On the two-filter approximations of marginal smoothing distributions in general state space models

Sylvain Le Corff

based on joint work with Thi Ngoc Minh Nguyen and Eric Moulines

U. Paris-Sud and CNRS

July 5, 2016

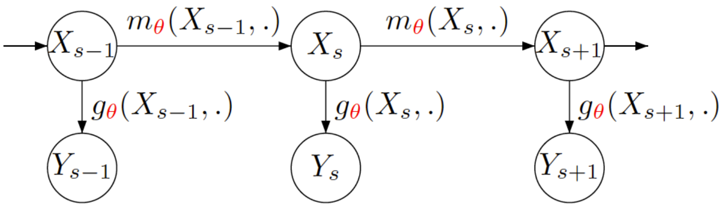
1. Framework
   - Hidden Markov models (HMM)
   - SMC to approximate smoothing distributions

2. Two-filter algorithms
   - Forward filter
   - Backward information filter
   - Recombinations to approximate marginal smoothing distributions

3. Exponential deviation inequalities

4. Asymptotic normality

Hidden Markov models (HMM)



- $\mathbf{Y} \stackrel{\text{def}}{=} \{\mathbf{Y}_t\}_{t\in\mathbb{Z}}$ is the observation process and $\mathbf{X} \stackrel{\text{def}}{=} \{\mathbf{X}_t\}_{t\in\mathbb{Z}}$ are the hidden states.
- The distribution of the HMM is specified by
  - Distribution of X_0 with probability density χ_θ.
  - Transition kernels with density m_θ on $\mathbb{X} \times \mathcal{B}(\mathbb{X})$ governing the transition of the hidden chain.
  - Transition kernels with density g_θ on $\mathbb{X} \times \mathcal{B}(\mathbb{Y})$, the conditional likelihood of the observations.

Examples of hidden Markov models (HMM)

- Simultaneous localization and mapping (SLAM)

$\{X_k\}_{k \geq 0}$ is the state (cartesian coordinates, bearing) of a mobile device.

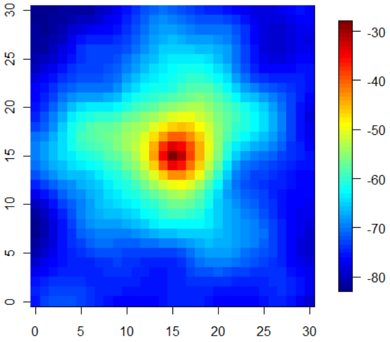
- Transition model with input u_k:

$$X_k = h(X_{k-1}, u_k, \varepsilon_k) \; .$$

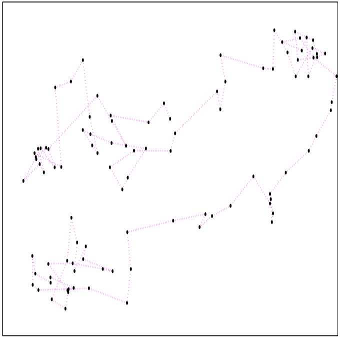
- Environment represented by a set of landmarks : $(\theta_j)_{j \in [\![1,p]\!]}$.
  Observations received according to the model:

$$Y_{k,i} = t_k(\theta_{c_k^i}, X_K) + \delta_{k,i} \; .$$

Power received from a WiFi access point.



Device position.



- X_k is the device position.
- At each time k, the device observes the power of signals transmitted by ℓ antennas.
- In this application, $f_\star$ is the mean propagation model :

$$Y_k \sim \mathcal{N}(f_\star(X_k), \sigma^2 I_\ell) \ .$$

Smoothing distributions

We are interested in estimating the joint smoothing distributions, defined, for any measurable function h on $\mathbb{X}^{t-s+1}$, $T \geq 0$ and $0 \leq s \leq t \leq T$, by:

$$\phi_{s:t|T}[h] = \frac{\int \chi(x_0)g_0(x_0)\prod_{u=1}^{T} m(x_{u-1}, x_u)g_u(x_u)h(x_{s:t})\mathrm{d}x_{0:T}}{\int \chi(x_0)g_0(x_0)\prod_{u=1}^{T} m(x_{u-1}, \mathrm{d}x_u)g_u(x_u)\mathrm{d}x_{0:T}} .$$

$$\underline{\phi_{s:t|T}[h] = \mathbb{E}\left[h(X_{s:t})\middle|Y_{0:T}\right]} .$$

These distributions are crucial for Inference of HMM:

- Statistical inference for the distributions of the X_k's and the Y_k's.

- Parameter estimation (EM algorithm, stochastic gradient, Particle MCMC).

The two-filter algorithms are designed to estimate $\phi_{s|T} = \phi_{s:s|T}$.

Auxiliary particle filter, Pitt and Shephard, J. Am. Statist. Assoc. '99

First step to estimate $\phi_{s|T}$ shared by common SMC smoothers.

$\phi_{s|s}[h]$ is approximated by particles and weights $\{(\xi_s^\ell, \omega_s^\ell)\}_{\ell=1}^N$:

$$\phi_s^N[h] = \frac{1}{\Omega_s^N} \sum_{\ell=1}^N \omega_s^\ell h(\xi_s^\ell) \; .$$

- Initialisation:

1. (Initial states) $\{\xi_0^\ell\}_{\ell=1}^N$ i.i.d. distributed according to ρ_0.
2. (Initial weights) $\omega_0^\ell = \chi(\xi_0^\ell)\, g_0(\xi_0^\ell) \,/ \rho_0(\xi_0^\ell)$.

- Iterations for $s \geq 1$:

1. (selection and propagation) Pairs $\{(I_s^\ell, \xi_s^\ell)\}_{\ell=1}^N$ of indices and particles are simulated independently from:
$$\pi_s(\ell, x) \propto \omega_{s-1}^\ell \vartheta_s(\xi_{s-1}^\ell) p_s(\xi_{s-1}^\ell, x) \; .$$
2. (weights) ξ_s^ℓ is associated with the importance weight defined by:
$$\omega_s^\ell = \frac{m(\xi_{s-1}^{I_s^\ell}, \xi_s^\ell) g_s(\xi_s^\ell)}{\vartheta_s(\xi_{s-1}^{I_s^\ell}) p_s(\xi_{s-1}^{I_s^\ell}, \xi_s^\ell)} .$$

Forward-backward smoothers

- Forward Filtering Backward Smoothing (FFBS), Doucet et al., Statist. and Comput. '00:
  - stores all filtering particles and weights discarding the genealogy of the particles.
  - keeps all the particles fixed but modifies all importance weights during a backward pass.
  - $\mathcal{O}(N^2)$ complexity for marginal smoothing distributions.
  - exponential deviation inequalities, CLT, $\mathrm{L_q}$-mean error...
  - forward only version for fixed smoothed expectations.
- Forward Filtering Backward Simulation (FFBSi), Godsill et al., J. Am. Statist. Assoc. '04:
  - stores all filtering particles and weights discarding the genealogy of the particles.
  - samples trajectories backward among all the possible paths made of filtering particles.
  - $\mathcal{O}(N^2)$ complexity for all smoothing distributions.
  - same results as the FFBS algorithm.
  - may be implemented with $\mathcal{O}(N)$ complexity if m is upper bounded.
- PaRIS, Olsson and Westerborn, Bernoulli '15:
  - combines the forward only version of the FFBS with the sampling procedure of the FFBSi.
  - may be implemented with $\mathcal{O}(N)$ complexity if m is upper bounded.
  - same results as the FFBS algorithm.

Auxiliary distribution

Kitagawa, J. Comput. Graph. Statist. '96 and Briers et al, AISM, '10

Let $\{\gamma_t\}_{t\geq 0}$ be positive measurable functions such that, for all $t \in \{0, \ldots, T\}$,

$$\int \gamma_t(x_t)\,\mathrm{d}x_t \left[\prod_{u=t+1}^{T} g_{u-1}(x_{u-1})\, m(x_{u-1}, x_u)\right] g_T(x_T)\mathrm{d}x_{t:T} < \infty\ .$$

Then, the backward information filter is given by

$$\psi_{\gamma,t|T}[h] \propto \int \gamma_t(x_t) \left[\prod_{u=t+1}^{T} g_{u-1}(x_{u-1})\, m(x_{u-1}, x_u)\right] g_T(x_T)h(x_t)\mathrm{d}x_{t:T}\ .$$

If X_t has pdf γ_t, then $\psi_{\gamma,t|T}$ is the conditional distribution of X_t given $Y_{t:T}$.

The marginal smoothing distribution may be expressed as

$$\phi_{s|T}[h] \propto \int \phi_{s-1}(\mathrm{d}x_{s-1})\psi_{\gamma,s+1|T}(\mathrm{d}x_{s+1})m(x_{s-1}, x_s)g_s(x_s)\frac{m(x_s, x_{s+1})}{\gamma_{s+1}(x_{s+1})}h(x_s)\mathrm{d}x_s\ .$$

Particle approximation of the backward information filter

$\psi_{\gamma,t|T}[h]$ is approximated by particles and weights $\{(\check{\xi}_t^\ell, \check{\omega}_t^\ell)\}_{\ell=1}^N$:

$$\psi_{\gamma,t|T}^N[h] = \frac{1}{\check{\Omega}_t^N}\sum_{\ell=1}^N \check{\omega}_t^\ell h(\check{\xi}_t^\ell)\,.$$

- Initialisation:

1. (Initial states) $\{\check{\xi}_{T|T}^i\}_{i=1}^N$ i.i.d. distributed according to $\check{\rho}_T$.
2. (Initial weights) $\check{\omega}_{T|T}^i = g_T(\check{\xi}_{T|T}^i)\gamma_T(\check{\xi}_{T|T}^i)/\check{\rho}_T(\check{\xi}_{T|T}^i)$.

- Iterations for $t \leq T-1$:

1. (selection and propagation) Pairs $\{(\check{I}_t^i, \check{\xi}_{t|T}^i)\}_{i=1}^N$ of indices and particles are simulated independently from:

$$\pi_{t|T}(i, x_t) \propto \frac{\check{\omega}_{t+1|T}^i \vartheta_{t|T}(\check{\xi}_{t+1|T}^i)}{\gamma_{t+1}(\check{\xi}_{t+1|T}^i)} r_{t|T}(\check{\xi}_{t+1|T}^i, x_t)\,.$$

2. (weights) ξ_s^ℓ is associated with the importance weight defined by:

$$\check{\omega}_{t|T}^i \stackrel{\text{def}}{=} \frac{\gamma_t(\check{\xi}_{t|T}^i) g_t(\check{\xi}_{t|T}^i) m(\check{\xi}_{t|T}^i, \check{\xi}_{t+1|T}^{\check{I}_t^i})}{\vartheta_{t|T}(\check{\xi}_{t+1|T}^{\check{I}_t^i}) r_{t|T}(\check{\xi}_{t+1|T}^{\check{I}_t^i}, \check{\xi}_{t|T}^i)}\,.$$

The TwoFilt$_{fwt}$ algorithm, Fearnhead et al., Biometrika '10

SMC approximations are plugged in

$$\phi_{s|T}[h] \propto \int \phi_{s-1}(\mathrm{d}x_{s-1})\psi_{\gamma,s+1|T}(\mathrm{d}x_{s+1})m(x_{s-1},x_s)g_s(x_s)\frac{m(x_s,x_{s+1})}{\gamma_{s+1}(x_{s+1})}h(x_s)\mathrm{d}x_s \;,$$

to obtain

$$\hat{\phi}^{\mathrm{tar}}_{s|T}(x_s) \propto \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\omega^i_{s-1}\check{\omega}^j_{s+1|T}}{\gamma_{s+1}(\check{\xi}^j_{s+1|T})} m(\xi^i_{s-1},x_s)g_s(x_s)q(x_s,\check{\xi}^j_{s+1|T}) \;.$$

1. (selection and propagation) Pairs $\{(I^\ell_s, \check{I}^\ell_s, \tilde{\xi}^\ell_{s|T})\}_{\ell=1}^N$ of indices and particles are simulated independently from:

$$\pi_{s|T}(i,j,x_s) \propto \frac{\omega^i_{s-1}\tilde{\vartheta}_{s|T}(\xi^i_{s-1},\check{\xi}^j_{s+1|T})\check{\omega}^j_{s+1|T}}{\gamma_{s+1}(\check{\xi}^j_{s+1|T})}\tilde{r}_{s|T}(\xi^i_{s-1},\check{\xi}^j_{s+1|T};x_s) \;.$$

2. (weights) $\tilde{\xi}^\ell_{s|T}$ is associated with the importance weight defined by:

$$\tilde{\omega}^\ell_{s|T} \stackrel{\mathrm{def}}{=} \frac{m(\xi^{I^\ell_s}_{s-1},\tilde{\xi}^\ell_{s|T})g_s(\tilde{\xi}^\ell_{s|T})m(\tilde{\xi}^\ell_{s|T},\check{\xi}^{\check{I}^\ell_s}_{s+1|T})}{\tilde{\vartheta}_{s|T}(\xi^{I^\ell_s}_{s-1},\check{\xi}^{\check{I}^\ell_s}_{s+1|T})\tilde{r}_{s|T}(\xi^{I^\ell_s}_{s-1},\check{\xi}^{\check{I}^\ell_s}_{s+1|T};\tilde{\xi}^\ell_{s|T})} \;.$$

The TwoFilt$_{bdm}$ algorithm, Briers et al, AISM, '10

Following instead Briers et al, AISM, '10, we may consider one of the *partial* auxiliary distributions:

$$\phi^{\mathrm{tar,f}}_{s|T}(i, x_s) \propto \omega^i_{s-1} m(\xi^i_{s-1}, x_s) g_s(x_s) \sum_{j=1}^{N} \frac{\breve{\omega}^j_{s+1|T}}{\gamma_{s+1}(\breve{\xi}^j_{s+1|T})} m(x_s, \breve{\xi}^j_{s+1|T}) \,,$$

$$\phi^{\mathrm{tar,b}}_{s|T}(j, x_s) \propto \frac{\breve{\omega}^j_{s+1|T}}{\gamma_{s+1}(\breve{\xi}^j_{s+1|T})} m(x_s, \breve{\xi}^j_{s+1|T}) g_s(x_s) \sum_{i=1}^{N} \omega^i_{s-1} m(\xi^i_{s-1}, x_s) \,.$$

1. (selection and propagation) Pairs $\{(I^\ell_s, \xi^\ell_s)\}_{\ell=1}^N$ or $\{(\breve{I}^\ell_s, \breve{\xi}^\ell_{s|T})\}_{\ell=1}^N$ of indices and particles are simulated independently from:

$$\pi^{\mathrm{f}}_{s|T}(i, x_s) \propto \omega^i_{s-1} \vartheta_s(\xi^i_{s-1}) p_s(\xi^i_{s-1}, x_s) \,,$$

$$\pi^{\mathrm{b}}_{s|T}(j, x_s) \propto \vartheta_{s|T}(\breve{\xi}^j_{s+1|T}) \breve{\omega}^j_{s+1|T} r_{s|T}(\breve{\xi}^j_{s+1|T}, x_s) / \gamma_{s+1}(\breve{\xi}^j_{s+1|T}) \,.$$

2. (weights) ξ^ℓ_s is associated with the importance weight defined by:

$$\tilde{\omega}^{i,\mathrm{f}}_{s|T} \stackrel{\mathrm{def}}{=} \omega^i_s \sum_{j=1}^{N} \breve{\omega}^j_{s+1|T} m(\xi^i_s, \breve{\xi}^j_{s+1|T}) / \gamma_{s+1}(\breve{\xi}^j_{s+1|T}) \,,$$

$$\tilde{\omega}^{j,\mathrm{b}}_{s|T} \stackrel{\mathrm{def}}{=} \breve{\omega}^j_{s|T} \sum_{i=1}^{N} \omega^i_{s-1} m(\xi^i_{s-1}, \breve{\xi}^j_{s|T}) / \gamma_s(\breve{\xi}^j_{s|T}) \,.$$

Exponential deviation inequality for TwoFilt_{fwt}

We first show that the weighted sample $\{(\omega_s^i \check{\omega}_{t|T}^j), (\xi_s^i, \check{\xi}_{t|T}^j)\}_{i,j=1}^N$ targets the product distribution $\phi_s \otimes \psi_{\gamma,t|T}$.

For all $0 \leq s < t \leq T$, there exist $0 < B_{s,t|T}, C_{s,t|T} < \infty$ such that for all $N \geq 1$, $\epsilon > 0$ and all bounded function h,

$$\mathbb{P}\left(\left|\sum_{i,j=1}^N \frac{\omega_s^i}{\Omega_s} \frac{\check{\omega}_{t|T}^j}{\check{\Omega}_{t|T}} h(\xi_s^i, \check{\xi}_{t|T}^j) - \phi_s \otimes \psi_{\gamma,t|T}[h]\right| > \varepsilon\right) \leq B_{s,t|T} \mathrm{e}^{-C_{s,t|T} N \epsilon^2 / \operatorname{osc}^2(h)} .$$

and there exist $0 < B_{s|T}, C_{s|T} < \infty$ such that $\{(\tilde{\omega}_{s|T}^i, \tilde{\xi}_{s|T}^\ell)\}_{\ell=1}^N$ satisfies:

$$\mathbb{P}\left(\left|\sum_{i=1}^N \frac{\tilde{\omega}_{s|T}^i}{\tilde{\Omega}_{s|T}} h(\tilde{\xi}_{s|T}^i) - \phi_{s|T}[h]\right| > \epsilon\right) \leq B_{s|T} \mathrm{e}^{-C_{s|T} N \epsilon^2 / \operatorname{osc}^2(h)} .$$

Exponential deviation inequality for TwoFilt_{bdm}

Similarly, we may derive an exponential inequality for the weighted samples $\{(\xi_s^i, \tilde{\omega}_{s|T}^{i,\mathrm{f}})\}_{i=1}^N$ and $\{(\check{\xi}_{s|T}^i, \tilde{\omega}_{s|T}^{i,\mathrm{b}})\}_{i=1}^N$ produced by the TwoFilt_{bdm} algorithm.

Then, for all $1 \leq s \leq T-1$, there exist $0 < B_{s|T}, C_{s|T} < \infty$ such that for all $N \geq 1$, $\varepsilon > 0$ and all bounded function h,

$$
\mathbb{P}\left(\left|\sum_{i=1}^N \frac{\tilde{\omega}_{s|T}^{i,\mathrm{f}}}{\tilde{\Omega}_{s|T}^{\mathrm{f}}} h(\xi_s^i) - \phi_{s|T}[h]\right| > \epsilon\right) \leq B_{s|T} \mathrm{e}^{-C_{s|T} N \epsilon^2 / \operatorname{osc}^2(h)},
$$

$$
\mathbb{P}\left(\left|\sum_{i=1}^N \frac{\tilde{\omega}_{s|T}^{i,\mathrm{b}}}{\tilde{\Omega}_{s|T}^{\mathrm{b}}} h(\check{\xi}_{s|T}^i) - \phi_{s|T}[h]\right| > \epsilon\right) \leq B_{s|T} \mathrm{e}^{-C_{s|T} N \epsilon^2 / \operatorname{osc}^2(h)}.
$$

Time uniform exponential inequalities may be obtained using *strong mixing* assumptions which are standard in the SMC literature:

A CLT may be derived for the weighted samples $\{(\xi_s^\ell, \omega_s^\ell)\}_{\ell=1}^N$ and $\{(\check{\xi}_{t|T}^i, \check{\omega}_{t|T}^i)\}_{i=1}^N$ which target respectively the filtering distribution ϕ_s and the backward information filter $\psi_{\gamma,t|T}$.

$$N^{1/2} \sum_{i=1}^N \frac{\omega_s^i}{\Omega_s} \left(h(\xi_s^i) - \phi_s[h] \right) \xrightarrow{\mathcal{D}}_{N\to\infty} \mathcal{N}\left(0, \Gamma_s\left[h - \phi_s[h]\right]\right) ,$$

$$N^{1/2} \sum_{j=1}^N \frac{\check{\omega}_{t|T}^j}{\check{\Omega}_{t|T}} \left(h(\check{\xi}_{t|T}^j) - \psi_{\gamma,t|T}[h] \right) \xrightarrow{\mathcal{D}}_{N\to\infty} \mathcal{N}\left(0, \check{\Gamma}_{\gamma,t|T}\left[h - \psi_{\gamma,t|T}[h]\right]\right) .$$

Then, for all $0 \le s < t \le T$ and all bounded function h,

$$\sqrt{N} \left(\sum_{i,j=1}^N \frac{\omega_s^i}{\Omega_s} \frac{\check{\omega}_{t|T}^j}{\check{\Omega}_{t|T}} h(\xi_s^i, \check{\xi}_{t|T}^j) - \phi_s \otimes \psi_{\gamma,t|T}[h] \right)$$

$$\xrightarrow{\mathcal{D}}_{N\to\infty} \mathcal{N}\left(0, \tilde{\Gamma}_{s,t|T}\left[h - \phi_s \otimes \psi_{\gamma,t|T}[h]\right]\right) ,$$

where

$$\tilde{\Gamma}_{s,t|T}[h] \stackrel{\mathrm{def}}{=} \Gamma_s\left[\int \psi_{\gamma,t|T}(\mathrm{d}x_t) h(\cdot, x_t)\right] + \check{\Gamma}_{\gamma,t|T}\left[\int \phi_s(\mathrm{d}x_s) h(x_s, \cdot)\right] .$$

Asymptotic normality of TwoFilt_{bdm}

$$\sqrt{N}\left(\sum_{i=1}^{N}\frac{\tilde{\omega}_{s|T}^{i,\mathrm{f}}}{\tilde{\Omega}_{s|T}^{\mathrm{f}}}h(\xi_s^i)-\phi_{s|T}[h]\right)\xrightarrow{\mathcal{D}}_{N\to\infty}\mathcal{N}\left(0,\Delta_{s|T}^{\mathrm{f}}\left[h-\phi_{s|T}[h]\right]\right),$$

where

$$\Delta_{s|T}^{\mathrm{f}}[h]\stackrel{\text{def}}{=}\tilde{\Gamma}_{s,s+1|T}\left[H_s^{\mathrm{f}}\right]/\{\phi_s\otimes\psi_{\gamma,s+1|T}[q\odot\gamma_{s+1}^{-1}]\}^2,$$

$$H_s^{\mathrm{f}}(x,x')\stackrel{\text{def}}{=}h(x)q(x,x')\gamma_{s+1}^{-1}(x').$$

Similarly,

$$\sqrt{N}\left(\sum_{i=1}^{N}\frac{\tilde{\omega}_{s|T}^{i,\mathrm{b}}}{\tilde{\Omega}_{s|T}^{\mathrm{b}}}h(\xi_s^i)-\phi_{s|T}[h]\right)\xrightarrow{\mathcal{D}}_{N\to\infty}\mathcal{N}\left(0,\Delta_{s|T}^{\mathrm{b}}\left[h-\phi_{s|T}[h]\right]\right),$$

where

$$\Delta_{s|T}^{\mathrm{b}}[h]\stackrel{\text{def}}{=}\tilde{\Gamma}_{s-1,s|T}\left[H_s^{\mathrm{b}}\right]/\{\phi_{s-1}\otimes\psi_{\gamma,s|T}[q\odot\gamma_s^{-1}]\}^2,$$

$$H_s^{\mathrm{b}}(x,x')\stackrel{\text{def}}{=}q(x,x')\gamma_s^{-1}(x')h(x').$$

Asymptotic normality of TwoFilt_{bdm}

- In the case where $\tilde{r}_{s|T}(x_s, x_{s+1}; x_s) = p_s(x_{s-1}, x_s)$ and $\tilde{\vartheta}_{s|T}(x, x') = \vartheta_s(x)\vartheta_{s|T}(x')$, the smoothing distribution approximation given by the TwoFilt_{fwt} algorithm is obtained by reweighting the particles obtained in the forward filtering pass.

- When $\vartheta_{s|T} = \gamma_{s+1}$, the asymptotic variance $\Upsilon_{s|T}[h]$ of the TwoFilt_{fwt} algorithm may be compared to $\Delta^{\mathrm{f}}_{s|T}[h]$ as both approximations of $\phi_{s|T}[h]$ are based on the same particles (associated with different importance weights):

$$\Upsilon_{s|T}[h] \geq \Delta^{\mathrm{f}}_{s|T}[h] \ .$$

- Under the strong mixing assumptions, time uniform bounds for the asymptotic variances of the two-filter approximations may be obtained.

Particle filter variance estimation, Lee and Whiteley, '15

- The asymptotic variance $\tilde{\Gamma}_{s,t|T}[h]$ is the sum of two variances: the forward filter asymptotic variance Γ_s and the backward information filter asymptotic variance $\check{\Gamma}_{\gamma,t|T}$.

- Lee and Whiteley, '15 introduced a weakly consistent estimator $\Gamma_s^N[h]$ of the asymptotic variance $\Gamma_s[h]$ based on $\{(\xi_r^\ell, \omega_r^\ell)\}_{\ell=1}^N$, $0 \leq r \leq s$, and may be computed on-the-fly.

- This algorithm may also be used to obtain an estimator $\check{\Gamma}_{\gamma,T|t}^N[h]$ of $\check{\Gamma}_{\gamma,t|T}[h]$.

- Let $(E_r)_{0 \leq r \leq s} \in \{1,\ldots,N\}^{s+1}$ be such that for all $i \in \{1,\ldots,N\}$ and all $0 \leq r \leq s$, E_r^i is the index of the time 0 ancestor of ξ_r^i.

  For all $i \in \{1,\ldots,N\}$, $E_0^i = i$ and for all $i \in \{1,\ldots,N\}$ and all $1 \leq r \leq s$, E_r^i,

$$E_r^i \overset{\text{def}}{=} E_{r-1}^{I_r^i} .$$

Particle filter variance estimation, Lee and Whiteley, '15

- For $i = 1$ to $i = N$, compute

$$\psi_{0,s}^i[h] \stackrel{\text{def}}{=} \sum_{\substack{j=1 \\ E_s^j = i}}^N \omega_s^j \left[h(\xi_s^j) - \phi_s^N[h]\right]$$

and

$$\check{\psi}_{t,T}^i[h] \stackrel{\text{def}}{=} \sum_{\substack{j=1 \\ \check{E}_t^j = i}}^N \check{\omega}_{t|T}^j \left[h(\check{\xi}_{t|T}^j) - \psi_{\gamma,t|T}^N[h]\right] .$$

- Set $\alpha_N = N/(N-1)$.
- Set

$$\Gamma_s^N\left[h - \phi_s[h]\right] \stackrel{\text{def}}{=} N\alpha_N^{s+1} \sum_{i=1}^N \left(\psi_{0,s}^i[h]/\Omega_s\right)^2$$

and

$$\check{\Gamma}_{\gamma,T|t}^N\left[h - \psi_{\gamma,t|T}[h]\right] \stackrel{\text{def}}{=} N\alpha_N^{T-t+1} \sum_{i=1}^N \left(\check{\psi}_{t,T}^i[h]/\check{\Omega}_{t|T}\right)^2 .$$

TwoFilt$_{bdm}$ variance estimation

Weakly consistent estimator of the asymptotic variance of the TwoFilt$_{bdm}$ algorithm:

$$\Delta^{\mathrm{f}}_{s|T}\left[h\right] \stackrel{\mathrm{def}}{=} \tilde{\Gamma}_{s,s+1|T}\left[H^{\mathrm{f}}_s\right] / \left\{\phi_s \otimes \psi_{\gamma,s+1|T}[q \odot \gamma^{-1}_{s+1}]\right\}^2 ,$$

$$H^{\mathrm{f}}_s(x,x') \stackrel{\mathrm{def}}{=} h(x) q(x,x') \gamma^{-1}_{s+1}(x') .$$

$$\tilde{\Gamma}_{s,s+1|T}\left[h\right] \stackrel{\mathrm{def}}{=} \Gamma_s\left[\int \psi_{\gamma,s+1|T}(\mathrm{d}x_{s+1}) h(\cdot, x_{s+1})\right] + \check{\Gamma}_{\gamma,s+1|T}\left[\int \phi_s(\mathrm{d}x_s) h(x_s,\cdot)\right] .$$

Define

$$H^{\mathrm{f},N}_{s,1}(x') \stackrel{\mathrm{def}}{=} \Omega^{-1}_s \sum_{\ell=1}^{N} \omega^{\ell}_s \{h(\xi^{\ell}_s) - \phi^{\mathrm{f},N}_{s|T}[h]\} q(\xi^{\ell}_s, x') \gamma^{-1}_{s+1}(x') ,$$

$$H^{\mathrm{f},N}_{s+1,2}(x) \stackrel{\mathrm{def}}{=} \check{\Omega}^{-1}_{s+1|T} \sum_{j=1}^{N} \check{\omega}^{j}_{s+1|T} \{h(x) - \phi^{\mathrm{f},N}_{s|T}[h]\} q(x, \check{\xi}^{j}_{s+1|T}) \gamma^{-1}_{s+1}(\check{\xi}^{j}_{s+1|T}) ,$$

Conclusions and extensions

- Extensions of the theoretical properties of the usual smoothers to the two-filter algorithms (FFBS, FFBSi, PaRIS).

  $\Rightarrow$ Nonasymptotic deviation inequalities, CLT, L_q-mean error.

- Asymptotic variance easier to estimate than variance of Forward-Backward smoothers.

- Theoretical analysis of sensitivity to the choice of the artificial distribution.

- Analysis of Rao blackwellised extensions of two-filer algorithms to regime switching models (several regimes in commodity markets).