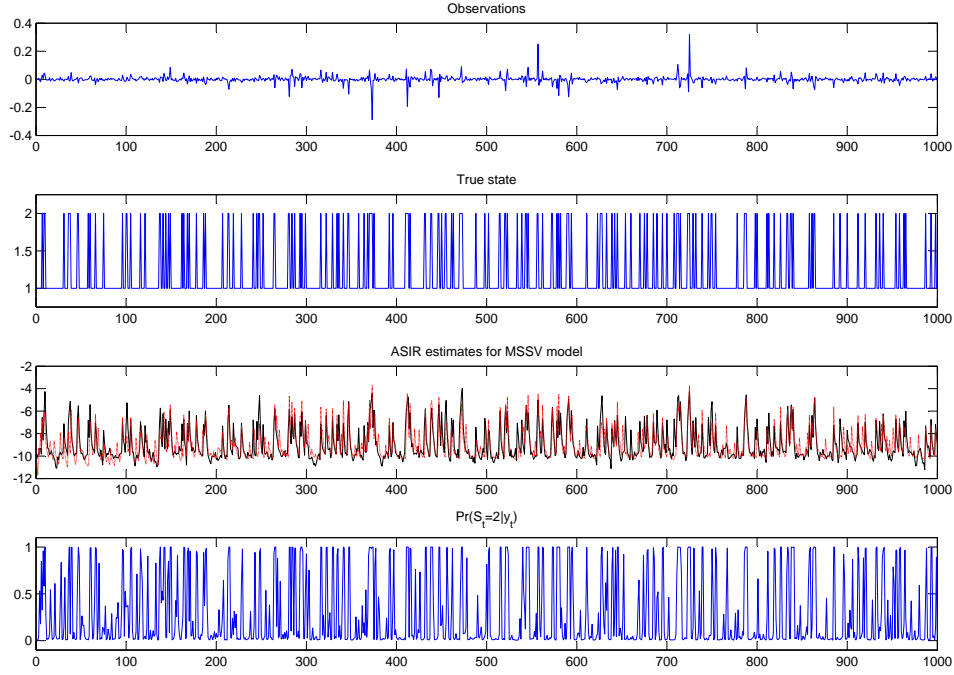


FIGURE 3. Simulated data 2: top graph—simulated time series  $y_t$ , second graph—true regime variables  $s_t$ , third graph—true (blank line) and estimated log-volatilities (red line), and bottom graph—estimated probability  $Pr(s_t = 2|D_t)$ .

TABLE 2. Posterior modes of the parameters

	Test 1	Test 2	Test 3	Test 4
	mode	mode	mode	mode
$\gamma_1$	-5.0604	-4.9303	-4.9163	-5.0800
$\gamma_2$	3.2833	2.9811	3.0863	3.1540
$\sigma^2$	0.0880	0.1292	0.1332	0.1149
$\phi$	0.5335	0.5170	0.4874	0.9012
$p_{11}$	0.9712	0.7841	0.5816	0.9782
$p_{22}$	0.9669	0.3690	0.5529	0.9747

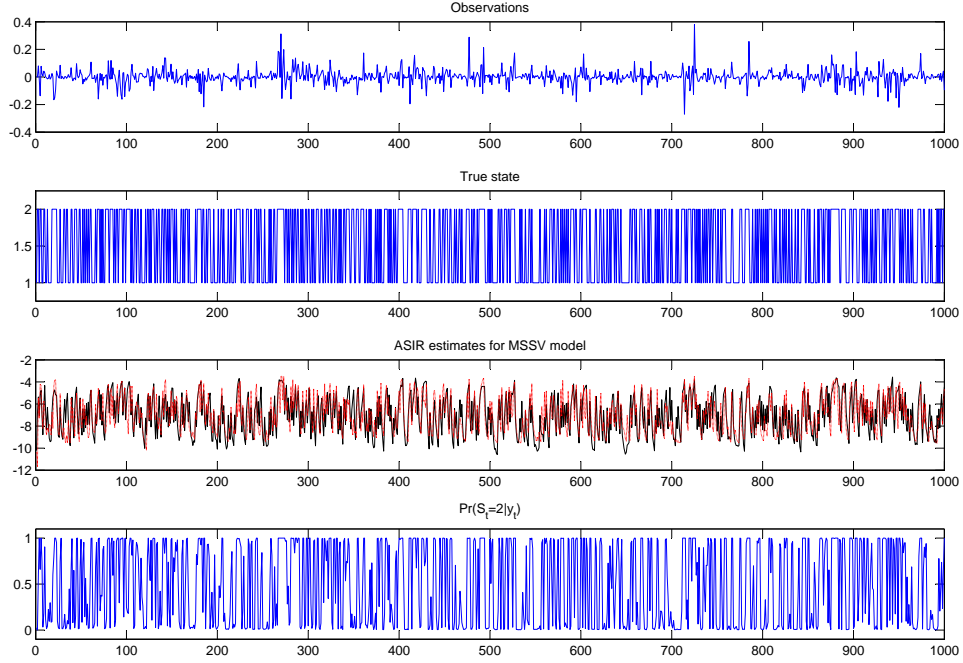


FIGURE 4. Simulated data 3: top graph—simulated time series  $y_t$ , second graph—true regime variables  $s_t$ , third graph—true (blank line) and estimated log-volatilities (red line), and bottom graph—estimated probability  $Pr(s_t = 2|D_t)$ .

## 5. APPLICATION TO REAL DATA

In this section, the proposed algorithms is applied to South Korea Won/AU Dollar exchange rate from 02/01/1995 to 31/12/1999 (1014 observations). This period includes the Asian financial crisis in 1997, where South Korea suffers a lot.

Figure 10 shows the log difference of the exchange rete, the estimated log volatility and the estimated probability that state is equal to 2. According to Figure 10, the volatility of the exchange rate becomes big during the Asian financial crisis, and switches regimes frequently afterwards. In other words, the stable movement of the exchange rate does not last for longer than that before the crisis.

The sequential estimation of the exchange rate is shown in Figure 11, and the updated values of mode and quantiles are shown in Table 3. The estimate of the persistence

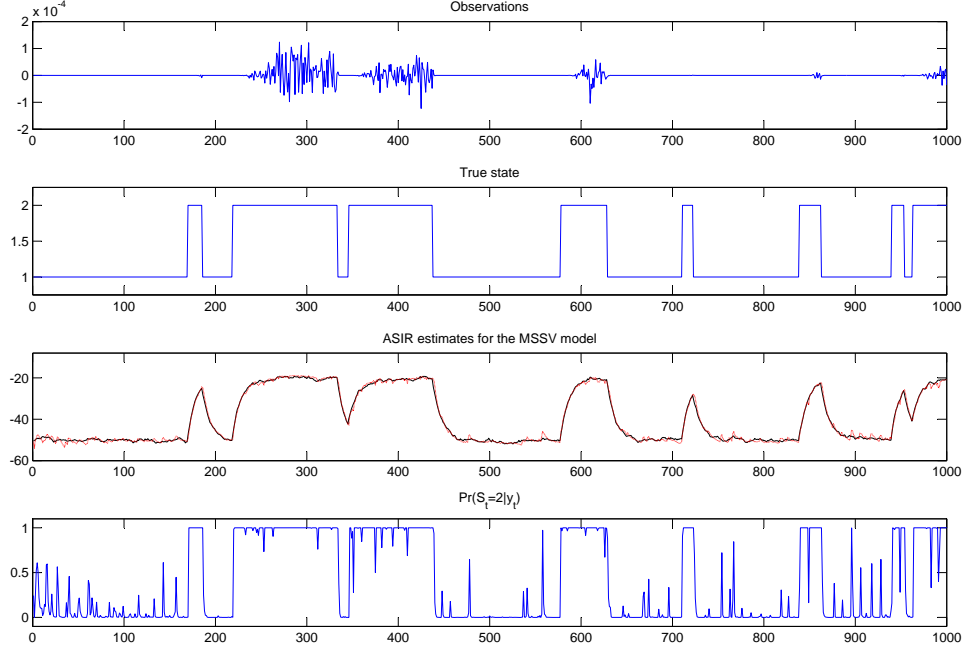


FIGURE 5. Simulated data 4: top graph—simulated time series  $y_t$ , second graph—true regime variables  $s_t$ , third graph—true (blank line) and estimated log-volatilities (red line), and bottom graph—estimated probability  $Pr(s_t = 2|D_t)$ .

parameter  $\phi$  is 0.5, which is not overestimated, according to So et al.'s (1998) and Carvalho and Lopes' (2006) findings. The diagonal elements of the transition probability descends over time, in particular after the Asian financial crisis. In other words, the exchange rate of Australian dollar vs South Korean won becomes more volatile after this crisis.

TABLE 3. The updated posterior modes, 10% and 90% quantiles of the parameters

	$\gamma_1$	$\gamma_2$	$\sigma^2$	$\phi$	$p_{11}$	$p_{22}$
Mode	-5.4294	2.2087	0.2533	0.5216	0.9629	0.9504
10%	-5.4885	2.1534	0.2428	0.5159	0.9548	0.9396
90%	-5.3708	2.2668	0.2648	0.5272	0.9699	0.9599

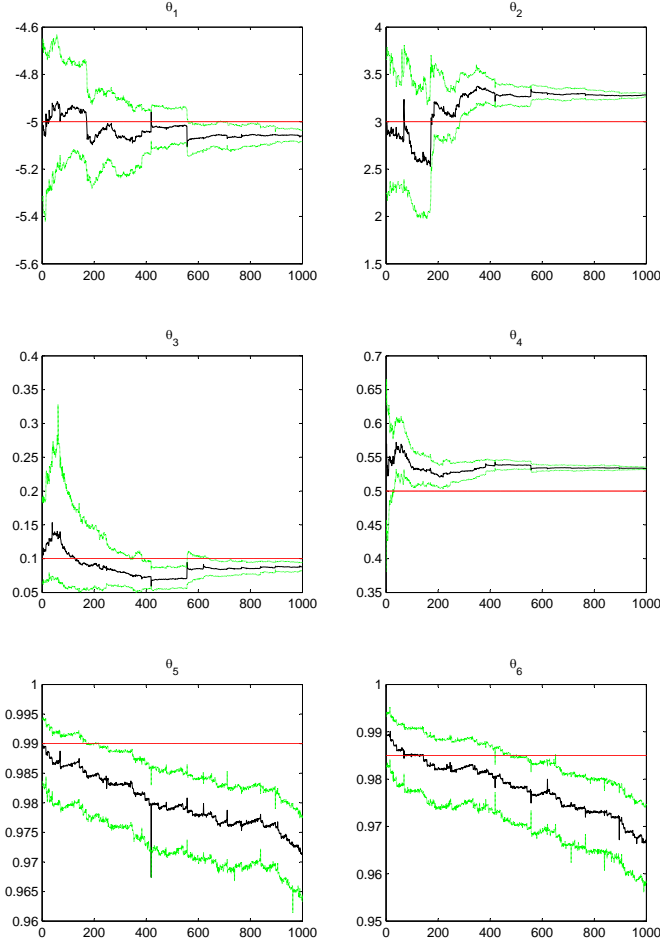


FIGURE 6. Posterior mode, 10 and 90 percent quantiles of  $\Theta$  for the first simulated data:  $\theta_1 = \gamma_1$ ,  $\theta_2 = \gamma_2$ ,  $\theta_3 = \sigma^2$ ,  $\theta_4 = \phi$ ,  $\theta_5 = p_{11}$ ,  $\theta_6 = p_{22}$ .

## 6. DIAGNOSTIC FOR SAMPLING IMPROVEMENT

Following Carpenter, Clifford and Fearnhead (1999), we implement effective sample size to assess the performance of the particle filtering. The comparison of the effective sample size of the ASIR with multi-normal kernel smoothing and the (proposed) ASIR with updated Dirichlet distribution is presented to show whether the proposed ASIR is robust rather than previous one.

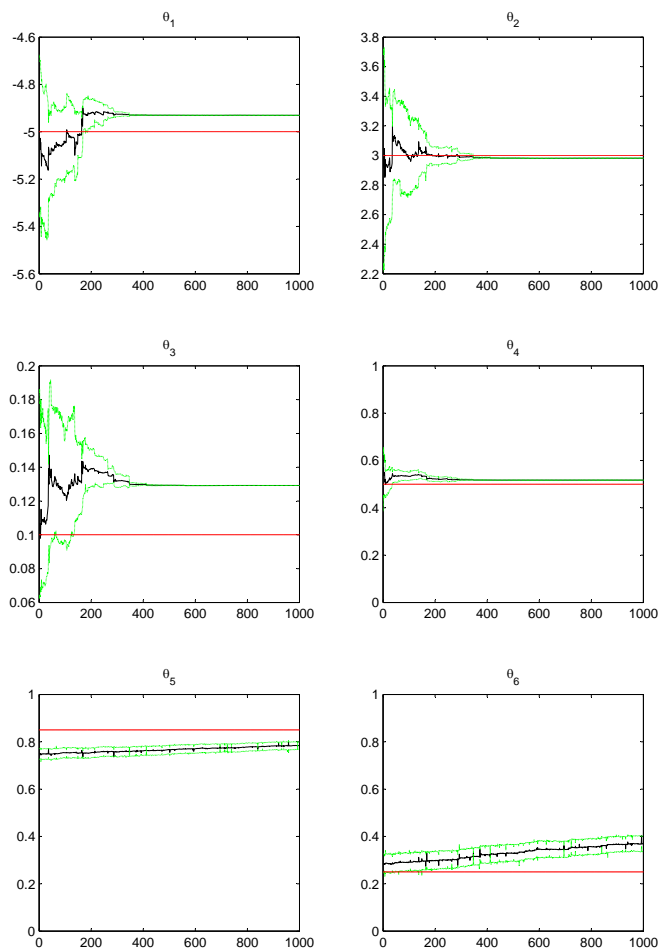


FIGURE 7. Posterior mode, 10 and 90 percent quantiles of  $\Theta$  for the second simulated data:  $\theta_1 = \gamma_1$ ,  $\theta_2 = \gamma_2$ ,  $\theta_3 = \sigma^2$ ,  $\theta_4 = \phi$ ,  $\theta_5 = p_{11}$ ,  $\theta_6 = p_{22}$ .

The algorithm of calculating the effective sample size is show below. Suppose  $g(x_t)$  is a measure of  $x_t$ , and its expectation is

$$\theta = \int g(x_t)p(x_t|y_t)dx_t.$$

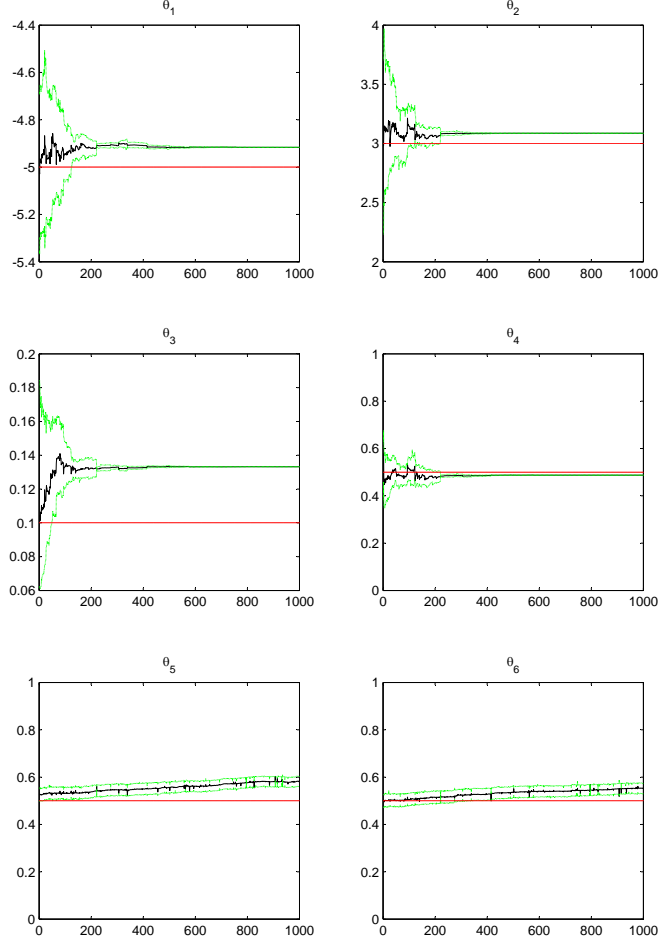


FIGURE 8. Posterior mode, 10 and 90 percent quantiles of  $\Theta$  for the third simulated data:  $\theta_1 = \gamma_1$ ,  $\theta_2 = \gamma_2$ ,  $\theta_3 = \sigma^2$ ,  $\theta_4 = \phi$ ,  $\theta_5 = p_{11}$ ,  $\theta_6 = p_{22}$ .

The discrete approximation of the  $\theta$  is given by

$$z_t = \sum_{i=1}^N \omega_t^i g(x_t^i),$$

and the variance of mature  $g(x_t)$  is given by

$$v_t = \sum_{i=1}^N \omega_t^i g^2(x_t^i) - z_k^2.$$



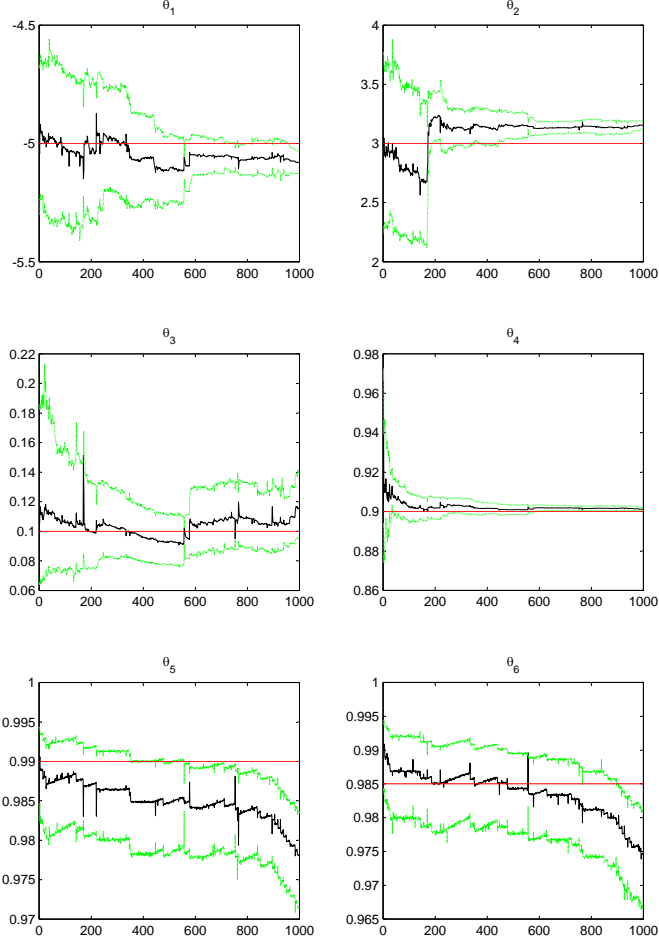


FIGURE 9. Posterior mode, 10 and 90 percent quantiles of  $\Theta$  for the fourth simulated data:  $\theta_1 = \gamma_1$ ,  $\theta_2 = \gamma_2$ ,  $\theta_3 = \sigma^2$ ,  $\theta_4 = \phi$ ,  $\theta_5 = p_{11}$ ,  $\theta_6 = p_{22}$ .

Suppose the independent filter has been run for  $M$  times, and the value of  $z_t$  and  $v_t$  can be calculated. Then the effective sample size is given by

$$N_t^* = \frac{\overline{v_t}}{\frac{1}{M} \sum_{j=1}^M (z_t^j - \bar{z}_t)}.$$

The greater the value of the effective sample size, the more likely the filter is reliable. The comparison of the proposed model and the model with kernel smoothing for transition

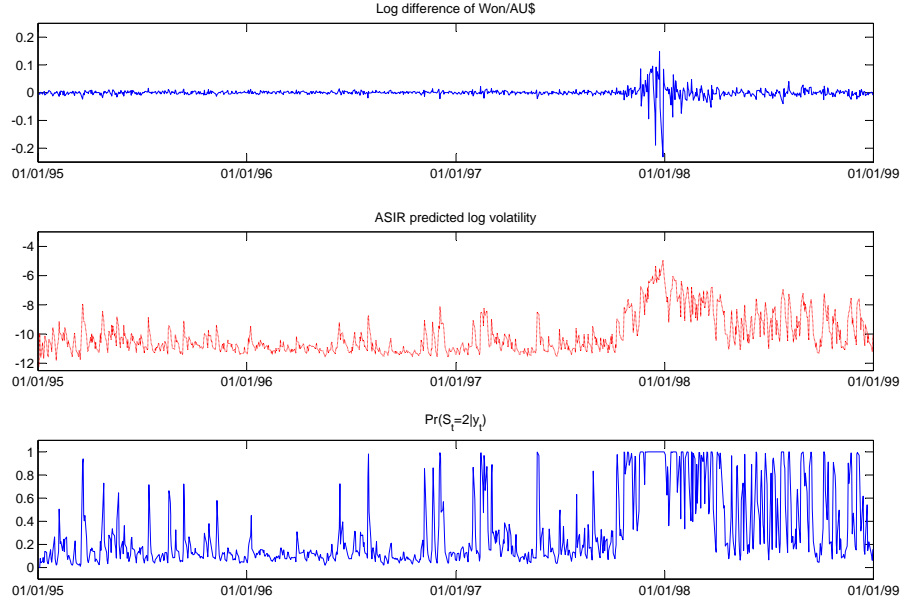


FIGURE 10. The exchange rate: top graph—the observed time series  $y_t$ , second graph—the estimated log-volatilities, and bottom graph—estimated probability  $Pr(s_t = 2|D_t)$ .

probability is shown in Table 4. According to the results, as time goes by, the use of Dirichlet updating is more reliable than the use of Kernel smoothing method.

TABLE 4. Comparison of effective sample sizes

Time horizon	100	200	300	400	500	600	700	800	900	1000
Dirichlet updating	32	33	37	2	3	27	15	18	21	8
Kernel smoothing	70	27	62	1	2	20	8	2	11	1

## 7. CONCLUSION

In this article we developed and implemented an auxiliary particle filter algorithm to estimate a univariate regime switching stochastic volatility model. The use of simulated examples was intended to show the performance of the proposed method.

In particular, in terms of estimating the transition probabilities of the Markov Chain, we modified the method in Carvalho and Lopes (2006) to use an updated Dirichlet distribution to search reliable transition probabilities rather than applying a multi-normal kernel smoothing which is only able to have a good estimate when the probability that the system state transits from one regime to a different regime is rather low. The combination of auxiliary particle filters and Dirichlet distribution of transition

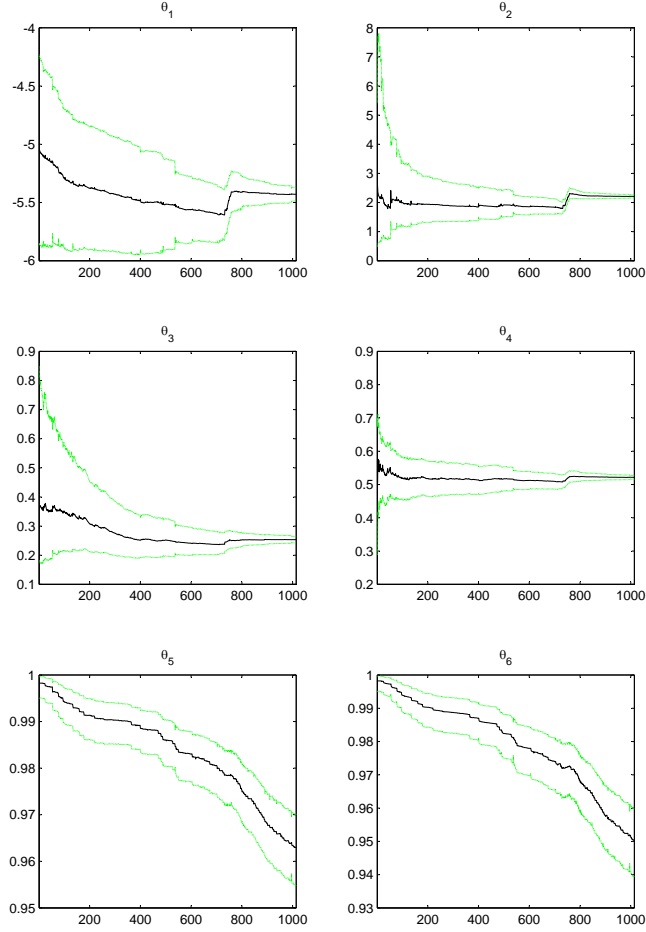


FIGURE 11. Posterior mode, 10% and 90% quantiles of  $\Theta$  for the exchange rate from 02/01/1995 to 31/12/1999:  $\theta_1 = \gamma_1$ ,  $\theta_2 = \gamma_2$ ,  $\theta_3 = \sigma^2$ ,  $\theta_4 = \phi$ ,  $\theta_5 = p_{11}$ ,  $\theta_6 = p_{22}$ .

probabilities allows for an updating path of transition probabilities over time and also will accommodate the cases that the probability that the system state transits from one regime to a different regime is relatively high. This feature is often observed in the Energy, Commodity or FX market.

## 8. APPENDIX: RESAMPLING

We adopt the systematic resampling method, which is mentioned in Ristic, Arulampalam and Gordon (2004). The algorithm of  $\{\Theta_t^{j^l}, \mu_{t+1}^{j^l}, x_t^{j^l}, s_t^{j^l}\} = \text{resample} \left\{ \Theta_t^j, \mu_{t+1}^j, x_t^j, s_t^j, \widetilde{\omega_{t+1}^{\mu^j}} \right\}$  can be shown below.

Step 1: Draw  $u_0 \sim \text{uniform}(0, 1)$ ,

construct a cdf of importance weights, i.e.,  $c_j = \sum_{j=1}^N \widetilde{\omega_{t+1}^{\mu^j}}$ ,

$i = 1$ .

Step 2: For  $j = 1$  to  $N$

$u_j = u_{j-1} + 1/n_{\text{parts}}$ ,

If  $u_j > c_j$ ,

$i = i + 1$ ,

else  $i = i$ ,

End if.

$\Theta_t^{j^l} = \Theta_t^i, \quad \mu_{t+1}^{j^l} = \mu_{t+1}^i, \quad x_t^{j^l} = x_t^i, \quad s_t^{j^l} = s_t^i,$

End for.

## REFERENCES

- [1] Carpenter, J., Clifford, P., and Fearnhead, P. (1999), "An Improved Particle Filter for Nonlinear Problems," *IEEE Proceedings-Radar, Sonar and Navigation*, 146(1), 2-7.
- [2] Carvalho, C.M. and Lopes, H.F. (2007), "Simulation-based Sequential Analysis of Markov Switching Stochastic Volatility Models," *Computational Statistics and Data Analysis*, 51(9), 4526-4542.
- [3] Chib, S. (1996), "Calculating Posterior Distributions and modal Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79-97.
- [4] Chou, R.Y. (1988), "Volatility Persistence and Stock Valuation: Some Empirical Evidence Using GARCH," *Journal of Applied Econometrics*, 3, 279-294.
- [5] Elliott, R.J., Hunter, W. and Jamieson, B.M. (1997), "Drift and Volatility Estimation in Discrete Time," *Journal of Economic Dynamics and Control*, 22, 209-218.
- [6] Elliott, R.J. and Malcolm, W.P. (2008), "Discrete-Time Expectation Maximization Algorithms for Markov-Modulated Poisson Processes," *IEEE Transactions on Automatic Control*, 53(1), 247-256.

- [7] French, K.R., Schwert, G.W. and Stambaugh, R.F. (1987), "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19, 3-30.
- [8] Frühwirth-Schnatter, S. (2006), *Finite Mixture and Markov Switching Models*. Springer, Berlin.
- [9] Jacquier, E., Polson, N.G. and Rossi, P.E. (1994), "Bayesian Analysis of Stochastic Volatility Models," (with discussion), *Journal of Business & Economics Statistics*, 12, 371-417.
- [10] James, M.R., Krishnamurthy, V. and Le Gland, F. (1996), "Time Discretization of Continuous-Time Filters and Smoothers for HMM Parameter Estimation," *IEEE Transactions on Information Theory*, 42(2), 593-605.
- [11] Kalimipalli, M. and Susmel, R. (2004), "Regime-switching stochastic volatility and short-term interest rates," *Journal of Empirical Finance*, 11, 309-329.
- [12] Lamoureux, C.G. and Lastrapes, W.D. (1990), "Persistence in Variance Structural Change, and the GARCH model," *Journal of Business & Economics Statistics*, 8(2), 225-234.
- [13] Liu, J. and West, M. (2001), "Combined Parameter and State Estimation in Simulation-based Filtering," In: de Freitas, N., Gordon, N. and Doucet, A. (Eds.), *Sequential Monte Carlo Methods in Practice*. Springer, Berlin.
- [14] Hahn, M., Frühwirth-Schnatter, S. and Sass, J. (2010), "Markov Chain Monte Carlo Methods for Parameter Estimation in Multidimensional Continuous Time Markov Switching Models," *Journal of Financial Econometrics*, 8(1), 88-121.
- [15] Pitt, M.K. and Shephard, N. (1999), "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 590-599.
- [16] Poon, S.H. and Taylor, S.J. (1992), "Stock returns and Volatility: An Empirical Study of U.K. Stock Market," *Journal of banking and Finance*, 16, 37-59.
- [17] Ristic, B., Arulampalam S. and Gordon N. (2004), *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Artech House, Norwood, MA, USA.
- [18] So, M.K.P., Lam, K., and Li, W.K. (1998), "A Stochastic Volatility Model with Markov Switching," *Journal of Business & Economics Statistics*, 16(2), 244-253.
- [19] West, M. (1993), "Approximating posterior distributions by mixture," *Journal of Royal Statistical Society*, 55, 409-422.