# Sequential Monte Carlo Filtering: an Example*

Jesús Fernández-Villaverde — University of Pennsylvania

Juan F. Rubio-Ramírez — Federal Reserve Bank of Atlanta

January 5, 2004

**Abstract**

This short note presents an example of how to use a Sequential Monte Carlo to evaluate the likelihood of a nonlinear and nongaussian process.

*Corresponding author: Jesús Fernández-Villaverde, Department of Economics, 160 McNeil Building, 3718 Locust Walk, University of Pennsylvania, Philadelphia PA 19104. E-mail: jesusfv@econ.upenn.edu. We must notice that any views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or of the Federal Reserve System.

This short note presents an example of how to evaluate the likelihood function of a nonlinear and nonnormal stochastic process using a Sequential Monte Carlo Filter. More details, including technical requirements can be found in Fernández-Villaverde and Rubio-Ramírez (2004).

Suppose that we are interested in evaluating the likelihood function $L\left(y^T\right)$ of the nonlinear, nonnormal process:

$$\begin{aligned} x_t &= 0.5 + 0.3\frac{x_{t-1}}{1+x_{t-1}^2} + w_t \\ y_t &= x_t + v_t \end{aligned}$$

where $w_t \sim \mathcal{N}(0,1)$ and $v_t \sim t(2)$ given some observables $y^T = \{y_t\}_{t=1}^T$.

The first equation, known as the transition equation, accounts for the evolution of the state $x_t$ while the second one, known as the observation equation, relates the state with the observable. For simplicity only the first equation is nonlinear and only the second one has nonnormal innovations. We can handle the more general case where nonlinearities and nonnormalities are present in both equations without further problem.

Both the nonlinear structure of the process and the presence of nonnormal innovations stop us from using the Kalman filter or other similar procedure to evaluate the likelihood function. However a Sequential Monte Carlo filter delivers a consistent evaluation of that function using simulation methods. See Fernández-Villaverde and Rubio-Ramírez (2004) for details.

For simplicity we assume that $x_0$ is known.[1] Then we can proceed as follows:

1. Fix a large integer $N$.

2. Draw $N$ samples from $\mathcal{N}(0,1)$. Call each of those $w^{1|0,i}$ for $i = 1, ..., N$. Use those values to generate $N$ values of $x^{1|0,i}$ with the transition equation:

$$x^{1|0,i} = 0.5 + 0.3\frac{x_0}{1+x_0^2} + w^{1|0,i}$$

3. Use $y_1$ and each of the $N$ $x^{1|0,i}$ from the previous step to find $N$ values of $v^{1|0,i}$:

$$v^{1|0,i} = y_1 - x^{1|0,i}$$

[1] We could solve the problem when $x_0$ is unknown paying a cost in terms of further notation. See again Fernández-Villaverde and Rubio-Ramírez (2004) for the general treatment.

4. Evaluate the relative probability of each of the $N$ $v^{1|0,i}$

$$q_1^i = \frac{p\left(v^{1|0,i}\right)}{\sum_{i=1}^{N} p\left(v^{1|0,i}\right)}$$

using the fact that $v_1 \sim t\,(2)$.

5. Resample with replacement $N$ values of $x^{1|0,i}$ with relative weights $q_1^i$. Call each of those sampled values $x_1^i$.

6. Go to 2. Draw $N$ samples from $\mathcal{N}\,(0,1)$. Call each of those $w^{2|1,i}$ for $i = 1, ..., N$. Use those values to generate $N$ values of $x^{2|1,i}$ with the transition equation

$$x^{2|1,i} = 0.5 + 0.3\frac{x_1^i}{1 + \left(x_1^i\right)^2} + w^{2|1,i}$$

and proceed as in steps 3-5.

7. Iterate until the end of the sample.

Now, since $N$ is large, a law of the large numbers delivers that:

$$p\left(y_t|\, y^{t-1}\right) \simeq \frac{1}{N}\sum_{i=1}^{N} p\left(v^{t|t-1,i}\right)$$

and with the factorization of the likelihood function we have:

$$p\left(y^T\right) \simeq \prod_{t=1}^{T}\frac{1}{N}\sum_{i=1}^{N} p\left(v^{t|t-1,i}\right)$$

The key of the algorithm is step 5. A naive extension of basic Monte Carlo techniques without this step will diverge as $T$ grows. Why? The naive approach will generate $N$ sequences $\left\{w^{t|t-1,i}\right\}_{t=1}^{T}$ collecting together independent draws from a normal distribution and then use the transition and measurement equation and the observables to find the corresponding $\left\{v^{t|t-1,i}\right\}_{t=1}^{T}$. But this means that as $T$ grows all the sequences of simulated states will wander away from the true (unobserved) states. To avoid this problem we do not carry over to the next period all the particles generated. We draw with replacement from them, giving a higher probability to those particles that are more likely given our process. The fitting criterion is critical to the convergence of the procedure and very similar to the intuition behind why genetic algorithms work: we allow randomness to generate new simulations but we favor the survival of the particles that are more informative.

Once we have the likelihood function standard inference is simple. For example, if instead of the process described above, we have the case

$$\begin{aligned} x_t &= \alpha + \beta \frac{x_{t-1}}{1 + x_{t-1}^2} + w_t \\ y_t &= \delta x_t + v_t \end{aligned}$$

where $w_t \sim \mathcal{N}(0, \sigma)$ and $v_t \sim t(2)$ where $\alpha, \beta, \delta$ and $\sigma$ are unknown parameters, we can apply our filter to compute $p\left(y^T \middle| \alpha, \beta, \delta, \sigma\right)$ for each particular value of the parameters. Then we can either maximize this function with respect to the four parameters to undertake classical inference or, after postulating some prior distribution $\pi(\alpha, \beta, \delta, \sigma)$, find the posterior distribution:

$$p\left(\alpha, \beta, \delta, \sigma \middle| y^T\right) \propto p\left(y^T \middle| \alpha, \beta, \delta, \sigma\right) \pi(\alpha, \beta, \delta, \sigma)$$

to perform Bayesian inference.

## References


[1] Fernández-Villaverde, J. and J. Rubio-Ramírez (2004), "Estimating Nonlinear Dynamic Equilibrium Economies: A Likelihood Approach". *Mimeo,* University of Pennsylvania. Available at `www.econ.upenn.edu/~jesusfv`