Faculdade de Economia da Universidade de Coimbra
Grupo de Estudos Monetários e Financeiros (GEMF)
Av. Dias da Silva, 165 – 3004-512 COIMBRA, PORTUGAL
gemf@fe.uc.pt
http://gemf.fe.uc.pt

J. Q. SMITH e ANTÓNIO A. F. SANTOS

Second Order Filter Distribution Approximations for Financial Time Series with Extreme Outliers

ESTUDOS DO GEMF

N.º 3 **2003**

PUBLICAÇÃO CO-FINANCIADA PELA
FUNDAÇÃO PARA A CIÊNCIA E TECNOLOGIA

Second order filter distribution approximations for financial time series with extreme outliers

J. Q. Smith and António A. F. Santos*

Abstract

Particle Filters are now regularly used to obtain the filter distributions associated with state space financial time series. The method most commonly used nowadays is the auxiliary particle filter method in conjunction with a first order Taylor expansion of the log-likelihood. We argue in this paper that, for series such as stock return, which exhibit fairly frequent and extreme outliers, filters based on this first order approximation can easily break down. However, the auxiliary particle filter based on the much more rarely used second order approximation appears to perform well in these circumstances. We demonstrate our results with a typical stock market series.

Keywords: Particle filters; Second order approximations; State space models; Stochastic volatility.

1 Introduction

Of the two most reported characteristics associated with financial returns time series the first is the fat tails in the unconditional distribution of returns. More observa-

*University of Warwick(Department of Statistics) and Faculty of Economics at the University of Coimbra.

tions appear in the tails than for Gaussian processes, giving rise to high kurtosis. The second is volatility clustering, indicating the need to model the variance evolution of the series. It is a well established fact from empirical as well as theoretical financial literature that with short term series, variances as measures of volatility in financial markets are time varying and present some degree of predictability (Bollerslev et al. 1994; Taylor 1994; Diebold and Lopez 1995; Engle 1995; Campbell et al. 1997; Christoffersen and Diebold 1997; Diebold et al. 1997; Ait-Sahalia 1998; Andersen et al. 1999). Variances are used as a measure of risk in a variety of senses: Value-at-Risk (VaR) calculations, portfolio allocation and pricing options.

To model variance dynamics it is usually necessary to use non-linear models (Gallant et al. 1993; Hsieh 1993; Bollerslev et al. 1994; Åsbrink 1997; Campbell et al. 1997), which, in turn, usually require numerical algorithms to make estimations and predictions. The two most common classes of models used in financial time series are the Auto-Regressive Conditional Heteroscedastic (ARCH) models and Stochastic Volatility (SV) models. The focus of this paper is on the prediction of the variance evolution in SV models. The method used here is the Particle Filter method as described in Kong et al. (1994), Carpenter et al. (1998), Fearnhead (1998), Liu and Chen (1998), Carpenter et al. (1999), Freitas (1999), Doucet (2000), Doucet et al. (2000), Godsill et al. (2000), Doucet et al. (2001) and Liu (2001).

The SV model (Taylor 1986) is a nonlinear state space model. Financial returns y_t are related to unobserved states which are serially correlated. Thus we may write

$$y_t = \beta \exp\left(\frac{\alpha_t}{2}\right) \varepsilon_t \qquad \varepsilon_t \sim N(0,1) \tag{1}$$

$$\alpha_t = \phi \alpha_{t-1} + \sigma_\eta \eta_t \qquad \eta_t \sim N(0,1) \tag{2}$$

where α_t are the states of the process for $t = 1, \dots, n$. Note that the model is characterized by the vector of parameters $\theta = (\beta, \phi, \sigma_\eta)$.

Assuming that the parameters are known or have been previously estimated, for example, by using Markov chain Monte Carlo (MCMC) techniques, the main aim of this paper is to present certain modifications of particle filter methods which have recently been proposed to predict the process. The predictions made by this model, in contrast with ARCH family models, are expressed through the posterior density of the states $f(\alpha_t|D_t)$ and the predictive density of returns $f(y_{t+1}|D_t)$ rather than through point predictions. Henceforth we assume a closed forecasting system and we let $D_t = \{y_0, y_1, \ldots, y_t\}$ represent the available information at time t. Our modifications are straightforward but nevertheless appear to improve the predictive performance dramatically when these models are applied to stock return series. We first review current particle filter methods.

2 Particle Filter methods

The Bayes' rule allows us to assert that the posterior density $f(\alpha_t|D_t)$ of states is related to the density $f(\alpha_t|D_{t-1})$ prior to y_t and the density $f(y_t|\alpha_t)$ of y_t given α_t by

$$f(\alpha_t|D_t) \propto f(y_t|\alpha_t) f(\alpha_t|D_{t-1}) \tag{3}$$

and the predictive density of y_{t+1} given D_t is

$$f(y_{t+1}|D_t) = \int f(y_{t+1}|\alpha_{t+1}) f(\alpha_{t+1}|D_t) \, d\alpha_{t+1} \tag{4}$$

Instead of numerically estimating these integrals, the particle filter approximates these densities using a simulated sample.

Particle filters approximate the posterior density of interest, $f(\alpha_t|D_t)$, through a set of m "particles" $\{\alpha_{t,1}, \ldots, \alpha_{t,m}\}$ and their respective weights $\{\pi_{t,1}, \ldots, \pi_{t,m}\}$ where $\pi_{t,j} \geq 0$ and $\sum_{j=1}^{m} \pi_{t,j} = 1$. To implement these filters, we must first be able to sample from the nonstandard density $f(\alpha_t|D_t)$. It is possible to develop simulation procedures to approximate the distribution of interest and to calculate certain statistics that characterize the distribution. Secondly, we must be able to implement these procedures sequentially as states evolve over time and new information becomes available. This implementation needs to be efficient and the approximations need to remain good as we move through the sequence of states.

There are several ways of sampling from $f(\alpha_t|D_t)$. Typically we simulate from an approximating density $g(\alpha_t|\cdot)$. After a draw is obtained from the approximating density, it is modified to make it a draw from $f(\alpha_t|D_t)$. The two most popular techniques for performing this modification are sampling importance resampling (SIR) and rejection sampling/Markov chain Monte Carlo (Gilks, Richardson, and Spiegelhalter 1996; Gamerman 1997; Robert and Casella 1999; Doucet, de Freitas, and Gordon 2001; Liu 2001).

There are always errors associated with the approximation of distributions with continuous support from a discrete mass function. However, ignoring this aspect of approximating error, the effective implementation of the particle filter depends on how well we approximate $f(\alpha_t|D_t)$ by $g(\alpha_t|\cdot)$. If we could sample directly from $f(\alpha_t|D_t)$, then the sample would be independent and identically distributed and the numerical approximation would depend only on the number of draws. On the other hand, when it is not possible to sample directly from $f(\alpha_t|D_t)$, it becomes crucial to define good approximations $g(\alpha_t|\cdot)$. Here we suggest how this might be done for a stock market return series.

We essentially use the SIR method to sample from the distribution of interest. Taking into account the structure of the model, equation (3) can be used to define the approximating density $g(\alpha_t|\cdot)$, and subsequently its associated modifications. When a set of particles is used to approximate $f(\alpha_{t-1}|D_{t-1})$, $\{\alpha_{t-1,1}, \ldots, \alpha_{t-1,m}\}$ with respective weights $\{\pi_{t-1,1}, \ldots, \pi_{t-1,m}\}$, each particle is used to define the density $f(\alpha_t|D_t) \propto f(y_t|\alpha_t) f(\alpha_t|\alpha_{t-1,j})$, $j = 1, \ldots, m$. The approximating density

$g(\alpha_t|\alpha_{t-1,j}) = f(\alpha_t|\alpha_{t-1,j})$ is used to define a tentative draw from $f(\alpha_t|D_t)$. The "plain" SIR algorithm can be expressed using the following steps:

1. Generate a new set of particles using the transition equation, $\alpha_{t,j} = g_t(\alpha_t|\alpha_{t-1,j})$, $j = 1, \ldots, m$. A new set of weights is calculated for each of the m particles using the formula

$$w_{t,j} = f(y_t|\alpha_{t,j}) \tag{5}$$

$$\pi_{t,j} = \frac{w_{t,j}}{\sum_{i=1}^{m} w_{t,i}}, \quad j = 1, \ldots, m \tag{6}$$

2. Resample from $\{\alpha_{t,1}, \ldots, \alpha_{t,m}\}$ using the weights $\{\pi_{t,1}, \ldots, \pi_{t,m}\}$, and thus obtain a new set of particles with equal weights. These are then used in the next iteration.

Although this method was regarded as a considerable breakthrough, it is now widely recognized that this algorithm suffers from several weaknesses:

1. sample impoverishment (so the quality of the approximation thus deteriorates as time passes);

2. a lack of robustness regarding outliers, and

3. typically poor approximation of the tails of the posterior distribution.

Improvements to the basic SIR algorithm focusing on robust filters to outliers were recently proposed by Pitt and Shephard (1999, 2001). This algorithm – called the auxiliary particle filter (APF) – is widely recognized as an important improvement on the basic algorithm when implemented in time series such as those in finance, where the weaknesses referred to above become critical.

From a sequential perspective the main objective is to update the particles at $t-1$, and the respective weights, $\{\alpha_{t-1,1}, \ldots, \alpha_{t-1,m}\}$ and $\{\pi_{t-1,1}, \ldots, \pi_{t-1,m}\}$. Using the structure of the model (1)-(2), due to the Gaussian characteristics of the transition density, $f(\alpha_t|\alpha_{t-1})$, this would be a natural candidate for the approximating density. However, as stated by Pitt and Shephard (1999, 2001), this is not the most efficient procedure because it constitutes a *blind* proposal that does not take into account the information contained in y_t. One way of improving forecasting procedures is to include this information in the approximating distribution. When this is done, the nonlinear/non-Gaussian component of the measurement equation starts to play an important role and certain algebraic manipulations need to be carried out in order to use a standard approximation.

The design of the samplers must approximate the target distribution well but another important aspect need to be taken into account. When states are updated, in the presence of extreme observations, there are many particles with negligible weight and it is extremely difficult to propagate such particles. More rudimentary procedures, that treat all previous particles equally, will imply that only a small set of the new particles have non-negligible weight.

3 Auxiliary Particle Filter procedures

To overcome the problems posed by more rudimentary particle filter procedures, Pitt and Shephard (1999, 2001) proposed the APF method. The basic idea is that only part of the particles available at $t-1$ are propagated. These particles are chosen randomly but take into account the information presented in y_t. Only particles with non-negligible likelihood are propagated.

This can be accomplished by sampling from a higher dimensional distribution. First an index k is sampled, which defines the particles at $t-1$ that are propagated to t. This corresponds to sampling from

$$f\left(\alpha_t, k|D_t\right) \propto f\left(y_t|\alpha_t\right) f\left(\alpha_t|\alpha_{t-1}\right) \pi_k, \quad k = 1, \ldots, m \tag{7}$$

where π_k represents the weight given to each particle. The aim is then to sample first from $f\left(k|D_t\right)$ and then from $f\left(\alpha_t|k, D_t\right)$, obtaining the sample $\left\{\left(\alpha_{t,j}, k_j\right); j = 1, \ldots, m\right\}$. The marginal density $f\left(\alpha_t|D_t\right)$ is obtained by dropping the index k.

This resolves the problem of too many states with negligible weight being carried forward. However, the problem of defining a good approximation to the target distribution still remains. One of the simplest approaches is to define

$$g\left(\alpha_t, k|D_t\right) \propto f\left(y_t|\mu_{t,k}\right) f\left(\alpha_t|\alpha_{t-1}\right) \pi_k \tag{8}$$

where $\mu_{t,k}$ is the mean, mode or a highly probable value associated to $f\left(\alpha_t|\alpha_{t-1}\right)$. It can easily be seen that

$$\begin{aligned} g\left(k|D_t\right) &\propto \int f\left(y_t|\mu_{t,k}\right) f\left(\alpha_t|\alpha_{t-1}\right) \pi_k \, d\alpha_t & (9) \\ &= f\left(y_t|\mu_{t,k}\right) & (10) \end{aligned}$$

This density is used to define the first stage weights. These are the ones used to sample the index that tell us which particles at $t-1$ are used to define the posterior distribution at t. Given a set of indexes, the states are drawn from $f\left(\alpha_t|\alpha_{t-1,k}\right)$ and the second stage weights are defined as

$$w_j = \frac{f\left(y_t|\alpha_{t,j}\right)}{f\left(y_t|\mu_{t,j}\right)} \tag{11}$$

The information contained in y_t is carried forward through first stage weights. After the particles $\alpha_{t-1,k}$, $k = 1, \ldots, m$ are chosen, the densities used, $f\left(\alpha_t|\alpha_{t-1,k}\right)$, $k = 1, \ldots, m$ do not depend any further on y_t.

4 Stochastic Volatility models and Particle Filters

To develop efficient particle filter procedures which may be applied to predict the variance evolution in financial markets, we propose to use the characteristics of the model to find better approximations of the target distribution described above.

It is straightforward to simulate from a Gaussian distribution and, for a given value $\alpha_{t-1,k}$, $k = 1, \ldots, m$, the transition density in (2) assumes the Gaussian form. To obtain the posterior distribution, this must be combined with the likelihood function, and in this case, the conjugate property does not apply. It is not possible to sample directly from the target distribution, but we are able to define an approximating Gaussian distribution from which it is easy to sample.

The way to implement these procedures is to perform a first or second order Taylor approximation of the log-likelihood. The log-likelihood function associated to model (1) as a function of α_t is

$$l\left(\alpha_t\right) = const - \frac{\alpha_t}{2} - \frac{y_t^2}{2\beta^2 \exp\left(\alpha_t\right)} \tag{12}$$

This function is concave in α_t and so first and second order Taylor series approximations may work.

Based on a first order Taylor approximation, Pitt and Shephard (1999, 2001) developed a rejection sampler which was used to implement the particle filter. If this approximation is defined around some arbitrary value $\mu_{t,k}$, $g\left(y_t|\alpha_t, \mu_{t,k}\right)$, it can easily be seen that $g\left(y_t|\alpha_t, \mu_{t,k}\right) \geq f\left(y_t|\alpha_t\right)$ due to the assumed log-concavity of $f\left(y_t|\alpha_t\right)$. This allows the definition of a perfect envelope of the target density and a rejection sampler can be implemented. These arguments can be summarized by the following equations,

$$\begin{aligned} f\left(\alpha_t, k|D_t\right) &\propto f\left(y_t|\alpha_t\right) f\left(\alpha_t|\alpha_{t-1,k}\right) && (13) \\ &\leq g\left(y_t|\alpha_t, \mu_{t,k}\right) f\left(\alpha_t|\alpha_{t-1,k}\right) && (14) \\ &= g\left(y_t|\mu_{t,k}\right) g\left(\alpha_t|\alpha_{t-1,k}, y_t, \mu_{t,k}\right) && (15) \\ &\propto g\left(\alpha_t, k|D_t\right) && (16) \end{aligned}$$

To implement rejection sampling, k is first sampled from a density proportional to $g\left(y_t|\mu_{t,k}\right)$, then α_t is drawn from $g\left(\alpha_t|\alpha_{t-1,k}, y_t, \mu_{t,k}\right)$, which is equivalent of sampling from $g\left(\alpha_t, k|D_t\right)$. As this is an approximating density, the pair (k, α_t) is accepted with probability

$$\frac{f\left(\alpha_t, k|D_t\right)}{g\left(\alpha_t, k|D_t\right)} \tag{17}$$

which can be rewritten as

$$\frac{f\left(y_t|\alpha_t\right)}{g\left(y_t|\alpha_t, \mu_{t,k}\right)} \tag{18}$$

Pitt and Shephard (1999, 2001) developed these results and applied them to the SV model in (1)-(2). They used $\mu_{t,k} = \phi\alpha_{t-1,k}$ and the first order approximation

$$\log g\left(y_t|\alpha_t, \mu_{t,k}\right) = const - \frac{\alpha_t}{2} - \frac{y_t^2}{2\beta^2 \exp\left(\mu_{t,k}\right)}\left(1 - \left(\alpha_t - \mu_{t,k}\right)\right) \tag{19}$$

which, combined with the Gaussian transition density, gives rise to an approximating density that can be factorized into two densities

$$g\left(\alpha_t|\alpha_{t-1}^k, y_t, \mu_{t,k}\right) = N\left(\mu_{t,k}^*, \sigma_\eta^2\right) \tag{20}$$

and

$$g\left(y_t|\mu_{t,k}\right) \propto \exp\left(\frac{\mu_{t,k}^{*2} - \mu_{t,k}^2}{2\sigma_\eta^2}\right) \exp\left(-\frac{y_t^2\left(1 + \mu_{t,k}\right)}{2\beta^2 \exp\left(\mu_{t,k}\right)}\right) \tag{21}$$

where

$$\mu_{t,k}^* = \mu_{t,k} + \frac{\sigma_\eta^2}{2}\left(\frac{y_t^2}{\beta^2 \exp\left(\mu_{t,k}\right)} - 1\right) \tag{22}$$

The probabilities of acceptance referred to in (17)-(18) can be rewritten as

$$\exp\left(-\frac{y_t^2}{2\beta^2\exp\left(\alpha_t\right)}+\frac{y_t^2}{2\beta^2\exp\left(\mu_{t,k}\right)}+\frac{y_t^2\left(\alpha_t-\mu_{t,k}\right)}{2\beta^2\exp\left(\mu_{t,k}\right)}\right) \tag{23}$$

However, in the presence of outliers, this first order approximation turns out to be a poor approximation of the target density. It is very difficult to accept any candidate and the procedure can take an excessive amount of time to update the posterior density. On the other hand, if instead of probabilities, (21) and (23) are used to define first and second stage weights in an SIR procedure, then only a small number of weights are non-negligible and so a continuous target is approximated by a small number of distinct particles.

A common characteristic associated with financial time series like stock market series is the presence of extreme observations. It is sometimes difficult to update the information contained in y_t when this represents an extreme observation. We demonstrate below that the procedure based on a first order approximation cannot cope with extreme observations. The approximating distribution is not close enough to the target distribution and the approximation of the posterior distribution is very poor.

Pitt and Shephard (1999, 2001) suggested the possibility of using a second order approximation without developing it. The main problems are that it can be more algebraically intensive and a perfect envelope cannot be defined. So SIR must be used instead. The main idea here is to perform a Gaussian approximation to the log-likelihood and combine it with a Gaussian transition density.

Within an APF approach the way forward is to develop an approximation to the likelihood function, $g\left(y_t|\alpha_t,\mu_{t,k}\right)$, and an approximation of the target density, $g\left(\alpha_t,k|D_t\right)$, which is factorized in $g\left(y_t|\mu_{t,k}\right)$ and $g\left(\alpha_t|\alpha_{t-1,k},y_t,\mu_{t,k}\right)$, from which first and second stage weights are defined. The approximation $g\left(y_t|\alpha_t,\mu_{t,k}\right)$ is defined through the second order Taylor approximation of $\log f\left(y_t|\alpha_t\right)$ around $\mu_{t,k}$,

$$\begin{aligned}\log \quad g\left(y_t|\alpha_t,\mu_{t,k}\right) &\propto l\left(\mu_{t,k}\right)+l'\left(\mu_{t,k}\right)\left(\alpha_t-\mu_{t,k}\right)+\frac{1}{2}l''\left(\mu_{t,k}\right)\left(\alpha_t-\mu_{t,k}\right)^2 && (24)\\ &= const-\frac{\alpha_t}{2}-\frac{y_t^2}{2\beta^2\exp\left(\mu_{t,k}\right)}\left(\left(\alpha_t-\mu_{t,k}\right)+\frac{\left(\alpha_t-\mu_{t,k}\right)^2}{2}-1\right) && (25)\end{aligned}$$

Using this second order approximation, the density $g\left(\alpha_t,k|D_t\right)$ is factorized as in equation (15) and the factors are

$$g\left(\alpha_t|\alpha_{t-1,k},y_t,\mu_{t,k}\right)=N\left(\mu_{t,k}^*,\sigma_{t,k}^2\right) \tag{26}$$

and

$$\begin{aligned}g\left(y_t|\mu_{t,k}\right) &= \exp\left(\frac{1}{2}\left(\frac{1}{\sigma_\eta^2}+\frac{y_t^2}{2\beta^2\exp\left(\mu_{t,k}\right)}\right)\left(\mu_{t,k}^{*^2}-\mu_{t,k}^2\right)\right) && (27)\\ &\times\exp\left(-\frac{y_t^2\left(1+\mu_{t,k}\right)}{2\beta^2\exp\left(\mu_{t,k}\right)}\right) && (28)\end{aligned}$$

where

$$\mu_{t,k}^*=\left(\frac{1}{\sigma_\eta^2}+\frac{y_t^2}{2\beta^2\exp\left(\mu_{t,k}\right)}\right)^{-1}\left(\frac{y_t^2\left(1+\mu_{t,k}\right)}{2\beta^2\exp\left(\mu_{t,k}\right)}+\frac{\mu_{t,k}}{\sigma_\eta^2}-\frac{1}{2}\right) \tag{29}$$

and

$$\sigma_{t,k}^2=\frac{2\beta^2\sigma_\eta^2}{2\beta^2+\exp\left(-\mu_{t,k}\right)\sigma_\eta^2y_t^2} \tag{30}$$

As we sample from $g\left(\alpha_t,k|D_t\right)$, an approximating sample, the elements in it must be resampled in order to obtain a sample that gives a better approximation of the target density $f\left(\alpha_t,k|D_t\right)$. The weights used in this resampling step are

$$\begin{aligned}\log w_j &= -\frac{y_t^2}{2\beta^2\exp\left(\alpha_{t,j}\right)}+\frac{y_t^2\left(1-\alpha_{t,j}\left(1-\frac{\alpha_{t,j}}{2}+\mu_{t,k}\right)+\left(\mu_{t,k}+\frac{\mu_{t,k}^2}{2}\right)\right)}{2\beta^2\exp\left(\mu_{t,k}\right)}\\ \pi_j &= \frac{w_j}{\sum_{i=1}^m w_i}, \qquad j=1,\ldots,m\end{aligned}$$

These are the so-called second stage weights that allow the modification of the approximating distribution towards the target distribution. Obviously, these weights must be more evenly distributed than those from the first order approximation because the second order approximation allows a better approximation of the target distribution.

5 Approximations based on maximum likelihood estimates

The rather cumbersome second order methods described above all use the approximation based on a Taylor series expansion around the point $\mu_{t,k} = \phi\alpha_{t-1,k}$ suggested by Pitt and Shephard (1999, 2001). For likelihoods associated with extreme observations, this is not where we expect the posterior density to centre its weight (Dawid 1973). For the class of SV models the weight should be more closely centred around the maximum, $\alpha_t^* = \log\left(\frac{y_t^2}{\beta^2}\right)$, of the likelihood function. We therefore propose using the Taylor series approximation above in (24), but around α_t^*.

There are two main advantages in using this approximation. Firstly, the algebra needed to implement the procedure is greatly simplified. Secondly, these procedures can be extended to include the cases where the likelihood is no longer log-concave. We will focus here on the first advantage. Using $\alpha_t^* = \log\left(y_t^2/\beta^2\right)$ in (24), as $l'\left(\alpha_t^*\right) = 0$, we have

$$\log g\left(y_t|\alpha_t, \alpha_t^*\right) = l\left(\alpha_t^*\right) + \frac{1}{2}l''\left(\alpha_t^*\right)\left(\alpha_t - \alpha_t^*\right)^2 \tag{31}$$

The algebra is simpler because we are able to combine the logarithm of the kernel of two Gaussian densities, one given by the transition density and the other given by $\frac{1}{2}l''\left(\alpha_t^*\right)\left(\alpha_t - \alpha_t^*\right)^2$, which is the log-kernel of a Gaussian density with mean α_t^* and variance $-1/l''\left(\alpha_t^*\right) = 2$. Furthermore, the definition of the first weights is also simplified,

$$g(y_t|\alpha_t^*) = \exp\left(-\frac{\alpha_t^*}{2}\left(1+\frac{\alpha_t^*}{2}\right)+\frac{\mu_{t,k}^{*2}}{4}+\frac{\mu_{t,k}^{*2}-\mu_{t,k}^2}{2\sigma_\eta^2}\right) \tag{32}$$

After sampling k from a distribution proportional to (32), the particle $\alpha_{t-1,k}$ is chosen, and the density, assuming the role of prior density, assumes a Gaussian form with mean $\mu_{t,k} = \phi\alpha_{t-1,k}$ and variance σ_η^2. This is combined with a Gaussian density with mean α_t^* and variance $\sigma^2 = 2$. The approximating density thus becomes:

$$g\left(\alpha_t|\alpha_{t-1,k},\alpha_t^*\right) = N\left(\mu_{t,k}^*, \sigma_{t,k}^2\right) \tag{33}$$

where

$$\mu_{t,k}^* = \frac{2\mu_{t,k}+\sigma_\eta^2\alpha_t^*}{2+\sigma_\eta^2} \tag{34}$$

and

$$\sigma_{t,k}^2 = \frac{2\sigma_\eta^2}{2+\sigma_\eta^2} \tag{35}$$

After the particles have been sampled, they must be resampled in order to take into account the target density. They are resampled using the second stage weights

$$\log w_j = -\frac{\alpha_{t,j}}{2} - \frac{y_t^2}{2\beta^2\exp\left(\alpha_{t,j}\right)} + \frac{\left(\alpha_{t,j}-\alpha_t^*\right)^2}{4} \tag{36}$$

$$\pi_{t,j} = \frac{w_j}{\sum_{j=1}^m w_j}, \qquad j = 1,\ldots,m \tag{37}$$

Following the resampling stage, an approximation of the target posterior distribution of the states at t is available, which will be used as a prior distribution to update the states at $t+1$.

To summarize, the particles at $t-1$ propagated to update the distribution of the states at t are chosen randomly according to the weights defined in (32). These weights are influenced by the information contained in y_t. By conditioning on each particle chosen through the first stage weights, new particles are sampled. As these come from an approximating distribution, a second step is necessary. The particles are resampled using the weights defined in (36)-(37). Our modification, outlined above, makes this second order APF straightforward and quick to implement.

6 An empirical demonstration

In this section we intend to demonstrate empirically that the first order APF approximation is not reliable when applied to a stock return series, namely, when it has to update the distribution of the states associated with extreme observations. We use a series of graphs displays and tables of results that try to illustrate that, with the first order approximation, there are situations where sample impoverishment is extreme.

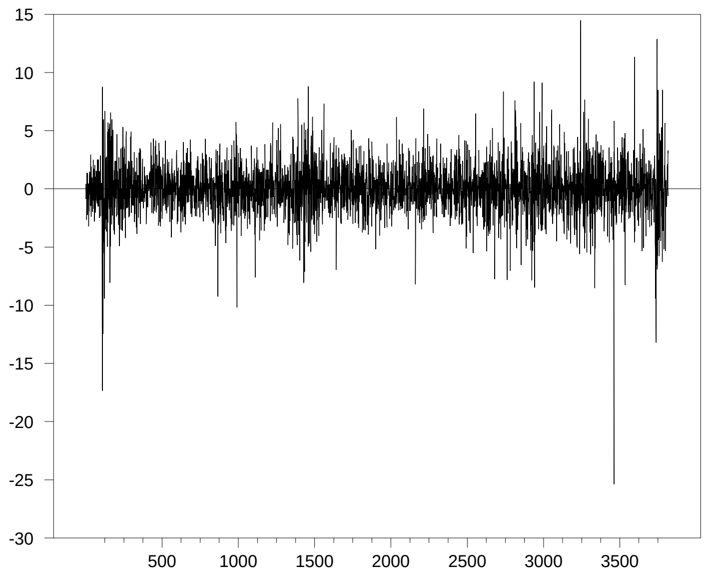

Figure 1: Rolls Royce stock returns

We analysed a series of daily returns regarding the Rolls Royce stock, traded on the London Stock Exchange. In this series we had 3818 observations. This series is depicted in Figure 1 and shows all the main characteristics found in the majority of financial time series, e.g., the presence of extreme observations and volatility clustering. We considered three sub-samples and analysed each one by adjusting an SV model to fit the data (Table 1). The parameters were estimated as being the

Rolls Royce Time series		**Sample 1**	**Sample 2**	**Sample 3**
Descriptive Statistics	Mean	-0.0315	0.0126	0.0104
	Variance	4.43597	3.7653	5.1998
	Skewness	-0.9808	0.0239	-0.6057
	Kurtosis	11.2178	4.8352	14.8458
	Min	-17.3518	-8.0899	-25.3876
	Max	11.2178	8.8179	14.4814
	Observations	1000	1000	1818
Stochastic Volatility Estimates	β	1.1405	1.1076	1.1535
	ϕ	0.9448	0.8358	0.9060
	σ_η	0.2078	0.3188	0.2784
Sample Impoverishment	First order	1	18	26
	Second order	0	0	0

Table 1: First section: Descriptive Statistics associated with each sub-sample of the series. Second section: Estimates of the parameters for each sub-sample, which were obtained as the mean of the posterior distribution defined through MCMC estimation techniques. Third section: Sample impoverishment figures, which were calculated as the number of times the range of the filter distribution was less than 0.2.

mean of the posterior distribution, which was obtained using Markov chain Monte Carlo techniques. By considering three sub-samples we were able to analyse the particle filter performance for different sets of parameters. A by-product of these estimation procedures is the smoothing distribution of the states. Despite being associated with different information sets, it is nonetheless convenient to compare the smoothed distributions with the filter distributions obtained using algorithms based on first and second order Taylor approximations to the log-likelihood. We ran the particle filter procedure from the first period in all three sub-samples to both algorithms and analysed the posterior distribution of each. A comparison was then made between these distributions and the respective smoothing distribution.

In general, rudimentarily applied particle filter procedures have certain difficulties in dealing with extreme observations. We therefore focused our attention on the distributions of the states associated with extreme observations. Pitt and Shephard (1999) applied the APF based on a first order approximation to the log-likelihood function to some less challenging financial time series such as that representing the evolution of the exchange rate of the Pound against the Dollar. However, stock return series typically exhibit more extreme observations, so even the APF based on a first order approximation to the log-likelihood starts to breakdown.

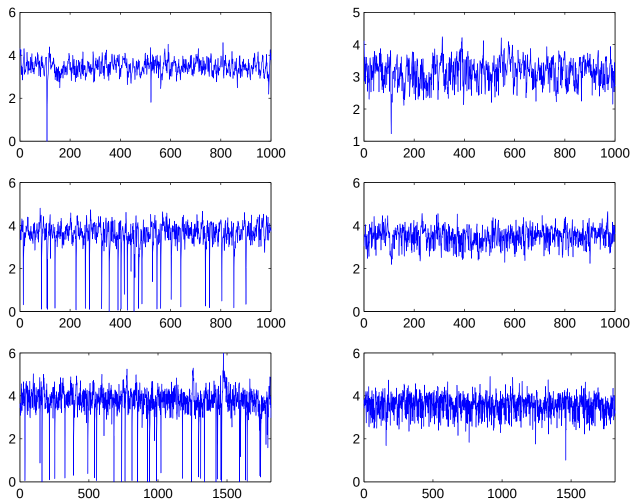

Figure 2: Difference between maximum and minimum value particles. Left-hand side:first order filter;Right-hand side:second order filter. Each row corresponds to a different sub-sample.

It can be seen that sample impoverishment for the APF, based on a first order approximation particle filter applied to the series presented here, is sometimes extreme. We measured this phenomena by counting the number of states for which the distribution was approximated using a small number of distinct particles. We ran the first order APF and analysed the distribution of the states for which the distance between the highest particle and the lowest was less than 0.2. In Figure 2 and the bottom row of Table 1 it can be seen that there were many such states in the first order APF but not in the second.

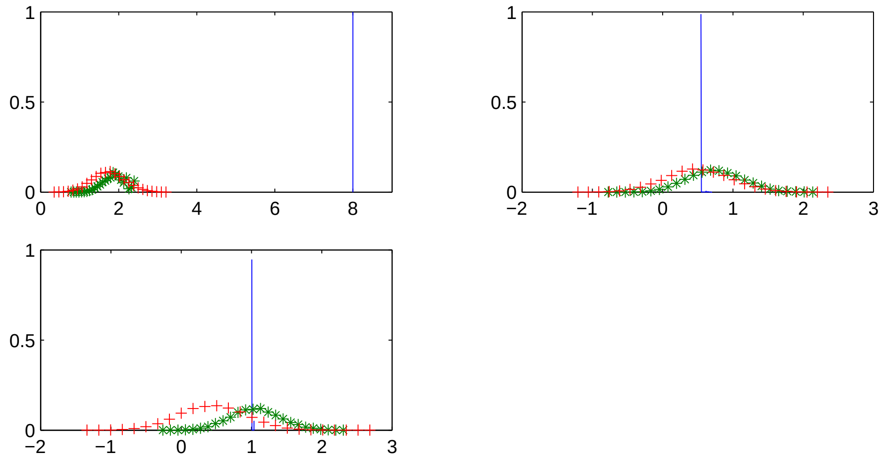

Figure 3: Histograms of the smoothing (+) distribution, first order filter distribution (-) and second order filter distribution(*). Each figure correspond to the respective sub-sample and the states analysed are for each sub-sample α_{108}, α_{852} and α_{1646}, respectively.

Finally in Figure 3, we compared the filter distribution using a first and second order approximation with the smoothing distribution associated with a given state. We obviously expected that the filter distributions to be different from the smoothing distributions because they were built from different information sets. However, the smoothing distribution is able to give us an initial idea regarding the location and dispersion of the filter distribution. In fact, it turns out that in the second and third sub-samples, these states were very near the end of the sample, which means that there was only a small difference between information sets. For example, in the second sub-sample we compared the density $f(\alpha_{852}|y_1,\ldots,y_{852},\theta)$ with $f(\alpha_{852}|y_1,\ldots,y_{1000},\theta)$, which we might have expected to be close. The difference between the filter distribution obtained using the first order particle filter was enormous in comparison with the smoothing distribution. The same did not occur with the second order particle filter, which is a clear indication of the feasibility of the procedures presented in this paper.

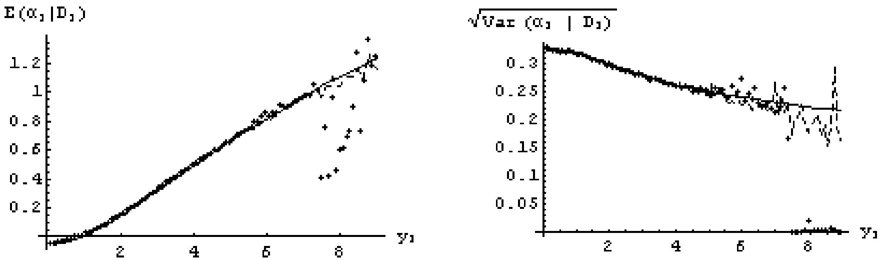

Figure 4: Comparison of particle filter procedures using the first and second order approximation to the log-likelihood function in a SV model. Left: mean evolution; Right: standard deviation evolution; exact path (solid line); first order (dotted line); second order (dashed line).

To better understand the infeasibility of using the first order APF to forecast the variance evolution associated with stock returns within a standard SV model, we present here a simple simulation that highlights the problems associated with the existence of outliers. We illustrate these problems applied to update the distribution of the first state, α_1, in a standard SV model. It is supposed that α_0 follows a Gaussian distribution with mean m_0 and variance C_0. In the univariate model, the updated distribution of $\alpha_1|D_1$, up to a normalizing constant c, has a known form with a density given by

$$f(\alpha_1|D_1) \propto \frac{1}{\exp\left(\frac{\alpha_1}{2}\right)} \exp\left(-\frac{y_1^2}{2\beta^2 \exp(\alpha_1)}\right) \exp\left(-\frac{(\alpha_1 - \phi m_0)^2}{2\left(\phi^2 C_0 + \sigma_\eta^2\right)}\right) \tag{38}$$

The constant c can be obtained using numerical integration. In this way the mean and standard deviation associated to the distribution of $\alpha_1|D_1$ can also be obtained. By varying the value of y_1, it can be appreciated how the results obtained using the APF with a first or second order approximation deviate from the exact results. In Figure 4 shows the mean and standard deviation evolution associated with the posterior distribution of $\alpha_1|D_1$ using the three techniques described above for different values of y_1, which vary between 0 and 9: values compatible with most financial time series. The parameters used were $\beta = 1.0$, $\phi = 0.97$, $\sigma_\eta = 0.15$, $m_0 = 0$ and $C_0 = 0.3$. It can easily be seen in Figure 4 that the APF based on a first order approximation of the log-likelihood is not robust to outliers when compared with an higher order approximation. We note that, with the first order approximation, for values of y_1 greater than 8 the estimated standard deviation is zero, indicating that the continuous density $f(\alpha_1|D_1)$ is approximated by a single point. This represents the extreme case of sample impoverishment.

7 Conclusion

We have demonstrated that it is possible to develop APFs based on a second order Taylor series approximation, which unlike their first order analogues perform well for series with extreme observations, which are fairly common in financial time series. We are now developing this procedure for time series whose likelihood is not log-concave. Preliminary results are encouraging and will be given in a future paper.

References

Ait-Sahalia Y. 1998. Dynamic equilibrium and volatility in financial asset markets. *Journal of Econometrics* **84**: 93–127.

Andersen TG, Bollerslev T, Lange S. 1999. Forecasting financial market volatility: Sample frequency vis-á-vis forecast horizon. *Journal of Empirical Finance* **6**: 457–477.

Åsbrink SE. 1997. *Nonlinearities and Regime shifts in Financial Time Series*. Ph. D. thesis, Stockholm School of Economics.

Bollerslev T, Engle RF, Nelson D. 1994. ARCH models. In *Handbook of Econometrics*, R. Engle and D. McFadden (eds.), Volume 4. Amsterdam: Elsevier.

Campbell JY, Lo AW, MacKinlay C. 1997. *The Econometrics of Financial Markets*. New Jersey: Princeton.

Carpenter J, Clifford P, Fearnhead P. 1998. An improved filter for non-linear problems. Technical report, Department of Statistics, University of Oxford.

Carpenter J, Clifford P, Fearnhead P. 1999. Building robust simulation-based filters for evolving data sets. Technical report, Department of Statistics, University of Oxford.

Christoffersen PF, Diebold FX. 1997. How relevant is volatility forecasting for financial risk management? Technical report, Wharton School, University of Pennsylvania.

Dawid AP. 1973. Posterior means for large observations. *Biometrika* **60**: 664–666.

Diebold FX, Hickman A, Inoue A, Schuermann T. 1997. Converting 1-day volatility to h-day volatility: Scaling by $\sqrt{h}$ is worse than you think. Technical report, Wharton School, University of Pennsylvania.

Diebold FX, Lopez JA. 1995. Modeling volatility dynamics. In *Macroeconomics: Developments, Tensions, and Prospects*, K. Hoover (ed.), pp. 427–472. Boston: Kluwer Academic Press.

Doucet A. 2000. On sequential simulation-based methods for bayesian filtering. *Statistics and Computing* **10**: 197–208.

Doucet A, de Freitas N, Gordon N. 2001. *Sequential Monte Carlo Methods in Practice*. New York: Springer.

Doucet A, Godsill S, Andrieu C. 2000. On sequential monte carlo sampling methods for bayesian filtering. Technical report, Signal Processing Group, Department of Engineering, University of Cambridge.

Engle RF. 1995. *ARCH: Selected Readings*. Oxford: Oxford University Press.

Fearnhead P. 1998. *Sequential Monte Carlo metods in filter theory*. Ph. D. thesis, University of Oxford.

Freitas JFG. 1999. *Bayesian Methods for Neural Networks*. Ph. D. thesis, Engineering Department, University of Cambridge.

Gallant AR, Rossi PE, Tauchen G. 1993. Nonlinear dynamics structures. *Econometrica* **61**: 871–907.

Gamerman D. 1997. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman & Hall.

Gilks WR, Richardson S, Spiegelhalter DJ. 1996. Introducing markov chain monte carlo. In *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (eds.). London: Chapman & Hall.

Godsill SJ, Doucet A, West M. 2000. Monte carlo smoothing for non-linear time series. Technical report, Institute of Statistics and Decision Sciences, Duke University.

Hsieh DA. 1993. Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis* **28**: 41–64.

Kong A, Liu JS, Wong WH. 1994. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association* **89**: 278–288.

Liu JS. 2001. *Monte Carlo Strategies in Scientific Computing*. Springer.

Liu JS, Chen R. 1998. Sequential monte carlo methods for dynamic systems. *Journal of the American Statistical Association* **93**: 1032–1044.

Pitt MK, Shephard N. 1999. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association* **94**: 590–599.

Pitt MK, Shephard N. 2001. Auxiliary variable based particle filters. In *Sequential Monte Carlo Methods in Practice*, A. Doucet, J. F. G. de Freitas, and N. J. Gordon (eds.), pp. 273–293. New York: Springer-Verlag.

Robert CP, Casella G. 1999. *Monte Carlo Statistical Methods*. New York: Springer-Verlag.

Taylor SJ. 1986. *Modelling Financial Time Series*. Chichester: Wiley.

Taylor SJ. 1994. Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* **4**: 183–204.

ESTUDOS DO G.E.M.F.

(Available on-line at http://gemf.fe.uc.pt)

2003

2002

N.º 3 *The consistency of optimal policy rules in stochastic rational expectations models*
- David Backus and John Driffill

N.º 2 *The term structure of the spreads between Portuguese and German interest rates during stage II of EMU*
- José Soares da Fonseca

N.º 1 *O processo desinflacionista português: análise de alguns custos e benefícios*
- António Portugal Duarte

2001

N.º 14 *Equity prices and monetary policy: an overview with an exploratory model*
- Fernando Alexandre e Pedro Bação

N.º 13 *A convergência das taxas de juro portuguesas para os níveis europeus durante a segunda metade da década de noventa*
- José Soares da Fonseca

N.º 12 *Le rôle de l'investissement dans l'éducation sur la croissance selon différentes spécifications du capital humain.*
- Adelaide Duarte e Marta Simões

N.º 11 *Ricardian Equivalence: An Empirical Application to the Portuguese Economy*
- Carlos Fonseca Marinheiro

N.º 10 *A Especificação da Função de Produção Macro-Económica em Estudos de Crescimento Económico.*
- Maria Adelaide Duarte e Marta Simões

N.º 9 *Eficácia da Análise Técnica no Mercado Accionista Português*
- Nuno Silva

N.º 8 *The Risk Premiums in the Portuguese Treasury Bills Interest Rates: Estimation by a cointegration method*
- José Soares da Fonseca

N.º 7 *Principais factores de crescimento da economia portuguesa no espaço europeu*
- Maria Adelaide Duarte e Marta Simões

N.º 6 *Inflation Targeting and Exchange Rate Co-ordination*
- Fernando Alexandre, John Driffill e Fabio Spagnolo

N.º 5 *Labour Market Transition in Portugal, Spain, and Poland: A Comparative Perspective*
- Paulino Teixeira

N.º 4 *Paridade do Poder de Compra e das Taxas de Juro: Um estudo aplicado a três países da UEM*
- António Portugal Duarte

N.º 3 *Technology, Employment and Wages*
- John T. Addison e Paulino Teixeira

N.º 2 *Human capital investment through education and economic growth. A panel data analysis based on a group of Latin American countries*
- Maria Adelaide Duarte e Marta Simões

N.º 1 *Risk Premiums in the Porutguese Treasury Bills Interest Rates from 1990 to 1998. An ARCH-M Approach*
- José Soares da Fonseca

2000

N.º 8 *Identificação de Vectores de Cointegração: Análise de Alguns Exemplos*
- Pedro Miguel Avelino Bação

N.º 7 *Imunização e M-quadrado: Que relação?*
- Jorge Cunha

N.º 6 *Eficiência Informacional nos Futuros Lisbor 3M*
- Nuno M. Silva

N.º 5 *Estimation of Default Probabilities Using Incomplete Contracts Data*
- J. Santos Silva e J. Murteira

N.º 4 *Un Essaie d'Application de la Théorie Quantitative de la Monnaie à l'économie portugaise, 1854-1998*
- João Sousa Andrade

N.º 3 *Le Taux de Chômage Naturel comme un Indicateur de Politique Economique? Une application à l'économie portugaise*
- Adelaide Duarte e João Sousa Andrade

N.º 2 *La Convergence Réelle Selon la Théorie de la Croissance: Quelles Explications pour l'Union Européenne?*
- Marta Cristina Nunes Simões

N.º 1 *Política de Estabilização e Independência dos Bancos Centrais*
- João Sousa Andrade

1999

N.º 9 *Nota sobre a Estimação de Vectores de Cointegração com os Programas CATS in RATS, PCFIML e EVIEWS*
- Pedro Miguel Avelino Bação

N.º 8 *A Abertura do Mercado de Telecomunicações Celulares ao Terceiro Operador: Uma Decisão Racional?*
- Carlos Carreira

N.º 7 *Is Portugal Really so Arteriosclerotic? Results from a Cross-Country Analysis of Labour Adjustment*
- John T. Addison e Paulino Teixeira

N.º 6 *The Effect of Dismissals Protection on Employment: More on a Vexed Theme*
- John T. Addison, Paulino Teixeira e Jean-Luc Grosso

N.º 5 *A Cobertura Estática e Dinâmica através do Contrato de Futuros PSI-20. Estimação das Rácios e Eficácia Ex Post e Ex Ante*
- Helder Miguel C. V. Sebastião

N.º 4 *Mobilização de Poupança, Financiamento e Internacionalização de Carteiras*
- João Sousa Andrade

N.º 3 *Natural Resources and Environment*
- Adelaide Duarte

N.º 2 *L'Analyse Positive de la Politique Monétaire*
- Chistian Aubin

N.º 1 *Economias de Escala e de Gama nos Hospitais Públicos Portugueses: Uma Aplicação da Função de Custo Variável Translog*
- Carlos Carreira

1998

N.º 11 *Equilíbrio Monetário no Longo e Curto Prazos - Uma Aplicação à Economia Portuguesa*
- João Sousa Andrade

N.º 10 *Algumas Observações Sobre o Método da Economia*
- João Sousa Andrade

N.º 9 *Mudança Tecnológica na Indústria Transformadora: Que Tipo de Viés Afinal?*
- Paulino Teixeira

N.º 8 *Portfolio Insurance and Bond Management in a Vasicek's Term Structure of Interest Rates*
- José Alberto Soares da Fonseca

1996