EABCN/CEPR/SNB Workshop

"Estimation and Empirical Validation of Structural Models for Business Cycle Analysis", Zurich, 29-30 August, 2006

discussion of the paper

"Estimating Macroeconomic models: a likelihood approach"

(by J. Fernández-Villaverde and J. F. Rubio-Ramírez)

Gianni Amisano (University of Brescia)

1 Main aim of the paper

- show how likelihood based methods can be used to estimate structural nonlinear dynamic macromodel for US data;

- show that the stochastic volatility features of the model are important in the explation of the behaviour of observed data

- show how to obtain smoothed estimates of the latent state variables

- run conterfactual experiments to ascertain relevance of exogenous changes in volatilities in determining changes of volatility of observed variables

2 Related literature

- literature on "great moderation"/time varying nature volatility of shocks driving the US economy (Justiniano and Primiceri, 3005)

- methodological literature on nonlinear state space models and in particular

- simulation based filtering literature

- macro papers in which these techniques are applied: contributions by F-VR-R (2005), Schorfheide and An (2005), An (2006), Amisano and Tristani (2006a, 2006b)

3 A non linear SSM

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{G}(\mathbf{s}_t, \mathbf{v}_t; \gamma) && (1)\\
\mathbf{s}_t &= \mathbf{H}(\mathbf{s}_{t-1}, \mathbf{w}_t; \gamma) && (2)\\
\text{dim}(\mathbf{v}_t) + \text{dim}(\mathbf{w}_t) &\geq \text{dim}(\mathbf{y}_t) && (3)\\
& \quad \text{define } \mathbf{w}_{2t} \text{ such that} \\
\text{dim}(\mathbf{v}_t) + \text{dim}(\mathbf{w}_{2t}) &= \text{dim}(\mathbf{y}_t) && (4)
\end{aligned}
$$

If linear+Gaussian: Kalman Filter

otherwise , a simulation based approach could be used for obtain likelihood of the model

(Doucet et al, 2001, Arulampalam et al 2002)

4 Particle filter (PF)

Main idea: given draw $\mathbf{s}_t^{(i)}, i = 1, 2, .., N$ from $p(\mathbf{s}_t|\underline{\mathbf{y}}_t; \boldsymbol{\gamma})$ perform projection and update steps via simulation (importance sampling):

- (projection): draw $\mathbf{s}_{t+1}^{(i)}$ from $p(\mathbf{s}_{t+1}|\mathbf{s}_t^{(i)}\underline{\mathbf{y}}_t; \boldsymbol{\gamma})$ ie draw $\mathbf{w}_t^{(i)}$(or $\mathbf{w}_{1t}^{(i)}$)$, i = 1, 2, ..., N$

- update: assign to each particle a weight (resample with probability) proportional to

$$w_{t+1}^{(i)} = p(\mathbf{y}_{t+1}|\mathbf{s}_t^{(i)}; \boldsymbol{\gamma})$$

- sample mean of weights is consistent (in N) estimate of $p(\mathbf{y}_{t+1}|\underline{\mathbf{y}_t};\boldsymbol{\gamma})$

This allow to obtain likelihood function, to be used for ML or for Bayesian inference.

5 Interpretation of the PF

Intuition: for Bayesians:

- $p(\mathbf{s}_{t+1}|\underline{\mathbf{y}}_t;\gamma)$ = "prior distribution" (prior to observing $\mathbf{y}_{t+1}$)

- $p(\mathbf{y}_{t+1}|\mathbf{s}_{t+1};\gamma)$ = "likelihood",

$\Rightarrow$ doing posterior simulation drawing from "prior" and using "likelihood" as weights

- $NEFF_t = \sum_{i=1}^{N}\left(w_t^{(i)}\right)^2$ ≈Herfindhal-Hirschmann index. Has to be safely far from 1 and as close as possible to 1/N

5.1 Problems with the PF

5.1.1 Adaption

Importance sampling (IS) interpretation .

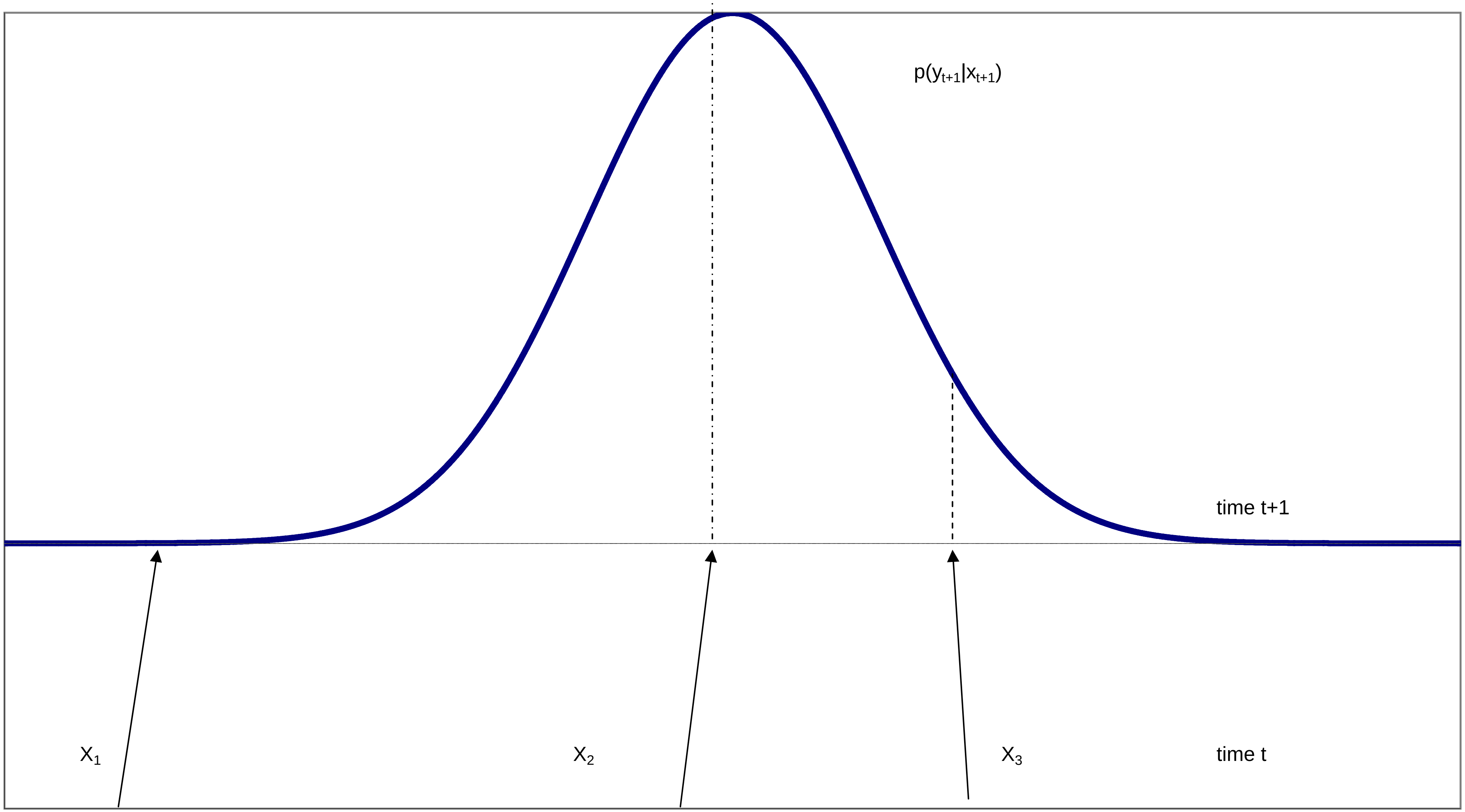
Importance function q should be more spread out than candidate distribution p: bounded weights $w_t^{(i)} = \frac{p_i}{q_i}$.

But if IS distribution too spread out, large number of draws given negligible weights $\Rightarrow$poor numerical accuracy properties.

$\Rightarrow$PF is based on a blind proposal

Figure (A): first particle will be killed either by reweighting or by resampling.

figure A: PF at work, N=3

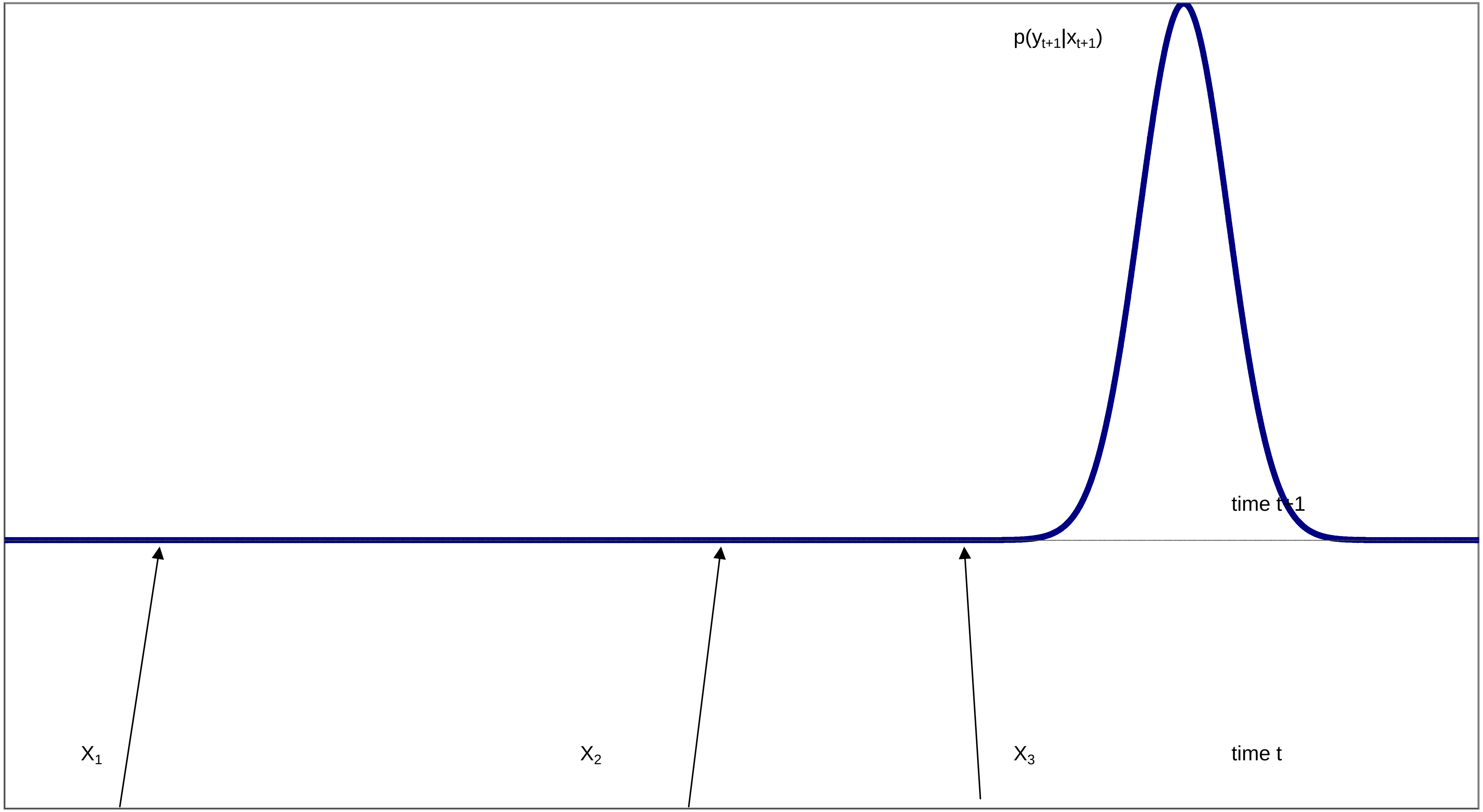


5.1.2 Sensitivity to outliers

see Figure (B)

particle 3 will get a unit weight. All the others killed by reweighting or by resampling

Figure B: PF at work, N=3



5.1.3 Rank deficiency in measurement error cov matrix

$p(\mathbf{y}_{t+1}|\mathbf{s}_{t+1}^{(j)};\gamma)$ to compute weights, but if no measurement error this becomes degenerate.

In F-V and R-R (2006) very clever modification of PF that works also in models with rank deficiencies in measurement errors covariance matrix but quite involved in general case

(in their model very easy: one observable is a linear function of a state variable).

5.2 Other filters

Many viable alternatives to PF

Adaption: use knowledge of $\underline{\mathbf{y}}_{t+1}$

- the conditional particle filter (CPF) of Ionides (2003);$\Rightarrow$used in Amisano and Tristani (2006a).

- the auxiliary variable particle filter (AVPF) (Pitt and Shephard, 1999).

5.2.1 The Conditional Particle Filter

Instead of drawing $\mathbf{s}_{t+1}^{(j)}$from $p(\mathbf{s}_{t+1}|\mathbf{s}_t^{(j)};\boldsymbol{\gamma})$ and resample (reweight) using $p(\mathbf{y}_{t+1}|\mathbf{s}_{t+1}^{(i)};\boldsymbol{\gamma})$ (PF)

- sample $\mathbf{s}_{t+1}^{(j)}$ from $p(\mathbf{s}_{t+1}|\mathbf{s}_t^{(j)},\mathbf{y}_{t+1};\boldsymbol{\gamma})$ (i.e. we condition explicitly on $\mathbf{y}_{t+1}^o$)

- resample using weights $p(\mathbf{y}_{t+1}|\mathbf{s}_t^{(i)};\boldsymbol{\gamma})$.

Works perfectly even in the absence of measurement error

$p(\mathbf{s}_{t+1}|\mathbf{s}_t^{(j)},\mathbf{y}_{t+1}^o;\boldsymbol{\gamma})$ and $p(\mathbf{y}_{t+1}|\mathbf{s}_t^{(i)};\boldsymbol{\gamma})$ are not known in closed form. $\Rightarrow$Use linearisation of measurement equation around $E(\mathbf{s}_t|\underline{\mathbf{y}}_t;\boldsymbol{\gamma})$.

Very easy to implement and works well

Visual intuition : Figure (C)

Figure D: CPF at work, N=3

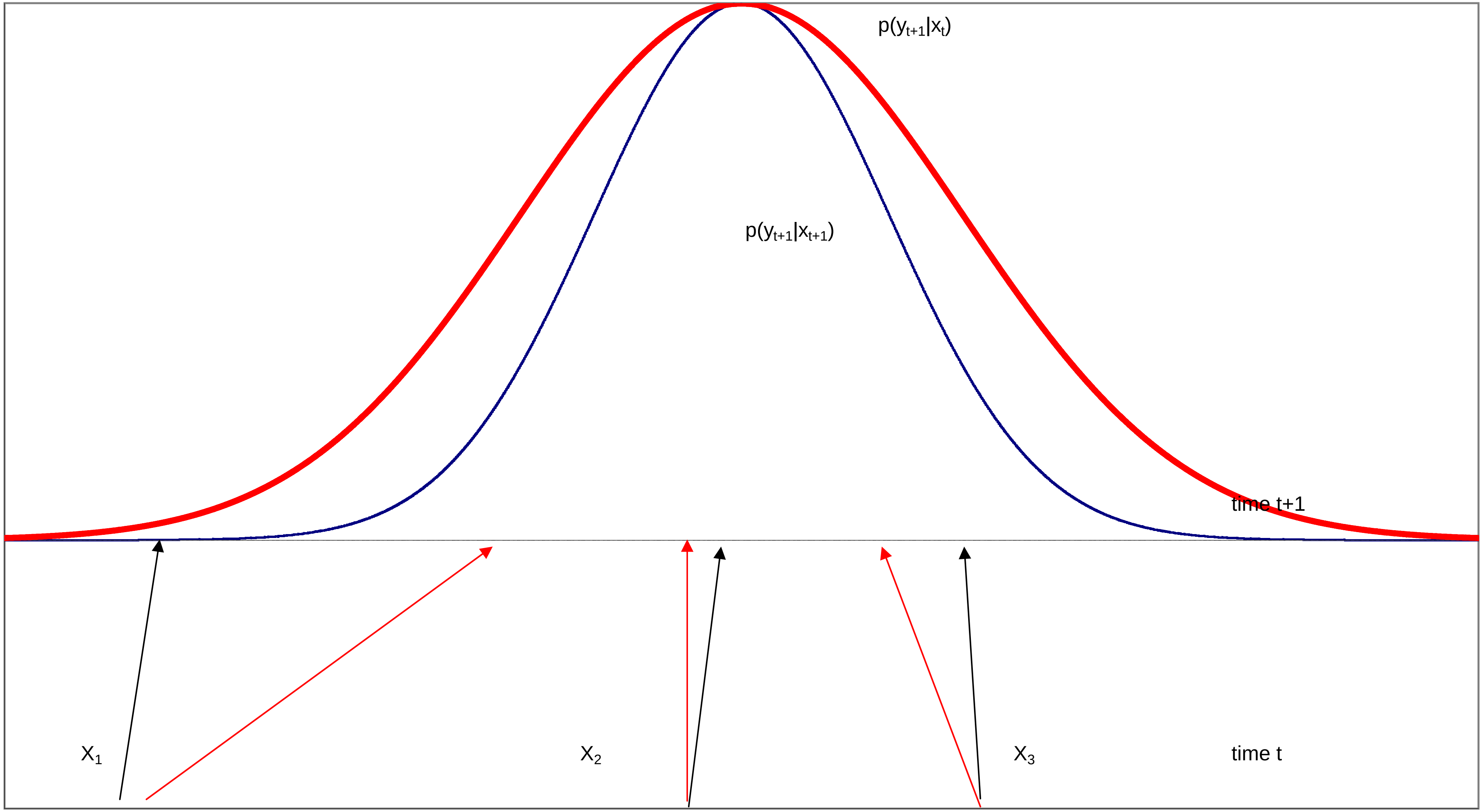
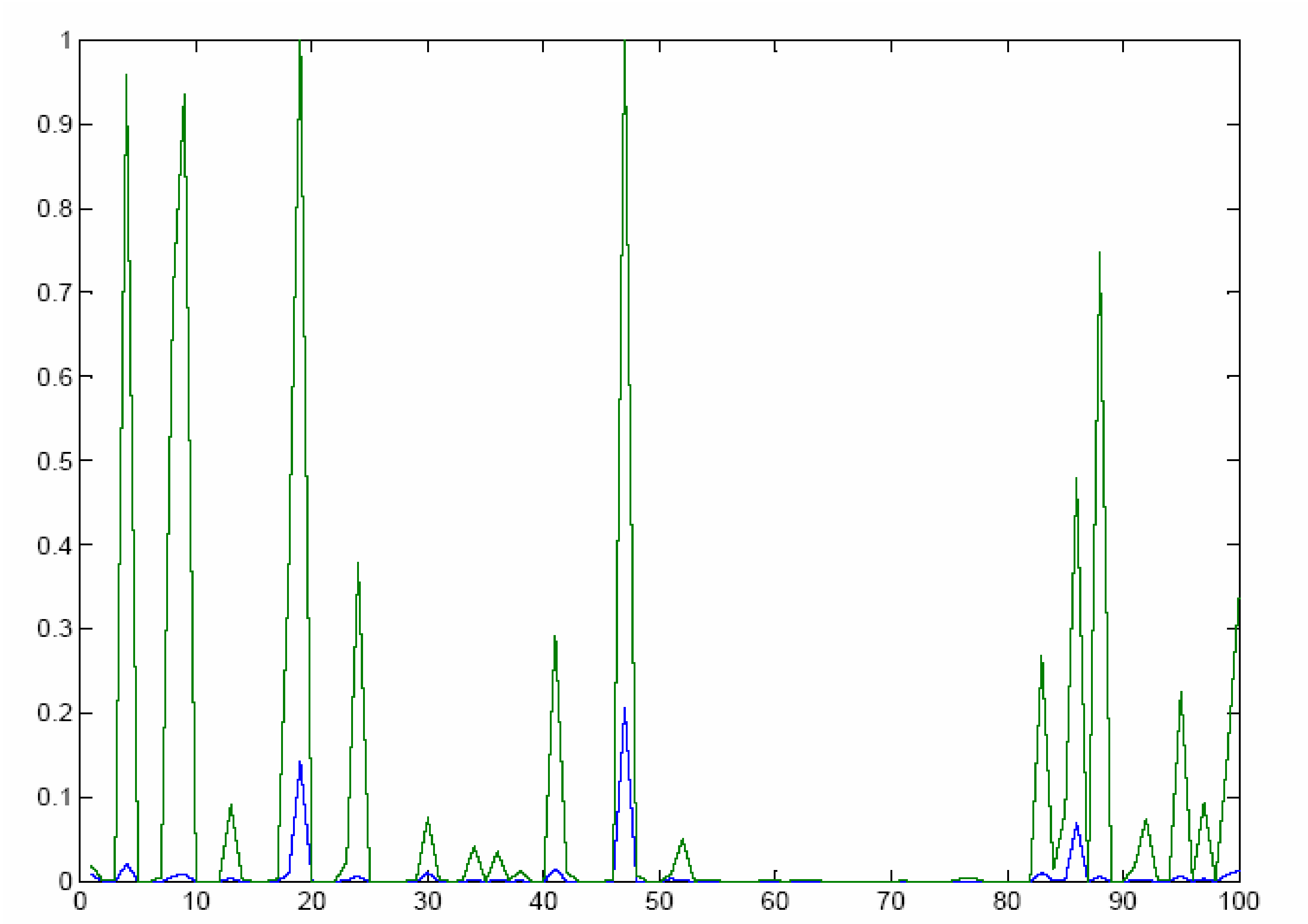


Figure C: NEFF for PF (green line) and CPF (blue line)



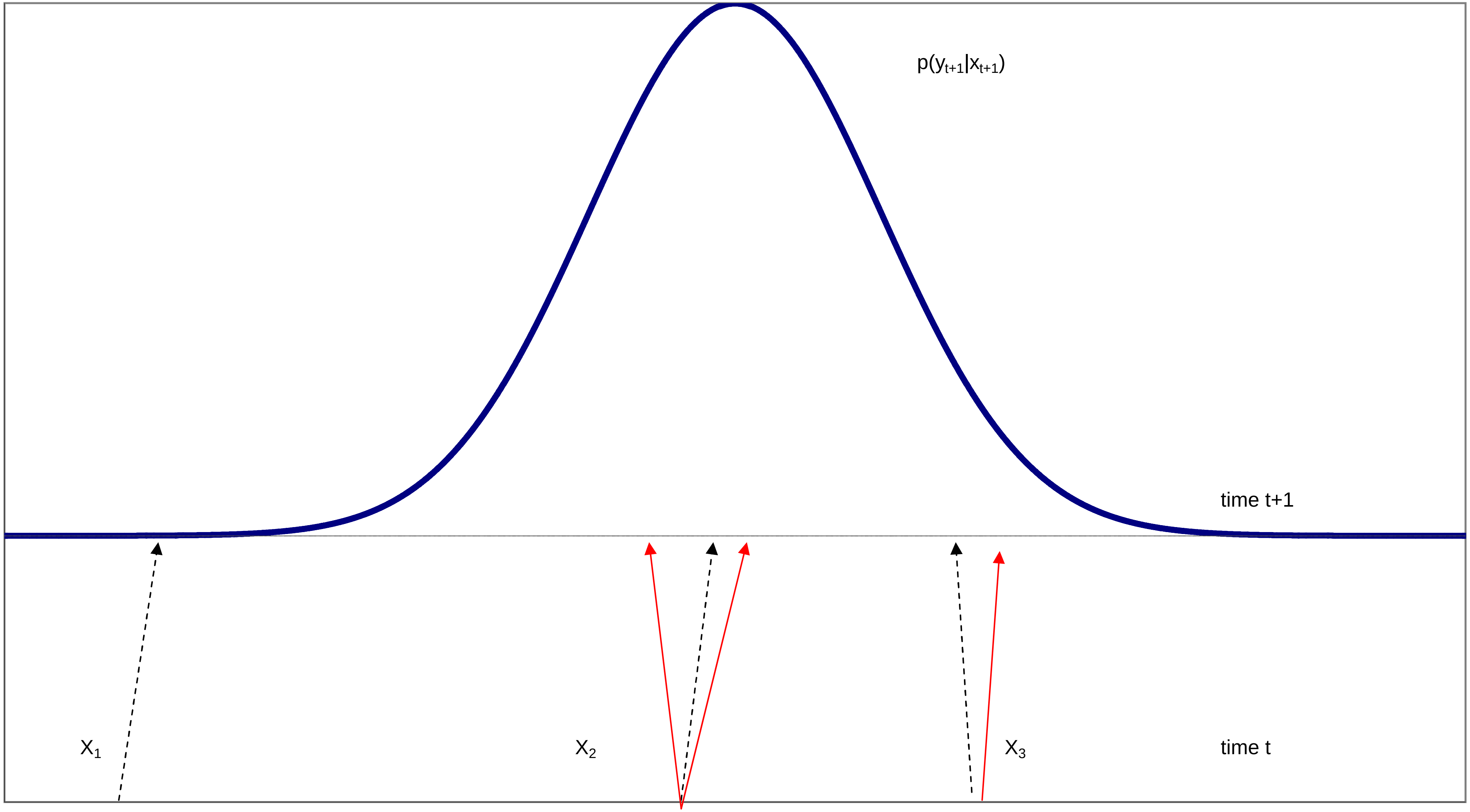
5.2.2 The Auxiliary Variable Particle Filter (Pitt and Shephard (1999).

Intuition: first select fittest particles (by reweighting) and allow only them to "breed "

Reweight them accordingly. Details in Amisano and Tristani (2006b). Figure (E)

Two resamplig steps at each t, so much slower than CPF.

Figure E: AVPF at work, N=3



6 The results in the paper

- Tight parameter estimates for (most of) the parameters: unconditional mean of log volatilities and std errors of log volatilities shocks less precisely estimated

- Very good fit

- Empirical relevance of SV and in particular

- changing volatility of preferences shocks especially relevant to explain changes in volatility in observable variables (counterfactual experiments)

- Using quadratic SSM is preferable to linear approx.

7 Some remarks

1. p.5: "Thus the KF not only induces an approximation error, but more important, it makes it impossibile to learn about time-varying volatility.": a bit excessive

2. Comparison with most interesting alternatives (linear+SV): can we use forecasting performance measures?

3. Role of other filters?

4. To assess goodness of fit: wouldn't it be better to use filtered state variables and one step ahead errors on observables?

5. Significance of results: all based on MLE of parameters. How would results change accounting for parameter uncertainty?

6. Evolution of vol: increase of $\sigma_v$during 70-75, decrease of σ_d in early 60s and decrease from late 1980s (figure 6.3)

7. Results concerning d_t: interpretation? Fig.6.3 and sample mean (very far from 0).