Inference for non-linear, non-Gaussian state-space models

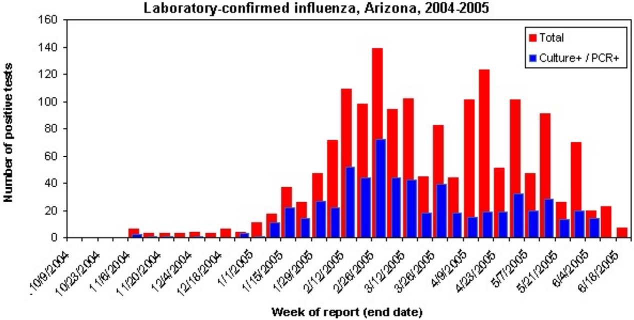
Universidad Complutense de Madrid
Seminarios de Investigación del Departamento de Economía Cuantitativa

November 19, 2014

Carles Bretó
Universidad Carlos III de Madrid (UC3M), Dep. de Estadística

Example of dynamic systems: financial crisis

Another dynamic system: Infectious Disease Epidemics



Wiley Online Library Biblioteca - Universidad Carlo

Home > Statistics > Probability & Mathematical Statistics > Wiley Series in Probability and Statistics > Statistical Advances ... Trials, Epidemiology, Survival Analysis, and Bioinformatics > Summary

BOOK TOOLS
- Save to My Profile
- Purchase a print copy
- Recommend to Your Librarian

BOOK MENU
- Book Home

GET ACCESS
- How to Get Online Access

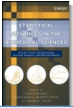
8. Modeling Disease Dynamics: Cholera as a Case Study

Atanu Biswas[4], Sujay Datta[5], Jason P. Fine[6] andMark R. Segal[7]

Edward L. Ionides[1], Carles Bretó[2] and Aaron A. King[3]

Published Online: 23 MAR 2007
DOI: 10.1002/9780470181218.ch8

Book Title



Statistical Advances in the Biomedical Sciences: Clinical Trials, Epidemiology, Survival Analysis, and Bioinformatics

ELSEVIER journal homepage: www.elsevier.com/locate/stapro

On idiosyncratic stochasticity of financial leverage effects

CrossMark

Carles Bretó *
Departamento de Estadística and Instituto Flores de Lemus, Universidad Carlos III de Madrid, C/ Madrid 126, Getafe, 28903, Madrid, Spain

ARTICLE INFO

Article history:
Received 12 December 2013
Received in revised form 28 March 2014
Accepted 1 April 2014
Available online 8 April 2014

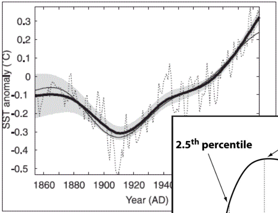
Keywords:
Stochastic leverage
Random-walk time-varying parameter
Non-linear non-Gaussian state-space model
Maximum likelihood estimation
Particle filter

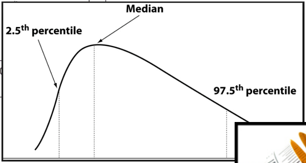
ABSTRACT

We model leverage as stochastic but independent of return shocks and of volatility and perform likelihood-based inference via the recently developed iterated filtering algorithm using S&P500 data, contributing new evidence to the still slim empirical support for random leverage variation.

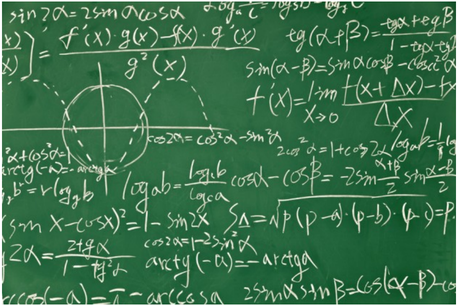
1. Introduction

Problem: realistic models are likely non-linear, non-Gaussian & partially observed

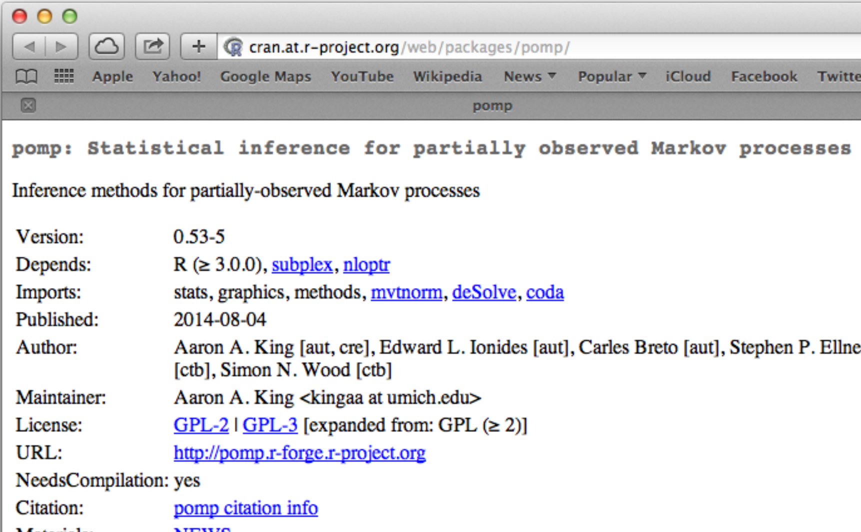




Hence: their statistical analysis is complicated & slows down scientific exploration

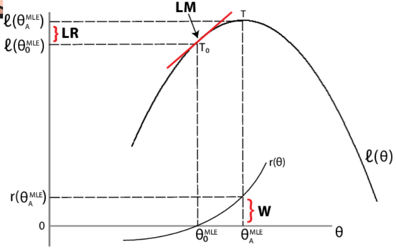


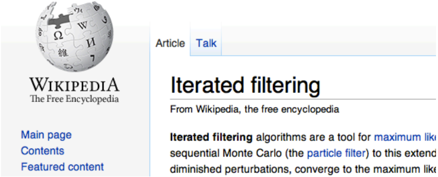
Contribution (I): R package POMP (Frequentist particle filter, PMCMC, ABC, etc)



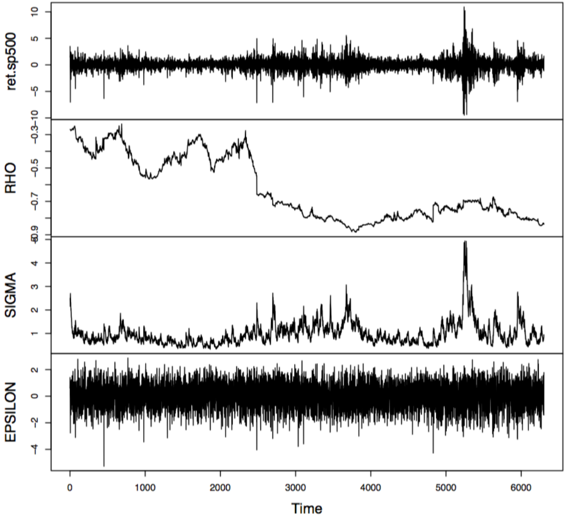
Contribution (II): iterated filtering algorithm

Plug-and-play likelihood-based inference on POMPs



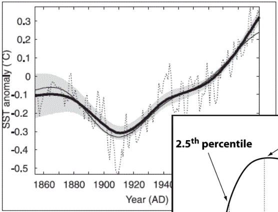


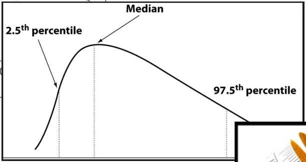
Contribution (III): Example from financial econometrics
Stochastic volatility with stochastic leverage



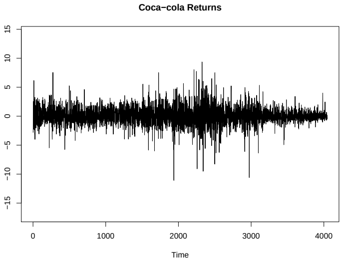
Take-home message: straightforward, likelihood-based inference is possible for general dynamic systems

Problem: realistic models are likely non-linear, non-Gaussian & partially observed



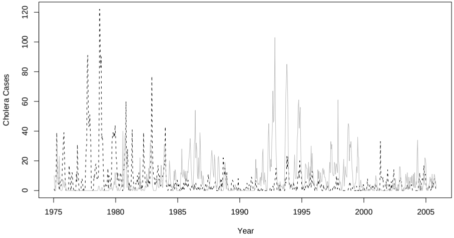


State-space: unobservable variables/mechanisms



- Stochastic volatility:

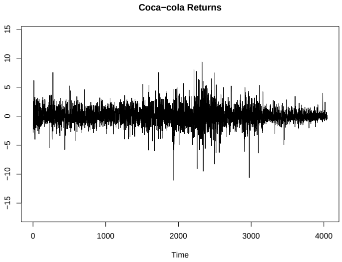
$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\sigma_t^2 &= \exp(h_t) \\
h_t &= \mu(1-\phi) + \phi h_{t-1} + \eta_t^*
\end{aligned}
$$



- SIR-type compartment models:

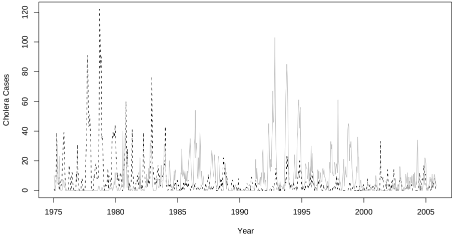
$$
\begin{aligned}
C_t &= \rho I_t + \sigma \epsilon_t \\
I_t &= \exp(\tilde{I}_t) = (\beta S_{t-1}) I_{t-1}^{\alpha} \eta_t^* \\
\tilde{I}_t &= (\tilde{\beta} + \tilde{S}_{t-1}) + \alpha \tilde{I}_{t-1} + \tilde{\eta}_t^*
\end{aligned}
$$

State-space: unobservable variables/mechanisms



- Stochastic volatility:

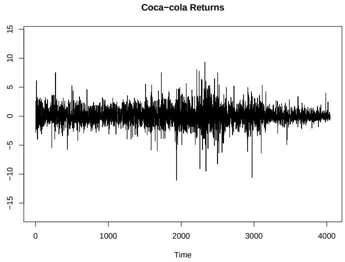
$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\sigma_t^2 &= \exp\left(h_t\right) \\
h_t &= \mu(1-\phi) + \phi h_{t-1} + \eta_t^*
\end{aligned}
$$



- SIR-type compartment models:

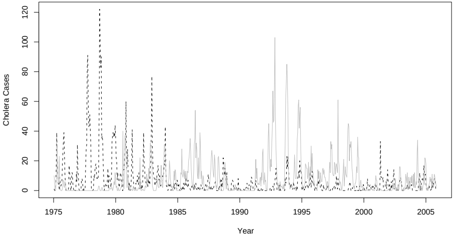
$$
\begin{aligned}
C_t &= \rho I_t + \sigma \epsilon_t \\
I_t &= \exp(\tilde{I}_t) = (\beta S_{t-1}) I_{t-1}^{\alpha} \eta_t^* \\
\tilde{I}_t &= (\tilde{\beta} + \tilde{S}_{t-1}) + \alpha \tilde{I}_{t-1} + \tilde{\eta}_t^*
\end{aligned}
$$

State-space: unobservable variables/mechanisms



- Stochastic volatility:

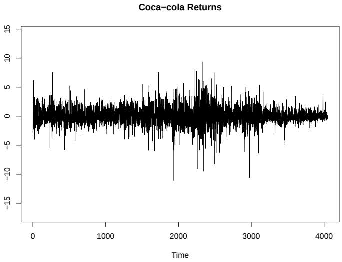
$$
\begin{aligned}
y_t &= \sigma_t \epsilon_t \\
\sigma_t^2 &= \exp(h_t) \\
h_t &= \mu(1-\phi) + \phi h_{t-1} + \eta_t^*
\end{aligned}
$$



- SIR-type compartment models:

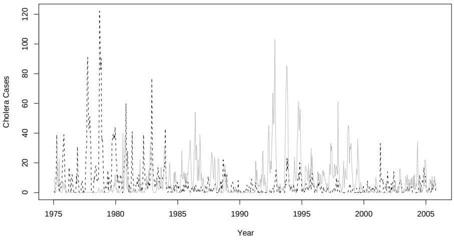
$$
\begin{aligned}
C_t &= \rho I_t + \sigma \epsilon_t \\
I_t &= \exp(\tilde{I}_t) = (\beta S_{t-1}) I_{t-1}^{\alpha} \eta_t^* \\
\tilde{I}_t &= (\tilde{\beta} + \tilde{S}_{t-1}) + \alpha \tilde{I}_{t-1} + \tilde{\eta}_t^*
\end{aligned}
$$

Linearity and Gaussianity: unusual but convenient



- Stochastic volatility:

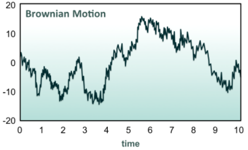
$$\begin{aligned} \log(y_t) &= \log(\sigma_t) + \log(\epsilon_t) \\ \sigma_t^2 &= \exp(h_t) \\ h_t &= \mu(1-\phi) + \phi h_{t-1} + \eta_t^* \end{aligned}$$

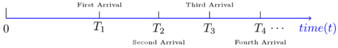


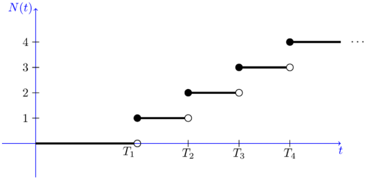
- SIR-type compartment models:

$$\begin{aligned} C_t &\sim \text{Pois}(\rho I_t) \\ I_t &\sim \text{bin}\left(S_{t-1}, e^{-\beta I_{t-1}^{\alpha}}\right) - \\ &- \text{bin}\left(I_{t-1}, e^{-\gamma}\right) \end{aligned}$$

Linearity and Gaussianity: unusual but convenient







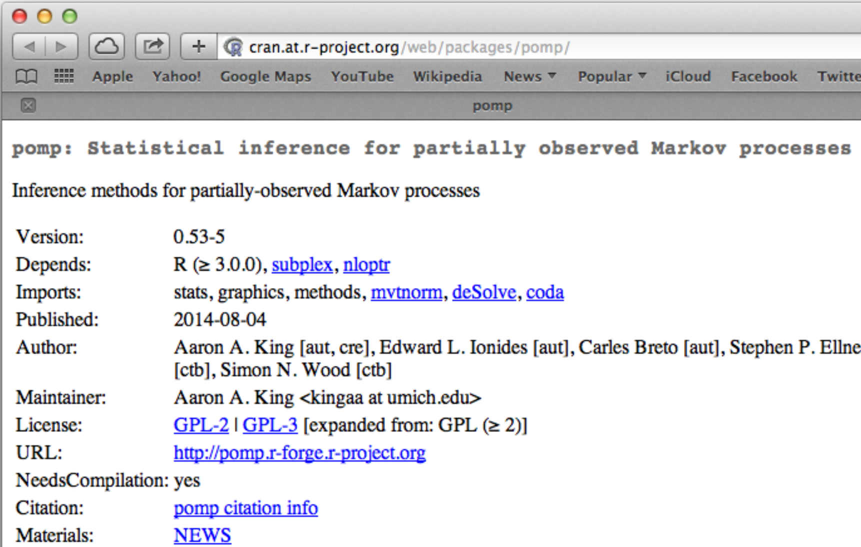
- Stochastic volatility:

$$
\begin{aligned}
dy^*(t) &= \mu + \beta\sigma^2(t)dt + \sigma(t)dB(t) \\
d\sigma^2(t) &= -\lambda\sigma^2(t)dt + dz(\lambda t)
\end{aligned}
$$

- SIR-type compartment models:

$$
\begin{aligned}
P(S \to S) &= 1 - \beta s i^\alpha h + o(h) \\
P(S \to I) &= \beta s i^\alpha h + o(h)
\end{aligned}
$$

Contribution (I): R package POMP (Frequentist particle filter, PMCMC, ABC, etc)



POMP object: notation

- Formal:
  - Markov unobservables: $\left(X_1(t), \ldots, X_{K_X}(t)\right)$
  - Unobservable time: either continuous $t \in \mathbb{R}_0^+$ or discrete $t \in \mathbb{N}_0$
  - Conditionally independent measurements: $\left(Y_1(t_n), \ldots, Y_{K_Y}(t_n)\right)$
  - Measurement time: discrete $t_1, \ldots, t_N$
  - Parameters θ
- Algorithmic (POMP code):
  - rprocess: a draw from $f_{\boldsymbol{X}(t_n)|\boldsymbol{X}(t_{n-1})}(\boldsymbol{x}(t_n)|\boldsymbol{x}(t_{n-1}), \boldsymbol{\theta})$
  - dprocess: evaluate $f_{\boldsymbol{X}(t_n)|\boldsymbol{X}(t_{n-1})}(\boldsymbol{x}(t_n)|\boldsymbol{x}(t_{n-1}), \boldsymbol{\theta})$
  - rmeasure: a draw from $f_{\boldsymbol{Y}(t_n)|\boldsymbol{X}(t_n)}(\boldsymbol{y}(t_n)|\boldsymbol{x}(t_n), \boldsymbol{\theta})$
  - dmeasure: evaluate $f_{\boldsymbol{Y}(t_n)|\boldsymbol{X}(t_n)}(\boldsymbol{y}(t_n)|\boldsymbol{x}(t_n), \boldsymbol{\theta})$
- Nuisances: Initial value parameters

POMP inference: difficult parameter estimation

- Alternative model-based inference approaches:
  - MoM (need to check moments, not full information)
  - MQLE (need to check Gaussian approx.)
  - Bayesian MCMC (Jaquier et al., 1994) (need to check priors)
  - EMM (Gallant & Tauchen, 1996) (need to check auxiliary model)
  - MCL (Sandmann and Koopman, 1998) (need to check approx.)
  - EIS (Lesenfeld and Richard, 2003) (need to check Imp. Sampler)
- Reasonable estimates on average across different samples (error compensation)
- We have only one (long) sample: importance of efficiency
- What about “plug-and-play” modelling?

POMP: Plug-and-play inference

- Plug-and-play algorithm:
  - rprocess but not dprocess
  - code simulating sample paths is “plugged” into inference software
- Not plug-and-play:
  - EM algorithm (dprocess)
  - MCMC (dprocess+dmeasure)
- Bayesian plug-and-play:
  - Artificial parameter evolution (Liu & West, 2001: posterior correction, rprocess+dmeasure)
  - ABC (Beaumont et al., 2002: sufficient statistics, rprocess+rmeasure)
  - PMCMC (Andrieu et al., 2010: SMC + MCMC, rprocess+dmeasure)
- Non-Bayesian plug-and-play:
  - Iterated filtering (Ionides et al., 2006: likelihood-based inference, rprocess+dmeasure)

Other settings: plug-and-play

- Optimization: Methods requiring only evaluation of the objective function to be optimized are sometimes called gradient-free. This is the same concept as plug-and-play: the code to evaluate the objective function can be plugged into the optimizer

- Complex systems: Methods to study the behavior of large numerical simulations (e.g., molecular models for phase transitions) that only employ the underlying code as a “black box” to generate simulations are called equation-free (Kevrekidis et al., 2003, 2004)

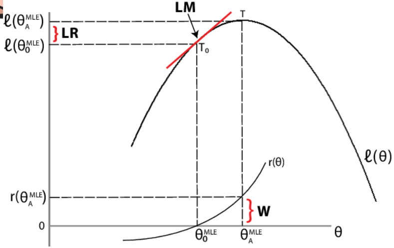
- ABC and PMCMC: Plug-and-play methods have recently been called likelihood-free. In this terminology, iterated filtering does likelihood-free likelihood-based inference

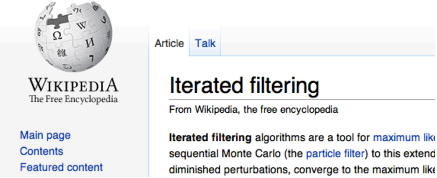
Cost: plug-and-play

- Efficiency: Approximate Bayesian methods and simulated moment methods lead to a loss of statistical efficiency
- Iterated filtering: enables (almost) exact likelihood-based inference
- Improvements: numerical efficiency may be possible when analytic properties are available (at the expense of plug-and-play). But many interesting dynamic models are analytically intractable—for example, it is standard to investigate systems of ordinary differential equations numerically

Contribution (II): iterated filtering algorithm

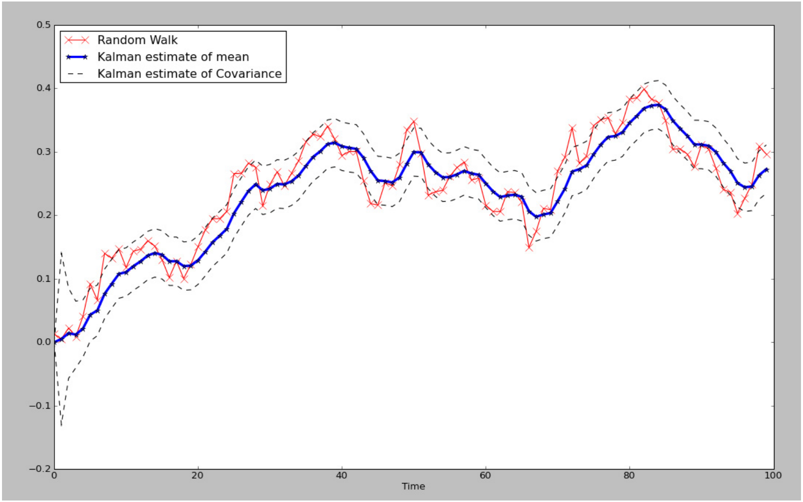
Plug-and-play likelihood-based inference on POMPs



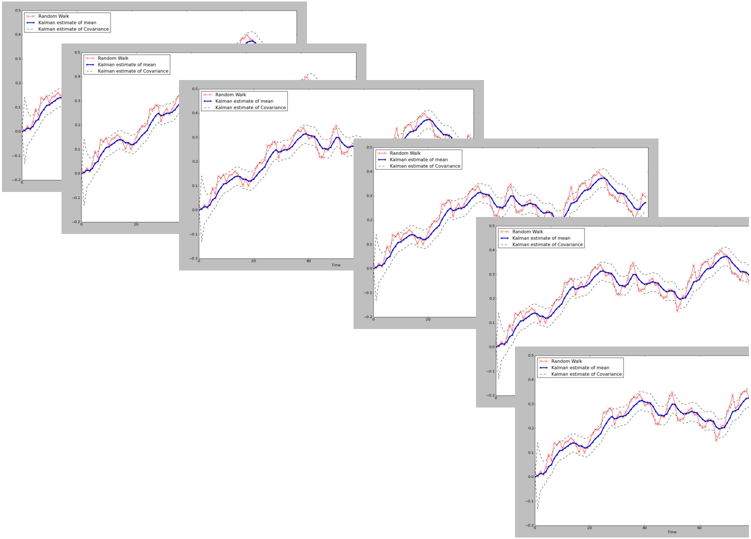


Iterated filtering. *Filtering Probl.*: extensively-studied

Cond. distr. of state $x(t_n)$ given obs. $y(t_1), \ldots, y(t_n)$

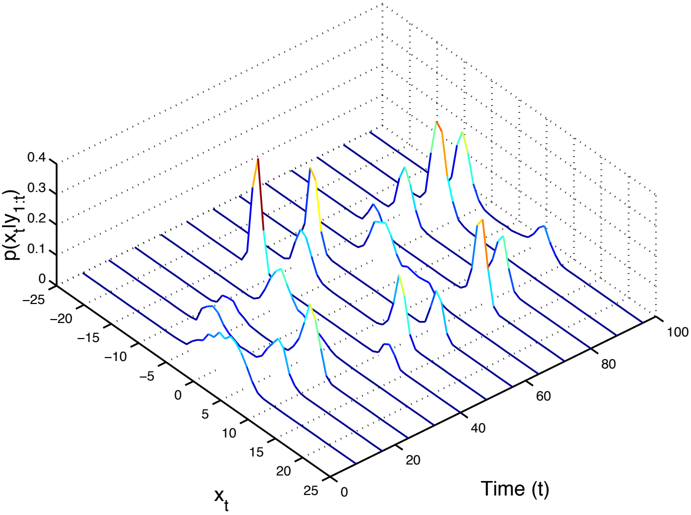


Iterated filtering: sequence of solutions to filtering to maximize likelihood over unknown parameters



Iterated filtering. Sequential Monte Carlo

Provides plug-and-play filter (for P&P IF and PMCMC)



Likelihood-based inference via iterated filtering

- MLE: asymptotically smallest estimator variance across different samples
- The properties of likelihood-based inference have been extensively studied:
  1. Invariant estimators
  2. Nested and non-nested hypothesis testing (via LR and AIC) (meaningful differences in the criterion)
  3. Computationally cheap standard errors (via FI)
  4. Likelihood profiles: robustness to identifiability issues

Toy example: AR(1) with noisy measurement

$$\begin{aligned} X_t &= \phi X_{t-1} + \epsilon_{xt} \\ Y_t &= X_t + \epsilon_{yt} \end{aligned}$$

$$\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{pmatrix} = N\left(0, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}\right)$$

Toy example: AR(1) with noisy measurement

$$X_t = \phi X_{t-1} + \epsilon_{xt}$$
$$Y_t = X_t + \epsilon_{yt}$$

$$\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{pmatrix} = N\left(0, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \right)$$

Toy example: AR(1) with noisy measurement

$$\begin{aligned} X_t &= \phi X_{t-1} + \epsilon_{xt} \\ Y_t &= X_t + \epsilon_{yt} \end{aligned}$$

$$\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{pmatrix} = N \left(0, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \right)$$

Extend the model: Random-walk time-varying parameters

$$
\begin{aligned}
X_t &= \text{logit}(\phi_t) X_{t-1} + \epsilon_{xt} \\
Y_t &= X_t + \epsilon_{yt} \\
\phi_t &= \phi_{t-1} + \epsilon_{\phi t}
\end{aligned}
$$

$$
\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \\ \epsilon_{\phi t} \end{pmatrix} = \mathrm{N} \left(0, \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\phi^2 \end{pmatrix} \right)
$$

Take a limit as $\sigma_\phi \downarrow 0$

Extend the model: Random-walk time-varying parameters

$$
\begin{aligned}
X_t &= \text{logit}(\phi_t) X_{t-1} + \epsilon_{xt} \\
Y_t &= X_t + \epsilon_{yt} \\
\phi_t &= \phi_{t-1} + \epsilon_{\phi t}
\end{aligned}
$$

$$
\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \\ \epsilon_{\phi t} \end{pmatrix} = N\left(0, \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\phi^2 \end{pmatrix} \right)
$$

Take a limit as $\sigma_\phi \downarrow 0$

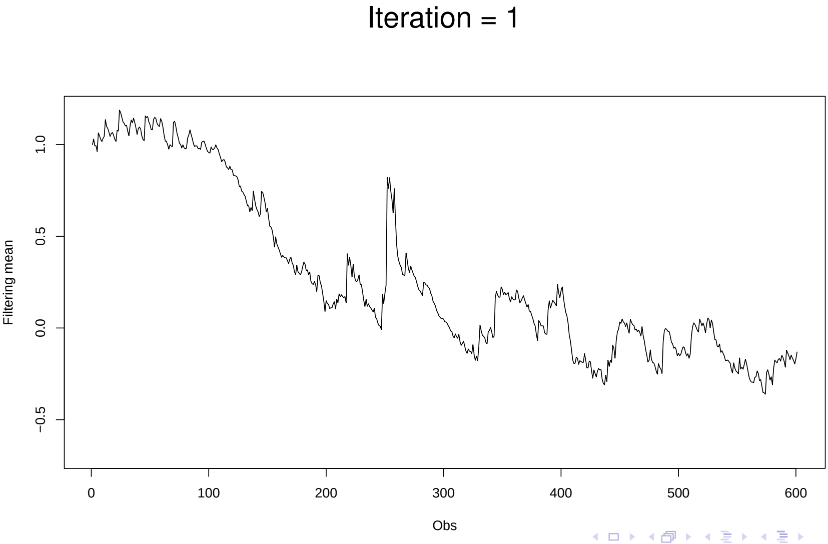
Extend the model: Random-walk time-varying parameters

$$
\begin{aligned}
X_t &= \text{logit}(\phi_t) X_{t-1} + \epsilon_{xt} \\
Y_t &= X_t + \epsilon_{yt} \\
\phi_t &= \phi_{t-1} + \epsilon_{\phi t}
\end{aligned}
$$

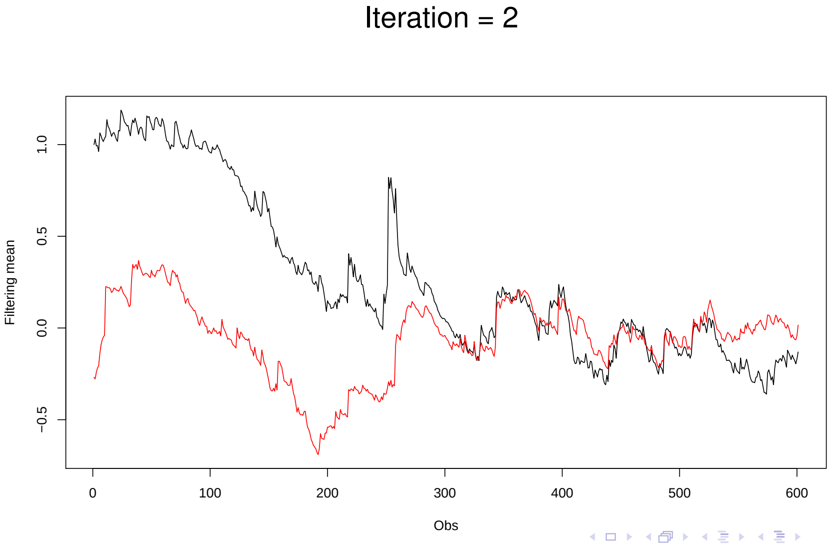
$$
\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \\ \epsilon_{\phi t} \end{pmatrix} = \text{N} \left(0, \begin{pmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_\phi^2 \end{pmatrix} \right)
$$

Take a limit as $\sigma_\phi \downarrow 0$

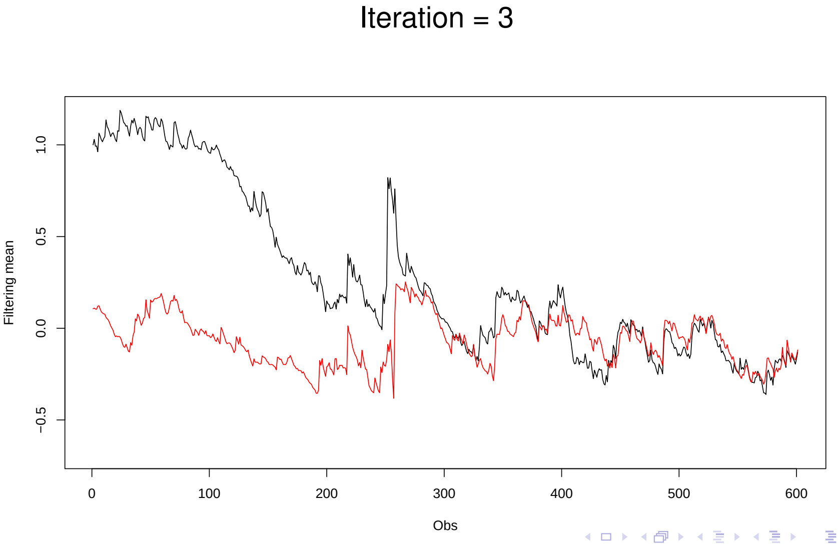
$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ
$V[\phi_t|y_{1:t}]$: weighted by local uncertainty



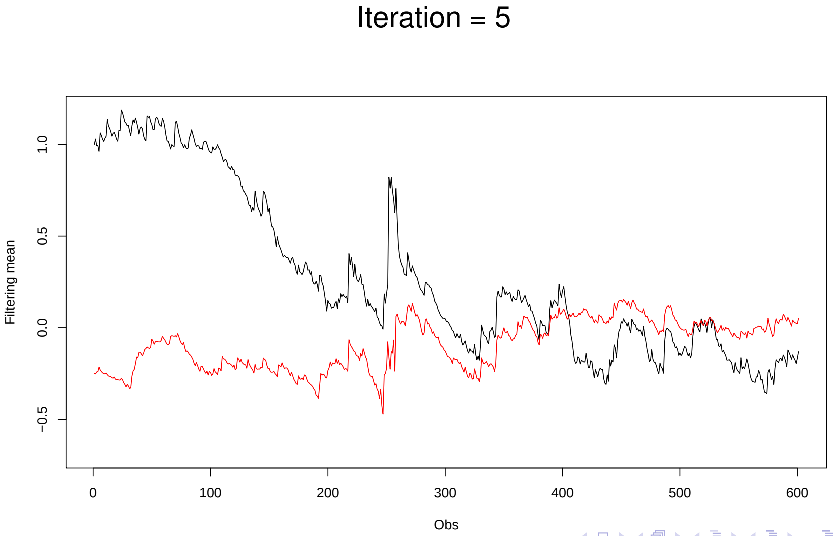
$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ $V[\phi_t|y_{1:t}]$: weighted by local uncertainty



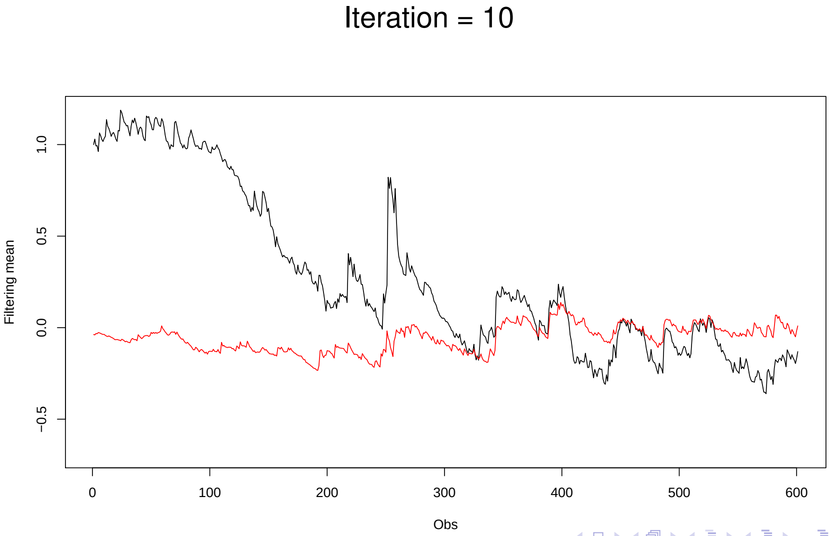
$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ
$V[\phi_t|y_{1:t}]$: weighted by local uncertainty



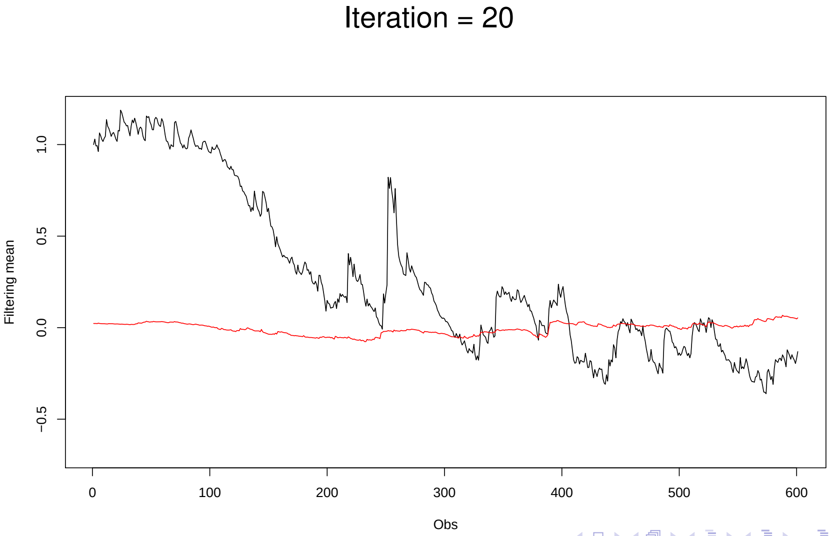
$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ $V[\phi_t|y_{1:t}]$: weighted by local uncertainty



$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ $V[\phi_t|y_{1:t}]$: weighted by local uncertainty



$E[\phi_t|y_{1:t}]$: Local estimates of fixed AR parameter ϕ
$V[\phi_t|y_{1:t}]$: weighted by local uncertainty



Easy to implement: Use POMP package!

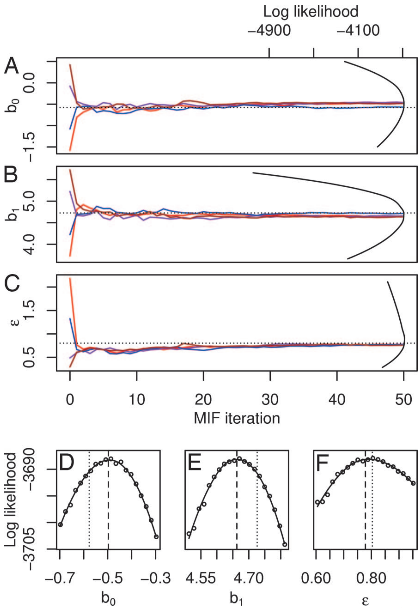
- SMC Filtering input:
  - $f(x_t|x_{t-1})$ or *rprocess*: rnorm(ϕx_{t-1}, σ_x^2)
  - $f(y_t|x_t)$ or *dmeasure*: dnorm(x_t, σ_y^2)

$$\begin{aligned} X_t &= \phi X_{t-1} + \epsilon_{xt} \\ Y_t &= X_t + \epsilon_{yt} \end{aligned}$$

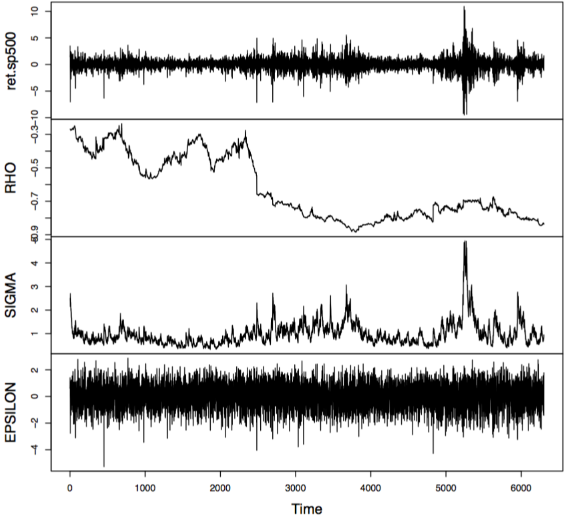
$$\begin{pmatrix} \epsilon_{xt} \\ \epsilon_{yt} \end{pmatrix} = N \left(0, \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \right)$$

- Algorithmic input:
  - Extended model variance (σ_ϕ^2): how big should the random walk variances in the extended model be
  - Speed of convergence ($\sigma_\phi^2 \to 0$): how fast should the extended model converge to the true model ($\propto$ number of iterations)

Accurate: Maximizing the likelihood



Contribution (III): Example from financial econometrics Stochastic volatility with stochastic leverage



Leverage equations: Harvey & Shephard, 1996

- Stochastic volatility with leverage (Harvey & Shephard, 1996)

$$
\begin{aligned}
y_t &= \sigma_t^2 \epsilon_t = \exp\left\{h_t/2\right\}\epsilon_t \\
h_t &= \mu_h(1-\phi) + \phi h_{t-1} + \sigma_\eta\sqrt{1-\phi^2}\left(\rho\epsilon_{t-1} + \omega_t\sqrt{1-\rho^2}\right)
\end{aligned}
$$

- Modifications to POMP code:
  - rprocess:
    $f(x_t|x_{t-1}) = \text{rnorm}(\phi x_{t-1}, \sigma_x^2) \longrightarrow \text{rnorm}(\mu_{x_t|x_{t-1}}, \sigma^2_{x_t|x_{t-1}})$
    $\mu_{x_t|x_{t-1}} = \mu(1-\phi) + \phi h_{t-1} + \sigma_\eta\sqrt{1-\phi^2}\rho\epsilon_{t-1}$
    $\sigma^2_{x_t|x_{t-1}} = \sigma_\eta\sqrt{1-\phi^2}\sqrt{1-\rho^2}$
  - *dmeasure*:
    $f(y_t|x_t) = \text{dnorm}(x_t, \sigma_y^2) \longrightarrow \text{dnorm}\ (0,\sigma_t^2)$

Idiosyncratic leverage: AR and RW equations

- Stochastic volatility with stochastic leverage (Bretó, 2013)

$$y_t = \sigma_t^2 \epsilon_t = \exp\left\{h_t/2\right\}\epsilon_t$$

$$h_t = \mu_h(1-\phi) + \phi h_{t-1} + \sigma_\eta \sqrt{1-\phi^2}\left(\rho_t \epsilon_{t-1} + \omega_t \sqrt{1-\rho^2}\right)$$

- Fisher-transformed correlation: $\rho_t = \frac{e^{2f_t}-1}{e^{2f_t}+1} \in (-1, 1)$
- AR(1) leverage: $f_t = \mu_f(1-\psi) + \psi f_{t-1} + \nu_t \sigma_\nu \sqrt{1-\psi^2}$
- RW leverage: $f_t = f_{t-1} + \nu_t \sigma_\nu$
- Stochastic volatility with random walk leverage: highly non-Gaussian, non-linear state-space model

Data analysis: S&P500 1988-2012 (25 y., 6302 obs.)

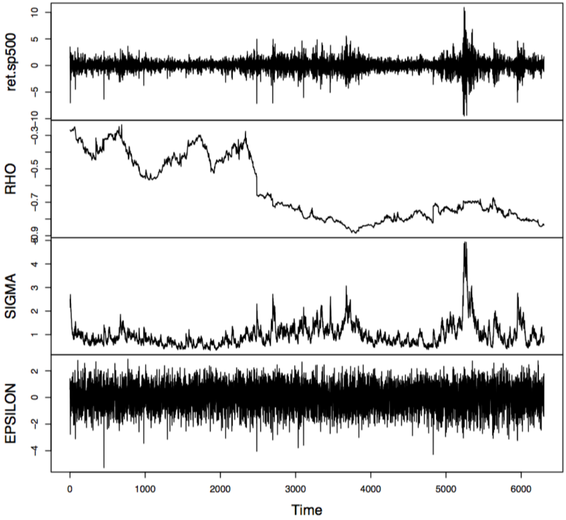
- Estimates for the volatility equation: usual & equal

Model	μ_h	ϕ	σ_η
Fixed leverage	-0.2506	0.9805	0.9003
	(0.0710)	(0.0017)	(0.0375)
Random-walk leverage	-0.2610	0.9818	0.9222
	(0.0776)	(0.0015)	(0.0406)

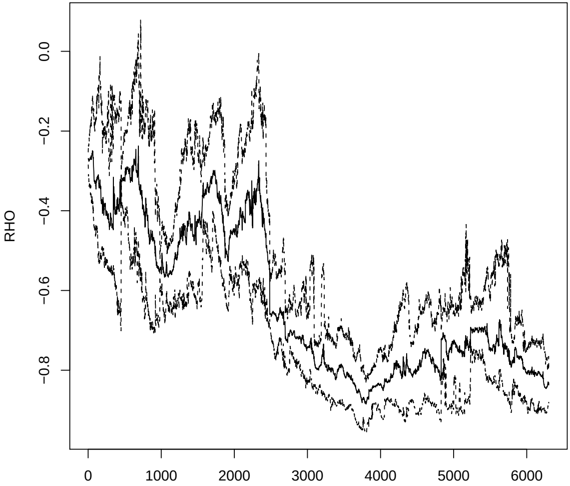
- ρ and likelihood (6.78 log-lik. units $\approx$ 7 parameters)

Model	ρ	σ_ν	ℓ
Fixed leverage	-0.6579	–	-8416.44
	(0.0599)	–	(0.0410)
Random-walk leverage	–	0.0086	-8409.06
	–	(0.0013)	(0.1333)

Random walk leverage (1988-2012, 25 y., 6302 obs.)



Random walk leverage (1988-2012, 25 y., 6302 obs.)



Flexibility: Extending stochastic volatility models

- Yu 2012

$$h_{t+1} = \varphi h_t + \gamma \sum_{i=1}^{m+1} (\rho_i \epsilon_t + \sqrt{1-\rho_i^2} w_t) \mathbf{1}(\tau_{i-1} \geq \epsilon_t > \tau_i)$$

- Bandi & Renó 2012

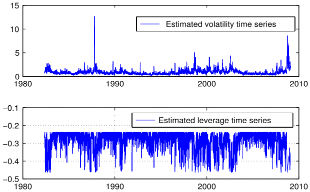
$$\begin{pmatrix} d \log p_t \\ d\xi(\sigma_t^2) \end{pmatrix} = \begin{pmatrix} \mu_t \\ m_t \end{pmatrix} dt + \begin{pmatrix} \sigma_t & 0 \\ 0 & \Lambda(\sigma_t^2) \end{pmatrix}$$
$$\times \begin{pmatrix} \rho(\sigma_t^2) & \sqrt{1-\rho^2(\sigma_t^2)} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} dW_t^1 \\ dW_t^2 \end{pmatrix} + \begin{pmatrix} dJ_t^r \\ dJ_t^\sigma \end{pmatrix}$$

- Veraart & Veraart 2012

$$dY_t = \left(\mu + b\sigma_t^2\right) dt + \sigma_t dX_t,$$
$$dX_t = \rho_t dW_t + \sqrt{1-\rho_t^2} d\widetilde{W}_t,$$
$$d\sigma_t^2 = \alpha(\beta - \sigma_t^2)dt + \gamma \sigma_t dW_t,$$

$$d\rho_t = ((2\zeta - \eta) - \eta\rho_t)\, dt + \theta\sqrt{(1+\rho_t)(1-\rho_t)} dW_t^V$$

Empirical evidence: Bandi and Renó (2012)



- Model: driven by mean-reverting spot volatility $\rho_t = \rho(\sigma_t^2)$
- Period: 1982-2008 (long: 27 years)
- Estimation: non-parametric & high-frequency data
- Results: $\rho_t \in (-0.45, -0.25) \pm$ s.e.
- Implication: $E[\rho_t] \approx -0.30$ (?)

Empirical evidence: Yu (2012)

- Model: white noise driven by return noise ϵ_{t-1}

- $\rho_t = \rho(\epsilon_{t-1})$:

$$\rho_t = \begin{cases} \rho_1 & \text{if } \tau_1 < \epsilon_{t-1} \leq \tau_0 \\ \vdots & \\ \rho_{m+1} & \text{if } \tau_{m+1} < \epsilon_{t-1} \leq \tau_m \end{cases}$$

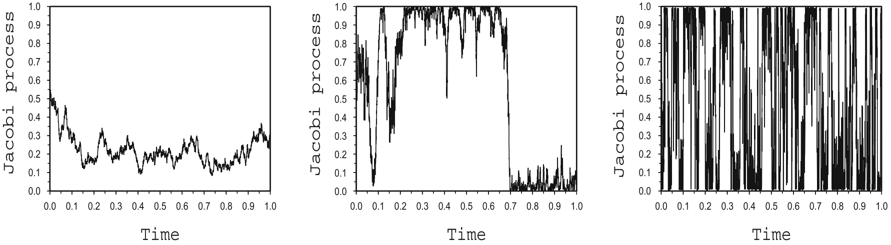
- $m = 1$ & $\tau_1 = 0$: simply two leverages

Empirical evidence: Yu (2012)

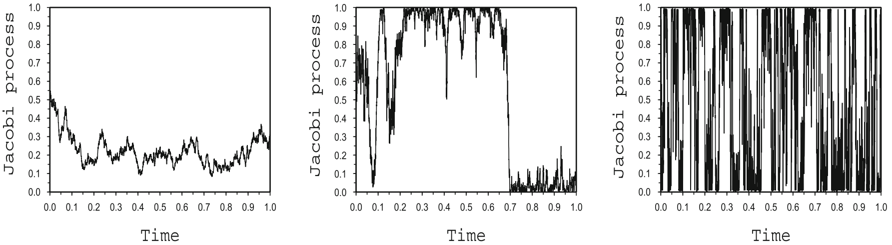
Data	Model	-Log MargLik	σ	φ	γ	ρ_1	ρ_2	ρ_3
S&P500	Basic	1730.26	0.9302	0.9426	0.2621			
			(0.067)	(0.019)	(0.042)			
			[0.001]	[0.0008]	[0.0021]			
	Leverage	1720.13	0.9336	0.9212	0.3066	−0.3691		
			(0.055)	(0.024)	(0.049)	(0.087)		
			[0.0014]	[0.0012]	[0.0030]	[0.0041]		
	Spline1	1700.40	2.077	0.9135	0.3689	−0.8386	0.1435	
			(0.3961)	(0.019)	(0.058)	(0.090)	(0.137)	
			[0.015]	[0.0008]	[0.0026]	[0.0133]	[0.005]	
	Spline2	1704.38	1.874	0.9157	0.3458	−0.8446	0.2059	0.1429
			(0.3537)	(0.019)	(0.051)	(0.1079)	(0.3484)	(0.1672)
			[0.014]	[0.0007]	[0.0022]	[0.0046]	[0.011]	[0.0066]

- Model: white noise driven by return noise ϵ_{t-1}
- Period: 1986-1989 (shorter: 4 years)
- Estimation: non-parametric & MCMC
- Results (m=2): $\rho_t \in (-1, -0.65) \cup (-0.13, 0.40)$
- Implication: $E[\rho_t] \approx -0.35$ ($\approx \rho$)

Idiosyncratic leverage: Veraart & Veraart (2012)



Idiosyncratic leverage: Veraart & Veraart (2012)



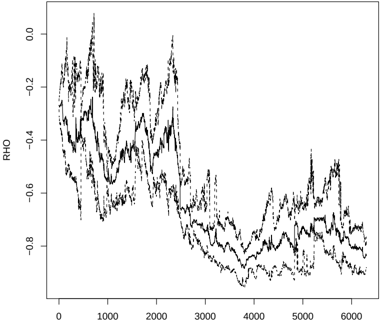
- **Model**: driven by idiosyncratic, independent Jacobi process V_t

$$dV_t = (\zeta - V_t)dt + \theta\sqrt{V_t(1 - V_t)}dW_t$$

$$\rho_t = 2V(t) - 1 \in [-1, 1]$$

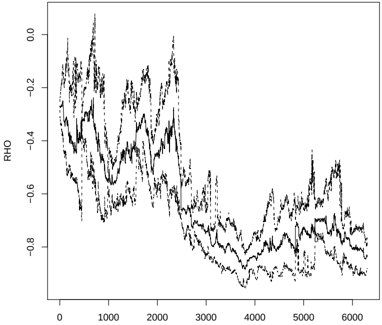
- **Period**, **Estimation**, **Results**: –
- **Implication**: $E[\rho_t] = 2\zeta - 1$
- Good candidate for new evidence!

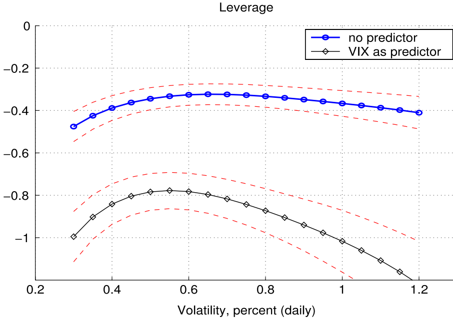
AR / RW evidence: consistent with Yu (2012)?



- Period 1988-1999:
  - RW: $\rho_t \in (-0.55, -0.10)$
  - Yu (2012): $\rho_1 \in (-0.5431, -0.1951)$

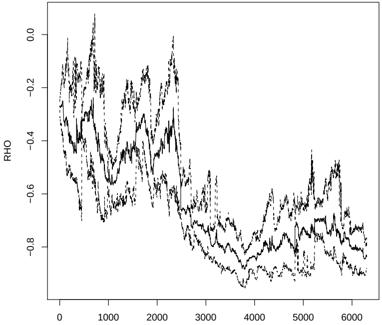
AR / RW evidence: consistent with Bandi (2012)?

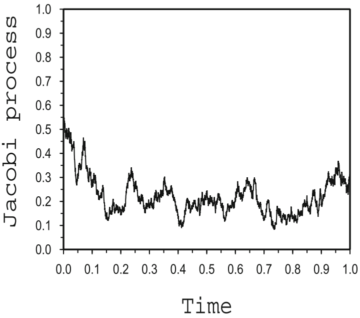




- Period 2003-2012:
  - RW: $\rho_t \in (-0.9, -0.6)$
  - Bandi (2012): ρ_t “more negative than those generally found in the literature” when using VIX as an instrumental variable for daily variance estimation

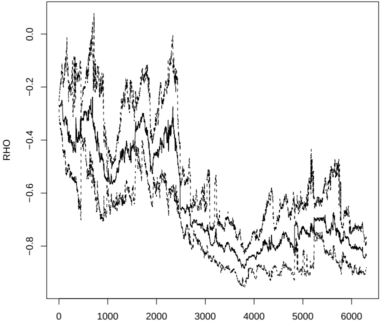
AR / RW evidence: consistent with Veraart (2012)?





- RW decade 1988-1997: $E[\rho_t] \approx -0.4$
- RW decade 2003-2012: $E[\rho_t] \approx -0.8$
- Fixed-leverage: $\rho \approx -0.6$

Idiosyncratic leverage: AR and RW evidence



- Model: driven by idiosyncratic AR/RW process
- Period: 1988-2012 (three sub-periods)
- Estimation: parametric & likelihood maximization
- Results: $\rho_t \in (-0.8, -0.3) \pm$ s.e. (consistent with Bandi, Yu, and Veraart?)
- Implication: $E[\rho_t] = ?$ $\quad E[\rho_t] \approx -0.6$ ($\approx \rho$) ?

Theory behind iterated filtering

- Update: $\hat{\theta}_0^{(n+1)} = V_{1,n}\Big(\sum_{t=1}^{T-1}(V_{t,n}^{-1} - V_{t+1,n}^{-1})\hat{\theta}_t^{(n)} + V_{T,n}^{-1}\hat{\theta}_T^{(n)}\Big)$
- Equivalently $\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + V_{1,n}\sum_{t=1}^{T} V_{t,n}^{-1}(\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)})$
- Assuming sufficient regularity conditions, a Taylor expansion gives

$$\lim_{\sigma\to 0}\sum_{t=1}^{T} V_t^{-1}(\hat{\theta}_t - \hat{\theta}_{t-1}) = \nabla \log f(y_{1:T}|\theta, \sigma=0)$$

Theory behind iterated filtering

- Update: $\hat{\theta}_0^{(n+1)} = V_{1,n}\Big(\sum_{t=1}^{T-1}(V_{t,n}^{-1} - V_{t+1,n}^{-1})\hat{\theta}_t^{(n)} + V_{T,n}^{-1}\hat{\theta}_T^{(n)}\Big)$
- Equivalently $\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + V_{1,n}\sum_{t=1}^{T} V_{t,n}^{-1}(\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)})$
- Assuming sufficient regularity conditions, a Taylor expansion gives

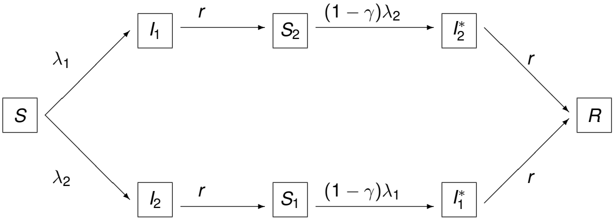
$$\lim_{\sigma\to 0}\sum_{t=1}^{T} V_t^{-1}(\hat{\theta}_t - \hat{\theta}_{t-1}) = \nabla \log f(y_{1:T}|\theta, \sigma{=}0)$$

Theory behind iterated filtering

- Update: $\hat{\theta}_0^{(n+1)} = V_{1,n}\Big(\sum_{t=1}^{T-1} (V_{t,n}^{-1} - V_{t+1,n}^{-1})\hat{\theta}_t^{(n)} + V_{T,n}^{-1}\hat{\theta}_T^{(n)}\Big)$
- Equivalently $\hat{\theta}^{(n+1)} = \hat{\theta}^{(n)} + V_{1,n}\sum_{t=1}^{T} V_{t,n}^{-1}(\hat{\theta}_t^{(n)} - \hat{\theta}_{t-1}^{(n)})$
- Assuming sufficient regularity conditions, a Taylor expansion gives

$$\lim_{\sigma\to 0} \sum_{t=1}^{T} V_t^{-1}(\hat{\theta}_t - \hat{\theta}_{t-1}) = \nabla \log f(y_{1:T}|\theta, \sigma{=}0)$$

A word of caution: Compartment models with random rates



$$\Delta N_{S_1 I_1^*}(t) = N_{S_1 I_1^*}(t+h) - N_{S_1 I_1^*}(t)$$

- Continuous Time Markov Chain (system of death processes)

$$\begin{aligned} P(\Delta N_{S_1 I_1^*}(t) = 0 | \boldsymbol{X}(t) = \boldsymbol{x}) &= 1 - (1-\gamma)\lambda_1 s_1 h + o(h) \\ P(\Delta N_{S_1 I_1^*}(t) = 1 | \boldsymbol{X}(t) = \boldsymbol{x}) &= (1-\gamma)\lambda_1 s_1 h + o(h) \\ P(\Delta N_{S_1 I_1^*}(t) > 1 | \boldsymbol{X}(t) = \boldsymbol{x}) &= o(h) \end{aligned}$$

Continuous-time & iterated filtering

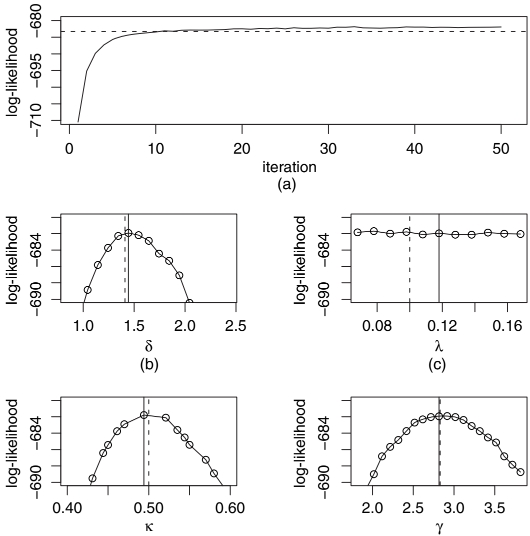
- IF has been applied to continuous-time stochastic volatility models
  - Comment to Andrieu et al., 2010, JRSSB: Particle MCMC and continuous-time Lévy-driven stochastic volatility model (Barndorff-Nielsen and Sheppard, 2001)

$$
\begin{aligned}
dy^*(t) &= \mu + \beta\sigma^2(t)dt + \sigma(t)dB(t) \\
d\sigma^2(t) &= -\lambda\sigma^2(t)dt + dz(\lambda t)
\end{aligned}
$$

  - Working paper - Bretó and Veiga: Forecasting performance and continuous-time log-linear one volatility factor model (Chernov et al., 2003)

$$
\begin{aligned}
dU_1(t) &= \alpha_{10}dt + \exp(\beta_{10} + \beta_{12}U_2(t))(\psi_{11}dW_1(t) + \psi_{12}dW_2(t)) \\
dU_2(t) &= \alpha_{22}U_2(t)dt + (1 + \beta_{22}U_2(t))dW_2(t).
\end{aligned}
$$

Continuous-time SV



Sequential Monte Carlo: Particle Filter

$t = 0$ $t = 1$ $t = 2$ $t = 3$

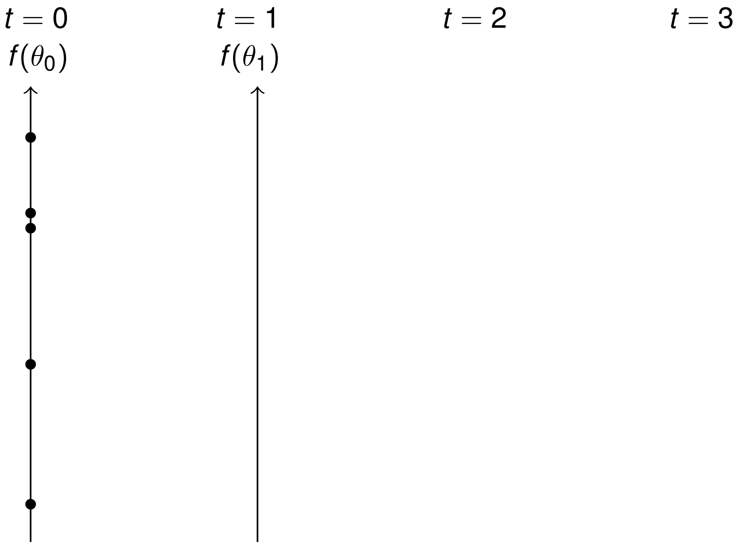
$f(\theta_0)$

Sequential Monte Carlo: Particle Filter

$t = 0$ $t = 1$ $t = 2$ $t = 3$

$f(\theta_0)$

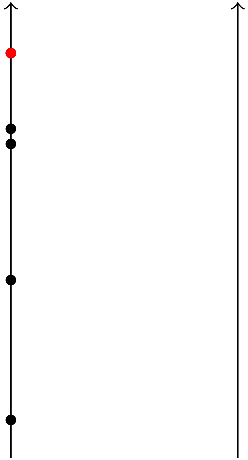
Sequential Monte Carlo: Particle Filter



Sequential Monte Carlo: Particle Filter

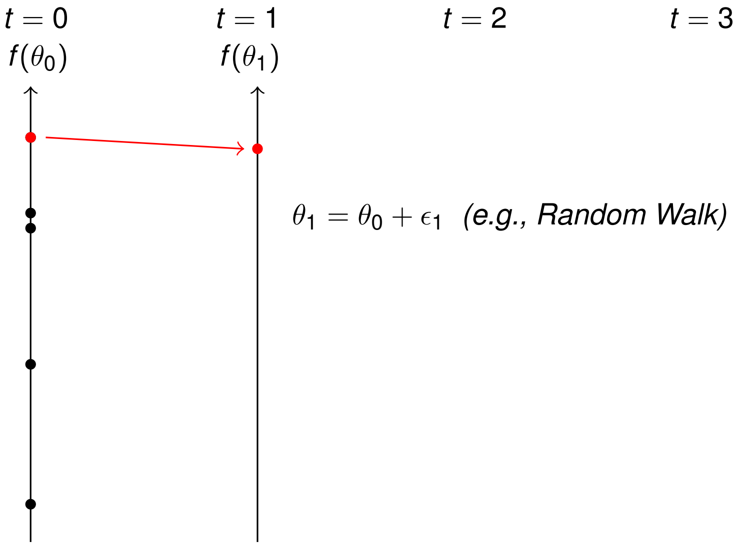
$t = 0$ $t = 1$ $t = 2$ $t = 3$

$f(\theta_0)$ $f(\theta_1)$



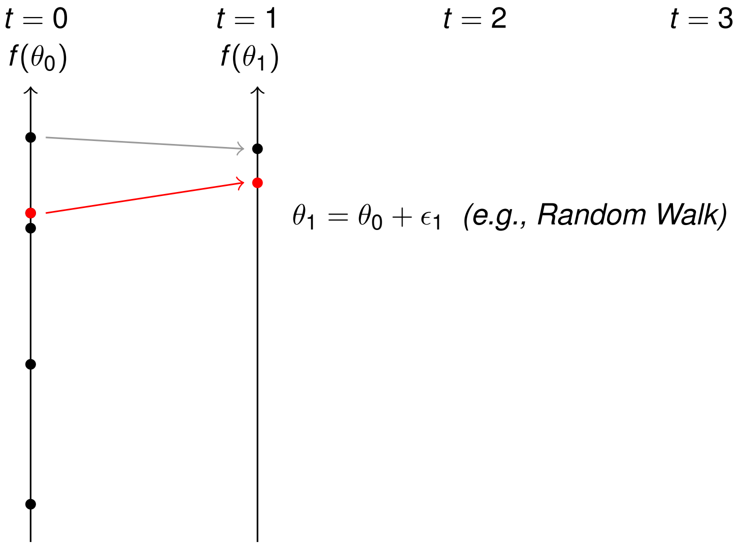
$\theta_1 = \theta_0 + \epsilon_1$ *(e.g., Random Walk)*

Sequential Monte Carlo: Particle Filter



$\theta_1 = \theta_0 + \epsilon_1$ *(e.g., Random Walk)*

Sequential Monte Carlo: Particle Filter

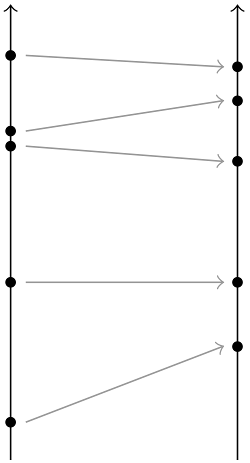


$\theta_1 = \theta_0 + \epsilon_1$ *(e.g., Random Walk)*

Sequential Monte Carlo: Particle Filter

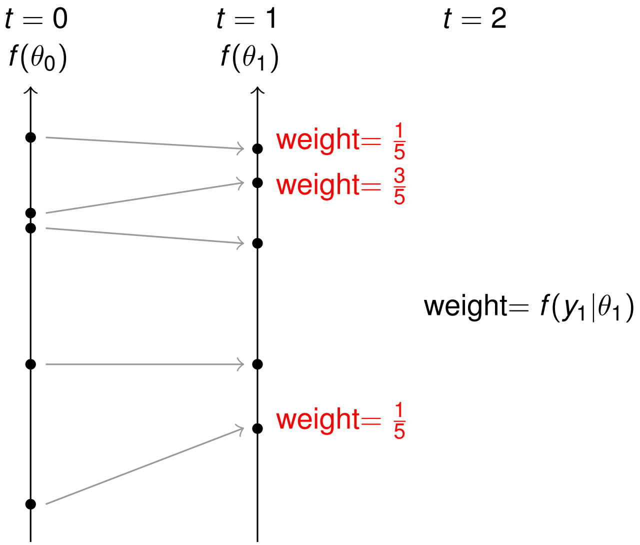
$t = 0$ $t = 1$ $t = 2$ $t = 3$

$f(\theta_0)$ $f(\theta_1)$

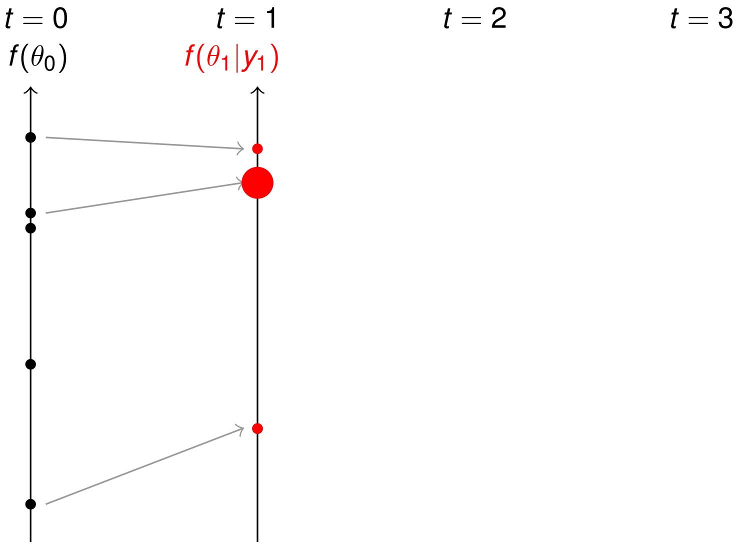


$\theta_1 = \theta_0 + \epsilon_1$ *(e.g., Random Walk)*

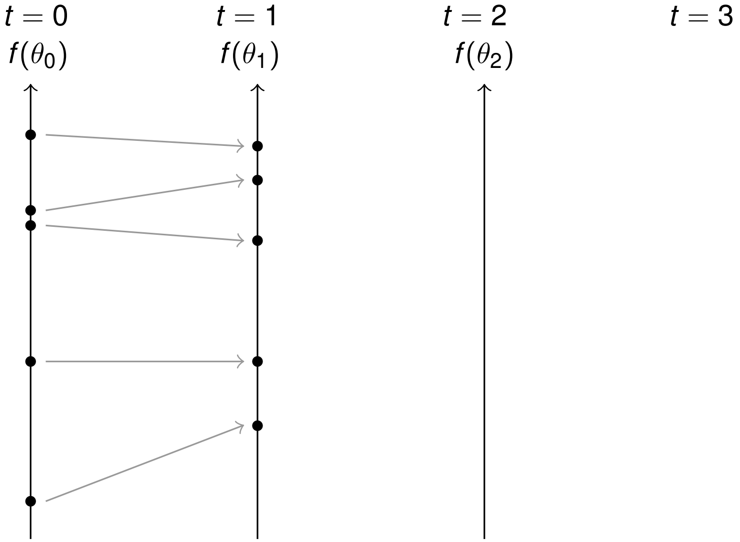
Sequential Monte Carlo: Particle Filter



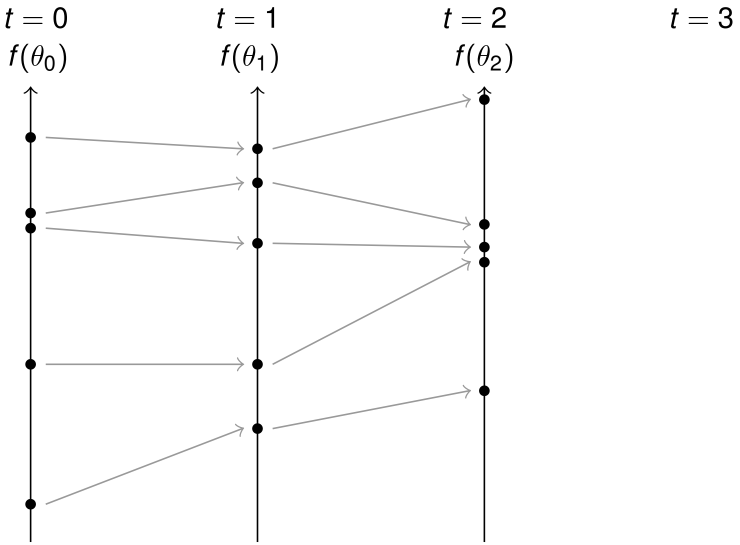
Sequential Monte Carlo: Particle Filter



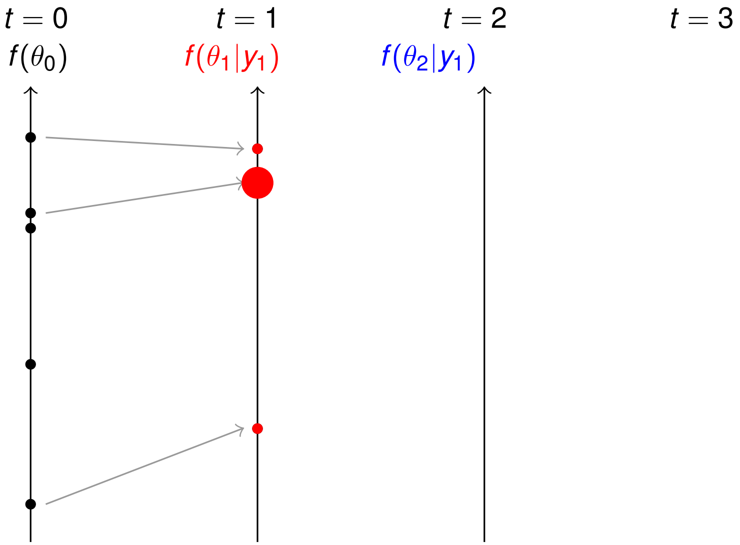
Sequential Monte Carlo: Particle Filter



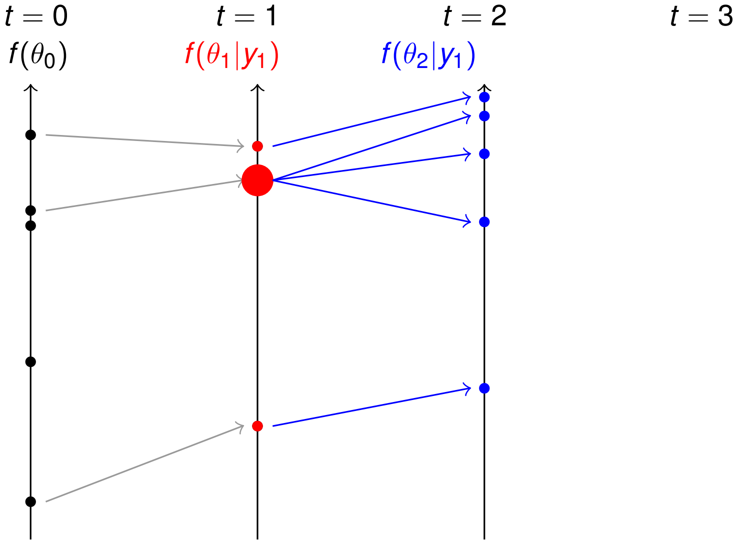
Sequential Monte Carlo: Particle Filter



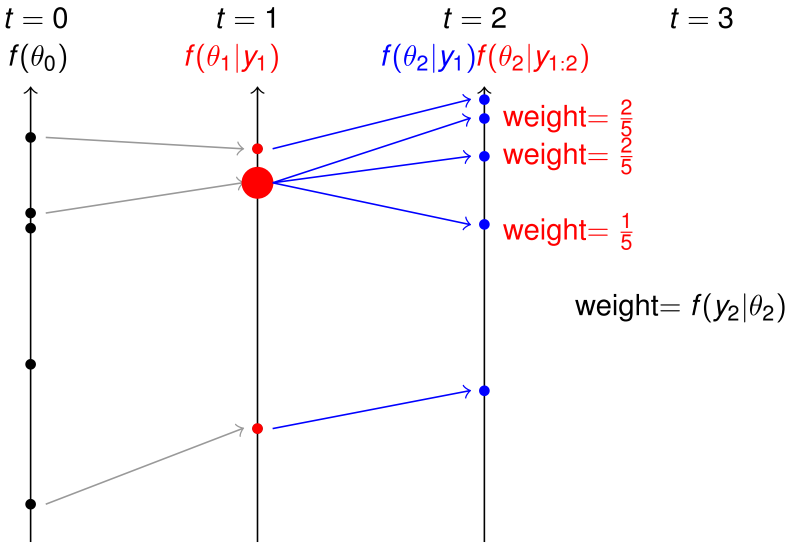
Sequential Monte Carlo: Particle Filter



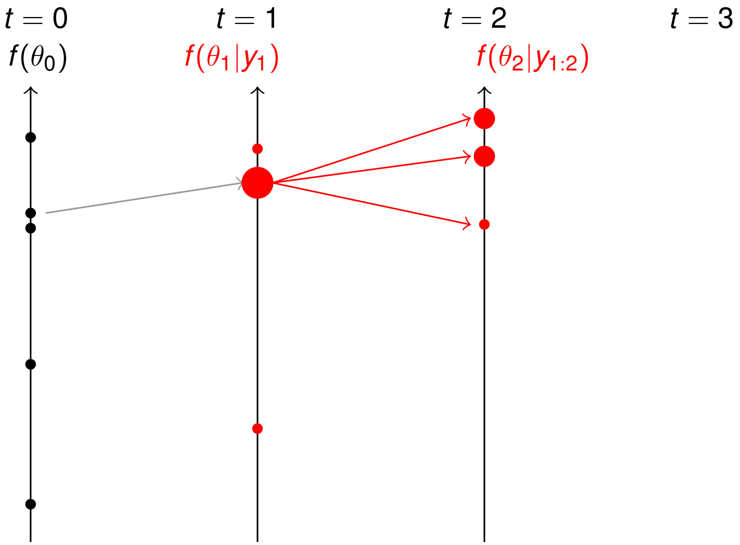
Sequential Monte Carlo: Particle Filter



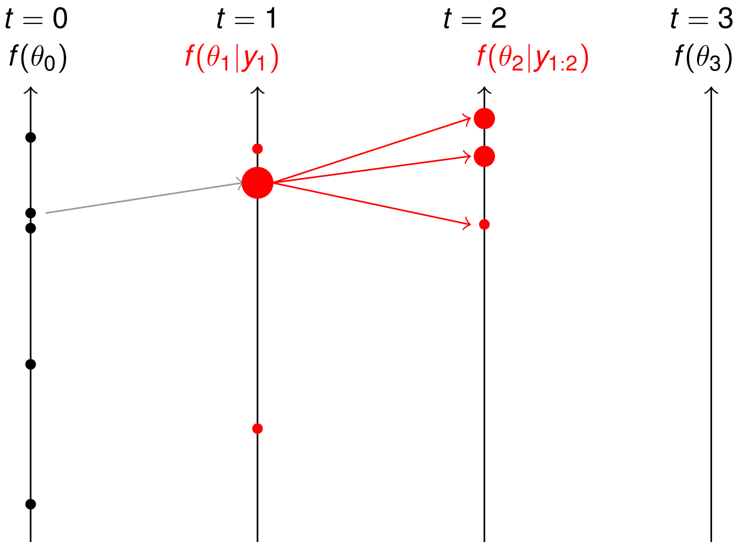
Sequential Monte Carlo: Particle Filter



Sequential Monte Carlo: Particle Filter



Sequential Monte Carlo: Particle Filter



Take-home message: straightforward, likelihood-based inference is possible for general dynamic systems

THANK YOU