Pseudo-Code for An Efficient Approach to Analyzing State-Space Representations

David N. DeJong *  
Department of Economics,  
University of Pittsburgh,  
Pittsburgh, PA 15260, USA

Hariharan Dharmarajan  
Department of Economics,  
University of Pittsburgh,  
Pittsburgh, PA 15260, USA

Roman Liesenfeld  
Department of Economics,  
Universität Kiel,  
24118 Kiel, Germany

Jean-François Richard  
Department of Economics,  
University of Pittsburgh,  
Pittsburgh, PA 15260, USA

First Version: April 2007  This Revision: October 2007

*Contact Author: D.N. DeJong, Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260, USA; Telephone: 412-648-2242; Fax: 412-648-1793; E-mail: dejong@pitt.edu. Documentation and code used to execute the EIS filter are available at www.pitt.edu/~dejong/wp.htm
1 Notation

(See the paper for extended descriptions)

- $s_t$ : State Variable, $S_t : \{s_j\}_{j=1}^t$.
- $y_t$ : Observable, $Y_t : \{y_j\}_{j=1}^t$.
- $f()$ : Generic notation for density/probability.
- $s_{t}^{t,j}$: $i^{th}$ draw of $s_t$ conditional on $(Y_{t-1}, S_{t-1})$.
- $s_{t}^{0,j}$: $i^{th}$ draw of $s_t$ conditional on $(Y_t, S_{t-1})$.

1.1 Main Equations

(See the paper for extended descriptions)

- State-transition equation:
  \[ s_t = \gamma(s_{t-1}, Y_{t-1}, u_t). \] (1)
- Observation (or measurement) equation:
  \[ y_t = \delta(s_t, Y_{t-1}, u_t). \] (2)
- PDF of $s_t$ conditional on $Y_t$:
  \[ f(s_t|Y_t) = \frac{f(y_t,s_t|Y_{t-1})}{f(y_t|Y_{t-1})} = \frac{f(y_t|s_t,Y_{t-1}) f(s_t|Y_{t-1})}{f(y_t|Y_{t-1})}. \] (3)
- PDF of $s_t$ conditional on $Y_{t-1}$:
  \[ f(s_t|Y_{t-1}) = \int f(s_t|s_{t-1}, Y_{t-1}) f(s_{t-1}|Y_{t-1}) ds_{t-1}. \] (4)
- Period-\(t\) likelihood:
  \[ f(y_t|Y_{t-1}) = \int f(y_t|s_t, Y_{t-1}) f(s_t|Y_{t-1}) ds_t. \] (5)
- Total likelihood, \(f(Y_T)\):
  \[ f(Y_T) = \prod_{t=1}^{T} f(y_t|Y_{t-1}), \] (6)

where $f(y_1|Y_0) \equiv f(y_1)$. 
Conditional Expectation of $h(s_t)$:

$$E_t(h(s_t)|Y_t) = \int h(s_t) f(s_t|Y_t) \, ds_t \quad (7)$$

$$= \frac{\int h(s_t) f(y_t|s_t, Y_{t-1}) f(s_t|Y_{t-1}) \, ds_t}{\int f(y_t|s_t, Y_{t-1}) f(s_t|Y_{t-1}) \, ds_t}.$$

2 Pseudo-Code for the Standard Particle Filter

$t=1$ initialization: Given $f(s_0)$, draw $N$ values from $f(s_0)$ to create the cluster $\{s_0^{0,i}\}_{i=1}^N$. Combine each realization of $s_0^{0,i}$ with a draw from the transition density $f(s_1|s_0^{0,i}, Y_0)$ to obtain a cluster $\{s_1^{1,i}\}_{i=1}^N$, which serves as a discrete approximation of $f(s_1|Y_0, S_0)$.

Step 1: (Prediction) At period $t$, we have the cluster $\{s_{t-1}^{0,i}\}_{i=1}^N$ (approximating $f(s_{t-1}|Y_{t-1})$) from the previous period. For each $s_{t-1}^{0,i}$, obtain a draw $s_{t}^{1,i}$ from the conditional density $f(s_t|s_{t-1}^{0,i}, Y_{t-1})$, creating the cluster $\{s_t^{1,i}\}_{i=1}^N$. This cluster serves as a discrete approximation of $f(s_t|Y_{t-1})$, in light of (4).

Step 2: (Likelihood Evaluation) Using $\{s_t^{1,i}\}_{i=1}^N$, approximate the likelihood in (5) as

$$\hat{f}_N(y_t|Y_{t-1}) = \frac{1}{N} \sum_{i=1}^N f(y_t|s_t^{1,i}, Y_{t-1}). \quad (8)$$

Step 3: (Filtering) To each $s_t^{1,i}$ in the cluster $\{s_t^{1,i}\}_{i=1}^N$, assign a weight

$$w_t^{0,i} = \frac{f(y_t|s_t^{1,i}, Y_{t-1})}{\sum_{j=1}^N f(y_t|s_t^{1,j}, Y_{t-1})}. \quad (9)$$

Sampling with replacement from the cluster $\{s_t^{1,i}\}_{i=1}^N$ using the weights $\{w_t^{0,i}\}_{i=1}^N$ yields the new cluster $\{s_t^{0,i}\}_{i=1}^N$ (which approximates $f(s_t|Y_t)$), in light of (3).

Step 4: (Conditional Expectation of $h(s_t)$) The cluster of draws obtained in the previous step, $\{s_t^{0,i}\}_{i=1}^N$, approximates $f(s_t|Y_t)$. Thus, from (7), the conditional expectation can be obtained as

$$E_t(h(s_t)|Y_t) = \frac{1}{N} \sum_{i=1}^N h(s_t^{0,i}). \quad (10)$$
Return to Step 1 using the cluster \( \{ s^0_i \}^N_{i=1} \), and repeat Steps 1-4 until period \( T \) has been reached.

3 Pseudo-Code for EIS Filter

There are two important choices to be made when using the EIS Particle Filter.

1. The family of importance sampling densities \( g(s_t; a_t) \) (e.g., gaussian, piecewise-continuous etc.).

2. The method for approximating \( f(s_t|Y_{t-1}) \) (see section 4.3 in the paper).

3.1 Gaussian-EIS Particle Filter

Step 1: (Initialize Sampler \( g(s_t; a_t) \)) At period \( t \), we have the EIS draws and their corresponding weights \( \{ s^0_{t-1}, \omega^k_{t-1} \}^N_{k=1} \) from the previous period (recall the process is initialized by the known density \( f(s_0) \)). Choose initial values of the auxiliary parameters \( a_0^t \) (the mean and variance of the Gaussian density \( g(s_t; a_t) \)).

Step 2: (Recursive Optimization) The objective is to obtain optimal values of the auxiliary parameters; the following steps are repeated until convergence.

1. Draw \( R \) values of \( s_t \) from \( g(s_t; a_t^1) \); denote these draws as \( \{ s^{i,t}_t \}^R_{i=1} \).

2. Obtain updated values of \( a^t_{t+1} \) as the solution to the least squares problem

\[
(a_t, c_t)^{t+1} = \arg \min_{a_t, c_t} \sum_{i=1}^{R} \left[ \ln \left( \frac{f(y_t|s^{i,t}_t) \hat{f}(s^{i,t}_t|Y_{t-1})}{g(s^{i,t}_t; a_t)} \right) - c_t - \ln g(s^{i,t}_t; a_t) \right]^2,
\]

where \( f(\cdot|\cdot) \) is the known distribution associated with (2) and \( c_t \) is an intercept meant to calibrate \( \ln \left( \frac{f(y_t|s_t)f(s_t|Y_{t-1})}{g(s_t; a_t)} \right) \). (Details on the least-squares problem are provided below.)

3. Check for convergence. Upon convergence, we have the optimal mean and variance of the EIS sampling density: \( \hat{a}_t \).

Step 3: (Likelihood Evaluation) Draw \( N \) values \( \{ s^i_t \}^N_{i=1} \) from the optimal EIS sampling density \( g(s_t; \hat{a}_t) \). The IS estimate of the period-t likelihood in (5) and the EIS weights are given by

\[
\hat{f}_N(y_t|Y_{t-1}) = \frac{1}{N} \sum_{i=1}^{N} \omega^i_t,
\]

\[
\omega^i_t = \frac{f(y_t|s^i_t) \hat{f}(s^i_t|Y_{t-1})}{g(s^i_t; \hat{a}_t)}.
\]
Step 4: (Conditional Expectation of \( h(s_t) \))

1. The MC estimate for \( E_t(h(s_t)|Y_t) \) is given by

\[
E_t(h(s_t)|Y_t) = \frac{\sum_{i=1}^{S} h(s_{t}^{0,i}) \cdot \omega(s_{t}^{0,i}; \tilde{a}_t)}{\sum_{i=1}^{S} \omega(s_{t}^{0,i}; \tilde{a}_t)}. \tag{14}
\]

2. When using separate IS densities \( g_n(s_t; \tilde{a}_t) \) for the integrand in the numerator of (7), and \( g_d(s_t; \tilde{b}_t) \) for the integrand in the denominator of (7), \( R \) draws of \( s_t \) can be obtained from each (using the same set of CRNs). Thus the conditional expectation \( E_t(h(s_t)|Y_t) \) is obtained as

\[
E_t(h(s_t)|Y_t) = \frac{\sum_{i=1}^{N} \omega_{t,\text{numerator}}^j}{\sum_{j=1}^{N} \omega_{t,\text{denominator}}^j}, \tag{15}
\]

where

\[
\omega_{t,\text{numerator}}^j = \frac{h(s_{t}^i) f(y_t|s_{t}^i) \hat{f}(s_{t}^i|Y_{t-1})}{g_n(s_{t}^i; \tilde{a}_t)}, \tag{16}
\]

and

\[
\omega_{t,\text{denominator}}^j = \frac{f(y_t|s_{t}^i) \hat{f}(s_{t}^i|Y_{t-1})}{g_d(s_{t}^i; \tilde{b}_t)}. \tag{17}
\]

The EIS draws and their corresponding weights \( \{s_{t}^{0,k}, \omega_{t}^k\}_{k=1}^{N} \) from the IS for period-t likelihood constitute the computational pre-requisites for the above steps in period-(t+1). Thus, return to Step 1 using these draws and weights, and repeat Steps 1-4 until period \( T \) has been reached.

Regarding the least squares problem, let the integrand (in (5) or (7)) be denoted by \( \varphi(s_t) \). Neglecting the time subscript, let \( s \) be a \( j \)-dimensional variable with elements \( (x_1, x_2, ..., x_j) \). Then the auxiliary parameters \( a_t \) are the \( j \times 1 \) vector of means and the \( j \times j \) covariance matrix. Since the covariance matrix is symmetric, the number of auxiliary parameters reduces to \( j + j(j + 1)/2 \). We take \( a^0_t \) as given, initialized by \( a^0_t \). Hereafter, we will drop the superscript \( t \) and the subscript \( t \), and describe a single iteration of the auxiliary regression within a period.

Let the mean vector associated with \( a \) be denoted by \( \mu \), and the precision matrix (the inverse of the covariance matrix) by \( H \). The setup of the problem arises from the approximation \( \ln \varphi(s) \)
by a gaussian kernel:

\[
\ln \varphi(s) \propto -\frac{1}{2} (s - \mu)' H (s - \mu)
\]

\[
\propto -\frac{1}{2} (s'H s - 2s'H \mu).
\]

The term \(s'H s\) can be written as

\[
\begin{pmatrix}
  x_1 & x_2 & \ldots & x_j
\end{pmatrix}
\begin{pmatrix}
  h_{11} & h_{21} & \ldots & h_{1j} \\
  h_{21} & h_{22} & \ldots & h_{2j} \\
  \vdots & \vdots & \ddots & \vdots \\
  h_{j1} & h_{j2} & \ldots & h_{jj}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_j
\end{pmatrix}
\]

\[
= h_{11} (x_1^2) + h_{22} (x_2^2) + \ldots + h_{jj} (x_j^2)
\]

\[
+ 2h_{21} (x_2x_1) + 2h_{31} (x_3x_1) + \ldots + 2h_{j1} (x_jx_1)
\]

\[
+ 2h_{32} (x_3x_2) + 2h_{42} (x_4x_2) + \ldots + 2h_{j2} (x_jx_2)
\]

\[
\vdots
\]

\[
+ 2h_{j(j-1)} (x_jx_{j-1}).
\]

Note that the coefficients of the squares, pairwise products and the individual components of \(s\) are in one-to-one correspondence with the means and precision matrix of the Gaussian approximation. Thus, the problem reduces to the regression of \(\ln \varphi(s)\) on \([1, x_1^2, x_2^2, \ldots, x_j^2, x_1x_2, x_1x_3, \ldots, x_{j-1}x_j, x_1, \ldots, x_j]\) For a \(j\)-dimensional variable \(s\), the number of regressors is \(\left(1 + j + \frac{j(j+1)}{2}\right)\).

Consider a 3-dimensional problem where \(s = [x_1, x_2, x_3]\). The regression reduces to

\[
\ln \varphi(s) = \beta_0 + \beta_1 (x_1^2) + \beta_2 (x_2^2) + \beta_3 (x_3^2)
\]

\[
+ \beta_4 (x_2x_1) + \beta_5 (x_3x_1) + \beta_6 (x_3x_2)
\]

\[
+ \beta_7 x_1 + \beta_8 x_2 + \beta_9 x_3.
\]

Having obtained the LS estimates \(\hat{\beta}\), the updated precision matrix is given by

\[
h_{11} = -2\hat{\beta}_1; \ h_{22} = -2\hat{\beta}_2; \ h_{33} = -2\hat{\beta}_3
\]

\[
h_{21} = -\hat{\beta}_4; \ h_{31} = -\hat{\beta}_5; \ h_{32} = -\hat{\beta}_6.
\]
The updated means can be obtained by using the coefficients \( \hat{\beta}_7, \hat{\beta}_8, \hat{\beta}_9 \):

\[
\mu = H^{-1} \begin{pmatrix} \hat{\beta}_7 \\ \hat{\beta}_8 \\ \hat{\beta}_9 \end{pmatrix}.
\]

When \( s_t \) is univariate, the LS problem reduces to the regression

\[
\ln \varphi(s_t) = \beta_0 + \beta_1 s_t + \beta_2 s_t^2.
\]

The updated mean and variance can be written as

\[
\sigma^2 = \frac{-1}{2\beta_2},
\]

\[
\mu = \frac{-1}{2\beta_2} \hat{\beta}_1.
\]

### 3.2 Piecewise-EIS Particle Filter

Here, the coefficients of the piecewise approximations and the location of the nodes are the auxiliary parameters. Let the product of densities \( f(y_t|s_t) \hat{f}(s_t|Y_{t-1}) \) be denoted by \( \varphi(s_t) \).

**Step 1: (Initial Approximation)** At period \( t \), we have the EIS draws and their corresponding weights \( \left\{ s_{t-1}^k, \omega_{t-1}^k \right\}_{k=1}^N \) from the previous period.

- Choose an equally spaced partition in \( s_t \) with \( R \) subintervals i.e., \( a' = (a_0, ..., a_R) \), with \( a_0 < a_1 < ... < a_R \). The interval \([a_0, a_R]\) is understood as being sufficiently wide to cover the support of the density kernel \( \varphi(s_t) \).
- At each of the \( R+1 \) grid points, compute \( \ln (\varphi(a_i)) \) and construct a linear approximation to it in each subinterval:

\[
\ln k_j(s; a) = \alpha_j + \beta_j s \quad \forall s \in [a_{j-1}, a_j],
\]

\[
\beta_j = \frac{\ln \varphi(a_j) - \ln \varphi(a_{j-1})}{a_j - a_{j-1}}, \quad \alpha_j = \ln \varphi(a_j) - \beta_j a_j,
\]

where \( k(s; a) \) is the kernel of the EIS sampling density.

- Compute the CDF defined by these linear segments:

\[
K_j(s; a) = \frac{\chi_j(s; a)}{\chi_n(a)}, \quad \forall s \in [a_{j-1}, a_j],
\]

\[
\chi_j(s; a) = \chi_{j-1}(a) + \frac{1}{\beta_j} \left[ k_j(s; a) - k_j(a_{j-1}; a) \right],
\]

\[
\chi_0(a) = 0, \quad \chi_j(a) = \chi_j(a_j; a).
\]
Step 2: (Refinement by Inversion) The objective is to obtain optimal values of the auxiliary parameters.

- Define a uniformly spaced partition in [0, 1] and invert the above CDF to obtain an equal-probability partition in $s_t$. The inverse of $K()$ is given by

$$s = \frac{1}{\beta_j} \left\{ \ln \left[ k_j(a_{j-1}; a) + \beta_j \left( u_{jR}(a) - \chi_{j-1}(a) \right) \right] - \alpha_j \right\}.$$  

- Generally, a one-step refinement is sufficient. However, as discussed in the paper, one may choose to iterate on the above procedure.

- We now have the optimal auxiliary parameters of the EIS sampling density: $\widehat{a}_t$

Step 3: (Likelihood Evaluation) Draw $N$ values from the sampling density $k(s_t; \widehat{a}_t)$. The IS estimate of the period-$t$ likelihood in (5) and the EIS weights are given by

$$\widehat{f}_N(y_t|Y_{t-1}) = \frac{1}{N} \sum_{i=1}^{N} \omega_t^i,$$

$$\omega_t^i = \frac{f(y_t|s_{0i}^t)\widehat{f}(s_{0i}^t|Y_{t-1})}{k(s_{0i}^t; \widehat{a}_t)}.$$  

Step 4: (Conditional Expectation of $h(s_t)$)

1. The MC estimate for $E_t(h(s_t)|Y_t)$ is given by

$$E_t(\widehat{h}(s_t)|Y_t) = \sum_{i=1}^{S} \frac{h(s_{0i}^t) \cdot \omega(s_{0i}^t; \widehat{a}_t)}{\sum_{i=1}^{S} \omega(s_{0i}^t; \widehat{a}_t)}.$$  

2. When using separate IS densities $k_n(s_t; \widehat{a}_t)$ for the integrand in the numerator of (7), and $k_d(s_t; \widehat{b}_t)$ for the integrand in the denominator of (7), $R$ draws of $s_t$ can be obtained from each (using the same set of CRNs). Thus the conditional expectation $E_t(h(s_t)|Y_t)$ is obtained as

$$E_t(\widehat{h}(s_t)|Y_t) = \frac{\sum_{i=1}^{N} \omega_{t,\text{numerator}}^i}{\sum_{j=1}^{N} \omega_{t,\text{denominator}}^j},$$

where

$$\omega_{t,\text{numerator}}^i = \frac{h(s_{0i}^t) f(y_t|s_{0i}^t)\widehat{f}(s_{0i}^t|Y_{t-1})}{g_n(s_{0i}^t; \widehat{a}_t)}.$$  

7
and

\[ \omega^i_{t, \text{denominator}} = \frac{f(y_t | s^i_t) \hat{f}(s^i_t | Y_{t-1})}{g_d(s^i_t; \hat{b}_t)}. \] (23)

The EIS draws and their corresponding weights \( \{s^0_i, \omega^i_k\}_{k=1}^N \) from the IS for period-\( t \) likelihood constitute the computational pre-requisites for the above steps in period-(\( t+1 \)). Thus, return to Step 1 using these draws and weights, and repeat Steps 1-4 until period \( T \) has been reached.

4 Overview of Auxiliary and Adapted Particle Filters

Description of the pseudo-code for the auxiliary and adapted particle filters requires a brief explanation of the methods. Section 5 provides the computational algorithms for implementing these methods. Pitt and Shephard (1999) tackle the problem of adaption in particle filters through the use of an Importance Sampling (IS) procedure. Consider the marginalization step of the filtering process. Faced with the problem of calculating

\[ f(y_t | Y_{t-1}) = \int f(y_t | s_t, Y_{t-1}) f(s_t | Y_{t-1}) ds_t, \]

but with \( f(s_t | Y_{t-1}) \) unknown, importance sampling achieves approximation via the introduction into the integral of an importance density \( g(s_t | Y_t) \):

\[ f(y_t | Y_{t-1}) = \int \frac{f(y_t | s_t, Y_{t-1}) f(s_t | Y_{t-1})}{g(s_t | Y_t)} g(s_t | Y_t) ds_t. \] (24)

Obtaining drawings \( s^0_t \) from \( g(s_t | Y_t) \), this integral is approximated as

\[ \hat{f}(y_t | Y_{t-1}) \approx \frac{1}{N} \sum_{i=1}^N \frac{f(y_t | s^0_i, Y_{t-1}) f(s^0_i | Y_{t-1})}{g(s^0_i | Y_t)}. \] (25)

Pitt and Shephard referred to the introduction of \( g(s_t | Y_t) \) in this context as adaption. Full adaption is achieved when \( g(s_t | Y_t) \) is constructed as being proportional to \( f(y_t | s_t, Y_{t-1}) f(s_t | Y_{t-1}) \), rendering the ratios in (25) as constants. Full adaption was viewed as being computationally infeasible, due to the cost of having to compute \( f(s^0_i | Y_{t-1}) \) for every value of \( s^0_i \) produced by the sampler. Instead they introduced the concept of an auxiliary particle filter designed to yield partial adaption.

4.1 Auxiliary Particle Filter

This algorithm takes as input in period \( t \) a period-(\( t-1 \)) approximation of \( f(s_t | Y_{t-1}) \) of the form

\[ \hat{f}_N(s_t | Y_{t-1}) = \sum_{i=1}^N \pi^i_{t-1} f(s_t | s^0_{t-1}, Y_{t-1}), \] (26)
where \( \left\{ s_{t-1}^{0,i} \right\}_{i=1}^{N} \) denotes the period-\((t-1)\) swarm of particles, and \( \left\{ \pi_{t-1}^{i} \right\}_{i=1}^{N} \) a vector of probabilities initialized by \( \pi_{0}^{i} = 1/N \), and recursively updated as described below. The corresponding approximation of \( f(y_t, s_t|Y_{t-1}) \) is then interpreted as the marginal of the mixed density

\[
\tilde{f}(y_t, s_t, s|Y_{t-1}) = \pi_{t-1}^{k} f(y_t|s_t, Y_{t-1}) f(s_t|s_{t-1}^{0,k}, Y_{t-1}),
\]

with an auxiliary discrete random variable \( k \in \{1, ..., N\} \) (omitting a subscript \( t \) for ease of notation).

The adaption step consists of constructing an importance sampler \( g(s_t, k|Y_{t}) \). In its simplest form this obtains by replacing \( s_t \) in \( f(y_t|s_t, Y_{t-1}) \) with its conditional expectation \( \mu_{t}^{k} = E \left( s_t|s_{t-1}^{0,k}, Y_{t-1} \right) \). The corresponding (normalized) importance sampler is then given by

\[
g(s_t, k|Y_{t}) = \lambda_{k} f(s_t|s_{t-1}^{0,k}, Y_{t-1}),
\]

with

\[
\lambda_{k} = \frac{1}{D} \pi_{t-1}^{k} f \left( y_t|\mu_{t}^{k}, Y_{t-1} \right),
\]

\[
D = \sum_{j=1}^{N} \pi_{t-1}^{j} f \left( y_t|\mu_{t}^{j}, Y_{t-1} \right).
\]

Draws from \( g(s_t, k|Y_{t}) \) are obtained as follows: first draw \( k^{i} \in \{1, ..., N\} \) with probabilities \( \left\{ \lambda_{k} \right\}_{k=1}^{N} \); next, conditionally upon \( k^{i} \), draw \( s_{t}^{0,i} \) from \( f(s_t|s_{t-1}^{0,k^{i}}, Y_{t-1}) \). The IS estimate of the period-\( t \) likelihood is then given by

\[
\tilde{f}_{N}(y_t|Y_{t-1}) = \frac{1}{N} \sum_{i=1}^{N} \frac{\tilde{f}\left(y_t, s_{t}^{0,i}, k^{i}|Y_{t-1}\right)}{g\left(s_{t}^{0,i}, k^{i}|Y_{t}\right)}
\]

\[
= \frac{D}{N} \sum_{i=1}^{N} \omega_{i}^{j},
\]

with

\[
\omega_{i}^{j} = \frac{f\left(y_t|s_{t}^{0,i}, Y_{t-1}\right)}{f\left(y_t|\mu_{t}^{k^{i}}, Y_{t-1}\right)},
\]

and \( \left\{ s_{t}^{0,i}, k^{i} \right\}_{i=1}^{N} \) denoting i.i.d. draws from \( g(s_t, k|Y_{t}) \).

Similarly, the density \( f(s_{t+1}|Y_{t}) \) is approximated by

\[
\tilde{f}_{N}(s_{t+1}|Y_{t}) = \frac{1}{f_{N}(y_t|Y_{t-1})} \int f(s_{t+1}|s_t, Y_{t}) \tilde{f}(y_t, s_t|Y_{t-1}) ds_t,
\]
whose IS estimate under $g()$ is given by

$$\tilde{f}_N(s_{t+1}|Y_t) = \sum_{i=1}^{N} \pi^i_t f \left( s_{t+1}|s^0_t, Y_t \right),$$

with

$$\pi^i_t = \frac{\omega^i_t}{\sum_{j=1}^{N} \omega^j_t}.$$ 

### 4.2 Adapted Particle Filter

Note from (31) that $g(s_t, k|Y_t)$ is based implicitly on a zero-order Taylor series expansion of $\ln f(y_t|s_t, Y_{t-1})$ around $s_t = \mu^k_t$. Further adaption obtains if a higher-order expansion can be implemented feasibly.

Let the exponential version of such an approximation be given by

$$f(y_t|s_t, Y_{t-1}) \simeq f \left( y_t|\mu^k_t, Y_{t-1} \right) h \left( s_t; \mu^k_t, Y_t \right).$$

Combining $h(\cdot)$ with $f(s_t|s^0_{t-1}, Y_{t-1})$, we obtain

$$f_*(s_t|\mu^k_{st}, Y_t) = \frac{f \left( s_t|s^0_{t-1}, Y_{t-1} \right) h \left( s_t; \mu^k_t, Y_t \right)}{\chi(\mu^k_{st}, Y_t)},$$

where $\chi(\cdot)$ and $\mu^k_{st}$ denote the integrating constant and reparameterization associated with the replacement of $f(\cdot)$ by $f_*(\cdot)$.

Two requirements must be met in order for $f_*(\cdot)$ to be implemented feasibly. First, it must be possible to draw from $f_*(\cdot)$; second, $\chi(\cdot)$ must be an analytical expression, since as shown below, it is used to construct adapted resampling weights. As explained, e.g., in Richard and Zhang (2007), these conditions are met in working within the exponential family of distributions, which are closed under multiplication. (Pitt and Shephard, 1999, present examples involving first-order approximations; and Smith and Santos, 2006, present examples involving second-order expansions.)

Given the use of $f_*(\cdot)$, the adapted importance sampler obtains by replacing $f(y_t|s_t, Y_{t-1})$ in (27) by its approximation in (32):

$$g_*(s_t, k|Y_t) = \lambda^k f_*(s_t|\mu^k_{st}, Y_t),$$
with

$$\chi_k^t = \frac{1}{D_*} \pi_{t-1}^k f \left( y_t | \mu_{t}^k, Y_{t-1} \right) \chi \left( \mu_{st}^k; Y_t \right),$$

(35)

$$D_* = \sum_{j=1}^{N} \pi_{t-1}^j f \left( y_t | \mu_{t}^j, Y_{t-1} \right) \chi \left( \mu_{st}^j; Y_t \right).$$

(36)

The IS weight in (31) is replaced by

$$\omega_{st}^i = \frac{f \left( y_t | s_{0,i}^t, Y_{t-1} \right)}{f \left( y_t | \mu_{t}^i, Y_{t-1} \right) h \left( s_t; \mu_{t}^i, Y_t \right)}.$$ 

(37)

Relative to \( \omega_{st}^i \) in (31), \( \omega_{st}^i \) has a smaller MC sampling variance, since it is based upon a higher-order Taylor series expansion of \( f (y_t|s_t,Y_{t-1}) \). The corresponding likelihood estimate is

$$\hat{f}_{sN} (y_t|Y_{t-1}) = \frac{D_*}{N} \sum_{i=1}^{N} \omega_{st}^i.$$ 

5 **Pseudo-Code for the Auxiliary Particle Filter**

\textbf{t=1 initialization:} Given \( f(s_0) \), draw \( N \) values from \( f(s_0) \) to create the cluster \( \left\{ s_{0,k}^0, \pi_{0,k}^0 \right\}_{k=1}^{N} \) \( (\pi_{0,k}^0 = \frac{1}{N} \forall k) \).

\textbf{Step 1: (Create the IS} \( g(s_t,k|Y_t) \)) At period \( t \), we have the cluster \( \left\{ s_{t-1,k}^0, \pi_{t-1,k}^0 \right\}_{k=1}^{N} \) from the previous period.

- For each \( s_{t-1,k}^0 \), compute the conditional expectation

$$\mu_{t}^k = E \left( s_t | s_{t-1,k}^0, Y_{t-1} \right),$$

(38)

using the known distribution \( f(s_t|s_{t-1}, Y_{t-1}) \) associated with (1).

- Compute weights \( \lambda_k \) using

$$\lambda_k = \frac{1}{D} \pi_{t-1}^k f(y_t | \mu_{t}^k),$$

(39)

$$D = \sum_{j=1}^{N} \pi_{t-1}^j f(y_t | \mu_{t}^j),$$

(40)

where \( f(y_t|\cdot) \) is the known distribution associated with (2).

- This gives us the "first-stage weights".

\textbf{Step 2: (Sample from the IS)} Draws from \( g(s_t,k|Y_t) \) are obtained as follows:
- Draw $k^i \in \{1, ..., N\}$ with replacement, using probabilities $\{\lambda_k\}_{k=1}^{N}$.
- Obtain the associated particle $s_{t-1}^{0, k^i}$ from the swarm $\{s_{t-1}^{0,k}\}_{k=1}^{N}$.
- Conditional upon $s_{t-1}^{0,k^i}$, draw $s_t^{0,i}$ from $f(s_t|s_{t-1}^{0,k^i}, Y_{t-1})$.

Step 3: (Likelihood Evaluation) The IS estimate of the period-$t$ likelihood in (5) is then given by

$$\hat{f}_N (y_t|Y_{t-1}) = \frac{D}{N} \sum_{i=1}^{N} \omega_t^i$$

$$\omega_t^i = \frac{f(y_t|s_t^{0,k^i}, Y_{t-1})}{f(y_t|\mu_k^t, Y_{t-1})}.$$  

Step 4: (Define $\pi_t^k$) The cluster $\{s_t^{0,k}, \pi_t^k\}_{k=1}^{N}$ constitutes the computational pre-requisites for the above steps in period $t+1$, with $\pi_t^k$ given by

$$\pi_t^k = \frac{\omega_t^k}{\sum_{j=1}^{N} \omega_t^j}.$$  

Step 5: (Conditional Expectation of $h(s_t)$) The MC estimate for $E_t(h(s_t)|Y_t)$ is given by

$$E_t(h(s_t)|Y_t) = \sum_{i=1}^{S} h(s_t^{0,j}) \cdot \pi_t^i.$$  

Thus, return to Step 1 using this new cluster $\{s_t^{0,k}, \pi_t^k\}_{k=1}^{N}$, and repeat Steps 1-5 until period $T$ has been reached.

6 Pseudo-Code for the Adapted Particle Filter

t=1 initialization: Given $f(s_0)$, draw $N$ values from $f(s_0)$ to create the cluster $\{s_0^{0,k}, \pi_0^k\}_{k=1}^{N}$ ($\pi_0^k = \frac{1}{N} \forall k$).

Step 1: (Create the IS $g(s_t,k|Y_t)$) At period $t$, we have the cluster $\{s_{t-1}^{0,k}, \pi_{t-1}^k\}_{k=1}^{N}$ from the previous period.

- For each $s_{t-1}^{0,k}$, compute the conditional expectation

$$\mu_t^k = E(s_t|s_{t-1}^{0,k}, Y_{t-1}),$$

using the known distribution $f(s_t|s_{t-1}^{0,k}, Y_{t-1})$ associated with (1).
• Compute weights \( \lambda_k \) using

\[
\lambda_k = \frac{1}{D} \pi_{t-1}^k f(y_t|\mu_t^k) \chi(\mu_t^k), \tag{46}
\]

\[
D = \sum_{j=1}^N \pi_{t-1}^j f(y_t|\mu_t^j) \chi(\mu_t^j), \tag{47}
\]

where \( f(y_t|\cdot) \) is the known distribution associated with (2).

• This gives us the "first-stage weights".

**Step 2: (Sample from the IS)** Draws from \( g(s_t, k|Y_t) \) are obtained as follows:

• Draw \( k^i \in \{1, \ldots, N\} \) with replacement, using probabilities \( \{\lambda_k\}_{k=1}^N \).

• Obtain the associated particle \( s_{t-1}^{0,k^i} \) from the swarm \( \{s_{t-1}^{0,k}\}_{k=1}^N \).

• Conditional upon \( s_{t-1}^{0,k^i} \), draw \( s_t^{0,i} \) from \( f_*(s_t) \).

**Step 3: (Likelihood Evaluation)** The IS estimate of the period-\( t \) likelihood in (5) is then given by

\[
\tilde{f}_N (y_t|Y_{t-1}) = \frac{D}{N} \sum_{i=1}^N \omega_i^*, \tag{48}
\]

\[
\omega_i^* = \frac{f(y_t|s_t^{0,i}, Y_{t-1})}{f(y_t|\mu_t^{k^i}, Y_{t-1}) h(s_t; \mu_t^{k^i})}. \tag{49}
\]

**Step 4: Define \( \pi_t^k \)** The cluster \( \{s_t^{0,k}, \pi_t^k\}_{k=1}^N \) constitutes the computational pre-requisites for the above steps in period-\( t+1 \), with \( \pi_t^k \) given by

\[
\pi_t^k = \frac{\omega_i^k}{\sum_{j=1}^N \omega_t^j}. \tag{50}
\]

**Step 5: (Conditional Expectation of \( h(s_t) \))** The MC estimate for \( E_t(h(s_t)|Y_t) \) is given by

\[
E_t(h(s_t)|Y_t) = \sum_{i=1}^S h(s_t^{0,i}) \cdot \pi_t^i \tag{51}
\]

Thus, return to Step 1 using this new cluster \( \{s_t^{0,k}, \pi_t^k\}_{k=1}^N \), and repeat Steps 1-5 until period \( T \) has been reached.