Hidden Markov Models and Sequential Monte-Carlo methods

January 13, 2018

Authors:
CUI Shasha
DING Wensi
Sun Quan

Professor:
Chopin Nicolas

École nationale
de la statistique
et de l'administration
économique

université
PARIS-SACLAY

Contents

1 Introduction

In this project, we applied a new **resampling** method introduced in the paper to have **continuous** likelihood evaluations within the **particle filtering** process. This method is implemented for four different models which are presented below.

2 AR(1) plus noise model

$$y_t = x_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2)$$
$$x_{t+1} = \mu + \phi(x_t - \mu) + \eta_t, \eta_t \sim N(0, \sigma_\epsilon^2)$$

First, we experiment on simulated dataset with $\sigma_\epsilon^2 = 2, \sigma_\eta^2 = 0.02, \phi = 0.975, \mu = 0.5, T = 150.$ and get the log likelihood as shown in figure 1. In order to control the stability of the estimations, the results of 50 runs with different random seeds are displayed below.

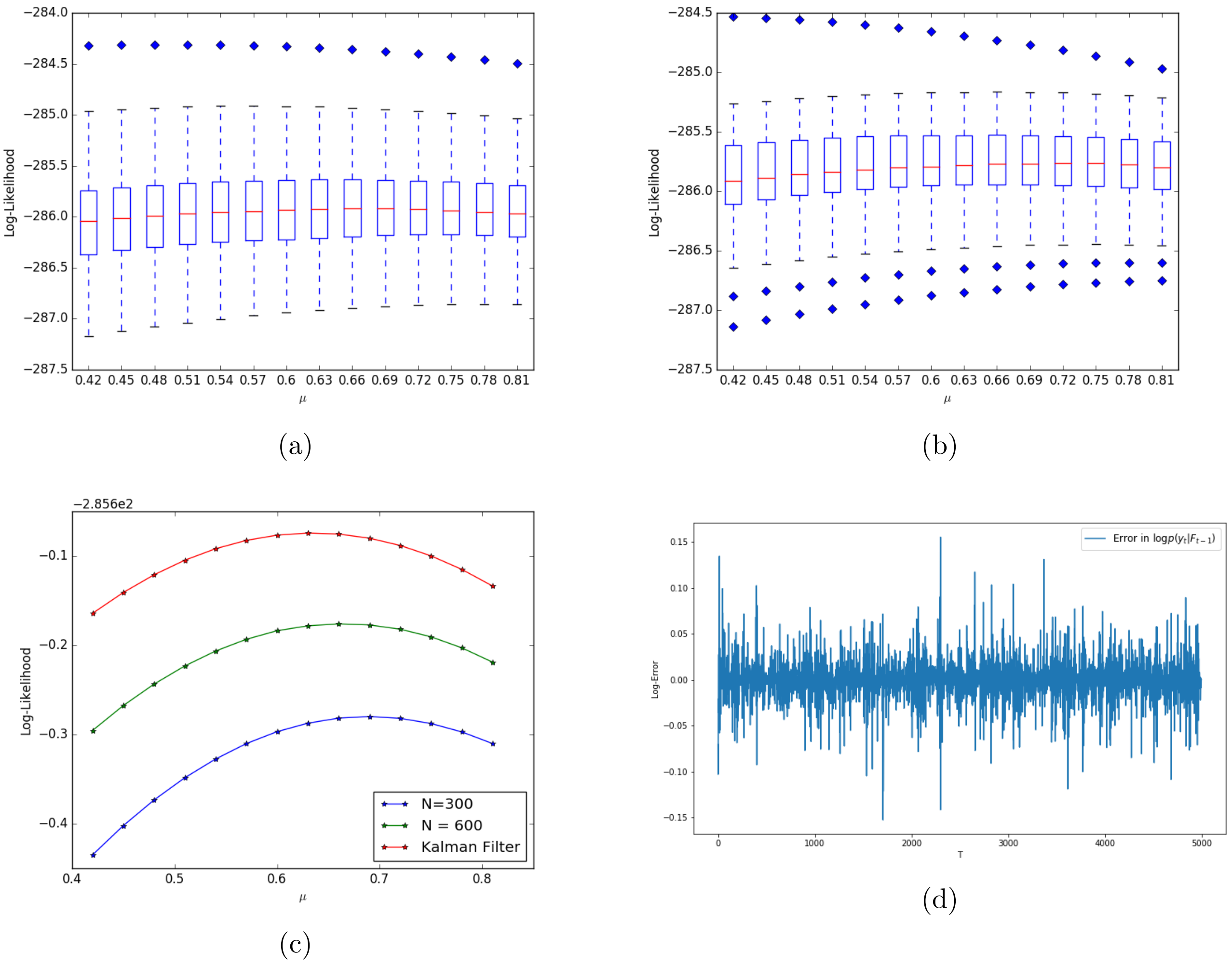


Figure 1: Fixed dataset, boxplot of 50 runs with different random seeds, T = 150, N = 300 (left), N = 600 (right)

In order to assess the performance of the CSIR method, we have implemented both CSIR and kalman filter for AR(1) model and considered output of kalman filter as true value of estimates. As suggested in the paper,we simulate one series of length $T = 5000$ and then evaluate the performance for $N = 300$. In figure 1d, we display the error between the log-likelihood estimate $\hat{l}_t$ from CSIR method and the corresponding true value $l_t = \log p(y_t|\mathcal{F}_{t-1})$ from kalman filter. The variation in this error, does not increase and appears apparently stable over time. Futhermore, we also examine sliced log-likelihood and likelihood for all three parameters (holding the other parameters at their true values), as shown in figure 2. Clearly, considering the fact that the length of series is relatively small and also a small $N = 300$, the estimation of both likelihoods and three model parameters by particle filtering are very good.

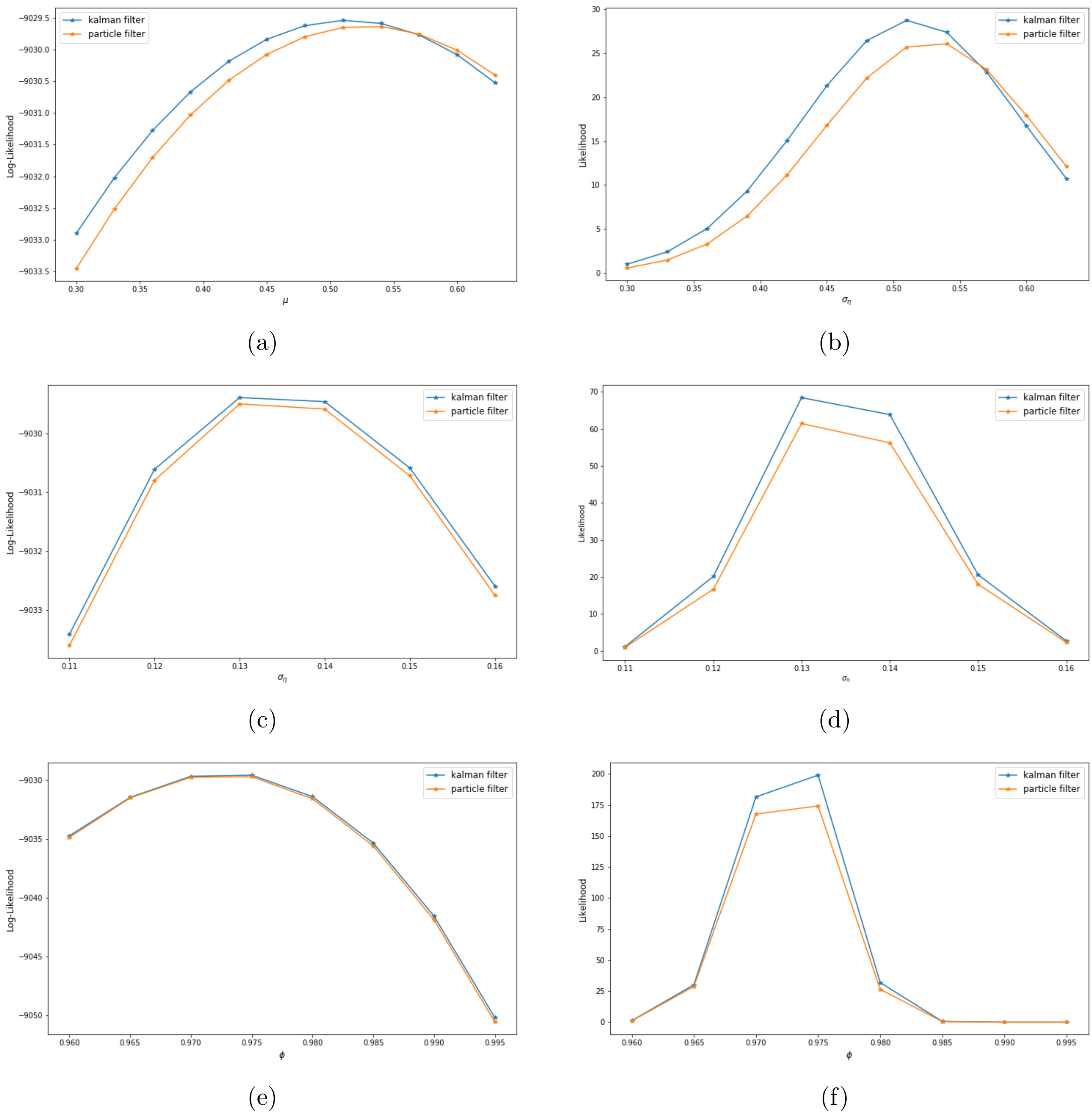
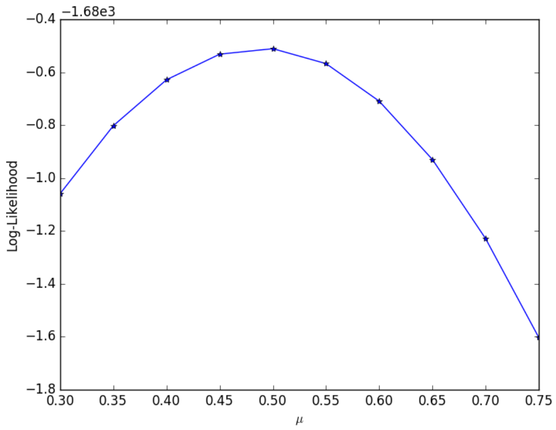


Figure 2: The true and estimated log-likelihood (left) and likelihood (right) profiles for the AR(1) model plus noise for (from top to bottom) σ_η, ϕ, μ. Length of series $T = 5000$. True parameters: $\sigma_\eta = \sqrt{0.02}, \mu = 0.5, \phi = 0.975$ and $N = 300$

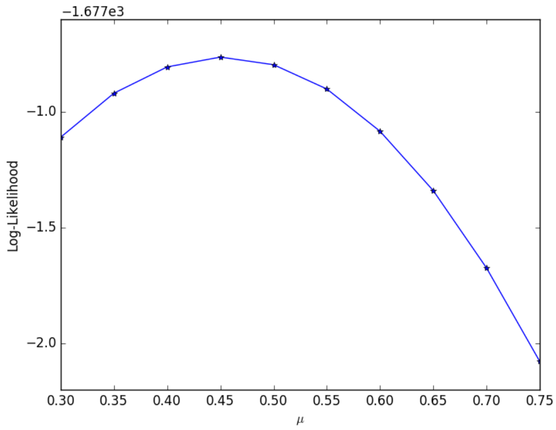
3 Stochastic volatility model with leverage

$$y_t = \epsilon_t \exp(x_t/2), t = 1, ..., T$$
$$x_{t+1} = \mu(1-\phi) + \phi x_t + \sigma_\eta \eta_t$$
$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N(0, \Sigma), \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

For stochastic volatility model with leverage, we firstly experiment on simulated dataset as before, with $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02, \rho = -0.8$ as shown in figure 3.



(a)



(b)

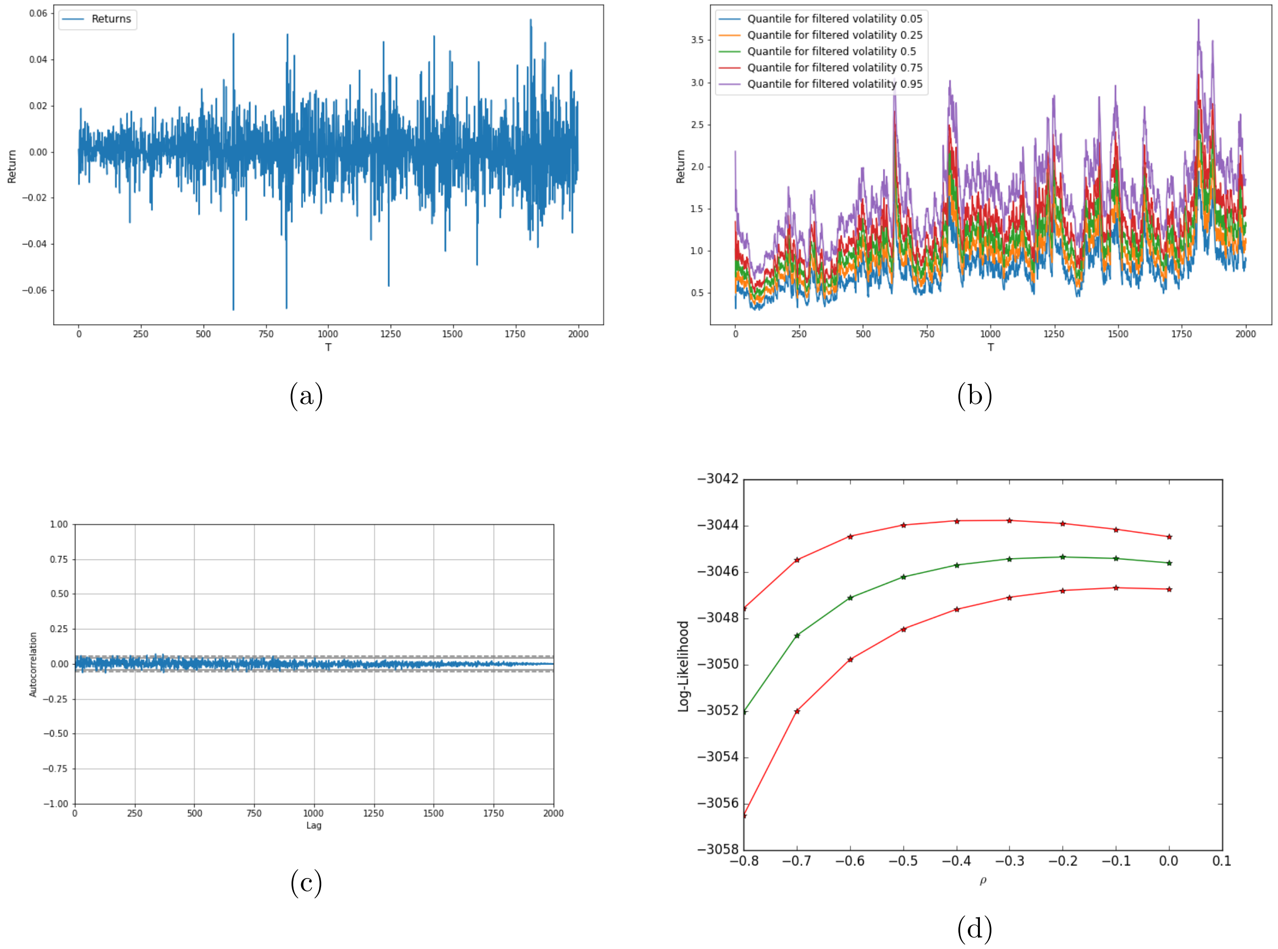
Figure 3: Fixed dataset, mean value of 50 runs with different random seeds, T = 1000, N = 300 (left), N = 600 (right)

μ	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75
N = 300	11.502	11.226	10.991	10.766	10.558	10.390	10.246	10.097	9.970	9.845
N = 600	3.529	3.468	3.408	3.356	3.314	3.282	3.270	3.253	3.236	3.226

Table 1: Empirical variance comparaison with N=300 and N=500

Figure 3 shows results of examining sliced log-likelihood and likelihood over μ and Table 1 displays the corresponding empirical variance of log-likelihood. We can find that the estimation of log-likelihood is maximzed at its true value $\mu = 0.5$ and the empirical variance is decreasing as N becomes larger.

Now, we turn to experimenting on S&P 500 and the result is displayed in figure 4 We can find that the autocorrelations of estimated distribution functions are small, which means they are independently uniformly distributed throught time, as expected.



(a) (b) (c) (d)

Figure 4: Results for SP 500 index returns. SV with leverage model. Top (left): 2000 daily returns from 16 May 1995 to 24 April 2003. Top (right): The quantiles for filtered volatility (0.05, 0.25, 0.5, 0.75, 0.95). Bottom (left): correlogram of $\Phi^{-1}(u_t)$

4 GARCH plus error model

This is a special case of the multivariate factor volatility model with the following form:

$$y_t|\sigma_t^2 \sim N(0, \sigma^2 + \sigma_t^2)$$
$$x_t|\sigma_t^2, y_t \sim N(\frac{b^2 y_t}{\sigma^2}; b^2)$$
$$\sigma_{t+1}^2 = \beta_0 + \beta_1 x_t^2 + \beta_2 \sigma_t^2$$

with $b^2 = \frac{\sigma^2\sigma_t^2}{\sigma^2+\sigma_t^2}$
First, we apply experiment on simulated dataset with $\beta_0 = 0.01, \beta_1 = 0.2, \beta_2 = 0.75, \sigma = 0.1$ and shown the sliced estimated log likelihood with garch plus error model in figure 5.

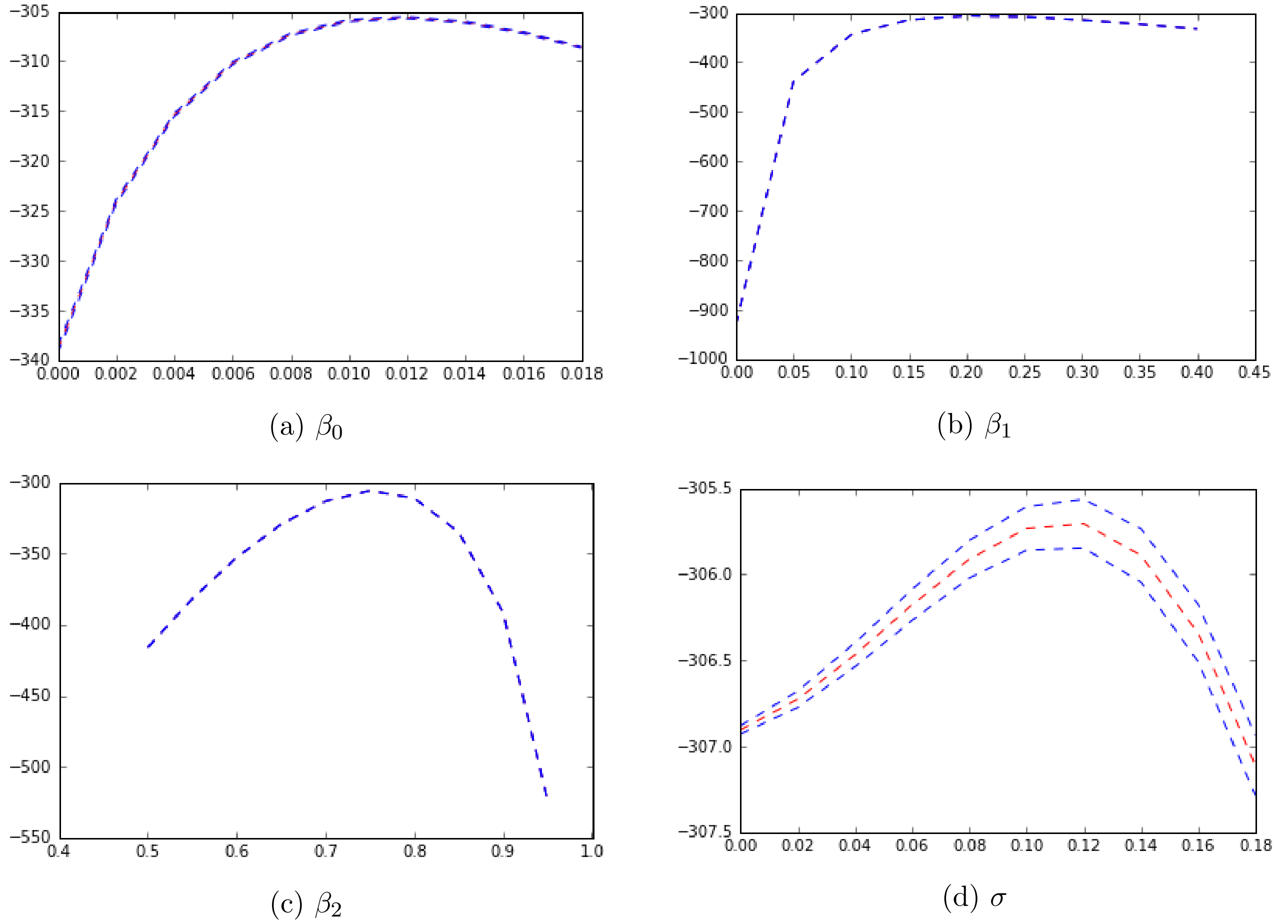


Figure 5: The sliced estimated log-likelihood for the GARCH model plus noise for (from top to bottom) $\beta_0, \beta_1, \beta_2, \sigma$. $T = 500, N = 500$.

Then, we aplly them on true data us uk compound daily return (with N=300) is shown in figure 6. When σ is zero, β_1 is estimated to be 0.02, when σ is not zero, β_1 is around 0.12, showing an apparent difference of the two models.

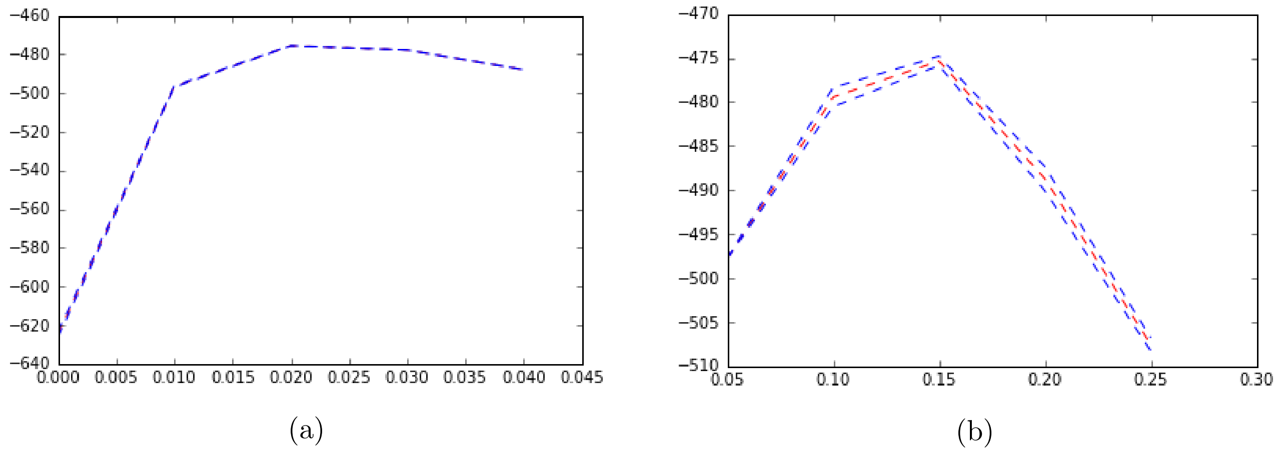


Figure 6: estimated likelihood of β_1 of us vs uk dataset,left: $\sigma = 0$, right: $\sigma = 0.55, T = 500$

We also apply a likelihood ratio test with garch model and garch plus error model: when $\sigma = 0$ the maximum log likelihood is -476.02, when σ is not forced to be zero, the maximum log likelihood is -473.92. We can reject the null hypythese that σ at the 0.01 level of significance and favour richer Garch(1,1) model for the dataset which is the same conclusion in the article.

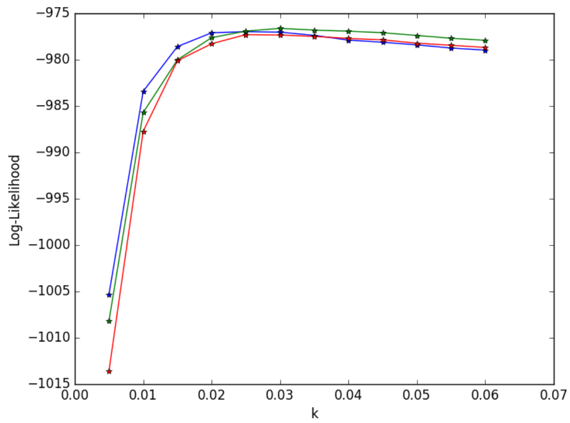
5 Continuous time models

In continuous time models, we focus on the specific example of a stochastic volatility model which may be used to value options.The log of a price, $y(t)$ is evolved as follows:

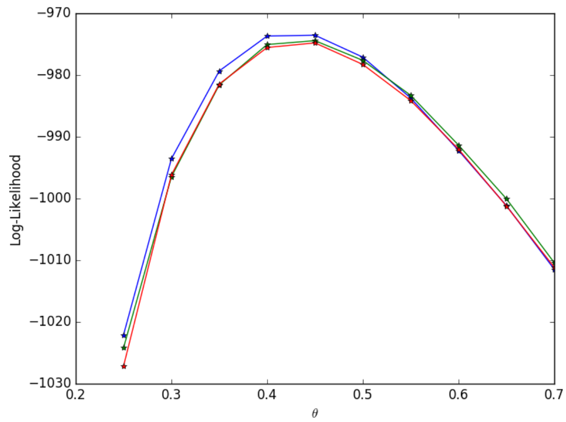
$$dy(t) = \{\mu + \beta\sigma^2(t)\}dt + \sigma(t)dW_1(t)$$

$$d\sigma^2(t) = a(\sigma^2(t))dt + b(\sigma^2(t))dW_2(t)$$

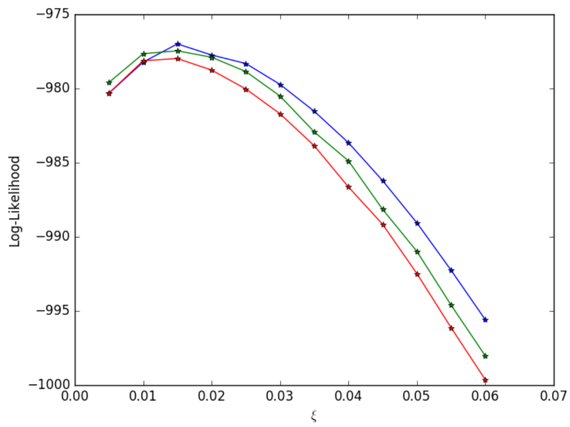
Since the log of the price is only observed with a certain frequency, the process between two observations can be simulated by a very fine Euler approximation. Conditional upon these trajectories, the distribution of the returns has a closed form which are used for our simulations. However, this distribution depend not only on the value of the hidden state variable on the observation times, but also on the whole trajectory between observations, an additional smoothing technique is applied in the resampling step to ensure that weights are a continuous function of the end points of the trajectories. Below is the simulation results of Nelson volatility process ($\mu = \beta = 0$, $a = k(\theta - \sigma^2)$, $b = \sqrt{\xi}\sigma^2$) with observation frequency equals to 1, and 20 points are placed within each interval.



(a)



(b)



(c)

Figure 7: The log-likelihood for Nelson volatility model. Three runs of a fixed sample, n=1000, N=600, true parameters k=0.02, θ=0.5, ξ=0.0178