Tempered and Conditionally-Optimal Particle Filtering

Frank Schorfheide
University of Pennsylvania, CEPR, NBER

January 6, 2018

ASSA Meetings 2018

This Talk is Based on

ACS Borağan Aruoba, Pablo Cuba-Borda, and F. Schorfheide (2018): “Solution and Estimation of Approximately Piecewise-Linear DSGE Models,” research in progress.

HS Ed Herbst and F. Schorfheide (2018): “Tempered Particle Filtering,” *Journal of Econometrics*, forthcoming.

Introduction

- Nonlinear State-Space Model

$$\begin{aligned} \text{Measurement Eq.} \quad &: \quad y_t = \Psi(s_t, t; \theta) + u_t, \quad u_t \sim F_u(\cdot; \theta) \\ \text{State Transition} \quad &: \quad s_t = \Phi(s_{t-1}, \epsilon_t; \theta), \quad \epsilon_t \sim F_\epsilon(\cdot; \theta). \end{aligned}$$

- Objects of interest:
  - Estimates of states: $p(s_t | Y_{1:t}, \theta)$
  - Likelihood function: $p(Y_{1:T}|\theta) = \prod_{t=1}^{T} p(y_t | Y_{1:t-1}, \theta)$.
- Construct numerical approximation by particle filtering (sequential Monte Carlo).
- In DSGE models with occasionally-binding constraints one can often approximate $\Phi(\cdot)$ by a piecewise linear function.

Particle Filtering – Idea

- Represent distribution $p(s_t|Y_{1:t})$ by swarm of particles $\{s_t^j, W_t^j\}_{j=1}^M$ such that

$$\frac{1}{M}\sum_{j=1}^{M} h(s_t^j)W_t^j \overset{\text{SLLN,CLT}}{\approx} \mathbb{E}[h(s_t)|Y_{1:t}]$$

- Iteration t, given $\{s_{t-1}^j, W_{t-1}^j\}_{j=1}^M$

  1. **Mutation:** Draw $\tilde{s}_t^j \sim g_t(\tilde{s}_t|s_{t-1}^j)$.
  2. **Correction:** Compute incremental weights and update/normalize weights

$$\tilde{w}_t^j = \frac{p(\tilde{s}_t|s_{t-1}^j)}{g_t(\tilde{s}_t|s_{t-1}^j)}p(y_t|\tilde{s}_t^j,\theta), \quad \tilde{W}_t^j \propto \tilde{w}_t^j W_{t-1}^j.$$

  3. **Selection:** Resampling.

Implementations

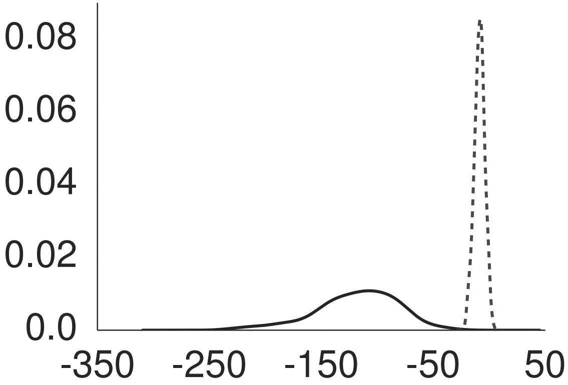
- Recall: incremental weights

$$\tilde{w}_t^j = \frac{p(\tilde{s}_t|s_{t-1}^j)}{g_t(\tilde{s}_t|s_{t-1}^j)} p(y_t|\tilde{s}_t^j, \theta)$$

- Bootstrap particle filter (BSPF): $g_t(\tilde{s}_t|s_{t-1}^j) = p(\tilde{s}_t|s_{t-1}^j)$.
- Conditionally-optimal particle filter (COPF): $g_t(\tilde{s}_t|s_{t-1}^j) \propto p(y_t|\tilde{s}_t)p(\tilde{s}_t|s_{t-1}^j)$.
- $(\ldots)$

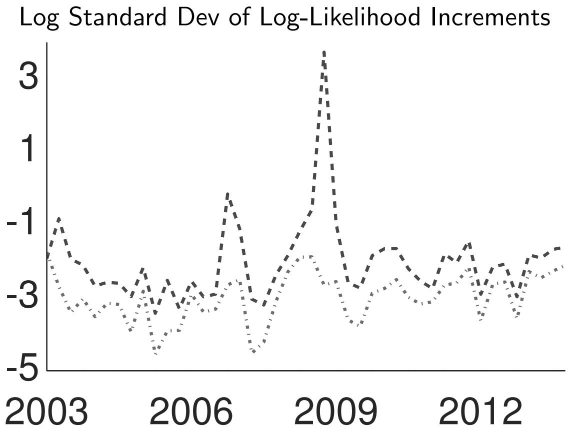
Example 1: Linearized Smets-Wouters Model

BS ($M = 400,000$) versus CO ($M = 4,000$)



Notes: Density estimates of $\hat{\Delta}_1 = \ln \hat{p}(Y|\theta) - \ln p(Y|\theta)$ based on $N_{run} = 100$. Solid densities summarize results for the bootstrap (BS) particle filter; dashed densities summarize results for the conditionally-optimal (CO) particle filter.

Example 2: Linearized Small-Scale NK DSGE Model



Notes: Solid lines represent results from Kalman filter. Dashed lines correspond to bootstrap particle filter ($M = 40,000$) and dotted lines correspond to conditionally-optimal particle filter ($M = 400$). Results are based on $N_{run} = 100$ runs of the filters.

[HS] A Generic Approach

- "Tempered Particle Filter"
  - Construct a sequence "bridge distributions" with inflated measurement errors.
  - Traverse these bridge distributions with "static" Sequential Monte Carlo method (Chopin, 2002).
- This PF has much better statistical properties than the naive bootstrap PF, *at little computational cost.*
- Unlike other versions of the PF, this algorithm is self-tuning and does not require the researcher to manually construct proposal densities.
- Some related concurrent work in statistics literature:
  Godsill and Clapp (2001), Johansen (2016)

[HS] The Key Idea

- Define

$$p_n(y_t|s_t,\theta) \propto \phi_n^{d/2}|\Sigma_u(\theta)|^{-1/2}\exp\left\{-\frac{1}{2}(y_t-\Psi(s_t,t;\theta))' \times\phi_n\Sigma_u^{-1}(\theta)(y_t-\Psi(s_t,t;\theta))\right\},$$

  where:

$$\phi_1 < \phi_2 < \ldots < \phi_{N_\phi} = 1.$$

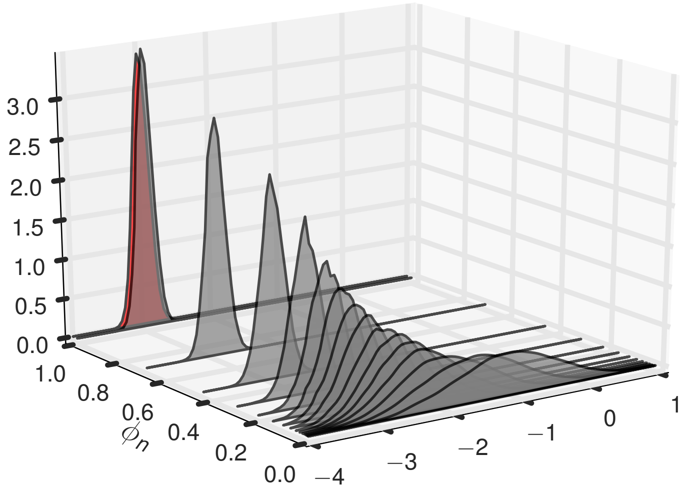
- Bridge posteriors given s_{t-1}:

$$p_n(s_t|y_t,s_{t-1},\theta) \propto p_n(y_t|s_t,\theta)p(s_t|s_{t-1},\theta).$$

- Bridge posteriors given $Y_{1:t}$:

$$p_n(s_t|Y_{1:t}) = \int p_n(s_t|y_t,s_{t-1},\theta)p(s_{t-1}|Y_{1:t-1})ds_{t-1}.$$

An Illustration [HS]: $p_n(s_t|Y_{1:t})$, $n = 1, \ldots, N_\phi$.



[HS] Algorithm Overview

- For each time period t, we embed a "static" SMC sampler used for parameter estimation [Chopin (2002), (...), Herbst and Schorfheide (2014, 2015), (...)]:

  Iterate over $n = 1, \ldots, N_\phi$:

  - **Goal:** approximate bridge distributions $p_n(y_t|Y_{1:t-1})$ and $p_n(s_t|Y_{1:t})$.
  - **Correction step**: change particle weights (importance sampling)
  - **Selection step**: equalize particle weights (resampling of particles)
  - **Mutation step**: change particle values (based on Markov transition kernel generated with Metropolis-Hastings algorithm)

[HS] Log-Likelihood Approx. Errors: Linearized Smets-Wouters Model

	BSPF	TPF	
Number of Particles M	40,000	2,000	2,800
Target Ineff. Ratio r^*		2	3
High Posterior Density: $\theta = \theta^m$			
MSE($\hat{\Delta}$)	63,882	1,164	1,135
$T^{-1}\sum_{t=1}^{T} N_{\phi,t}$	1	6	5
Average Run Time (sec)	3	3	3
Low Posterior Density: $\theta = \theta^l$			
MSE($\hat{\Delta}$)	69,613	1,490	1,994
$T^{-1}\sum_{t=1}^{T} N_{\phi,t}$	1	6	5
Average Run Time (sec)	3	3	3

[ACS] Conditionally-Optimal Filtering for Piecewise-Linear Approximations

- Not possible to directly sample from CO proposal in general nonlinear models.
- However, it can be done in piece-wise linear approximations.

[ACS] A Small-Scale Model

- Households:

$$\mathbb{E}_t \left[\sum_{s=0}^{\infty} \beta^s \left(\frac{C_{t+s}^{1-\tau} - 1}{1-\tau} - \chi_H \frac{H_{t+s}^{1+1/\eta}}{1+1/\eta} + \chi_M V\left(\frac{M_{t+s}}{P_{t+s} A_{t+s}} \right) \right) \right]$$

- Production of intermediate good j:

$$Y_t(j) = H_t(j), \quad AC_t(j) = \frac{\phi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - \bar{\pi} \right)^2 Y_t(j).$$

- Resource constraint: (g_t is a generic demand shock):

$$C_t + AC_t + G_t = Y_t, \quad G_t = \left(1 - \frac{1}{g_t}\right) Y_t, \quad \log g_t = (1-\rho_g) \log g^* + \rho_g \log g_{t-1} + \sigma_g \epsilon_{g,t}$$

- Monetary Policy:

$$R_t = \max \left\{ 1, \left[r\pi^* \left(\frac{\pi_t}{\pi^*} \right)^{\psi_1} \right]^{1-\rho_R} R_{t-1}^{\rho_R} e^{\epsilon_{R,t}} \right\}$$

[ACS] Approximate Solution

- State variables: $\mathbb{X} = (\epsilon_R, \hat{g}, \hat{R}_{-1})$.
- Policy functions:

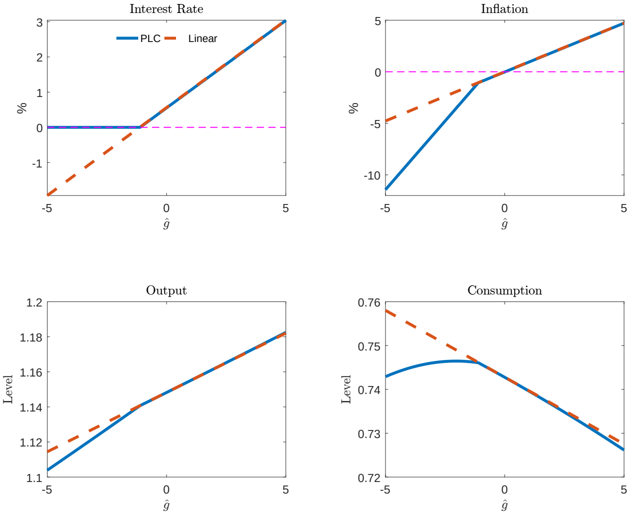
$$
\begin{aligned}
\hat{\pi} &= f_\pi(\mathbb{X}) = ? \\
\hat{y} &= f_y(\mathbb{X}) = ?
\end{aligned}
$$

- Equilibrium conditions:

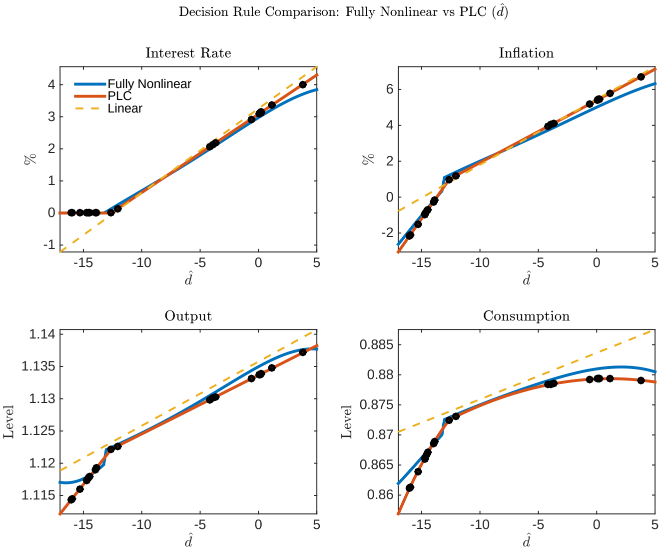
$$
\begin{aligned}
F(\mathbb{X}) &= 0 \\
h(\mathbb{X}) &\geq 0
\end{aligned}
$$

- Construct approximate solution by making policy functions piecewise linear and continuous (PLC).

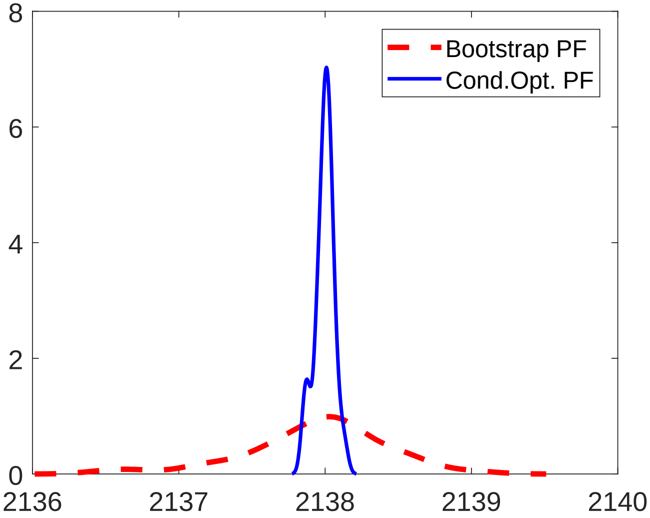
[ACS] Illustration (Baseline Model): PLC vs. Linear



[ACS] Illustration (Richer Model): PLC vs. Linear vs. Nonlinear



[ACS] Filtering



Conclusion

- Structural macroeconometrics faces many computational challenges:
  - model solution,
  - likelihood computation
  - Posterior sampling or maximization of extremum estimator objective function.
- Potential shortcuts to keep computations fast and feasible:
  - less accurate model solution
  - cruder state extraction / likelihood approximation
  - non-likelihood-based parametrization of the model.
- In this talk: Slightly less accurate solution enables efficient evaluation of likelihood function.