

Parameter Estimation for Infinite Lévy Processes  
via Sequential Monte Carlo

Liew Kuang Chen Joel

February 26, 2016

## Acknowledgements

## Abstract

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Preliminaries</b>	<b>5</b>
2.1	The Lévy Process . . . . .	5
2.1.1	Infinite Activity Processes - Variance Gamma . . . . .	6
2.2	Sequential Monte Carlo . . . . .	7
2.2.1	Perfect Monte Carlo . . . . .	7
2.2.2	Importance Sampling . . . . .	8
2.2.3	Sequential Importance Sampling (SIS) . . . . .	10
2.2.4	Sequential Monte Carlo Samplers (SMC) . . . . .	11
<b>3</b>	<b>Literature Review</b>	<b>12</b>
3.1	The Model . . . . .	12
3.2	Derivation of Posterior Density . . . . .	13
3.3	Simulation using a Sequence of Densities . . . . .	15
3.4	Simulation Procedure . . . . .	15
3.5	Resampling . . . . .	16
3.6	The SMC Algorithm . . . . .	17

# 1 Introduction

## 2 Preliminaries

### 2.1 The Lévy Process

In this paper, we will be using the Variance Gamma model from the family of Infinite Activity Lévy Processes. For the uninitiated, we will first go through the basics.

**Definition 2.1.** Lévy Process

A Lévy Process,  $X_t$ , is a cadlag stochastic process on  $(\Omega, \mathcal{F}_t, P)$ , with  $X_0 = 0$  and has the following properties:

1. Independent increments:  $X_{t+j+1} - X_{t+j} \perp X_{t+k+1} - X_{t+k}, \forall k, j \geq 0, k \neq j$
2. Stationary increments: The Law of  $X_{t+h} - X_t$  does not change for any  $t$ .
3. Stochastic continuity: For any  $\epsilon > 0, \lim_{h \rightarrow 0} P(|X_{t+h} - X_t| > \epsilon) = 0$

The Lévy Process also satisfies the Infinite Divisibility property.

**Definition 2.2.** Infinite Divisibility

For  $n \geq 2$ , a probability distribution  $F$  is infinitely divisible if there exists  $n$  i.i.d. random variables  $Y_1, Y_2, \dots$  such that  $\sum_{i=1}^n Y_i$  has distribution  $F$ .

One common distribution satisfying the infinite divisibility condition is the Normal Distribution. If  $X \sim N(\mu, \sigma^2)$ , then  $X = \sum_{i=1}^n Y_i$ , where  $Y_i \sim N(\frac{\mu}{n}, \frac{\sigma^2}{n})$ .

Due to fact that Lévy Processes are semimartingales and that their increments are independent, the process is Markovian and therefore in line with the Efficient Market Hypothesis. Hence, such processes are suitable to model asset prices (Li 2006).

### 2.1.1 Infinite Activity Processes - Variance Gamma

The probability densities of most Lévy processes are not known in closed forms. However, their characteristics function  $\phi_{X_t}(u)$  is as follows:

$$\phi_{X_t}(u) = E[\exp^{iuX_t}] = \exp^{-t\psi_x(u)}, t \geq 0$$

where  $\psi_x(u)$  is the characteristic exponent of  $X$  (Tankov) and satisfies the follow Lévy-Khintchine formula

$$\psi_x \equiv -i\mu u + \frac{\sigma^2 u^2}{2} + \int_{R_0} (1 - \exp^{iux} + iux1_{|x|<1})\pi(dx)$$

where  $R_0 = R \setminus \{0\}$  and  $\pi$  is a  $R_0$  Radon measure and

$$\int_{R_0} |x|^2 \pi(dx) < \infty \quad \int_{R_0} \pi(dx) < \infty$$

Unlike the Finite Activity models (such as the Merton Jump Model) under the family of Lévy Processes, the infinite-activity models allow infinite jumps within a finite time interval. Hence,

$$\int_{R_0} \pi(dx) \not< \infty$$

Within the category of Infinite Activity Lévy Processes, the jump process can have either finite or infinite variation. In the case of the finite/infinite variation, the sum of the absolute distance is finite/infinite over any finite time interval. For this paper, we will only be focussing on the Variance

Gamma Process - a member from the Infinite Activity with Finite Variation. This process,  $Z_t$ , is created using a Brownian Motion,  $B_t$ , with drift  $\gamma$  and variance  $\sigma_j^2$ , subordinated by an independent gamma process,  $G_t$ , with parameter  $\alpha$ .

$$dZ_t = \gamma dG_t + \sigma_j \sqrt{dG_t} dB_t$$

$$G_{t+t_0} - G_t \sim Ga(\frac{t}{\alpha}, \frac{1}{\alpha})$$

with the Lévy measure given by

$$\pi(dx) = \frac{A_{\pm}^2 \exp(-\frac{A_{\pm}}{B_{\pm}}|x|)}{B_{\pm}|x|} (dx)$$

where  $A_{\pm} = \frac{1}{v} \sqrt{\frac{\gamma^2 v^2}{4} + \frac{\sigma^2}{2}} \pm \frac{\gamma v}{2}$ ,  $B_{\pm} = A_{\pm}^2 v$ . In the case where  $\gamma = 0$ , the jumps are symmetric about 0.

## 2.2 Sequential Monte Carlo

### 2.2.1 Perfect Monte Carlo

Suppose we are able to simulate  $N$  i.i.d. random samples/particles  $\{\theta_{0:t}^i, i = 1, 2, \dots, N\}$  from a particular distribution (which in this paper, we are referring to the posterior distribution) -  $p(\theta_{0:t}|y_{0:t})$ . Then, the empirical estimate of this distribution is

$$P_N(d\theta_{0:t}|y_{0:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{0:t}^i}(d\theta_{0:t}) \quad (1)$$

where  $\delta_{\theta_{0:t}^i}(d\theta_{0:t})$  refers to the delta-dirac mass located at  $\theta_{0:t}^i$ . One can then obtain the estimate of  $E[f(\theta_{0:t})]$  via

$$E[f(\theta_{0:t})]_{est} = \int f(\theta_{0:t}^i) P_N(d\theta_{0:t}|y_{0:t}) = \frac{1}{N} \sum_{i=1}^N f(\theta_{0:t}^i) \quad (2)$$

If the posterior variance of  $f(\theta_{0:t})$  satisfies  $\sigma^2 := E[f(\theta_{0:t})^2] - E[f(\theta_{0:t})]^2 < +\infty$ ,

$$\text{var}(E[f(\theta_{0:t})]_{est}) = \frac{\sigma^2}{N} \quad (3)$$

and by the Strong Law of Large Numbers,

$$E[f(\theta_{0:t})]_{est} \xrightarrow[N \rightarrow \infty]{\text{a.s.}} E[f(\theta_{0:t})] \quad (4)$$

Also, if  $\sigma^2 < +\infty$ , the Central Limit Theorem holds

$$\sqrt{N}(E[f(\theta_{0:t})]_{est} - E[f(\theta_{0:t})]) \xrightarrow[N \rightarrow \infty]{} N(0, \sigma^2) \quad (5)$$

Unfortunately, it often is difficult to sample from the posterior distribution, especially in high dimensional cases. Markov Chain Monte Carlo (MCMC) methods is a possible alternative, but are unsuited for recursive problems. Hence, we will introduce a new method in this paper.

### 2.2.2 Importance Sampling

One classical method is the use of Importance Sampling method (Geweke 1989). Suppose we are interested in evaluating  $E[f(\theta_{0:t})]$ . Using a proposal density  $\pi(\theta_{0:t}|y_{0:t})$ , from which we can easily sample from, we get the



following:

$$\begin{aligned}
E[f(\theta_{0:t})]_{est} &= \int f(\theta_{0:t}^i) P_N(d\theta_{0:t}|y_{0:t}) \\
&= \int f(\theta_{0:t}^i) p(\theta_{0:t}|y_{0:t}) d\theta_{0:t} \\
&= \int f(\theta_{0:t}^i) \frac{p(\theta_{0:t}|y_{0:t})}{\pi(\theta_{0:t}|y_{0:t})} \pi(\theta_{0:t}|y_{0:t}) d\theta_{0:t} \\
&= \int f(\theta_{0:t}^i) w(\theta_{0:t}^i) \pi(\theta_{0:t}|y_{0:t}) d\theta_{0:t}
\end{aligned} \tag{6}$$

Since we can easily simulate from  $N$  i.i.d. from  $\pi(\theta_{0:t}|y_{0:t})$ , we can get an estimate of  $E[f(\theta_{0:t})]$

$$\begin{aligned}
\hat{E}[f(\theta_{0:t})]_{est} &= \int f(\theta_{0:t}^i) \hat{P}(d\theta_{0:t}) \\
&= \frac{\frac{1}{N} \sum_{i=1}^N f(\theta_{0:t}^i) w(\theta_{0:t}^i)}{\frac{1}{N} \sum_{i=1}^N w(\theta_{0:t}^i)} \\
&= \sum_{i=1}^N f(\theta_{0:t}^i) W(\theta_{0:t}^i)
\end{aligned} \tag{7}$$

where  $w(\theta_{0:t}^i)$  and  $W(\theta_{0:t}^i)$  is known as the importance weights and normalised weights respectively. Note that we are now operating under the new measure

$$\hat{P}(\theta_{0:t}) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{0:t}^i}(d\theta_{0:t}) W(\theta_{0:t}^i) \tag{8}$$

The Importance Sampling method does not work well in this form (especially in high dimensional cases). In addition, under  $\mathcal{F}_{t+1}$ , one has to recompute the weights/normalised weights despite having the weights/normalised weights at  $\mathcal{F}_t$  - the recursive problem has yet to be solved! However, this has

spawned new innovations, one which we will discuss and use in this paper.

### 2.2.3 Sequential Importance Sampling (SIS)

The Importance Sampling Method can be improved to include new information,  $\mathcal{F}_{t+1}$ , without recomputing  $\{\theta_{0:t}^i, i = 1, 2, \dots, N\}$ . Using the following fact,

$$P(A_n, \dots, A_1 | B) = P(A_n | A_{n-1}, \dots, B) P(A_{n-1} | A_{n-2}, \dots, B) \dots P(A_1 | B) \quad (9)$$

we get,

$$\pi(\theta_{0:t+1} | y_{1:t+1}) = \pi(\theta_{t+1} | \theta_{0:t}, y_{1:t+1}) \pi(\theta_{0:t} | y_{1:t}) \quad (10)$$

After iterating,

$$\pi(\theta_{0:t+1} | y_{1:t+1}) = \pi(\theta_0) \prod_{i=1}^{t+1} \pi(\theta_i | \theta_{0:i-1}, y_{1:i}) \quad (11)$$

Our weights can then be recursively updated using the following

$$W(\theta_{0:t+1}^i) \propto W(\theta_{0:t}^i) \frac{p(y_{t+1} | \theta_{t+1}^i) p(\theta_{t+1}^i | \theta_t^i)}{\pi(\theta_{t+1}^i | \theta_{0:t}^i, y_{1:t+1})} \quad (12)$$

One importance case (which we will use) is when we adopt the prior distribution as the importance distribution at  $t = 0$ . However, one drawback of using the SIS method is that one is unable to get the Importance distribution

$$\pi_t(\theta_{1:t})$$

$$\pi_t(\theta_{1:t}) = \int \pi_1(\theta_1) \prod_{k=2}^t K_k(\theta_{k-1}, \theta_k) d\theta_{1:t-1} \quad (13)$$

Although there are some approximations by used for local random-walk moves, the complexity of the algorithm is  $O(N^2)$  and one cannot compute  $K_t(\theta_{t-1}, \theta_t)$  pointwise (Del Moral, Doucet, Jasra 2006).

#### 2.2.4 Sequential Monte Carlo Samplers (SMC)

The main idea of the Sequential Monte Carlo Samplers is to introduce artificial backward Markov Kernels,  $L_{t-1}$  and propose an auxiliary variable technique. Given

$$\begin{aligned} \tilde{p}_t(\theta_{1:t}) &= \frac{\tilde{\gamma}_t(\theta_{1:t})}{C_t} \\ \tilde{\gamma}_t(\theta_{1:t}) &= \gamma_t(\theta_t) \prod_{k=1}^{t-1} L_k(\theta_{k+1}, \theta_k) \end{aligned}$$

where  $C_t$  is the normalisation constant for the unnormalised probability density  $\tilde{\gamma}_t(\theta_{1:t})$  and  $\gamma_t(\theta_t)$  is the actual unnormalised probability density function of the distribution in question. Suppose, at  $\mathcal{F}_t$ , we have  $\{W_t^i, \theta_{0:t}^i, i = 1, 2, \dots, N\}$  approximating  $\tilde{p}_t(\theta_{1:t})$

$$\begin{aligned} \widetilde{p}_t^N(d\theta_{1:t}) &= \sum_{i=1}^N W_t^i \delta_{\theta_{1:t}}(d\theta_{1:t}) \\ W_t^i(\theta_{1:t}^i) &= \frac{w_t^i(\theta_{1:t}^i)}{\sum_{i=1}^N w_t^i(\theta_{1:t}^i)} \end{aligned}$$

where  $w_t^i(\theta_{1:t}^i)$  is the standard Importance weights in (7). Now, at  $\mathcal{F}_{t+1}$ , we bring the particles forwards via the Markov Kernel  $K_{t+1}(\theta_t, \theta_{t+1})$ . The

unnormalised weights is computed using

$$w_{t+1}^i(\theta_{1:t+1}) = \frac{\gamma_{t+1}(\theta_{1:t+1})}{\pi_{t+1}(\theta_{1:t+1})} = w_t^i(\theta_{1:t}) \tilde{w}_{t+1}^i(\theta_t, \theta_{t+1}) \quad (14)$$

where the incremental weight,  $\tilde{w}_{t+1}^i(\theta_t, \theta_{t+1})$  equals to

$$\tilde{w}_{t+1}^i(\theta_t, \theta_{t+1}) = \frac{\gamma_{t+1}(\theta_{t+1}) L_t(\theta_{t+1}, \theta_t)}{\gamma_t(\theta_t) K_{t+1}(\theta_t, \theta_{t+1})} \quad (15)$$

In the case of Markov Chain Monte Carlo Kernels, we can define the backward Markov Kernel suboptimally, by

$$L_t(\theta_{t+1}, \theta_t) = \frac{p_{t+1}(\theta_t) K_{t+1}(\theta_t, \theta_{t+1})}{p_{t+1}(\theta_t)} \quad (16)$$

If  $p_t \approx p_{t+1}$ , then

$$L_t(\theta_{t+1}, \theta_t) \approx L_t^{sub}(\theta_{t+1}, \theta_t) = \frac{p_t(\theta_t) K_{t+1}(\theta_t, \theta_{t+1})}{\int p_t(\theta_t) K(\theta_t, \theta_{t+1}) d\theta_t}$$

Thus, we will have the following incremental weights

$$\tilde{w}_{t+1}^i(\theta_t, \theta_{t+1}) = \frac{\gamma_{t+1}(\theta_t)}{\gamma_t(\theta_t)} \quad (17)$$

### 3 Literature Review

#### 3.1 The Model

For the purposes of this paper, we will be focusing on the Stochastic Volatility with Variance Gamma Jumps in Returns (SVVG). The model is as follows:

$$\begin{aligned}
dY_t &= \mu dt + \sqrt{v_t}[\rho dW_{1t} + \sqrt{1 - \rho^2} dW_{2t}] + dZ_t \\
dv_t &= \kappa(v - v_t)dt + \sigma_v \sqrt{v_t} dW_{1t} \\
dZ_t &= \gamma dG_t + \sigma_j \sqrt{dG_t} dB_t
\end{aligned} \tag{18}$$

Using Euler Discretization, we have the following:

$$\begin{aligned}
Y_{t+1} &= Y_t + \mu \Delta + \sqrt{v_t} \Delta \epsilon_{t+1}^y + Z_{t+1} \\
v_{t+1} &= v_t + \kappa(v - v_t) \Delta + \sigma_v \sqrt{v_t} \Delta \epsilon_{t+1}^v \\
Z_{t+1} &= \gamma G_{t+1} + \sigma_j \sqrt{G_{t+1}} \epsilon_{t+1}^Z
\end{aligned} \tag{19}$$

where both  $\epsilon_{t+1}^y$  and  $\epsilon_{t+1}^v$  follow  $N(0, 1)$  with  $\text{corr}(\epsilon_{t+1}^y, \epsilon_{t+1}^v) = \rho$  while  $\epsilon_{t+1}^y, \epsilon_{t+1}^v \perp G_{t+1}$ . The Jump Process  $Z_{t+1}$  follows a variance gamma process where  $\epsilon_{t+1}^Z$  follows  $N(0, 1)$  and  $\epsilon_{t+1}^Z \perp \epsilon_{t+1}^y, \epsilon_{t+1}^v, G_{t+1}$ .  $G_{t+1}$  follows a Gamma Distribution,  $\Gamma(\alpha, \beta)$ .

Hence, we have observations  $(Y_t)_{t=0}^T$ ; latent variables  $(v_t)_{t=0}^T, (Z_t)_{t=1}^T, (G_t)_{t=1}^T$ ; and parameters  $\Theta = \{\mu, \rho, \kappa, v, \sigma_v, \gamma, \sigma_j\}$ . For convenience, we let  $\theta = \{\Theta, (v_t)_{t=0}^T, (Z_t)_{t=1}^T, (G_t)_{t=1}^T\}$ .

### 3.2 Derivation of Posterior Density

Since  $\text{corr}(\epsilon_{t+1}^y, \epsilon_{t+1}^v) = \rho$  and  $\epsilon_{t+1}^y, \epsilon_{t+1}^v \sim N(0, 1)$ , conditioning on the values of  $Z_{t+1}, v_t$  and  $\Theta$ , we get

$$\begin{pmatrix} Y_{t+1} - Y_t \\ v_{t+1} - v_t \end{pmatrix} | v_t, Z_{t+1}, \Theta \sim N \left( \begin{pmatrix} \mu \Delta + Z_{t+1} \\ \kappa(\theta - v) \Delta \end{pmatrix}, v_t \Delta \begin{pmatrix} 1 & \rho \sigma_v \\ \rho \sigma_v & \sigma_v^2 \end{pmatrix} \right) \tag{20}$$

As for the variance gamma process, Li et. al (2006) shows that conditioning on  $G_{t+1}$  and  $\Theta$ , we get

$$\begin{aligned} J_{t+1}|G_{t+1}, \Theta &\sim N(\gamma G_{t+1}, \sigma_j^2 G_{t+1}) \\ G_{t+1}|\Theta &\sim \Gamma(\frac{\Delta}{v}, v) \end{aligned}$$

However, in Jasra (2011), it is shown that  $G_{t+1}$  can be integrated out:

$$\begin{aligned} p(Z_{t+1}|\Theta) &= \int p(Z_{t+1}|G_{t+1}, \Theta) p(G_{t+1}|\Theta) dG_{t+1} \\ &= \frac{2 \exp(\frac{\gamma Z_{t+1}^2}{\sigma})}{\alpha^{\frac{t-u}{\alpha}} \sqrt{2\pi} \Gamma(\frac{t-u}{\alpha}, \frac{1}{\alpha})} \left( \frac{Z_{t+1}^2}{\gamma^2 + 2\frac{\sigma^2}{\alpha}} \right)^{\frac{t-u}{2\alpha} - \frac{1}{4}} K_{\frac{t-u}{\alpha} - \frac{1}{2}} \left( \frac{\sqrt{Z_{t+1}^2(\gamma^2 + \frac{2\sigma^2}{\alpha})}}{\sigma^2} \right) \end{aligned} \quad (21)$$

where  $K_a(\cdot)$  is the modified Bessel Function of the second kind, and  $\sigma = \sigma_j \sqrt{t-u}$ . In this paper, we will let  $\alpha = t-u$ . Thus, the simulation procedure is reduced by one dimension, and this reduces the simulation complexity, since there is one lesser latent variable to update.

As for the priors, we will follow the ones subscribed by Li (2006).

$$p(\Theta) = p(\mu)p(\gamma)p(\sigma_j)p(\kappa)p(\sigma_v, \rho)p(v)$$

$$\mu, \gamma, \sigma_j \sim N(0, 1)$$

$$\kappa, v \sim N(0, 1) \text{ truncated at } 0$$

$$\text{Reparameterize } (\rho, \sigma_v) \text{ to } (\phi, w), \text{ where } \phi = \sigma_v \rho, w = \sigma_v^2(1 - \rho^2)$$

$$w \sim IG(1.0, 0.5), \phi|w \sim N(0, 0.5w)$$

Now, the posterior distribution is as follows:

$$p(\Theta, v_{1:n}, Z_{1:n} | Y_{0:n}) \propto p(\Theta) \prod_{i=0}^n p(Y_{i+1}, v_{i+1} | Y_i, v_i, Z_{i+1}, \Theta) p(Z_{i+1} | \Theta) \quad (22)$$

### 3.3 Simulation using a Sequence of Densities

Using the following densities,

$$\pi_k(\theta_{1:n} | y_{0:n}) \propto p(\Theta) \prod_{i=1}^n [p(y_i, v_i | y_{i-1}, v_{i-1}, z_i, \Theta)^{\zeta_k} p(z_i | \Theta)] \quad (23)$$

where  $0 \leq \zeta_1 < \dots < \zeta_p = 1$ . The idea is to simulate from an 'easy' density, before shifting the particles to a much more difficult density. At  $\zeta_1$ , the tempered posterior density focusses more on the priors than when compared to  $\zeta_k$  where  $k > 1$ . Overtime, as the algorithm iterates, the tempered posterior focusses less on the prior and more on the likelihood. Hence, when  $\zeta_p = 1$ , the particles now exist in the target posterior density.

### 3.4 Simulation Procedure

To start off the simulation, we first sample from the prior distribution  $\pi_0$  and Importance sampling is then conducted as below

$$\begin{aligned} w_1(\theta^i) &= \frac{\pi_1(\theta_1^i | y_{0:n})}{\pi_0(\theta_1^i | y_{0:n})} \\ W_1(\theta^i) &= \frac{w_1^i}{\sum_{i=1}^M w_1^i} \end{aligned} \quad (24)$$

At the second iteration, the particles are shifted from  $\pi_1$  to  $\pi_2$  via a kernel of invariant distribution  $P_2(\theta_1^i, \theta_2^i)$

$$w_2(\theta^i) = \frac{\pi_2(\theta_2^i | y_{0:n})}{\int \pi_1(\theta_1^i | y_{0:n}) P(\theta_1^i, \theta_2^i) d\theta_1} \quad (25)$$

There are many choices of kernels, but in this paper we will adopt the Random Walk Metropolis Kernel. Hence, for  $t \geq 2$ ,

$$w_k(\theta^i) = W_{k-1}^i(\theta^i) \prod_{i=1}^n p(y_i | y_{i-1}, v_i, z_i, \Theta)^{\zeta_k - \zeta_{k-1}} \quad (26)$$

For the Random Walk Metropolis Kernel, one has to specify the proposal variance  $\psi_{k+1}$  in order for one to update the parameter values. Using adaptive MCMC techniques (Andrieu, Moulines 2006), we approximate the mean and therefore the variance of the parameters at iteration  $k$ , for the next iteration  $k + 1$ . Also, if the acceptance rate of the parameter is above 0.85, we tune up the variability by a factor of 5. Likewise, if the acceptance rate of the parameter is below 0.15, we tune down the variability by a factor of 1/5.

### 3.5 Resampling

Like the problem encountered in the Sequential Importance Sampling, the variability of the weights increases as the iteration increases. This is termed as weight degeneracy (Doucet et. al 2001). Hence, to counter this problem, we resample the particles according to the normalised weights within the cloud of particles. After resampling, the weights of the particles are reset



to 1. In this paper, we will be using the Multinomial Resampling method, even though more sophisticated methods exist. One point to note is that the resampling should occur too often. When resampling occurs, the number of unique particles fall, hence reducing the particles' approximation of the target density. One criterion to measure the variability of the weights is the Effective Sample Size (ESS) (Doucet et al. 2001).

$$ESS_k = \frac{(\sum_{i=1}^M w_k^i)^2}{\sum_{i=1}^M (w_k^i)^2} = \frac{(\sum_{i=1}^M W_k^i \prod_{i=1}^n p(y_i | y_{i-1}, v_i, z_i, \Theta)^{\zeta_k - \zeta_{k-1}})^2}{\sum_{i=1}^M (W_k^i \prod_{i=1}^n p(y_i | y_{i-1}, v_i, z_i, \Theta)^{\zeta_k - \zeta_{k-1}})^2} \quad (27)$$

The idea of this criterion is to give insight on the approximate number of samples relative to an independent simulation approach (Jasra et. al 2006). When  $ESS_k$  drops below some threshold  $l$ , we resample.

In this paper, at iteration  $k$  of the algorithm, we let  $ESS_k = 0.95 ESS_{k-1}$ . Using the bisection method, we are able to calculate  $\zeta_k$  using (18).

### 3.6 The SMC Algorithm

Therefore, the algorithm is as follows: At iteration  $k=1$ , sample  $\theta_1^i \sim \pi_1$  for  $i = 1, \dots, M$  and compute

$$w_1(\theta^i) = \frac{\pi_1(\theta_1^i | y_{0:n})}{\pi_0(\theta_1^i | y_{0:n})}$$

$$W_1(\theta^i) = \frac{w_1^i}{\sum_{i=1}^M w_1^i}$$

If  $ESS_k \leq l$ , resample and set  $W_1(\theta^i) = 1/M$ . Set  $\psi_{k+1}$  for each kernel.

At iteration  $k = 2, \dots, p$ , set  $\zeta_k$  and compute

$$w_k(\theta^i) = \frac{\pi_1(\theta_1^i | y_{0:n})}{\pi_0(\theta_1^i | y_{0:n})}$$

$$W_k(\theta^i) = \frac{w_k^i}{\sum_{i=1}^M w_k^i}$$

If  $ESS_k \leq l$ , resample and set  $W_k(\theta^i) = 1/M$ . If  $k < l$ , set  $\psi_{k+1}$  for each kernel.