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.
- $s_t^{1,i}$: i^{th} draw of s_t conditional on (Y_{t-1}, S_{t-1}).
- $s_t^{0,i}$: 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:
$$s_t = \gamma(s_{t-1}, Y_{t-1}) + \upsilon_t, \tag{1}$$
where $\gamma(s_{t-1}, Y_{t-1}) = E(s_t|s_{t-1}, Y_{t-1})$, and υ_t is a vector of structural shocks.

- Observation (or measurement) equation:
$$y_t = \delta\left(s_t, Y_{t-1}\right) + u_t, \tag{2}$$
where $\delta\left(s_t, Y_{t-1}\right) = E(y_t|s_t, Y_{t-1})$, and thus u_t is a vector of measurement errors.

- The likelihood function $f(Y_T)$:
$$f\left(Y_T\right) = \prod_{t=1}^{T} f\left(y_t|Y_{t-1}\right), \tag{3}$$
where $f\left(y_1|Y_0\right) \equiv f(y_1)$.

- The individual factors in the likelihood:
$$f\left(y_t|Y_{t-1}\right) = \int f\left(y_t|s_t, Y_{t-1}\right) f\left(s_t|Y_{t-1}\right) ds_t. \tag{4}$$

- Recursive evaluation of densities $f(s_t|Y_{t-1})$:
$$f\left(s_t|Y_{t-1}\right) = \int f\left(s_t|s_{t-1}, Y_{t-1}\right) f\left(s_{t-1}|Y_{t-1}\right) ds_{t-1}, \tag{5}$$
where
$$f\left(s_t|Y_t\right) = \frac{f\left(y_t, s_t|Y_{t-1}\right)}{f\left(y_t|Y_{t-1}\right)} = \frac{f\left(y_t|s_t, Y_{t-1}\right) f\left(s_t|Y_{t-1}\right)}{f\left(y_t|Y_{t-1}\right)}. \tag{6}$$

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 $\left\{s_0^{0,i}\right\}_{i=1}^N$. 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 $\left\{s_1^{1,i}\right\}_{i=1}^N$, which serves as a discrete approximation of $f(s_1|Y_0, S_0)$.

Step 1: (Prediction) At period t, we have the cluster $\left\{s_{t-1}^{0,i}\right\}_{i=1}^N$ (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 $\left\{s_t^{1,i}\right\}_{i=1}^N$. This cluster serves as a discrete approximation of $f(s_t|Y_{t-1})$, in light of (5).

Step 2: (Likelihood Evaluation) Using $\left\{s_t^{1,i}\right\}_{i=1}^N$, approximate the likelihood in (4) as

$$\widehat{f}_N(y_t|Y_{t-1}) = \frac{1}{N}\sum_{i=1}^N f(y_t|s_t^{1,i}, Y_{t-1}). \tag{7}$$

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

$$w_t^{0,i} = \frac{f(y_t|s_t^{1,i}, Y_{t-1})}{\sum_{j=1}^N f(y_t|s_t^{1,j}, Y_{t-1})}. \tag{8}$$

Sampling with replacement from the cluster $\left\{s_t^{1,i}\right\}_{i=1}^N$ using the weights $\left\{w_t^{0,i}\right\}_{i=1}^N$ yields the new cluster $\left\{s_t^{0,i}\right\}_{i=1}^N$ (which approximates $f(s_t|Y_t)$), in light of (6). Return to Step 1 using this new cluster, and repeat Steps 1-3 until period T has been reached.

3 Pseudo-Code for the Auxiliary Particle Filter

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

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

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

$$\mu_t^k = E\left(s_t|s_{t-1}^{0,k}, Y_{t-1}\right), \tag{9}$$

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

- Compute weights λ_k using

$$\lambda_k = \frac{1}{D}\pi_{t-1}^k f(y_t|\mu_t^k), \tag{10}$$

$$D = \sum_{j=1}^N \pi_{t-1}^j f(y_t|\mu_t^j), \tag{11}$$

  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 $g\left(s_t, k|Y_t\right)$ 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 $\left\{s_{t-1}^{0,k}\right\}_{k=1}^N$.
- 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 (4) is then given by

$$\widehat{f}_N\left(y_t|Y_{t-1}\right) = \frac{D}{N}\sum_{i=1}^N \omega_t^i \tag{12}$$

$$\omega_t^i = \frac{f\left(y_t|s_t^{0,k^i}, Y_{t-1}\right)}{f\left(y_t|\mu_t^{k^i}, Y_{t-1}\right)}. \tag{13}$$

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

$$\pi_t^k = \frac{\omega_t^k}{\sum_{j=1}^N \omega_t^j}. \tag{14}$$

Thus, return to Step 1 using this new cluster, and repeat Steps 1-4 until period T has been reached.

4 Pseudo-Code for the Adapted Particle Filter

4.1 Pre-Requisites

For the adapted particle filter, we require additional notation. As above, $\mu_t^k = E\left(s_t|s_{t-1}^{0,k}, Y_{t-1}\right)$. Also, let $h(s_t;\mu_t^k)$ be the first-order term in the Taylor expansion of $\ln f(y_t|s_t, Y_{t-1})$ around μ_t^k (second-order expansions are also feasible). The product $f(s_t|s_{t-1}^{0,k})h(s_t;\mu_t^k)$ is then transformed into a new density $f_*(s_t)$ for s_t centered around a new mean μ_{*t}^k, and a remainder term $\chi(\mu_{*t}^k)$.

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

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

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

$$\mu_t^k = E\left(s_t|s_{t-1}^{0,k}, Y_{t-1}\right), \tag{15}$$

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

- Compute weights λ_k using

$$\lambda_k = \frac{1}{D}\pi_{t-1}^k f(y_t|\mu_t^k)\chi(\mu_{*t}^k), \tag{16}$$

$$D = \sum_{j=1}^N \pi_{t-1}^j f(y_t|\mu_t^j)\chi(\mu_{*t}^k), \tag{17}$$

  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 $k^i \in \{1, ..., N\}$ with replacement, using probabilities $\{\lambda_k\}_{k=1}^N$.
- Obtain the associated particle s_{t-1}^{0,k^i} from the swarm $\left\{s_{t-1}^{0,k}\right\}_{k=1}^N$.
- 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 (4) is then given by

$$\widehat{f}_N(y_t|Y_{t-1}) = \frac{D}{N}\sum_{i=1}^N \omega_t^i, \tag{18}$$

$$\omega_t^i = \frac{f\left(y_t|s_t^{0,i}, Y_{t-1}\right)}{f\left(y_t|\mu_t^{k^i}, Y_{t-1}\right)h(s_t;\mu_t^{k^i})}. \tag{19}$$

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

$$\pi_t^k = \frac{\omega_t^k}{\sum_{j=1}^N \omega_t^j}. \tag{20}$$

Thus, return to Step 1 using this new cluster, and repeat Steps 1-4 until period T has been reached.

5 Pseudo-Code for EIS Particle 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).

5.1 Gaussian-EIS Particle Filter

With the IS densities being gaussian, the mean and variance of the sampling densities become the auxiliary parameters.

Step 1: (Initialize Sampler $g(s_t; a_t)$) At period t, we have the EIS draws and their corresponding weights $\left\{s_{t-1}^{0,k}, \omega_{t-1}^{k}\right\}_{k=1}^{N}$ from the previous period. Choose the initial value of the auxiliary parameters a_t^l, $l = 0$: 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.

- Repeat the following steps until convergence.
    1. Draw R values of s_t from $g(s_t; a_t^l)$; denote these draws as $\left\{s_t^{i,l}\right\}_{i=1}^{R}$.
    2. Obtain updated values of a_t^{l+1} as the solution to the least squares problem,

    $$(a_t, c_t)^{l+1} = \arg\min_{a_t, c_t} \sum_{i=1}^{R} \left[\ln\left(f(y_t|s_t^{i,l})\widehat{f}(s_t^{i,l}|Y_{t-1})\right) - c_t - \ln g(s_t^{i,l}; a_t)\right]^2, \quad (21)$$

    where $f(y_t|\cdot)$ is the known distribution associated with (2) and c_t is an intercept meant to calibrate $\ln\left(\frac{f(y_t|s_t)\widehat{f}(s_t|Y_{t-1})}{g(s_t;a_t)}\right)$. (Details on the least-squares problem are provided below.)
    3. Check for convergence.
- Once convergence is reached, we have the optimal mean and variance of the EIS sampling density: $\widehat{a}_t$.

Step 3: (Likelihood Evaluation) Draw N values $\{s_t^i\}_{i=1}^{N}$ from the optimal EIS sampling density $g\left(s_t; \widehat{a}_t\right)$. The IS estimate of the period-t likelihood in (4) and the EIS weights are given by

$$\widehat{f}_N\left(y_t|Y_{t-1}\right) = \frac{1}{N}\sum_{i=1}^{N}\omega_t^i, \quad (22)$$

$$\omega_t^i = \frac{f(y_t|s_t^i)\widehat{f}(s_t^i|Y_{t-1})}{g\left(s_t^i; \widehat{a}_t\right)}. \quad (23)$$

The EIS draws and their corresponding weights $\left\{s_t^{0,k}, \omega_t^k\right\}_{k=1}^{N}$ 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 1-3 until period T has been reached.

5.1.1 Details on the Auxiliary Regression in Gaussian-EIS

Let the product of densities $f(y_t|s_t)\widehat{f}(s_t|Y_{t-1})$ 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_t^l as given, initialized by a_t^0. Hereafter, we will drop the superscript l 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 μ, and the precision matrix (the inverse of the covariance matrix) by H. The setup of the LS problem arises from the approximation $\ln \varphi(s)$ by a gaussian kernel:

$$\begin{aligned} \ln \varphi(s) &\propto -\frac{1}{2}(s-\mu)'H(s-\mu) \\ &\propto -\frac{1}{2}\left(s'Hs - 2s'H\mu\right). \end{aligned}$$

The term $s'Hs$ can be written as,

$$\begin{pmatrix} x_1 & x_2 & . & . & x_j \end{pmatrix} \begin{pmatrix} h_{11} & h_{21} & . & . & h_{1j} \\ h_{21} & h_{22} & . & . & h_{2j} \\ . & . & . & . & . \\ . & . & . & . & . \\ h_{j1} & h_{j2} & . & . & h_{jj} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_j \end{pmatrix}$$

$$\begin{aligned} = \; & h_{11}\left(x_1^2\right) + h_{22}\left(x_2^2\right) + ... + h_{jj}\left(x_j^2\right) \\ & +2h_{21}\left(x_2x_1\right) + 2h_{31}\left(x_3x_1\right) + ... + 2h_{j1}\left(x_jx_1\right) \\ & +2h_{32}\left(x_3x_2\right) + 2h_{42}\left(x_4x_2\right) + ... + 2h_{j2}\left(x_jx_2\right) \\ & ... \\ & +2h_{j(j-1)}\left(x_jx_{j-1}\right). \end{aligned}$$

From the above decomposition it is evident 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 LS problem reduces to the regression of $\ln \varphi(s)$ on

$\left[1, x_1^2, x_2^2, ..., x_j^2, \;\; x_1x_2, x_1x_3, ..., x_{j-1}x_j, \;\; x_1, ..., x_j\right]$. For a j-dimensional variable s, the number of regressors is $\left(1 + j + \frac{j(j+1)}{2}\right)$.

As a concrete example, consider a 3-dimensional problem where $s = [x_1, x_2, x_3]$. The regression reduces to

$$\begin{aligned} \ln \varphi(s) &= \beta_0 + \beta_1\left(x_1^2\right) + \beta_2\left(x_2^2\right) + \beta_3\left(x_3^2\right) \\ &\quad +\beta_4\left(x_2x_1\right) + \beta_5\left(x_3x_1\right) + \beta_6\left(x_3x_2\right) \\ &\quad +\beta_7x_1 + \beta_8x_2 + \beta_9x_3. \end{aligned}$$

Having obtained the LS estimates, the updated precision matrix is given by

$$\begin{aligned} h_{11} &= -2\beta_1;\; h_{22} = -2\beta_2;\; h_{33} = -2\beta_3 \\ h_{21} &= -\beta_4;\; h_{31} = -\beta_5;\; h_{32} = -\beta_6. \end{aligned}$$

The updated means can be obtained by using the coefficients $(\beta_7, \beta_8, \beta_9)$:

$$\mu = H^{-1} \begin{pmatrix} \beta_7 \\ \beta_8 \\ \beta_9 \end{pmatrix}.$$

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

$$\begin{aligned} \sigma^2 &= -\frac{1}{2\beta_2}, \\ \mu &= -\frac{1}{2}\frac{\beta_1}{\beta_2}. \end{aligned}$$

5.2 Piecewise-EIS Particle Filter

With the IS densities being piecewise-continuous, the coefficients of the piecewise approximations and the location of the nodes become the auxiliary parameters.

Notation: Let the product of densities $f(y_t|s_t)\widehat{f}(s_t|Y_{t-1})$ be denoted by $\varphi(s_t)$.

Step 1: (Initial Approximation) At period t, we have the EIS draws and their corresponding weights $\left\{s_{t-1}^{0,k}, \omega_{t-1}^{k}\right\}_{k=1}^{N}$ 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 (as in eqn. 28 in the paper):

$$\begin{aligned} \ln k_j(s;a) &= \alpha_j + \beta_j s \qquad \forall s \in [a_{j-1}, a_j], \\ \beta_j &= \frac{\ln\varphi(a_j) - \ln\varphi(a_{j-1})}{a_j - a_{j-1}}, \quad \alpha_j = \ln\varphi(a_j) - \beta_j a_j, \end{aligned}$$

  where $k(s;a)$ is the kernel of the EIS sampling density.
- Compute the CDF defined by these linear segments:

$$\begin{aligned} K_j(s;a) &= \frac{\chi_j(s;a)}{\chi_n(a)}, \qquad \forall s \in [a_{j-1}, a_j], \\ \chi_j(s;a) &= \chi_{j-1}(a) + \frac{1}{\beta_j}\left[k_j(s;a) - k_j(a_{j-1};a)\right], \\ \chi_0(a) &= 0, \quad \chi_j(a) = \chi_j(a_j;a). \end{aligned}$$

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

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

- 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 optimal EIS sampling density $k(s_t;\widehat{a}_t)$. The IS estimate of the period-t likelihood in (4) and the EIS weights are given by

$$\widehat{f}_N(y_t|Y_{t-1}) = \frac{1}{N}\sum_{i=1}^{N}\omega_t^i, \tag{24}$$

$$\omega_t^i = \frac{f(y_t|s_t^i)\widehat{f}(s_t^i|Y_{t-1})}{k(s_t^i;\widehat{a}_t)}. \tag{25}$$

The EIS draws and their corresponding weights $\left\{s_t^{0,k}, \omega_t^k\right\}_{k=1}^{N}$ 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 1-3 until period T has been reached.