An introduction to Particle Approximate Bayesian Computation

P. Del Moral

UNSW, School of Mathematics and Statistics

ABC in Sydney, UNSW july 3rd 2014

Some hyper-refs

- The Monte-Carlo Method for filtering with discrete-time observations. with J. Jacod and Ph. Protter. (Purdue University, no.17-1998) & (PTRF-2001).
- The Monte-Carlo Method for filtering with discrete time observations. Central Limit Theorems. The Fields Institute Communications (2002).
- On parallel implementation of Sequential Monte Carlo methods: the island particle model, with C. Vergé, C. Dubarry, and E. Moulines. (Statistics and Computing-2013).
- On Feynman-Kac and particle Markov chain Monte Carlo models, with R. Kohn and F. Patras (ArXiv-2014).
- Feynman-Kac formulae, Genealogical & Interacting Particle Systems with appl., Springer (2004)
- Mean field simulation for Monte Carlo integration. Chapman - Hall (2013) [+ Refs]
- More references on website http://web.maths.unsw.edu.au/∼peterdel-moral/simulinks.html [+ Links]

ABC-Particle filters

Bayes' & Feynman-Kac path integration

ABC - Smoothing/Path-estimation

ABC - Island Particle models

ABC - PMCMC

ABC in Static/HMM inference

Some Self-tuning models

Origin [D-Jacod-Protter/Fields Inst.-02 + PTRF-01]

Nonlinear filtering model:

$$\begin{cases} dX_t &= a_t(X_t)dt + b_t(X_t)\ dW_t \\ dY_t &= a'_t(X_t, Y_t)dt + b'_t(X_t, Y_t)\ dW'_t \end{cases}$$

Origin [D-Jacod-Protter/Fields Inst.-02 + PTRF-01]

Nonlinear filtering model:

$$\begin{cases} dX_t &= a_t(X_t)dt + b_t(X_t)\ dW_t \\ dY_t &= a'_t(X_t, Y_t)dt + b'_t(X_t, Y_t)\ dW'_t \end{cases}$$

Filtering problem

$$\mathrm{Law}\left(X_{t_n} \mid (Y_{t_0}, \ldots, Y_{t_n})\right)$$

Origin [D-Jacod-Protter/Fields Inst.-02 + PTRF-01]

Nonlinear filtering model:

$$\left\{\begin{array}{lcl} dX_t & = & a_t(X_t)dt + b_t(X_t)\ dW_t \\ dY_t & = & a'_t(X_t, Y_t)dt + b'_t(X_t, Y_t)\ dW'_t \end{array}\right.$$

Filtering problem

$$\mathrm{Law}\left(X_{t_n} \mid (Y_{t_0}, \ldots, Y_{t_n})\right)$$

Main difficulty (using Particle Filters) = Intractable likelihoods

$$p_{t_n}(y_{t_n} \mid x_{t_n}, (y_{t_0}, \ldots, y_{t_{n-1}}))$$

Solution

ABC style model

$$\begin{cases} \mathcal{X}_t &= (X_t, Y_t) \\ \mathcal{Y}_{t_n} &= Y_{t_n} + \epsilon\, V_n \quad \text{with} \quad V_n \text{ i.i.d. } \sim g(v)\, dv \end{cases}$$

Solution

ABC style model

$$\left\{ \begin{array}{lcl} \mathcal{X}_t & = & (X_t, Y_t) \\ \mathcal{Y}_{t_n} & = & Y_{t_n} + \epsilon\, V_n \quad \text{with} \quad V_n \text{ i.i.d. } \sim g(v)\; dv \end{array} \right.$$

Single hypothesis:

$$\int v\; g(v)\; dv = 0 \qquad \int \|v\|^3\; g(v)\; dv < \infty$$

Solution

ABC style model

$$\left\{ \begin{array}{lcl} \mathcal{X}_t & = & (X_t, Y_t) \\ \mathcal{Y}_{t_n} & = & Y_{t_n} + \epsilon \; V_n \quad \text{with} \quad V_n \text{ i.i.d.} \sim g(v) \; dv \end{array} \right.$$

Single hypothesis:

$$\int v \; g(v) \; dv = 0 \qquad \int \|v\|^3 \; g(v) \; dv < \infty$$

Tractable ϵ-likelihoods $(y \in \mathbb{R}^d)$

$$p_{t_n}(\mathcal{Y}_{t_n} \mid \mathcal{X}_{t_n}, (\mathcal{Y}_{t_0}, \ldots, \mathcal{Y}_{t_{n-1}})) = \epsilon^{-d} \; g\left(\epsilon^{-1} \; (\mathcal{Y}_{t_n} - Y_{t_n})\right)$$

Solution

ABC style model

$$\begin{cases} \mathcal{X}_t &= (X_t, Y_t) \\ \mathcal{Y}_{t_n} &= Y_{t_n} + \epsilon\ V_n \quad \text{with} \quad V_n \text{ i.i.d. } \sim g(v)\ dv \end{cases}$$

Single hypothesis:

$$\int v\ g(v)\ dv = 0 \qquad \int \|v\|^3\ g(v)\ dv < \infty$$

Tractable ϵ-likelihoods $(y \in \mathbb{R}^d)$

$$p_{t_n}(\mathcal{Y}_{t_n} \mid \mathcal{X}_{t_n}, (\mathcal{Y}_{t_0}, \ldots, \mathcal{Y}_{t_{n-1}})) = \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right)$$

$$\Downarrow$$

ABC-Filter

$$\text{Law}\left(X_{t_n} \mid \mathcal{Y}_{t_p} = y_{t_p},\ p \leq n\right) \simeq_{\epsilon \downarrow 0} \text{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \leq n\right)$$

A couple of examples

$$g(v) = 1_{\|v\| \leq 1} \Rightarrow \epsilon^{-d}\, g\left(\epsilon^{-1}\, (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto 1_{\|\mathcal{Y}_{t_n} - Y_{t_n}\| \leq \epsilon}$$

A couple of examples

$$g(v) = 1_{\|v\| \leq 1} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto 1_{\|\mathcal{Y}_{t_n} - Y_{t_n}\| \leq \epsilon}$$

and

$$g(v) \overset{d=1}{\propto} e^{-\frac{v^2}{2}} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto e^{-\frac{(\mathcal{Y}_{t_n} - Y_{t_n})^2}{2\epsilon^2}}$$

A couple of examples

$$g(v) = 1_{\|v\| \leq 1} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto 1_{\|\mathcal{Y}_{t_n} - Y_{t_n}\| \leq \epsilon}$$

and

$$g(v) \overset{d=1}{\propto} e^{-\frac{v^2}{2}} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto e^{-\frac{(\mathcal{Y}_{t_n} - Y_{t_n})^2}{2\epsilon^2}}$$

the last avoids particle filters possible extinction

A couple of examples

$$g(v) = 1_{\|v\| \leq 1} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto 1_{\|\mathcal{Y}_{t_n} - Y_{t_n}\| \leq \epsilon}$$

and

$$g(v) \overset{d=1}{\propto} e^{-\frac{v^2}{2}} \Rightarrow \epsilon^{-d}\ g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \propto e^{-\frac{(\mathcal{Y}_{t_n} - Y_{t_n})^2}{2\epsilon^2}}$$

the last avoids particle filters possible extinction

$\oplus$ **many other density estimation kernels!**

ABC-Particle filters

Particle filters $\underline{\mathcal{X}_t^i = (X_t^i, Y_t^i)}$**-particles!!**

ABC-Particle filters

Particle filters $\mathcal{X}_t^i = (X_t^i, Y_t^i)$-particles!!

Functional Central Limit Theorem

$$\widehat{\mathrm{Law}}\,(X_{t_n} \mid (Y_{t_0}, \ldots, Y_{t_n})) := \frac{\mathbf{1}}{\mathbf{N}} \sum_{\mathbf{1 \leq i \leq N}} \delta_{\mathbf{X_t^i}}$$

$$=$$

$$\mathrm{Law}\left(X_{t_n} \mid \mathcal{Y}_{t_p} = y_{t_p},\ p \leq n\right) \ + \ \frac{\mathbf{1}}{\sqrt{\mathbf{N}}}\ \frac{1}{\epsilon^{\mathbf{d/2}}}\ \mathcal{W}_{t_n}^N$$

with

$$\mathcal{W}_{t_n}^N \to \mathcal{W}_{t_n} \ = \ \text{Centered Gaussian field} \ \left(\epsilon\textbf{-unif. } \mathbb{L}_2\textbf{-control}\right)$$

The bias/fluctuation

Taylor expansion - First order signed measure $\mathcal{D}_{\mathbf{t_n}}$

$$\mathrm{Law}\left(X_{t_n} \mid \mathcal{Y}_{t_p} = y_{t_p},\ p \leq n\right)$$

$$= \mathrm{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \leq n\right) + \epsilon^{\mathbf{2}}\ \mathcal{D}_{\mathbf{t_n}} + \mathrm{O}\left(\epsilon^3\right)$$

The bias/fluctuation

Taylor expansion - First order signed measure $\mathcal{D}_{\mathbf{t_n}}$

$$\mathrm{Law}\left(X_{t_n} \mid \mathcal{Y}_{t_p} = y_{t_p},\ p \le n\right)$$

$$= \mathrm{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \le n\right) + \epsilon^{\mathbf{2}}\ \mathcal{D}_{\mathbf{t_n}} + \mathrm{O}\left(\epsilon^3\right)$$

$$\frac{1}{N}\sum_{1\le i\le N}\ \delta_{X_t^i}$$

$$=$$

$$\mathrm{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \le n\right)\ +\ \frac{1}{\sqrt{\mathbf{N}}}\ \frac{1}{\epsilon^{\mathbf{d/2}}}\ \mathcal{W}_{t_n}^N + \epsilon^{\mathbf{2}}\ \mathcal{D}_{\mathbf{t_n}} + \mathrm{O}\left(\epsilon^3\right)$$

The bias/fluctuation

Taylor expansion - First order signed measure $\mathcal{D}_{\mathbf{t_n}}$

$$\mathrm{Law}\left(X_{t_n} \mid \mathcal{Y}_{t_p} = y_{t_p},\ p \leq n\right)$$

$$= \mathrm{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \leq n\right) + \epsilon^{\mathbf{2}}\ \mathcal{D}_{\mathbf{t_n}} + \mathrm{O}\left(\epsilon^3\right)$$

$$\frac{1}{N}\sum_{1\leq i\leq N}\ \delta_{X_t^i}$$

$$=$$

$$\mathrm{Law}\left(X_{t_n} \mid Y_{t_p} = y_{t_p},\ p \leq n\right)\ +\ \frac{1}{\sqrt{\mathbf{N}}}\ \frac{1}{\epsilon^{\mathbf{d/2}}}\ \mathcal{W}_{t_n}^N + \epsilon^{\mathbf{2}}\ \mathcal{D}_{\mathbf{t_n}} + \mathrm{O}\left(\epsilon^3\right)$$

$\Rightarrow$ **Theo: Optimization $\oplus$ CLT:**

$$\frac{1}{\sqrt{N}}\ \frac{1}{\epsilon^{d/2}} = \epsilon^2 \qquad \Rightarrow \qquad \epsilon(N) = N^{-\frac{1}{d+4}}$$

$$\epsilon(N)^{-2}\left(\frac{1}{\sqrt{N}}\ \frac{1}{\epsilon(N)^{d/2}}\ \mathcal{W}_{t_n}^N + \epsilon^2(N)\ \mathcal{D}_{\mathbf{t_n}}\right) \quad \rightarrow_{N\uparrow\infty} \quad \mathcal{W}_{t_n} + \mathcal{D}_n$$

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.
- $\uparrow$ Universal tools/easy to combine.
- ...

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.
- $\uparrow$ Universal tools/easy to combine.
- ...

Drawbacks:

- $\uparrow$ Sophisticated maths tools.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.
- $\uparrow$ Universal tools/easy to combine.
- ...

Drawbacks:

- $\uparrow$ Sophisticated maths tools.
- $\neq$ Notation/Language between disciplines.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.
- $\uparrow$ Universal tools/easy to combine.
- ...

Drawbacks:

- $\uparrow$ Sophisticated maths tools.
- $\neq$ Notation/Language between disciplines.
- $\downarrow$ Number of readers.

Abstraction - Math model

Advantages:

- Simplified/Universal perspective ($\oplus$ toy models).
- $\uparrow$ Range of application.
- $\uparrow$ Impact of theorems/cv results.
- $\downarrow$ Repetition/multiplication of "known"/existing/published results.
- $\uparrow$ Universal tools/easy to combine.
- ...

Drawbacks:

- $\uparrow$ Sophisticated maths tools.
- $\neq$ Notation/Language between disciplines.
- $\downarrow$ Number of readers.
- $\uparrow$ Sleepy audience in conferences ...

Bayes' & Feynman-Kac path integration

$$p((x_0,\ldots,x_n) \mid (y_0,\ldots,y_n)) \propto \underbrace{p((y_0,\ldots,y_n) \mid (x_0,\ldots,x_n))}_{\prod_{0\leq k\leq n} p(y_k|x_k) \leftarrow \text{likelihood functions } G_k(x_k)} \times p(x_0,\ldots,x_n)$$

Bayes' & Feynman-Kac path integration

$$p((x_0,\ldots,x_n) \mid (y_0,\ldots,y_n)) \propto \underbrace{p((y_0,\ldots,y_n) \mid (x_0,\ldots,x_n))}_{\prod_{0\leq k\leq n} p(y_k|x_k) \leftarrow \text{likelihood functions } G_k(x_k)} \times p(x_0,\ldots,x_n)$$

$$\Downarrow$$

Feynman-Kac models (∃¬!)/Path estimation/Smoothing :

$$G_n(x_n) := p(y_n|x_n) \quad \& \quad \mathbb{P}_n := \mathrm{Law}\,(X_0,\ldots,X_n)$$

$$\Downarrow$$

$$\mathbb{P}\left((X_0,\ldots,X_n) \in d(x_0,\ldots,x_n) \mid Y_p = y_p,\ p < n\right)$$

$$= \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0\leq p<n} G_p(x_p) \right\} \mathbb{P}_n(d(x_0,\ldots,x_n))$$

Bayes' & Feynman-Kac path integration

$$p((x_0,\ldots,x_n) \mid (y_0,\ldots,y_n)) \propto \underbrace{p((y_0,\ldots,y_n) \mid (x_0,\ldots,x_n))}_{\prod_{0\leq k\leq n} p(y_k|x_k) \leftarrow \text{likelihood functions } G_k(x_k)} \times p(x_0,\ldots,x_n)$$

$$\Downarrow$$

Feynman-Kac models (∃¬!)/Path estimation/Smoothing :

$$G_n(x_n) := p(y_n|x_n) \quad \& \quad \mathbb{P}_n := \mathrm{Law}\,(X_0,\ldots,X_n)$$

$$\Downarrow$$

$$\mathbb{P}\,((X_0,\ldots,X_n) \in d(x_0,\ldots,x_n) \mid Y_p = y_p,\ p < n)$$

$$= \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0\leq p<n} G_p(x_p) \right\} \mathbb{P}_n(d(x_0,\ldots,x_n))$$

and the normalizing constant: $\mathcal{Z}_n = p(y_0,\ldots,y_n)$.

2 key formulae

$$\mathbb{Q}_n(d(x_0,\ldots,x_n)) = \frac{1}{\mathcal{Z}_n}\left\{\prod_{0\leq p<n} G_p(x_p)\right\}\ \mathbb{P}_n(d(x_0,\ldots,x_n))$$

Notation

$$\eta_n := n\text{-th marginal}\,(= dp(x_n \mid y_k,\ k<n))$$

2 key formulae

$$\mathbb{Q}_n(d(x_0,\ldots,x_n)) = \frac{1}{\mathcal{Z}_n}\left\{\prod_{0\leq p<n} G_p(x_p)\right\}\ \mathbb{P}_n(d(x_0,\ldots,x_n))$$

Notation

$$\eta_n := n\text{-th marginal}\,(= dp(x_n \mid y_k,\ k<n))$$

$$\Downarrow$$

1) Product formulae/Particle approximation

$$\mathcal{Z}_n \quad = \quad \prod_{0\leq k<n} \eta_k(G_k) \quad \text{with} \quad \eta_k(G_k) = \int G_k(x_k)\ \eta_k(dx_k)$$

2 key formulae

$$\mathbb{Q}_n(d(x_0,\ldots,x_n)) = \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0\leq p<n} G_p(x_p) \right\} \mathbb{P}_n(d(x_0,\ldots,x_n))$$

Notation

$$\eta_n := n\text{-th marginal}\,(= dp(x_n \mid y_k,\ k<n))$$

$$\Downarrow$$

1) Product formulae/Particle approximation

$$\mathcal{Z}_n = \prod_{0\leq k<n} \eta_k(G_k) \quad \text{with} \quad \eta_k(G_k) = \int G_k(x_k)\,\eta_k(dx_k)$$

$$\stackrel{\text{unbias}}{=} \prod_{0\leq k<n} \frac{1}{N}\sum_{1\leq i\leq N} G_k(X_k^i) \quad \text{if} \quad \eta_k = \frac{1}{N}\sum_{1\leq i\leq N} \delta_{X_k^i}$$

2 key formulae

Hypothesis

$$M_{k+1}(x_k, dx_{k+1}) = H_{k+1}(x_k, x_{k+1})\ \lambda(dx_{k+1}) \overset{ex.}{\propto}\ e^{-\frac{1}{2}(x_{k+1}-a(x_k))^2}\ dx_{k+1}$$

$$\Downarrow$$

2) Backward formulae/Backward Particle chain

$$\mathbb{Q}_n(d(x_0,\ldots,x_n)) = \eta_n(dx_n)\ \mathbb{K}_{n,\eta_{n-1}}(x_n, dx_{n-1})\ \ldots\ \mathbb{K}_{1,\eta_0}(x_1, dx_0)$$

with

$$\mathbb{K}_{k+1,\eta_k}(x_{k+1}, dx_k)$$

$$= \frac{\eta_k(dx_k)\ G_k(x_k)H(x_k, x_{k+1})}{\int\ \eta_k(dx'_k)\ G_k(x'_k)H(x'_k, x_{k+1})}$$

$$= \sum_{1\leq i\leq N} \frac{G_k(X^i_k)H(X^i_k, x_{k+1})}{\sum_{1\leq j\leq N}\ G_k(X^j_k)H(X^j_k, x_{k+1})}\ \delta_{X^i_k}(dx_k) \quad \text{if} \quad \eta_k = \frac{1}{N}\sum_{1\leq i\leq N}\delta_{X^i_k}$$

ABC-Smoothing/Path-estimation

Exchange/Transfer rule

$$\begin{array}{rcccl} X_n & \rightsquigarrow & \mathcal{X}_n & := & (X_{t_n}, Y_{t_n}) \rightsquigarrow \text{ pair-particles } \mathcal{X}^i_n = \left(X^i_{t_n}, Y^i_{t_n}\right) \\ G_n(x_n) & \rightsquigarrow & \mathcal{G}_n(\mathcal{X}_n) & := & \epsilon^{-d}\, g\left(\epsilon^{-1}\, (\mathcal{Y}_{t_n} - Y_{t_n})\right) \end{array}$$

ABC-Smoothing/Path-estimation

Exchange/Transfer rule

$$\begin{array}{rcccl} X_n & \rightsquigarrow & \mathcal{X}_n & := & (X_{t_n}, Y_{t_n}) \rightsquigarrow \text{ pair-particles } \mathcal{X}_n^i = \left(X_{t_n}^i, Y_{t_n}^i\right) \\ G_n(x_n) & \rightsquigarrow & \mathcal{G}_n(\mathcal{X}_n) & := & \epsilon^{-d}\, g\left(\epsilon^{-1}\ (\mathcal{Y}_{t_n} - Y_{t_n})\right) \end{array}$$

$$\Downarrow$$

$$\mathbb{Q}_n = \text{Law}\left(\mathcal{X}_0, \ldots, \mathcal{X}_n \mid \mathcal{Y}_p = y_{t_p},\ p < n\right)$$

ABC-Smoothing/Path-estimation

Exchange/Transfer rule

$$\begin{array}{ccccc} X_n & \rightsquigarrow & \mathcal{X}_n & := & (X_{t_n}, Y_{t_n}) \rightsquigarrow \text{ pair-particles } \mathcal{X}_n^i = \left(X_{t_n}^i, Y_{t_n}^i\right) \\ G_n(x_n) & \rightsquigarrow & \mathcal{G}_n(\mathcal{X}_n) & := & \epsilon^{-d}\, g\left(\epsilon^{-1}\, (\mathcal{Y}_{t_n} - Y_{t_n})\right) \end{array}$$

$$\Downarrow$$

$$\mathbb{Q}_n = \text{Law}\left(\mathcal{X}_0, \ldots, \mathcal{X}_n \mid \mathcal{Y}_p = y_{t_p},\ p < n\right)$$

1) Product formulae/ABC-Particle approximation

$$\mathcal{Z}_n \stackrel{\text{N}-\text{unbias}}{=} \prod_{0 \leq k < n} \frac{1}{N} \sum_{1 \leq i \leq N} \epsilon^{-d}\, g\left(\epsilon^{-1}\, \left(\mathcal{Y}_{t_n}^i - y_{t_n}\right)\right)$$

ABC-Smoothing/Path-estimation

Exchange/Transfer rule

$$\begin{array}{rcccl} X_n & \rightsquigarrow & \mathcal{X}_n & := & (X_{t_n}, Y_{t_n}) \rightsquigarrow \text{ pair-particles } \mathcal{X}_n^i = \left(X_{t_n}^i, Y_{t_n}^i\right) \\ G_n(x_n) & \rightsquigarrow & \mathcal{G}_n(\mathcal{X}_n) & := & \epsilon^{-d}\; g\left(\epsilon^{-1}\; (\mathcal{Y}_{t_n} - Y_{t_n})\right) \end{array}$$

$$\Downarrow$$

$$\mathbb{Q}_n = \text{Law}\left(\mathcal{X}_0, \ldots, \mathcal{X}_n \mid \mathcal{Y}_p = y_{t_p},\; p < n\right)$$

1) Product formulae/ABC-Particle approximation

$$\mathcal{Z}_n \overset{\mathbf{N}-\text{unbias}}{=} \prod_{0 \leq k < n} \frac{1}{N} \sum_{1 \leq i \leq N} \epsilon^{-d}\; g\left(\epsilon^{-1}\; \left(\mathcal{Y}_{t_n}^i - y_{t_n}\right)\right)$$

2) Backward formulae/Backward Particle chain
(When the transition density of $\mathcal{X}_n$ is known!!)

$$\mathbb{Q}_n^{\mathbf{N}}(d(\mathcal{X}_0, \ldots, \mathcal{X}_n)) = \eta_n^{\mathbf{N}}(d\mathcal{X}_n)\; \mathbb{K}_{n,\eta_{n-1}^{\mathbf{N}}}(\mathcal{X}_n, d\mathcal{X}_{n-1})\; \ldots\; \mathbb{K}_{1,\eta_0^{\mathbf{N}}}(\mathcal{X}_1, d\mathcal{X}_0)$$

Unbias property for Feynman-Kac models

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n}G_k(X_k)\right)=\mathbb{E}\left(\eta_n^N(f)\prod_{0\leq k<n}\eta_k^N(G_k)\right)$$

with

$$\eta_n^N(f)=\frac{1}{N}\sum_{1\leq i\leq N}f(X_n^i)$$

Unbias property for Feynman-Kac models

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\eta_n^N(f)\prod_{0\leq k<n} \eta_k^N(G_k)\right)$$

with

$$\eta_n^N(f) = \frac{1}{N}\sum_{1\leq i\leq N} f(X_n^i) := \mathbf{f}(\mathbf{X_n}) \quad \text{with} \quad \mathbf{X_n} = (X_n^i)_{1\leq i\leq N}$$

Unbias property for Feynman-Kac models

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\eta_n^N(f)\prod_{0\leq k<n} \eta_k^N(G_k)\right)$$

with

$$\eta_n^N(f) = \frac{1}{N}\sum_{1\leq i\leq N} f(X_n^i) := \mathbf{f}(\mathbf{X_n}) \quad \text{with} \quad \mathbf{X_n} = (X_n^i)_{1\leq i\leq N}$$

$$\Downarrow$$

Many-body Feynman-Kac model (cf. MPRF-1996/Arxiv-2014)

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k})\right)$$

Unbias property for Feynman-Kac models

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\eta_n^N(f)\prod_{0\leq k<n}\eta_k^N(G_k)\right)$$

with

$$\eta_n^N(f) = \frac{1}{N}\sum_{1\leq i\leq N} f(X_n^i) := \mathbf{f}(\mathbf{X_n}) \quad \text{with} \quad \mathbf{X_n} = (X_n^i)_{1\leq i\leq N}$$

$$\Downarrow$$

Many-body Feynman-Kac model (cf. MPRF-1996/Arxiv-2014)

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n}\mathbf{G_k}(\mathbf{X_k})\right)$$

$$\Downarrow$$

Particle Filters $\rightsquigarrow$ Island Particle Filters (SC-13) $\oplus$ ABC-Island PF ($X \rightsquigarrow \mathcal{X}$)

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k})\right)$$

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k})\right)$$

with

$$\widehat{\mathcal{Z}}_n(\mathbf{X}) := \prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k}) = \prod_{0\leq k<n} \frac{1}{N}\sum_{1\leq i\leq N} G_k(X_k^i) \simeq_{N\uparrow\infty} \mathcal{Z}_n$$

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k})\right)$$

with

$$\widehat{\mathcal{Z}}_n(\mathbf{X}) := \prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k}) = \prod_{0\leq k<n} \frac{1}{N}\sum_{1\leq i\leq N} G_k(X_k^i) \simeq_{N\uparrow\infty} \mathcal{Z}_n$$

Independent Metropolis-Hasting model (Particle filters proposals)

$$(\mathbf{X_0},\ldots,\mathbf{X_n}) \rightsquigarrow (\mathbf{X'_0},\ldots,\mathbf{X'_n}) \quad \text{with} \quad \text{acceptance rate} \quad 1\wedge\frac{\widehat{\mathcal{Z}}_n(\mathbf{X})}{\widehat{\mathcal{Z}}_n(\mathbf{X'})}$$

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n} G_k(X_k)\right) = \mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k})\right)$$

with

$$\widehat{\mathcal{Z}}_n(\mathbf{X}) := \prod_{0\leq k<n} \mathbf{G_k}(\mathbf{X_k}) = \prod_{0\leq k<n} \frac{1}{N}\sum_{1\leq i\leq N} G_k(X_k^i) \simeq_{N\uparrow\infty} \mathcal{Z}_n$$

Independent Metropolis-Hasting model (Particle filters proposals)

$$(\mathbf{X_0},\ldots,\mathbf{X_n}) \rightsquigarrow (\mathbf{X'_0},\ldots,\mathbf{X'_n}) \quad \text{with} \quad \text{acceptance rate} \quad 1\wedge \frac{\widehat{\mathcal{Z}}_n(\mathbf{X})}{\widehat{\mathcal{Z}}_n(\mathbf{X'})}$$

- **ABC-PMH**
  $X_n \rightsquigarrow \mathcal{X}_n = (X_n, Y_n) \rightsquigarrow$ **ABC-Particle Filters proposals**

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n}G_k(X_k)\right)=\mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n}\mathbf{G_k}(\mathbf{X_k})\right)$$

with

$$\widehat{\mathcal{Z}}_n(\mathbf{X}):=\prod_{0\leq k<n}\mathbf{G_k}(\mathbf{X_k})=\prod_{0\leq k<n}\frac{1}{N}\sum_{1\leq i\leq N}G_k(X_k^i)\simeq_{N\uparrow\infty}\mathcal{Z}_n$$

Independent Metropolis-Hasting model (Particle filters proposals)

$$(\mathbf{X_0},\ldots,\mathbf{X_n})\rightsquigarrow(\mathbf{X'_0},\ldots,\mathbf{X'_n})\quad\text{with}\quad\text{acceptance rate}\quad 1\wedge\frac{\widehat{\mathcal{Z}}_n(\mathbf{X})}{\widehat{\mathcal{Z}}_n(\mathbf{X'})}$$

- **ABC-PMH**

  $X_n\rightsquigarrow\mathcal{X}_n=(X_n,Y_n)\rightsquigarrow$ **ABC-Particle Filters proposals**
- **ABC-Particle Gibbs** *(When the transition density of $\mathcal{X}_n$ is known!!)*

Target = Normalized many-body Feynman-Kac model

$$\mathbb{E}\left(f(X_n)\prod_{0\leq k<n}G_k(X_k)\right)=\mathbb{E}\left(\mathbf{f}(\mathbf{X_n})\prod_{0\leq k<n}\mathbf{G_k}(\mathbf{X_k})\right)$$

with

$$\widehat{\mathcal{Z}}_n(\mathbf{X}):=\prod_{0\leq k<n}\mathbf{G_k}(\mathbf{X_k})=\prod_{0\leq k<n}\frac{1}{N}\sum_{1\leq i\leq N}G_k(X_k^i)\simeq_{N\uparrow\infty}\mathcal{Z}_n$$

Independent Metropolis-Hasting model (Particle filters proposals)

$$(\mathbf{X_0},\ldots,\mathbf{X_n})\rightsquigarrow(\mathbf{X}'_0,\ldots,\mathbf{X}'_n)\quad\text{with}\quad\text{acceptance rate}\quad 1\wedge\frac{\widehat{\mathcal{Z}}_n(\mathbf{X})}{\widehat{\mathcal{Z}}_n(\mathbf{X}')}$$

- **ABC-PMH**
  $X_n\rightsquigarrow\mathcal{X}_n=(X_n,Y_n)\rightsquigarrow$ **ABC-Particle Filters proposals**
- **ABC-Particle Gibbs** *(When the transition density of $\mathcal{X}_n$ is known!!)*
- **Trivial extension(s) to fixed parameter estimation in HMM.**

$$p(y\mid\theta)\simeq\widehat{\mathcal{Z}}_{n,\theta}(\mathbf{X})$$

ABC in Static/HMM inference

$$p(\theta \mid y) \;\propto\; p(y \mid \theta)\, p(\theta)$$

ABC in Static/HMM inference

$$p(\theta \mid y) \;\propto\; p(y \mid \theta)\; p(\theta)$$

Case 1: Direct observations / $p(y \mid \theta)$ known:

$$y = (y_1, \ldots, y_n) \rightsquigarrow p(y \mid \theta) \;=\; \prod_{1 \le k \le n} p(y_k \mid \theta)$$

ABC in Static/HMM inference

$$p(\theta \mid y) \;\propto\; p(y \mid \theta)\; p(\theta)$$

Case 1: Direct observations / $p(y \mid \theta)$ known:

$$y = (y_1, \ldots, y_n) \rightsquigarrow p(y \mid \theta) \;=\; \prod_{1 \leq k \leq n} p(y_k \mid \theta)$$

$\rightsquigarrow$ **Product target measure**

$$\pi(d\theta) \propto \left\{ \prod_{1 \leq k \leq n} h_k(\theta) \right\} \lambda(d\theta)$$

with

$$\lambda(d\theta) = p(\theta)\; d\theta \quad \text{and} \quad h_k(\theta) = p(y_k \mid \theta)$$

ABC version of $p(\theta \mid y) \propto p(y \mid \theta)\, p(\theta)$

$$p((\theta, y) \mid \mathcal{Y}) \quad \propto \quad g(\epsilon^{-1}\,(\mathcal{Y} - y)) \times [p(y \mid \theta)\; p(\theta)]$$

ABC version of $p(\theta \mid y) \propto p(y \mid \theta)\, p(\theta)$

$$p((\theta, y) \mid \mathcal{Y}) \quad \propto \quad g(\epsilon^{-1}\,(\mathcal{Y} - y)) \times [p(y \mid \theta)\; p(\theta)]$$

Case 1′: $g(\epsilon^{-1}\,(\mathcal{Y} - y))$ known:

$$y = (y_k)_k \;\&\; \epsilon = (\epsilon_k)_k \;\rightsquigarrow g(\epsilon^{-1}\,(\mathcal{Y} - y)) \;=\; \prod_{1 \le k \le n} g(\epsilon_k^{-1}\,(\mathcal{Y}_k - y_k))$$

ABC version of $p(\theta \mid y) \propto p(y \mid \theta)\, p(\theta)$

$$p((\theta, y) \mid \mathcal{Y}) \quad \propto \quad g(\epsilon^{-1}\, (\mathcal{Y} - y)) \times [p(y \mid \theta)\, p(\theta)]$$

Case 1′: $g(\epsilon^{-1}\, (\mathcal{Y} - y))$ known:

$$y = (y_k)_k \;\&\; \epsilon = (\epsilon_k)_k \;\rightsquigarrow g(\epsilon^{-1}\, (\mathcal{Y} - y)) \;=\; \prod_{1 \leq k \leq n} g(\epsilon_k^{-1}\, (\mathcal{Y}_k - y_k))$$

$\rightsquigarrow$ **Product target measure**

$$\pi(d\overline{\theta}) \propto \left\{ \prod_{1 \leq k \leq n} \overline{h}_k(\overline{\theta}) \right\} \overline{\lambda}(d\overline{\theta})$$

with $\overline{\theta} = (\theta, y)$ and

$$\overline{\lambda}(d\overline{\theta}) = p(y \mid \theta)\, p(\theta)\, d\theta dy \quad \text{and} \quad \overline{h}_k(\overline{\theta}) = g(\epsilon_k^{-1}\, (\mathcal{Y}_k - y_k))$$

HMM models $\oplus$ ξ-Particle filters

Unbias property

$$p((\theta \mid y) \; \propto \; \stackrel{\theta-\text{marginal}}{\longleftarrow} \; \left\{ \prod_{1 \leq k \leq n} \frac{1}{N} \sum_{1 \leq i \leq N} p(y_k \; |\xi_k^i \; \theta) \right\} \; p(\xi \mid y \; \theta) \; p(\theta)$$

HMM models $\oplus$ ξ-Particle filters

Unbias property

$$p((\theta \mid y) \propto \overset{\theta-\text{marginal}}{\longleftarrow} \left\{ \prod_{1\le k\le n} \frac{1}{N} \sum_{1\le i\le N} p(y_k \mid \xi_k^i \; \theta) \right\} \; p(\xi \mid y \; \theta) \; p(\theta)$$

Case 1″: $\rightsquigarrow$ $\rightsquigarrow$ **Product target measure**

$$\pi(d\overline{\theta}) \propto \left\{ \prod_{1\le k\le n} \overline{h}_k(\overline{\theta}) \right\} \overline{\lambda}(d\overline{\theta})$$

with $\overline{\theta} = (\theta, \xi)$ and

$$\overline{\lambda}(d\overline{\theta}) = p(\xi \mid y \; \theta) \; p(\theta) \; d\theta d\xi \quad \text{and} \quad \overline{h}_k(\overline{\theta}) = \frac{1}{N} \sum_{1\le i\le N} p(y_k \mid \xi_k^i \; \theta)$$

ABC - HMM models $\oplus\ \xi = (X^i, Y^i)$-Particle filters

Unbias property

$$p(\theta \mid \mathcal{Y}) \propto \overset{\theta-\text{marginal}}{\longleftarrow} \left\{ \prod_{1 \le k \le n} \frac{1}{N} \sum_{1 \le i \le N} g(\epsilon_k^{-1} (\mathcal{Y}_k - Y_k^i)) \right\} p(\xi \mid \theta, \mathcal{Y})\, p(\theta)$$

ABC - HMM models $\oplus\ \xi = (X^i, Y^i)$-Particle filters

Unbias property

$$p(\theta \mid \mathcal{Y}) \propto \overset{\theta-\text{marginal}}{\longleftarrow} \left\{ \prod_{1\leq k\leq n} \frac{1}{N} \sum_{1\leq i\leq N} g(\epsilon_k^{-1}\,(\mathcal{Y}_k - Y_k^i)) \right\} p(\xi \mid \theta, \mathcal{Y})\, p(\theta)$$

Case 1‴: ⇝ Product target measure

$$\pi(d\overline{\theta}) \propto \left\{ \prod_{1\leq k\leq n} \overline{h}_k(\overline{\theta}) \right\} \overline{\lambda}(d\overline{\theta})$$

with $\overline{\theta} = (\theta, \xi)$ and

$$\overline{\lambda}(d\overline{\theta}) = p(\xi \mid \mathcal{Y}\ \theta)\, p(\theta)\, d\theta d\xi \quad \text{and} \quad \overline{h}_k(\overline{\theta}) = \frac{1}{N} \sum_{1\leq i\leq N} g(\epsilon_k^{-1}\,(\mathcal{Y}_k - Y_k^i))$$

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$ or Nb of observations,...

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$ or Nb of observations,...

As the filtering equation:

$$\pi_n \xrightarrow{\text{Correction/Updating}} d\pi_{n+1} \propto h_n \, d\pi_n \xrightarrow{\pi_{n+1}\text{-MCMC/Prediction}} \pi_{n+1}$$

$$\Downarrow$$

$\subset$ n-th marginals of a Feynman-Kac model

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1\leq k<n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$ or Nb of observations,...

As the filtering equation:

$$\pi_n \xrightarrow{\text{Correction/Updating}} d\pi_{n+1} \propto h_n \, d\pi_n \xrightarrow{\pi_{n+1}\text{-MCMC/Prediction}} \pi_{n+1}$$

$$\Downarrow$$

$\subset$ n-th marginals of a Feynman-Kac model

(Physics $\rightsquigarrow$ termed Crook/Jarzinsky formula;

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$ or Nb of observations,...

As the filtering equation:

$$\pi_n \xrightarrow{\text{Correction/Updating}} d\pi_{n+1} \propto h_n \, d\pi_n \xrightarrow{\pi_{n+1}\text{-MCMC/Prediction}} \pi_{n+1}$$

$$\Downarrow$$

$\subset$ n-th marginals of a Feynman-Kac model
(Physics $\rightsquigarrow$ termed Crook/Jarzinsky formula; Rare event $\rightsquigarrow$ subset sampling/multi-level splitting;

Product target measures

Interpolating path measures $\pi_0 \rightsquigarrow \ldots \rightsquigarrow \pi_T$

$$\pi_n(d\theta) \propto \left\{ \prod_{1 \leq k < n} h_k(\theta) \right\} \lambda(d\theta)$$

Tuning parameter: $\prod$ w.r.t. dimensions or $\epsilon \downarrow$ or Nb of observations,...

As the filtering equation:

$$\pi_n \xrightarrow{\text{Correction/Updating}} d\pi_{n+1} \propto h_n \; d\pi_n \xrightarrow{\pi_{n+1}\text{-MCMC/Prediction}} \pi_{n+1}$$

$$\Downarrow$$

$\subset$ n-th marginals of a Feynman-Kac model
(Physics $\rightsquigarrow$ termed Crook/Jarzinsky formula; Rare event $\rightsquigarrow$ subset sampling/multi-level splitting; Operation Research $\rightsquigarrow$ Interacting simulated annealing ...)

Acceptance probabilities ($\oplus$ recycling/resampling)

$$h_k = e^{-(\beta_{k+1}-\beta_k)\ V} \quad \overset{\beta_k \uparrow}{\Longrightarrow} \quad \pi_n(d\theta) \propto e^{-\beta_n V(\theta)}\ \lambda(d\theta)$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$h_k = e^{-(\beta_{k+1}-\beta_k)\ V} \quad \overset{\beta_k\uparrow}{\Longrightarrow} \quad \pi_n(d\theta) \propto e^{-\beta_n V(\theta)}\ \lambda(d\theta)$$

$$h_k = 1_{A_{k+1}} \quad \overset{A_k\downarrow}{\Longrightarrow} \quad \pi_n(d\theta) \propto 1_{A_n}(\theta)\ \lambda(d\theta)$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$
\begin{array}{rcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} \quad \stackrel{\beta_k \uparrow}{\Longrightarrow} \quad \pi_n(d\theta) \ \propto \ e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
h_k & = & 1_{A_{k+1}} \quad \stackrel{A_k \downarrow}{\Longrightarrow} \quad \pi_n(d\theta) \ \propto \ 1_{A_n}(\theta)\ \lambda(d\theta) \\
h_k(\theta) & = & \prod_{m_k < l \leq m_{k+1}} p(y_l \mid \theta) \quad \stackrel{m_k \uparrow}{\Longrightarrow} \quad \pi_n(d\theta) \ \propto \ p(\theta \mid y_0, \ldots, y_{m_n})
\end{array}
$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$
\begin{array}{rclcrcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} & \overset{\beta_k \uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
h_k & = & 1_{A_{k+1}} & \overset{A_k \downarrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & 1_{A_n}(\theta)\ \lambda(d\theta) \\
h_k(\theta) & = & \prod_{m_k < l \leq m_{k+1}} p(y_l \mid \theta) & \overset{m_k \uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & p(\theta \mid y_0, \ldots, y_{m_n})
\end{array}
$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$\begin{array}{rclcrcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} & \stackrel{\beta_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
h_k & = & 1_{A_{k+1}} & \stackrel{A_k\downarrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & 1_{A_n}(\theta)\ \lambda(d\theta) \\
h_k(\theta) & = & \prod_{m_k<l\leq m_{k+1}} p(y_l \mid \theta) & \stackrel{m_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & p(\theta \mid y_0, \ldots, y_{m_n})
\end{array}$$

At each step (after the shaking transition)

Acceptance probabilities ($\oplus$ recycling/resampling)

$$
\begin{array}{rclcrcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} & \stackrel{\beta_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
h_k & = & 1_{A_{k+1}} & \stackrel{A_k\downarrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & 1_{A_n}(\theta)\ \lambda(d\theta) \\
h_k(\theta) & = & \prod_{m_k<l\leq m_{k+1}} p(y_l \mid \theta) & \stackrel{m_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & p(\theta \mid y_0,\ldots,y_{m_n})
\end{array}
$$

At each step (after the shaking transition)

$$
\beta_{n+1} \in [\beta_n, \infty[\quad \text{s.t.} \quad \frac{1}{N}\sum_{1\leq i\leq N} e^{-(\beta_{k+1}-\beta_k)\ V(\theta_k^i)} \stackrel{\text{ex.}}{\simeq} 70\%
$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$
\begin{array}{rclcrcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} & \stackrel{\beta_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
h_k & = & 1_{A_{k+1}} & \stackrel{A_k\downarrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & 1_{A_n}(\theta)\ \lambda(d\theta) \\
h_k(\theta) & = & \prod_{m_k<l\leq m_{k+1}} p(y_l \mid \theta) & \stackrel{m_k\uparrow}{\Longrightarrow} & \pi_n(d\theta) & \propto & p(\theta \mid y_0,\ldots,y_{m_n})
\end{array}
$$

At each step (after the shaking transition)

$$\beta_{n+1} \in [\beta_n, \infty[\quad \text{s.t.} \quad \frac{1}{N}\sum_{1\leq i\leq N} e^{-(\beta_{k+1}-\beta_k)\ V(\theta_k^i)} \stackrel{\text{ex.}}{\simeq} 70\%$$

or when $A_k = \{\theta\ :\ V(\theta) \leq \epsilon_k\}$

$$\epsilon_{n+1} \in [0, \epsilon_n] \quad \text{s.t.} \quad \frac{1}{N}\sum_{1\leq i\leq N} 1_{V(\theta_k^i)\leq\epsilon_{k+1}} \stackrel{\text{ex.}}{\simeq} 70\%$$

Acceptance probabilities ($\oplus$ recycling/resampling)

$$
\begin{array}{rcl}
h_k & = & e^{-(\beta_{k+1}-\beta_k)\ V} \\
h_k & = & 1_{A_{k+1}} \\
h_k(\theta) & = & \prod_{m_k<l\leq m_{k+1}} p(y_l \mid \theta)
\end{array}
\quad
\begin{array}{c}
\overset{\beta_k\uparrow}{\Longrightarrow} \\
\overset{A_k\downarrow}{\Longrightarrow} \\
\overset{m_k\uparrow}{\Longrightarrow}
\end{array}
\quad
\begin{array}{rcl}
\pi_n(d\theta) & \propto & e^{-\beta_n V(\theta)}\ \lambda(d\theta) \\
\pi_n(d\theta) & \propto & 1_{A_n}(\theta)\ \lambda(d\theta) \\
\pi_n(d\theta) & \propto & p(\theta \mid y_0, \ldots, y_{m_n})
\end{array}
$$

At each step (after the shaking transition)

$$\beta_{n+1} \in [\beta_n, \infty[\quad \text{s.t.} \quad \frac{1}{N} \sum_{1\leq i\leq N} e^{-(\beta_{k+1}-\beta_k)\ V(\theta_k^i)} \overset{\text{ex.}}{\simeq} 70\%$$

or when $A_k = \{\theta \ : \ V(\theta) \leq \epsilon_k\}$

$$\epsilon_{n+1} \in [0, \epsilon_n] \quad \text{s.t.} \quad \frac{1}{N} \sum_{1\leq i\leq N} 1_{V(\theta_k^i)\leq \epsilon_{k+1}} \overset{\text{ex.}}{\simeq} 70\%$$

or

$$m_{n+1} \in [m_n, \infty[\quad \text{s.t.} \quad \frac{1}{N} \sum_{1\leq i\leq N} \prod_{m_k<l\leq m_{k+1}} p(y_l \mid \theta_l^i) \overset{\text{ex.}}{\simeq} 70\%$$

Updating/Adaptive criteria

Keep taking products until the weight is too degenerate

- Variance $\simeq$ Effective sample size

Updating/Adaptive criteria

Keep taking products until the weight is too degenerate

- Variance $\simeq$ Effective sample size
- Relative entropy weighted w.r.t uniform

Updating/Adaptive criteria

Keep taking products until the weight is too degenerate

- Variance $\simeq$ Effective sample size
- Relative entropy weighted w.r.t uniform
- Proportion of success/acceptance

Updating/Adaptive criteria

Keep taking products until the weight is too degenerate

- Variance $\simeq$ Effective sample size
- Relative entropy weighted w.r.t uniform
- Proportion of success/acceptance
- ...

Some Ref. = Modeling/Optimization level-sets $\oplus$ CLT

- **Genealogical Models in Entrance Times Rare Event Analysis**, DM+Cerou, Le Gland, Lezaud. Alea, Latin American Journal of Probability And Mathematical Statistics (2006).
- **Sequential Monte Carlo for Rare event estimation**. DM+Cerou, Furon, Guyader. (HAL-INRIA RR-6792 (2009)) Statistics and Computing (2011).
- **On Adaptive Resampling Procedures for Sequential Monte Carlo Methods.** DM+Doucet+Jasra. (HAL-INRIA RR-6700/2008). Bernoulli, Vol. 18, No. 1, pp. 252-278 (2012).
- **An Adaptive Sequential Monte Carlo Method for Approximate Bayesian Computation** DM+Doucet+Jasra. Statistics and Computing (online 2011).