APF, SIR & LW

- Consider the following state space model

$$\begin{aligned} \text{Observation equation} &: \quad p(y_{t+1}|x_{t+1},\theta) \\ \text{State equation} &: \quad p(x_{t+1}|x_t,\theta) \\ \text{Fixed parameters} &: \quad p(\theta) \end{aligned}$$

- For a given time t

$$\{(x_t,\theta,\omega_t)^{(i)}\}_{i=1}^N$$

is a particle representation of $p(x_t,\theta|y^t)$.

- $y^t = (y_1,\ldots,y_n)$

APF with known θ

- Pitt and Shephard (1999) propose augmenting

$$p^N(x_{t+1}, k|y^{t+1}) \propto p(x_{t+1}|x_t^{(k)})p(y_{t+1}|x_{t+1})\omega_t^{(k)}$$

- Proposal

$$q(x_{t+1}, k|y^{t+1}) \propto p(x_{t+1}|x_t^{(k)})p(y_{t+1}|g(x_t^{(k)}))\omega_t^{(k)}$$

from which $\{(\tilde{k}, \tilde{x}_{t+1})\}_{i=1}^N$ are generated.

- New weights

$$\tilde{\omega}_{t+1} \propto \frac{p(y_{t+1}|\tilde{x}_{t+1})}{p(y_{t+1}|g(x_t^{\tilde{k}}))}$$

SIR argument with known θ

The objective is to obtain

$$\{(x_t, x_{t+1}, \omega_{t+1})^{(i)}\}_{i=1}^N \sim p(x_t, x_{t+1}|y^{t+1})$$

based on $\{(x_t, \omega_t)^{(i)}\}_{i=1}^N \sim p(x_t|y^t)$.

- Resample $x_t^* \sim q_1(x_t|y_{t+1})$
- Propagate $\tilde{x}_{t+1} \sim q_2(x_{t+1}|x_t^*, y_{t+1})$
- Weight

$$\tilde{\omega}_{t+1} \propto \frac{\omega_t p(y_{t+1}|\tilde{x}_{t+1})}{q_1(x_t^*|y_{t+1})} \frac{p(\tilde{x}_{t+1}|x_t^*)}{q_2(\tilde{x}_{t+1}|x_t^*, y_{t+1})}$$

- Notes:
  - How to choose q_1 and q_2?
  - APF: $q_1 = \omega_t p(y_{t+1}|g(x_t^*))$ and $q_2 = p(\tilde{x}_{t+1}|x_t^*)$.

Perfect adaptation

- Weights are constant when

$$\begin{aligned} q_1(x_t|y_{t+1}) &= p(x_t|y_{t+1}) \\ q_2(x_{t+1}|x_t, y_{t+1}) &= p(x_{t+1}|x_t, y_{t+1}). \end{aligned}$$

- Both q_1 and q_2 depend on y_{t+1}.

SIR with learning of θ

The objective is to obtain

$$\{(x_t, x_{t+1}, \theta, \omega_{t+1})^{(i)}\}_{i=1}^N \sim p(x_t, x_{t+1}, \theta | y^{t+1})$$

based on $\{(x_t, \theta, \omega_t)^{(i)}\}_{i=1}^N \sim p(x_t, \theta | y^t)$.

- Resample $(x_t^*, \theta^*) \sim q_1(x_t, \theta | y_{t+1})$
- Propagate $(\tilde{x}_{t+1}, \tilde{\theta}) \sim q_2(x_{t+1}, \theta | x_t^*, \theta^*, y_{t+1})$
- Weight

$$\tilde{\omega}_{t+1} \propto \frac{\omega_t p(y_{t+1} | \tilde{x}_{t+1}, \tilde{\theta})}{q_1(x_t^*, \theta^* | y_{t+1})} \frac{p(\tilde{x}_{t+1}, \tilde{\theta} | x_t^*, \theta^*)}{q_2(\tilde{x}_{t+1}, \tilde{\theta} | x_t^*, \theta^*, y_{t+1})}$$

- Notes:
  - How to choose q_1 and q_2?
  - Is $p(x_{t+1}, \theta | x_t^*, \theta^*) = p(x_{t+1} | \theta, x_t^*) p(\theta | x_t^*, \theta^*)$?
  - If so, then what is $p(\theta | x_t^*, \theta^*)$?

Liu and West (2001)

They chose $p(\theta|x_t^*, \theta^*)$, q_1 and q_2 to be

$$
\begin{aligned}
p(\theta|x_t^*, \theta^*) &= f_N(\theta|m(\theta^*), h^2 V) \\
q_1(x_t^*, \theta^*|y_{t+1}) &= \omega_t p(y_{t+1}|g(x_t^*), m(\theta^*)) \\
q_2(x_{t+1}, \theta|x_t^*, \theta^*, y_{t+1}) &= q_{21}(\theta|x_t^*, \theta^*, y_{t+1}) q_{22}(x_{t+1}|\theta, x_t^*, \theta^*, y_{t+1}) \\
&= f_N(\theta|m(\theta^*), h^2 V) p(x_{t+1}|\theta, x_t^*)
\end{aligned}
$$

so weights are

$$
\tilde{\omega}_{t+1} \propto \frac{p(y_{t+1}|\tilde{x}_{t+1}, \tilde{\theta})}{p(y_{t+1}|g(x_t^*), m(\theta^*))}
$$

- Notes:
  - Neither q_{21} or q_{22} depend on y_{t+1}.

Simulation exercise

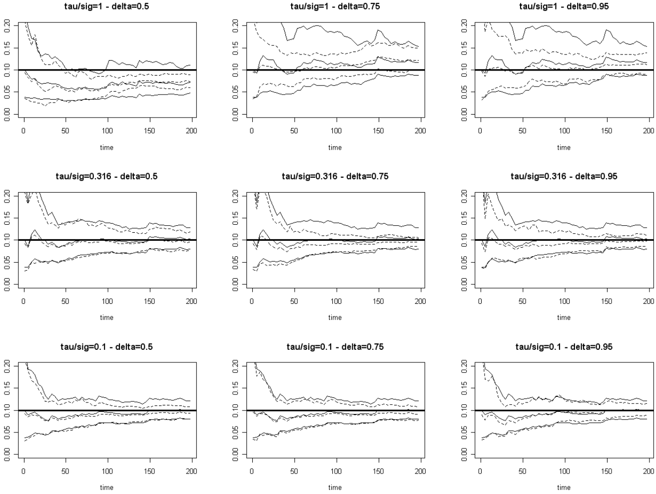
- Simulation set up

$$
\begin{aligned}
y_t &= x_t + \nu_t \qquad \nu_t \sim N(0, 0.1) \\
x_t &= x_{t-1} + \omega_t \qquad \omega_t \sim N(0, \tau^2) \\
x_0 &= 25 \\
\tau^2 &= (0.1, 0.01, 0.001) \\
\delta &= (0.5, 0.75, 0.95) \\
\text{Sample sizes} &= (200, 500) \\
\text{Particles} &= (2000, 5000)
\end{aligned}
$$

- Model set up

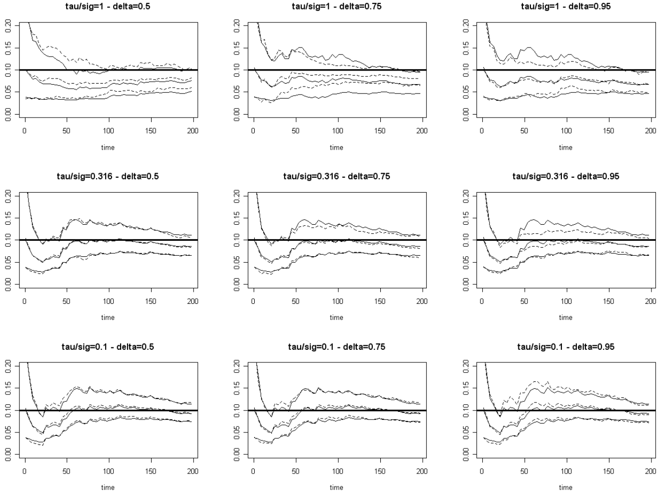
$$
\begin{aligned}
y_t &= x_t + \nu_t \qquad \nu_t \sim N(0, \sigma^2) \\
x_t &= x_{t-1} + \omega_t \qquad \omega_t \sim N(0, \tau^2) \\
\sigma^2 &\sim IG(5, 0.4) \\
x_0 &\sim N(25, 100)
\end{aligned}
$$

T=200, M=2000

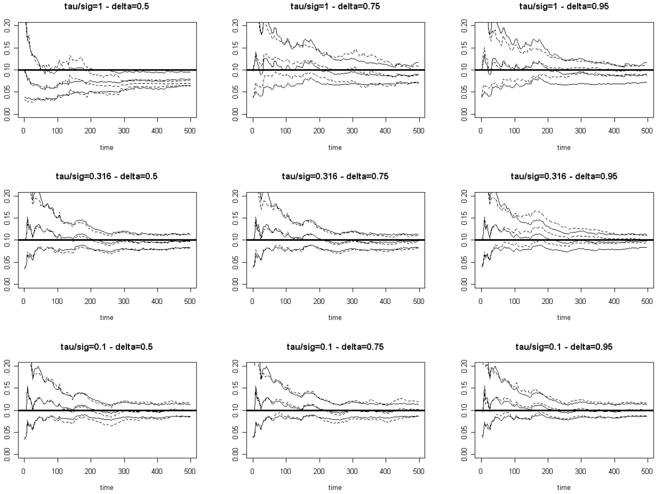


tau/sig=1 - delta=0.5
tau/sig=1 - delta=0.75
tau/sig=1 - delta=0.95
tau/sig=0.316 - delta=0.5
tau/sig=0.316 - delta=0.75
tau/sig=0.316 - delta=0.95
tau/sig=0.1 - delta=0.5
tau/sig=0.1 - delta=0.75
tau/sig=0.1 - delta=0.95
0.00
0.05
0.10
0.15
0.20
0
50
100
150
200
time

T=200, M=5000


tau/sig=1 - delta=0.5
tau/sig=1 - delta=0.75
tau/sig=1 - delta=0.95
tau/sig=0.316 - delta=0.5
tau/sig=0.316 - delta=0.75
tau/sig=0.316 - delta=0.95
tau/sig=0.1 - delta=0.5
tau/sig=0.1 - delta=0.75
tau/sig=0.1 - delta=0.95
0.00 0.05 0.10 0.15 0.20
0 50 100 150 200
time

T=500, M=5000


tau/sig=1 - delta=0.5
tau/sig=1 - delta=0.75
tau/sig=1 - delta=0.95
tau/sig=0.316 - delta=0.5
tau/sig=0.316 - delta=0.75
tau/sig=0.316 - delta=0.95
tau/sig=0.1 - delta=0.5
tau/sig=0.1 - delta=0.75
tau/sig=0.1 - delta=0.95
0.00 0.05 0.10 0.15 0.20
0 100 200 300 400 500
time

Thoughts

- $\delta = 0.75$ looks robust across T, M and τ/σ;
- The difference between PL and APF decreases with τ/σ.
  Conjecture: Probably partially because of the choice

$$q_{22}(x_{t+1}|\theta, x_t^*, \theta^*, y_{t+1}) = p(x_{t+1}|\theta^*, x_t^*)$$

  which is the usual *blind* propagation rule used in standard SIS-based filters.
- If so, then maybe

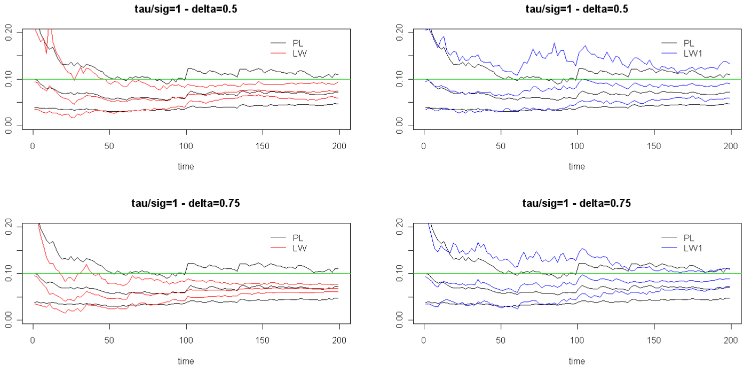
$$q_{21}(\theta|x_t^*, \theta^*, y_{t+1}) = f_N(\theta|m(\theta^*), h^2 V)$$

  is not such a bad choice!
- This will be my simulation exercise for next week.

M=2000

$$LW1: \quad q_{22}(x_{t+1}|x_t, \theta, y_{t+1}) = f_N(x_{t+1}|m_{t+1}, C_{t+1})$$

where $C_{t+1}^{-1} = \tau^{-2} + \sigma^{-2}$ and $m_{t+1} = C_{t+1}(\sigma^{-2} y_{t+1} + \tau^{-2} x_t)$.



References

- Godsill, S. and Clapp, T. (2001). Improvement strategies for Monte Carlo particle filters. In *Sequential Monte Carlo Methods in Practice*. Springer-Verlag New York.
- Liu, J. and West, M. (2001). Combined parameters and state estimation in simulation-based filtering. In *Sequential Monte Carlo Methods in Practice*. Springer-Verlag New York.
- Johansen, A. and Doucet, A. (2008). A note on auxiliary particle filters. *Statistics and Probability Letters* (to appear).
- Pitt, M. and Shephard, N. (1999) Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, 94, 590-599.