Island particle

C. Dubarry, P. Del Moral, E. Moulines, C. Vergé

Institut Mines-Télécom, Télécom ParisTech/ Télécom SudParis, INRIA Bordeaux, ONERA

IHP, May 15, 2013

Outline

Outline

Notations

- $(\mathbb{E}_n, \mathcal{E}_n)_{n\geq 0}$: a sequence of measurable sets.
- $(X_n)_{n\geq 0}$: a non-homogenous Markov chain with initial distribution η_0, and Markov kernels $(M_n)_{n\geq 1}$.
- $(g_n)_{n\geq 0}$: a sequence of potential functions, $g_n : \mathbb{E}_n \mapsto \mathbb{R}^+$
- The Feynman-Kac flow associated to $(M_n, g_n)_{n\geq 0}$ is defined by

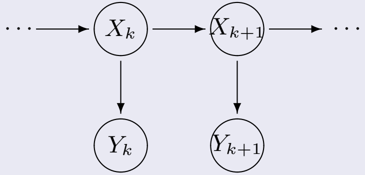
$$\eta_n(f_n) \stackrel{\text{def}}{=} \gamma_n(f_n)/\gamma_n(1) \ ,$$

$$\gamma_n(f_n) \stackrel{\text{def}}{=} \mathbb{E}\left[f_n(X_n) \prod_{0\leq p<n} g_p(X_p) \right] \ .$$

Nonlinear State-Space models or HMM

An HMM $(X_k, Y_k)_{k\geq 0}$ is a Markov process such that the conditional distribution of (X_k, Y_k) given $(X_i, Y_i)_{0\leq i\leq k-1}$ only depends on X_{k-1}.

Bayesian Network (Directed Graphical Model) Representation



Transition Kernel/ Likelihood

- The state dynamic is represented by a transition kernel, M_k, defined by

$$\mathbb{P}(X_{k+1} \in A \mid X_k = x_k) = M_k(x_k, A) \ .$$

- The measurement equation is specified by

$$\mathbb{P}(Y_k \in A \mid X_k = x_k) = G_k(x_k, A) = \int_A g(x_k, y)\nu(\mathrm{d}y) \ ,$$

- Filtering relates to the task of inferring the state X_k from the observations $Y_{0:k} = Y_0, \dots, Y_k$.
- The filtering distribution is an example of Feynman-Kac model with transition M_k [the prior transition] and potential $g_k(x) = g(x, Y_k)$, the likelihood of the observations.

Stochastic Optimization

1. Problem: Find $x_* = \max_{x \in \mathbb{X}} V(x)$ where V is a function on $\mathbb{X}$.
2. Let $(\beta_n)_{n \geq 0}$ be a nondecreasing sequence of positive numbers such that $\lim_{n \to \infty} \beta_n = \infty$. and let $(M_n)_{n \geq 0}$ be a sequence of Markov kernels such that $\mu_n M_n = \mu_n$ where

$$\mu_n(\mathrm{d}x) \propto \mathrm{e}^{-\beta_n V(x)} \lambda(\mathrm{d}x) \quad \text{and} \quad g_n(x) = \mathrm{e}^{-(\beta_{n+1} - \beta_n) V(x)}$$

3. Then, $\eta_n = \mu_n \propto \exp(-\beta_n V)$ which (under appropriate assumptions) converge weakly to a distributed to distribution concentrated on the set of local maxima.

Some Other Applications of Feynman-Kac Formulae

- Signal processing and automatic control
  - Open loop optimal control, optimal regulation.
  - Interacting Kalman-Bucy filters.
  - Stochastic and adaptive grid approximation-models
- Statistics/Probability:
  - Markov chains with constraints (w.r.t terminal values, visiting regions, constraints simulation problems,...)
  - Analysis of Boltzmann-Gibbs type distributions (simulation, partition functions, localization models...).
  - Combinatorial optimization, counting, graph-coloring
- Rare events analysis:
  - Multisplitting and branching particle models (Restart type methods).
  - Importance sampling and twisted probability measures.
  - Genealogical tree based simulations (default tree sampling models).

Feynman-Kac Flow

- The sequence $(\eta_n)_{n\geq 0}$ satisfies the recursion:

$$\eta_{n+1}(A_{n+1}) = \frac{\int \eta_n(\mathrm{d}x_n) g_n(x_n) M_{n+1}(x_n, \mathrm{d}x_{n+1})}{\int \eta_n(\mathrm{d}x_n) g_n(x_n)} = \Psi_n(\eta_n) M_{n+1}(A_{n+1}) ,$$

where $\Psi_n : \mathcal{P}(\mathbb{E}_n) \to \mathcal{P}(\mathbb{E}_n)$ is the non-linear mapping:

$$\Psi_n(\eta_n)(A_n) \stackrel{\text{def}}{=} \frac{1}{\eta_n(g_n)} \int_{A_n} g_n(x_n)\ \eta_n(\mathrm{d}x_n) , \quad A_n \in \mathcal{E}_n .$$

Outline

Bootstrap Particle Approximation I

- Denote by N_1 the number of particles in a given island.
- The evolution of the population of an island $\boldsymbol{X}_n = (X_n^1, \ldots, X_n^{N_1})$ is in two steps
  1. the particles are multinomially resampled with probabilities proportional to their potential $\{g_n(X_n^i)\}_{i=1}^{N_1}$;
  2. new particle positions are then sampled conditionally independently from the prior kernel M_{n+1}.

$$\left(X_n^i\right) \xrightarrow{\text{selection}} \left(\hat{X}_n^i\right) \xrightarrow{\text{mutation}} \left(X_{n+1}^i\right)$$

Bootstrap Particle Approximation II

The island evolution defines a Markov chain specified by the transition $\boldsymbol{M}_{n+1}$ from $(\mathbf{E}_n, \boldsymbol{\mathcal{E}}_n) = (\mathbb{E}_n^{N_1}, \mathcal{E}_n^{\otimes N_1})$ to $(\mathbf{E}_{n+1}, \boldsymbol{\mathcal{E}}_{n+1}) = (\mathbb{E}_n^{N_1}, \mathcal{E}_n^{\otimes N_1})$ by

$$\begin{aligned}\boldsymbol{M}_{n+1}(\boldsymbol{X}_n, \mathbf{A}_{n+1}) &= \prod_{1 \le i \le N_1} \sum_{j=1}^{N_1} \frac{g_n(X_n^j)}{\sum_{\ell=1}^{N_1} g_n(X_n^\ell)} M_{n+1}(X_n^j, A_{n+1}^j) \\ &= \prod_{1 \le i \le N_1} \Psi_n(m[\boldsymbol{X}_n]) M_{n+1}(A_{n+1}^i) \ ,\end{aligned}$$

where $m[\boldsymbol{X}_n]$ denotes the empirical measure of an island

$$m[\boldsymbol{X}_n] = m(X_n^1, \ldots, X_n^{N_1}) \stackrel{\text{def}}{=} \frac{1}{N_1} \sum_{i=1}^{N_1} \delta_{X_n^i} \ .$$

Particle approximation

- Denote by $\{\boldsymbol{X}_n = (X_n^1, \ldots, X_n^{N_1})\}_{n\geq 0}$ a Markov Chain with initial distribution $\boldsymbol{\eta}_0 \stackrel{\mathrm{def}}{=} \eta_0^{\otimes N_1}$ and transition kernel $\boldsymbol{M}_{n+1}$.
- The island approximation of the sequences Feynman-Kac measures $\{(\eta_n, \gamma_n)\}_{n\geq 1}$ is given by

$$\eta_n^{N_1}(f_n) \stackrel{\mathrm{def}}{=} m(\boldsymbol{X}_n, f_n)$$

$$\gamma_n^{N_1}(f_n) \stackrel{\mathrm{def}}{=} \eta_n^{N_1}(f_n) \prod_{0\leq p<n} \eta_p^{N_1}(g_p) \; .$$

Unbiasedness of the particle approximation

Theorem

For any $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, $\gamma_n^{N_1}(f_n)$ *is an unbiased estimator of* $\gamma_n(f_n)$.

Unbiasedness of the particle approximation

Theorem

For any $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, $\gamma_n^{N_1}(f_n)$ is an unbiased estimator of $\gamma_n(f_n)$.

$$\mathbb{E}\left[\eta_p^{N_1}(f_p)\middle|\mathcal{F}_{p-1}^{N_1}\right] = \frac{1}{N_1}\sum_{i=1}^{N_1}\mathbb{E}\left[f_p(X_p^i)\middle|\mathcal{F}_{p-1}^{N_1}\right] = \mathbb{E}\left[f_p(X_p^1)\middle|\mathcal{F}_{p-1}^{N_1}\right]$$
$$= \frac{\sum_{i=1}^{N_1} g_{p-1}(X_{p-1}^i)M_p f_p(X_{p-1}^i)}{\sum_{i=1}^{N_1} g_{p-1}(X_{p-1}^i)} = \frac{\eta_{p-1}^{N_1}(Q_p f_p)}{\eta_{p-1}^{N_1}(g_{p-1})} ,$$

where $Q_p(x_{p-1}, \mathrm{d}x_p) = g_{p-1}(x_{p-1})M_p(x_{p-1}, \mathrm{d}x_p)$.

Unbiasedness of the particle approximation

Theorem

For any $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, $\gamma_n^{N_1}(f_n)$ is an unbiased estimator of $\gamma_n(f_n)$.

$$\begin{aligned}
\mathbb{E}\left[\gamma_n^{N_1}(f_n)\right] &= \mathbb{E}\left[\mathbb{E}\left[\eta_n^{N_1}(f_n)\middle|\mathcal{F}_{n-1}^{N_1}\right] \prod_{0 \le p < n} \eta_p^{N_1}(g_p)\right] \\
&= \mathbb{E}\left[\frac{\eta_{n-1}^{N_1}(Q_n f_n)}{\eta_{n-1}^{N_1}(g_{n-1})} \prod_{0 \le p < n} \eta_p^{N_1}(g_p)\right] \\
&= \mathbb{E}\left[\eta_{n-1}^{N_1}(Q_n f_n) \prod_{0 \le p < n-1} \eta_p^{N_1}(g_p)\right] .
\end{aligned}$$

The island Feynman-Kac model

- For $\mathbf{x}_n = (x_n^1, \cdots, x_n^{N_1}) \in \mathbb{E}_n^{N_1}$ define the sample averaged potential

$$\boldsymbol{g}_n(\mathbf{x}_n) \stackrel{\text{def}}{=} \frac{1}{N_1} \sum_{i=1}^{N_1} g_n(x_n^i) \ .$$

- The Feynman-Kac model associated to $(\boldsymbol{M}_n, \boldsymbol{g}_n)_{n \geq 0}$ is given by

$$\boldsymbol{\eta}_n(\boldsymbol{f}_n) = \boldsymbol{\gamma}_n(\boldsymbol{f}_n) / \boldsymbol{\gamma}_n(1)$$
$$\boldsymbol{\gamma}_n(\boldsymbol{f}_n) = \mathbb{E}\left[\boldsymbol{f}_n(\mathbf{X}_n) \prod_{0 \leq p < n} \boldsymbol{g}_p(\mathbf{X}_p)\right] \ ,$$

where $(\boldsymbol{X}_n)_{n \geq 0}$ is a Markov chain with transition $(\boldsymbol{M}_n)_{n \geq 0}$.

Unbiasedness

Since $\boldsymbol{g}_n(\boldsymbol{X}_p) = \eta_n^{N_1}(g_p)$, the unbiasedness property implies that for any $\boldsymbol{f}_n$ of the form $\boldsymbol{f}_n(\mathbf{x}_n) = N_1^{-1}\sum_{i=1}^{N_1} f_n(x_n^i)$

$$\mathbb{E}\left[f_n(X_n)\prod_{0\leq p<n} g_p(X_p)\right] = \mathbb{E}\left[\boldsymbol{f}_n(\mathbf{X}_n)\prod_{0\leq p<n} \boldsymbol{g}_p(\mathbf{X}_p)\right] ,$$

or equivalently

$$\boldsymbol{\gamma}_n(\boldsymbol{f}_n) = \gamma_n(f_n) \quad \text{and} \quad \boldsymbol{\eta}_n(\boldsymbol{f}_n) = \eta_n(f_n) .$$

For functions defined as sample mean, the Feynman-Kac models $(\eta_n, \gamma_n)_{n\geq 0}$ and $(\boldsymbol{\eta}_n, \boldsymbol{\gamma}_n)_{n\geq 0}$ coincide !

Outline

Double Bootstrap

- Idea: we may apply the interacting particle system approximation of the Feynman-Kac semigroups both within each island but also between islands.
- We now describe the double bootstrap algorithm where the bootstrap algorithm is applied both within an island but also across the islands.
- Of course, many other options are available (more to come !)

Feynman-Kac at the island level

- Define by $\mathcal{P}(\mathbb{E}_n)$ the set of probabilities measures on $(\mathbb{E}_n, \boldsymbol{\mathcal{E}}_n)$.
- The sequence of measures $(\boldsymbol{\eta}_n)_{n\geq 0}$ satisfies the following recursion

$$\boldsymbol{\eta}_{n+1} = \boldsymbol{\Psi}_n(\boldsymbol{\eta}_n)\boldsymbol{M}_{n+1} \ ,$$

where $\boldsymbol{\Psi}_n : \mathcal{P}(\mathbb{E}_n) \to \mathcal{P}(\mathbb{E}_n)$ is defined by

$$\boldsymbol{\Psi}_n(\boldsymbol{\eta}_n)(\mathrm{d}\mathbf{x}) \stackrel{\text{def}}{=} \frac{\boldsymbol{g}_n(\mathbf{x}) \ \boldsymbol{\eta}_n(\mathrm{d}\mathbf{x})}{\boldsymbol{\eta}_n(\boldsymbol{g}_n)} \ .$$

The double bootstrap algorithm

$$\left(\boldsymbol{X}_n^i\right) \xrightarrow{\text{selection}} \left(\widehat{\boldsymbol{X}}_n^i\right) \xrightarrow{\text{mutation}} \left(\boldsymbol{X}_{n+1}^i\right)$$

- Let N_2 be the number of interacting islands.
- During the selection stage, we select randomly N_2 islands $\left(\widehat{\boldsymbol{X}}_n^i\right)_{1 \leq i \leq N_2}$ among the current islands $\left(\boldsymbol{X}_n^i\right)_{1 \leq i \leq N_2} \in \mathbb{E}_n^{N_2}$ with probability proportional to the empirical mean of the individuals in each island

$$\boldsymbol{g}_n(\boldsymbol{X}_n^i) = N_1^{-1} \sum_{j=1}^{N_1} g_n(X_n^{i,j}) \ , 1 \leq i \leq N_2 \ .$$

- During the mutation transition, selected islands $(\widehat{\boldsymbol{X}}_n^i)_{i=1}^{N_2}$ evolve randomly to a new configuration $\boldsymbol{X}_{n+1}^i$ according to the Markov transition $\boldsymbol{M}_{n+1}$.

The double bootstrap algorithm

for p from 0 to $n-1$ **do**

  selection between islands: Sample $\boldsymbol{I}_p = (I_p^i)_{i=1}^{N_2}$ multinomially with proba. prop. to $\left(\frac{1}{N_1}\sum_{j=1}^{N_1} g_p(X_p^{i,j})\right)_{i=1}^{N_2}$.

  for i from 1 to N_2 **do**

    selection within island: Sample $\boldsymbol{J}_p^i = (J_p^{i,j})_{j=1}^{N_1}$ multinomially with proba. prop. to $\left(g_p(X_p^{I_p^i,j})\right)_{j=1}^{N_1}$.

    For $1 \leq j \leq N_1$, sample independently $X_{p+1}^{i,j}$ according to $M_{p+1}(X_p^{I_p^i,J_p^j},\cdot)$.

  end for

end for

Bootstrap approximation: bias and variance

Theorem

For any time horizon $n \geq 0$ and any bounded function $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, we have

$$\lim_{N_1 \to \infty} N_1 \mathbb{E}\left[\eta_n^{N_1}(f_n) - \eta_n(f_n)\right] = B_n(f_n) ,$$

$$\lim_{N_1 \to \infty} N_1 \mathbb{V}\mathrm{ar}\left(\eta_n^{N_1}(f_n)\right) = V_n(f_n) ,$$

where $B_n(f_n)$ and $V_n(f_n)$ can be computed explicitly.

Double bootstrap approximation: bias and variance

Theorem

For any time horizon $n \geq 0$ and any $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, we have

$$\lim_{N_1\to\infty} \lim_{N_2\to\infty} N_1 N_2 \mathbb{E}\left[\boldsymbol{\eta}_n^{N_2}(m(\cdot, f_n)) - \boldsymbol{\eta}_n(m(\cdot, f_n))\right] = B_n(f_n) + \widetilde{B}_n(f_n) ,$$

$$\lim_{N_1\to\infty} \lim_{N_2\to\infty} N_1 N_2 \mathbb{V}\mathrm{ar}\left(\boldsymbol{\eta}_n^{N_2}(m(\cdot, f_n))\right) = V_n(f_n) + \widetilde{V}_n(f_n) ,$$

where $B_n(f_n)$, $\widetilde{B}_n(f_n)$, $V_n(f_n)$, $\widetilde{V}_n(f_n)$ can be computed explicitly.

- The rate of the interacting island (N_2 islands each with N_1 individuals) is the same as the one of the single island model with $N_1 N_2$ particles.
- Even though the constant terms may be worst in the interacting island model, it allows to use parallel implementations.

Independent islands

Theorem

For any time horizon $n \geq 0$ *and any* $f_n \in \mathcal{B}_b(\mathbb{E}_n)$*, we have*

$$\lim_{N_1 \to \infty} N_1 \left\{ \mathbb{E}\left[\widetilde{\boldsymbol{\eta}}_n^{N_2}(m(\cdot, f_n)) \right] - \eta_n(f_n) \right\} = B_n(f_n) \ ,$$

$$\lim_{N_1 \to \infty} N_1 N_2 \mathbb{V}\mathrm{ar} \left(\widetilde{\boldsymbol{\eta}}_n^{N_2}(m(\cdot, f_n)) \right) = V_n(f_n) \ ,$$

where $B_n(f_n)$ *and* $V_n(f_n)$ *are the same than for the single island model.*

Although the variance of the particle approximation is inversely proportional to $N_1 N_2$, the bias is independent of N_2 and is inversely proportional to N_1.

How to choose between interacting and independent islands?

	Independent islands	Interacting islands
Squared bias	$\dfrac{B_n(f_n)^2}{N_1^2}$	$\dfrac{\left(B_n(f_n) + \widetilde{B}_n(f_n)\right)^2}{N_1^2 N_2^2}$
Variance	$\dfrac{V_n(f_n)}{N_1 N_2}$	$\dfrac{V_n(f_n) + \widetilde{V}_n(f_n)}{N_1 N_2}$
Sum	$\dfrac{V_n(f_n)}{N_1 N_2} + \dfrac{B_n(f_n)^2}{N_1^2}$	$\dfrac{V_n(f_n) + \widetilde{V}_n(f_n)}{N_1 N_2}$

$$\frac{V_n(f_n)}{N_1 N_2} + \frac{B_n(f_n)^2}{N_1^2} < \frac{V_n(f_n) + \widetilde{V}_n(f_n)}{N_1 N_2} \quad \Leftrightarrow \quad N_1 > \frac{B_n(f_n)^2}{\widetilde{V}_n(f_n)} N_2 \; .$$

Example

1. Linear Gaussian Model

- $X_{p+1} = \phi X_p + \sigma_u U_p$,
- $Y_p = X_p + \sigma_v V_p$,

Computing the predictive distribution of the state X_n given the observations $Y_{0:n-1} = y_{0:n-1}$ up to time $n-1$ can be cast into the framework of Feynman-Kac model by setting for all $p \geq 0$

$$M_{p+1}(x_p, \mathrm{d}x_{p+1}) = \frac{1}{\sqrt{2\pi}\sigma_u} \exp\left[-(x_{p+1} - \phi x_p)^2/(2\sigma_u^2)\right] \mathrm{d}x_{p+1} ,$$

$$g_p(x_p) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left[-(y_p - x_p)^2/(2\sigma_v^2)\right] .$$

Results for the LGSS model

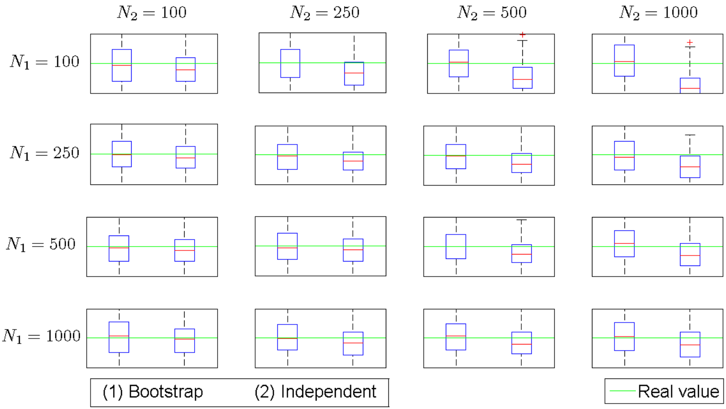


Figure: Interacting versus independent island renormalized estimators.

Outline

Espilon-Interaction Bootstrap I

- ϵ-bootstrap interaction is a variant of the bootstrap, in which only a fraction of the particles are resampled.
- ϵ_n be a nonnegative constant such that $\epsilon_n \ \|g_n\|_\infty \in [0,1]$, where $\|g_n\|_\infty = \sup_{x_n \in \mathbb{E}_n} |g_n(x_n)|$.
- At iteration n, a particle X_n^i is kept with a probability equal to $\epsilon_n \ g_n(epart[i]n)$ or resampled with a probability $1 - \epsilon_n \ g_n(X_n^i)$. Resampling a particle consists in replacing it by a particle selected at random in the current population with weights proportional to their potential $(g_n(X_n^1), \ldots, g_n(X_n^{N_1}))$.
- Then, each selected particle is independently updated according to the Markov kernel M_{n+1}.
- When $\epsilon_n = 0$, all the particles are resampled, which correspond to the bootstrap filter.

Espilon-Interaction Bootstrap II

- For any measure $\mu_n \in \mathcal{P}(\mathbb{E}_n)$, define S_{n,μ_n} the Markov kernel on $(\mathbb{E}_n, \mathcal{E}_n)$ given for $x_n \in \mathbb{E}_n$ and $A_n \in \mathcal{E}_n$ by

$$S_{n,\mu_n}(x_n, A_n) \stackrel{\text{def}}{=} \epsilon_n \; g_n(x_n)\delta_{x_n}(A_n) + (1 - \epsilon_n \; g_n(x_n))\, \Psi_n(\mu_n)(A_n) \;,$$

- Define the Markov kernel $\boldsymbol{M}_{n+1}(\mathbf{x}_n, d\mathbf{x}_{n+1})$ from $\mathbf{E}_n$ into $\mathbf{E}_{n+1}$ by

$$\boldsymbol{M}_{n+1}(\mathbf{x}_n, \mathrm{d}\mathbf{x}_{n+1}) \stackrel{\text{def}}{=} \prod_{1 \le i \le N_1} S_{n,\eta_n^{N_1}} M_{n+1}(x_n^i, \mathrm{d}x_{n+1}^i) \;.$$

- Consider the Feynman-Kac model associated to $(\boldsymbol{M}_n, \boldsymbol{g}_n)_{n \ge 0}$

$$\boldsymbol{\eta}_n(\boldsymbol{f}_n) = \boldsymbol{\gamma}_n(\boldsymbol{f}_n)/\boldsymbol{\gamma}_n(1)$$
$$\boldsymbol{\gamma}_n(\boldsymbol{f}_n) = \mathbb{E}\left[\boldsymbol{f}_n(\mathbf{X}_n) \prod_{0 \le p < n} \boldsymbol{g}_p(\mathbf{X}_p)\right] \;,$$

where

$$\boldsymbol{g}_n(\mathbf{x}_n) \stackrel{\text{def}}{=} m(\mathbf{x}_n, g_n) = \frac{1}{N_1} \sum_{i=1}^{N_1} g_n(x_n^i) \;.$$

between Island ϵ-interaction

- Idea: Apply the ϵ-interaction at the island level
- Algorithm
  - *Selection step*: each island $\boldsymbol{X}_n^i$ is kept with a probability equal to $\epsilon_n\ \boldsymbol{g}_n(\boldsymbol{X}_n^i)$ or resampled with a probability $1-\epsilon_n\ \boldsymbol{g}_n(\boldsymbol{X}_n^i)$. Resampling an island consists in replacing it by an island selected at random in the current population with weights proportional to their average potential $(\boldsymbol{g}_n(\boldsymbol{X}_n^1)),\ldots,\boldsymbol{g}_n(\boldsymbol{X}_n^{N_1})$.
  - *Mutation step:* each selected island is updated independently according to the Markov transition $\boldsymbol{M}_{n+1}$.
- It is not required to use ϵ-interaction both within and across the islands.

Effective Sample Size Interaction

- Idea: Perform the selection step of the current particles only when the importance weights do not satisfy some appropriately defined criterion.
- Contrary to the bootstrap filter, we now keep both the particles and the weights.
- For a weighted sample $\{(w_n^i, x_n^i)\}_{i=1}^{N_1}$, the criterion

$$\left(\sum_{i=1}^{N_1} w_n^i g_n(x_n^i)\right)^2 / \sum_{i=1}^{N_1} \left(w_n^i g_n(x_n^i)\right)^2$$

  is the effective sample size (ESS).
- Roughly speaking, the ESS is the way to quantify the dependence in the particle swarm.

Effective Sample Size Interaction

1. When the ESS is less than αN_1, for some $\alpha \in (0, 1)$, the particles are multinomially resampled with probabilities proportional to their weights times their potential functions; the weights are all reset to 1.
2. When the ESS is greater than αN_1, then the weights are simply multiplied by the potential function
3. The particle positions are then updated according to the transition kernel M_{n+1}.

This algorithm defines a Markov chain $\{\boldsymbol{X}_n\}_{n\geq 0}$ where for each $n \in \mathbb{N}$,

$$\boldsymbol{X}_n = \left[(X_n^1, \omega_n^1), \ldots, (X_n^{N_1}, \omega_n^{N_1})\right] \in \mathbb{E}_n \ ,$$

ESS: particle approximation

N_1-particle approximations of the measures η_n and γ_n

$$\eta_n^{N_1}(f_n) \stackrel{\text{def}}{=} m(\boldsymbol{X}_n, f_n) = \frac{1}{\sum_{i=1}^{N_1} \omega_n^i} \sum_{i=1}^{N_1} \omega_n^i f_n\left(X_n^i\right) ,$$

$$\gamma_n^{N_1}(f_n) \stackrel{\text{def}}{=} \eta_n^{N_1}(f_n) \prod_{0 \le p < n} \eta_p^{N_1}(g_p) .$$

Theorem

For any $f_n \in \mathcal{B}_b(\mathbb{E}_n)$, $\gamma_n^{N_1}(f_n)$ is an unbiased estimator of $\gamma_n(f_n)$:

$$\mathbb{E}\left[\gamma_n^{N_1}(f_n)\right] = \mathbb{E}\left[\eta_n^{N_1}(f_n) \prod_{0 \le p < n} \eta_p^{N_1}(g_p)\right] = \mathbb{E}\left[f_n(X_n) \prod_{0 \le p < n} g_p(X_p)\right] .$$

ESS: Feynman-Kac approximation

- For $\mathbf{x}_n = (x_n^1, w_n^1, \cdots, x_n^{N_1}, w_n^{N_1}) \in \mathbb{E}_n$ we set

$$\boldsymbol{g}_n(\mathbf{x}_n) \stackrel{\text{def}}{=} m(\mathbf{x}_n, g_n) = \frac{1}{\sum_{i=1}^{N_1} w_n^i} \sum_{i=1}^{N_1} w_n^i g_n\left(x_n^i\right) \ .$$

- The Feynman-Kac associated to $\{(\boldsymbol{M}_n, \boldsymbol{g}_n)\}_{n\geq 0}$ is

$$\boldsymbol{\eta}_n(\boldsymbol{f}_n) = \boldsymbol{\gamma}_n(\boldsymbol{f}_n)/\boldsymbol{\gamma}_n(1)$$
$$\boldsymbol{\gamma}_n(\boldsymbol{f}_n) = \mathbb{E}\left[\boldsymbol{f}_n(\mathbf{X}_n) \prod_{0\leq p<n} \boldsymbol{g}_p(\mathbf{X}_p)\right] \ ,$$

Feynman-Kac approximation

Since $\boldsymbol{g}_n(\boldsymbol{X}_n) = \eta_n^{N_1}(g_n)$, for any $\boldsymbol{f}_n$ of the form
$\boldsymbol{f}_n(\mathbf{x}_n) = \left(\sum_{i=1}^{N_1} w_n^i\right)^{-1} \sum_{i=1}^{N_1} w_n^i f_n\left(x_n^i\right)$ where $f_n \in \mathcal{B}_b(\mathbb{E}_n)$,

$$\mathbb{E}\left[f_n(X_n) \prod_{0 \leq p < n} g_p(X_p)\right] = \mathbb{E}\left[\boldsymbol{f}_n(\mathbf{X}_n) \prod_{0 \leq p < n} \boldsymbol{g}_p(\mathbf{X}_p)\right] ,$$

Therefore, the unbiasedness theorem implies

$$\boldsymbol{\gamma}_n(\boldsymbol{f}_n) = \gamma_n(f_n)$$
$$\boldsymbol{\eta}_n(\boldsymbol{f}_n) = \eta_n(f_n) .$$

Between island ESS: Principles

- Idea Apply the ESS at the island level.
- Let $(\boldsymbol{X}_n^1, \ldots, \boldsymbol{X}_n^{N_2}) \in \mathbb{E}_n^{N_2}$ be a population of N_2 islands each of N_1 individuals.
- We now associate to each island a weight denoted Ω_n^i, for $i \in \{1, \ldots, N_2\}$.
- At each iteration, we will assess the degeneracy of the population of islands using the ESS (at the island level !)

The N_2-particle approximation of the measures $\boldsymbol{\eta}_n$ and $\boldsymbol{\gamma}_n$ is given by

$$\boldsymbol{\eta}_n^{N_2}(\boldsymbol{f}_n) \stackrel{\text{def}}{=} \frac{1}{\sum_{i=1}^{N_2} \Omega_n^i} \sum_{i=1}^{N_2} \Omega_n^i \boldsymbol{f}_n(\boldsymbol{X}_n^i) ,$$

$$\boldsymbol{\gamma}_n^{N_2}(\boldsymbol{f}_n) \stackrel{\text{def}}{=} \boldsymbol{\eta}_n^{N_2}(\boldsymbol{f}_n) \prod_{0 \le p < n} \boldsymbol{\eta}_p^{N_2}(\boldsymbol{g}_p) = \boldsymbol{\eta}_n^{N_2}(\boldsymbol{f}_n)\, \boldsymbol{\gamma}_n^{N_2}(1) .$$

Between island ESS: algorithm

- Selection step:
  1. if the ESS criterion

$$\left(\sum_{i=1}^{N_2} \Omega_n^i \boldsymbol{g}_n(\boldsymbol{X}_n^i)\right)^2 / \sum_{i=1}^{N_2} \left(\Omega_n^i \boldsymbol{g}_n(\boldsymbol{X}_n^i)\right)^2$$

     is larger than βN_2 for one $\beta \in (0,1)$, the islands are kept and the weights are updated:

$$\Omega_{n+1}^i = \Omega_n^i \boldsymbol{g}_n(\boldsymbol{X}_n^i)$$

     ;
  2. otherwise, the islands are resampled multinomially with probability proportional to $\{\Omega_n^i \boldsymbol{g}_n(\boldsymbol{X}_n^i)\}_{i=1}^{N_2}$ and the weights are all reset to 1.
- Mutation step: Each selected island is updated independently according to the Markov transition $\boldsymbol{M}_{n+1}$.

Results

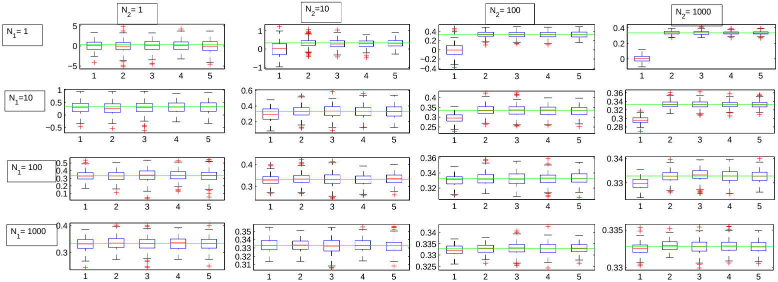


Figure: LGM model. Comparison of different interactions across the islands. The bootstrap is used within each island for the LGM (1) ESS/independent; (2) ESS/ESS; (3) ESS/Bootstrap; (4) ESS/$(1/g_n)$-bootstrap; (5) ESS/$\text{essup}_{\eta_p^{N_1}}(g_n)$-bootstrap

Number of interactions

N_1 \ N_2	1	10	100	1000
1	100	40.04	26.08	13.50
10	100	81.90	78.56	77.54
100	100	97.26	95.86	95.02
1000	100	99.86	100	100

Table: Gain in the number of interactions between islands for the ESS within ESS estimator as a percentage of the one of the ESS within bootstrap estimator in the LGM.