

Island particle algorithms and application to rare event analysis

Christelle Vergé

(joint work with P. Del Moral, E. Moulines, J. Morio, J. Olsson, C. Dubarry)

08/04/2015 - MASCOT-NUM







Plan

Introduction to island particle models : a way to parallelize sequential Monte Carlo (SMC) algorithms

An example of island particle algorithm for rare event analysis

Conclusion







Plan

Introduction to island particle models : a way to parallelize sequential Monte Carlo (SMC) algorithms

An example of island particle algorithm for rare event analysis

Conclusion







Feynman-Kac models

- $\{Q_n\}_{n\in\mathbb{N}}$: a set of unnormalized transition kernels on the measurable space (X, \mathcal{X})
- \blacksquare η_0 : a probability distribution

Definition:

A sequence $\{\eta_n\}_{n\in\mathbb{N}}$ of Feynman-Kac measures is defined, for all $h\in\mathsf{F}_\mathrm{b}(\mathsf{X})$ by

$$\eta_n h = \frac{\int \cdots \int h(x_n) \, \eta_0(\mathrm{d}x_0) \prod_{p=0}^{n-1} \, Q_p(x_p, \mathrm{d}x_{p+1})}{\int \cdots \int \eta_0(\mathrm{d}x_0) \prod_{p=0}^{n-1} \, Q_p(x_p, \mathrm{d}x_{p+1})} \quad (n \in \mathbb{N}).$$

▶ The Feynman-Kac measures satisfy the nonlinear recursion

$$\eta_{n+1} = \eta_n Q_n / \eta_n Q_n \mathbb{1}_{X}.$$

 Numerous applications (nonlinear filtering, rare event sampling, hidden Markov chain parameter estimation, stochastic control problems, financial mathematics...)







Interacting particle systems (IPS)

Approximation of Feynman-Kac measures $\{\eta_n\}_{n\in\mathbb{N}}$:

- Kalman filter: exact simulation for linear Gaussian models.
- Otherwise, SMC methods : empirical approximation of η_n thanks to an interacting particle systems $\{(\xi_n^N(i),\omega_n^N(i))\}_{i=1}^N$. The weighted sample evolves recursively according to selection and mutation steps :

$$\{(\xi_{n}^{N}(i), \omega_{n}^{N}(i))\}_{i=1}^{N} \xrightarrow{\text{Selection}} \{(\xi_{n}^{N}(I_{n}^{N}(i)), \omega_{n}^{N}(I_{n}^{N}(i)))\}_{i=1}^{N} \xrightarrow{\text{Mutation}} \{(\xi_{n+1}^{N}(i), \omega_{n+1}^{N}(i))\}_{i=1}^{N}$$

$$I_{n}^{N}(i) \in [1, N] \xrightarrow{\xi_{n+1}^{N}(i) \sim R_{n}(\xi_{n}^{N}(I_{n}^{N}(i)), \cdot)}$$

$$\omega_{n+1}^{N}(i) = \omega_{n}^{N}(I_{n}^{N}(i)) w_{n}(\xi_{n}^{N}(I_{n}^{N}(i)), \xi_{n+1}^{N}(i))$$
where R is a proposal kernel and w is the importance function such that

where R_n is a proposal kernel and w_n is the importance function such that $Q_n(x, dy) = w_n(x, y)R_n(x, dy) \quad ((x, y) \in X^2).$

- Different kinds of selection may be considered :
 - ▶ systematic resampling : $\forall i, I_n^N(i) \sim \text{Mult}(\{\omega_n^N(k)\}_{k=1}^N) \hookrightarrow \text{bootstrap}$ algorithm [Ref : N.J. Gordon, D. J. Salmond, A. F. M. Smith (1993)]
 - resampling of all the particles only when their weights are skewed (Effective Sample size, coefficient of variation) [Ref: J. Liu, R. Chen (1993)]







The bootstrap algorithm

```
/* Initialization
for i \leftarrow 1 to N do
    \xi_0^N(i) \sim \eta_0;

\xi_1^N(i) \sim R_0(\xi_0^N(i), \cdot) \quad \omega_1^N(i) \leftarrow w_0(\xi_0^N(i), \xi_1^N(i));
end
for p \leftarrow 1 to n-1 do
     /* Selection step
                                                                                                      */
     Sample \{I_n^N(i)\}_{i=1}^N \sim_{i,i,d} \text{Mult}(\{\omega_n^N(k)\}_{k=1}^N);
     /* Mutation step
     for i \leftarrow 1 to N do
           Sample conditionally independently
           \xi_{p+1}^{N}(i) \sim R_{p}(\xi_{p}^{N}(I_{p}^{N}(i)), \cdot);
           Update the weights \omega_{n+1}^N(i) \leftarrow w_n(\xi_n^N(I_n^N(i)), \xi_{n+1}^N(i));
     end
end
```



Analysis of the bootstrap algorithm

Proposition: [Ref: P. Del Moral, Feynman-Kac formulae, 2004] [Ref: R. Douc and E. Moulines, 2008]

For all $n \in \mathbb{N}$ and $h \in F_b(X)$, set $\eta_n^N h \triangleq \sum_{i=1}^N \omega_n^N(i) h(\xi_n^N(i)) / \sum_{k=1}^N \omega_n^N(k)$. Then,

$$\eta_n^N h \xrightarrow[N \to +\infty]{} \eta_n h$$
 almost surely,

$$\sqrt{N}(\eta_n^N h - \eta_n h) \xrightarrow[N \to +\infty]{\mathcal{D}} N(0, V_n(h)),$$
 where

$$V_0(h) = \eta_0 \{ (h - \eta_0 h)^2 \} \quad \text{and} \quad V_n(h) = \sum_{\ell=0}^{n-1} \frac{\eta_\ell R_\ell \{ w_\ell^2 Q_{\ell+1} \cdots Q_{n-1} (h - \eta_n h)^2 \}}{(\eta_\ell Q_\ell \cdots Q_{n-1} \mathbb{1}_{\mathsf{X}})^2}.$$

$$N \mathbb{E} \left[\eta_n^N h - \eta_n h \right] \xrightarrow[N \to +\infty]{} B_n(h).$$

- The precision of the estimation depends upon the size *N* of the particle swarm ⇒ critical for online applications
- Develop new SMC methods to reduce the size of the particle swarm, while ensuring good estimates ⇒ parallelization of SMC methods.







Parallelization of SMC methods

- Spread the total number $N \triangleq N_1 N_2$ of particles into N_1 batches of N_2 particles each.
- $lue{}$ Each batch is called an island. Each island evolves independently as a standard SMC algorithm with N_2 particles.
- The N_1 islands may be considered in a parallel architecture or may interact through a selection step on the island level, when assigning as island weight, the average of the particle weights in an island.
 - N₁ independent bootstraps
 - Double bootstrap (B²)
 - ► Double bootstrap with adaptive selection on the island level (B²ASIL)







The B²ASIL algorithm

```
/* Initialization
For all i \in \llbracket 1, N_1 \rrbracket and j \in \llbracket 1, N_2 \rrbracket, \xi_0^N(i,j) \sim \eta_0: \xi_1^N(i,j) \sim R_0(\xi_0^N(i,j),\cdot); \omega_1^N(i,j) \leftarrow w_0(\xi_0^N(i,j),\xi_1^N(i,j));
\Omega_1^N(i) \leftarrow \sum_{i=1}^{N_2} \omega_1^N(i,j)/N_2;
for p \leftarrow 1 to n-1 do
          /* Island selection
                                                                                                                                                                               */
         \text{if } \mathsf{CV}^2(\{\Omega_p^N(i)\}_{i=1}^{N_1}) \triangleq \mathsf{N}_1 \sum\nolimits_{i=1}^{N_1} \left(\Omega_p^N(i) / \sum\nolimits_{i'=1}^{N_1} \Omega_p^N(i')\right)^2 - 1 > \tau \text{ then }
                   For all i \in [1, N_1], I^N(i) \sim \text{Mult}(\{\Omega_n^N(i')\}^{\bar{N}_1}, );
          else
                   For all i \in [1, N_1], I^N(i) \leftarrow i;
          /* Island mutation
          for i \leftarrow 1 to N_1 do
                   /* Individual selection
                   For all j \in [1, N_2], J^N(i, j) \sim \text{Mult}(\{\omega_p^N(I^N(i), j')\}_{i'=1}^{N_2});
                   /* Mutation
                                                                                                                                                                              */
                   For all j \in [1, N_2], \xi_{p+1}^N(i, j) \sim R_p(\xi_p^N(I^N(i), J^N(i, j)), \cdot);
                   \omega_{p+1}^{N}(i,j) \leftarrow w_{p}(\xi_{p}^{N}(I^{N}(i),J^{N}(i,j)),\xi_{p+1}^{N}(i,j));
                  \Omega_{p+1}^{N}(i) \leftarrow \sum_{i=1}^{N_2} \omega_{p+1}^{N}(i,j)/N_2;
          end
end
```





Analysis of the B²ASIL algorithm

Denote by $\eta_n^N h = \sum_{i=1}^{N_1} \frac{\Omega_n^N(i)}{\sum_{i'=1}^{N_1} \Omega_n^N(i')} \sum_{j=1}^{N_2} \frac{\omega_n^N(i,j)}{\sum_{j'=1}^{N_2} \omega_n^N(i,j')} h(\xi_n^N(i,j))$ the estimators returned by the B²ASIL algorithm.

Theorem: [Ref: C. Vergé, P. Del Moral, E. Moulines, J. Olsson, preprint]

Let $n\in\mathbb{N}$ and $h\in\mathsf{F}_{\mathrm{b}}(\mathsf{X})$. Then, $\eta_n^Nh\xrightarrow{\mathbb{P}}\eta_nh$. Impose that for all $\beta>0$, $N_1\exp(-\beta N_2)\xrightarrow{N\to+\infty}0$. Then, for all $n\in\mathbb{N}$, the random variable $\mathbbm{1}\left\{\mathsf{CV}^2(\{\Omega_n^N(i)\}_{i=1}^{N_1})>\tau\right\}$ has a deterministic limit ε_n in probability. Moreover,

$$\sqrt{N}(\eta_n^N h - \eta_n h) \stackrel{\mathcal{D}}{\longrightarrow} N(0, V_n(h) + \widetilde{V}_n(h)),$$
 where

$$V_0(h) = \eta_0\{(h - \eta_0 h)^2\}, \ \widetilde{V}_0 = 0, \ \text{and} \ V_n h = \sum_{\ell=0}^{n-1} \frac{\eta_\ell R_\ell \{w_\ell^2 Q_{\ell+1} \cdots Q_{n-1} (h - \eta_n h)^2\}}{(\eta_\ell Q_\ell \cdots Q_{n-1} \mathbb{I}_X)^2},$$

$$\widetilde{V}_n h = \sum_{\ell=0}^{n-1} \sum_{n=\ell+1}^{n-1} \varepsilon_p \frac{\eta_\ell R_\ell \{w_\ell^2 Q_{\ell+1} \cdots Q_{n-1} (h - \eta_n h)^2\}}{(\eta_\ell Q_\ell \cdots Q_{n-1} \mathbb{1}_X)^2}.$$







Proof (sketch)

Proof (sketch)

The proof is by induction on n.

We decompose one iteration of the B²ASIL algorithm into elementary operations

- selection on the island level
- selection on the particle level
- mutation,

and show that each of them preserves a law of large numbers, a Hoeffding-type inequality and a central limit theorem.

This general framework allows to derive a law of large numbers and a central limit theorem for any algorithm that may be decomposed into these elementary operations.







A criterion to determine when islands should interact

	Bootstrap	N_1 independent bootstraps	B ² ASIL
Bias	$\frac{B_n}{N}$	$\frac{B_n}{N_2}$	$\frac{B_n + \widetilde{B}_n}{N_1 N_2}$
Variance	$\frac{V_n}{N}$	$\frac{V_n}{N_1 N_2}$	$\frac{V_n + \widetilde{V}_n}{N_1 N_2}$

Explicit expressions of $B_n, \widetilde{B}_n, V_n$ and \widetilde{V}_n can be found in :

[Ref : C. Vergé, C. Dubarry, P. Del Moral, E. Moulines, Statistics and Computing, 2015]
[Ref : C. Vergé, P. Del Moral, E. Moulines, J. Olsson, preprint]

Use the mean squared error to make a compromise between bias and variance : island interaction is beneficial when

$$\frac{V_n}{N_1N_2} + \frac{B_n^2}{N_2^2} > \frac{V_n + \widetilde{V}_n}{N_1N_2} \quad \Leftrightarrow \quad N_2 < \frac{B_n^2}{\widetilde{V}_n} N_1.$$

When $N_2 \ll N_1$, the interaction is beneficial, but prevents a total parallelization.







Stability of the double bootstrap (B^2)

Note that the B^2 algorithm, which selects systematically the islands, is a particular case of B^2ASIL algorithm for which $\tau=0$ (and hence $\varepsilon_n=1$ for all $n\in\mathbb{N}^*$). We may hence furnish the asymptotic variance of the B^2 algorithm :

$$\sigma_n^2(h) = \sum_{\ell=0}^{n-1} (n-\ell) \frac{\eta_\ell R_\ell \{ w_\ell^2 Q_{\ell+1} \cdots Q_{n-1} (h-\eta_n h)^2 \}}{(\eta_\ell Q_\ell \cdots Q_{n-1} \mathbb{1}_X)^2}.$$

Theorem [Ref : C. Vergé, P. Del Moral, E. Moulines, J. Olsson, preprint]

Suppose the standard strong mixing conditions :

- (i) There exist constants $0 < \sigma_- < \sigma_+ < \infty$ and $\varphi \in M_1(X)$ such that for all $p \in \mathbb{N}$, $x \in X$, and $A \in \mathcal{X}$, $\sigma_-\varphi(A) \leq M_p(x,A) \leq \sigma_+\varphi(A)$.
- (ii) $w_+ \triangleq \sup_{p \in \mathbb{N}} \|w_p\|_{\infty} < \infty$.
- (iii) $c_- \triangleq \inf_{(\rho,x) \in \mathbb{N} \times X} Q_\rho \mathbb{1}_X(x) > 0.$

Then for all $n \in \mathbb{N}$ and $h \in F_b(X)$, $\sigma_n^2(h) \leq w_+ \frac{\operatorname{osc}^2(h)}{(1-\rho)^2(1-\rho^2)^2c_-}$, where $\rho \triangleq 1 - \sigma_-/\sigma_+$.





Plan

Introduction to island particle models : a way to parallelize sequential Monte Carlo (SMC) algorithms

An example of island particle algorithm for rare event analysis

Conclusion







Context

Consider a black box model :



- A rare event is often modeled by the exceedence of a threshold $S: \{\phi(X) > S\}$ such that $\mathbb{P}(\phi(X) > S) < 10^{-4}$.
- Risk analysis is not just evaluating a risk or a probability of failure, but estimating the law of random phenomena that leads to a critic event.
- Some parameters θ of the model or density parameters of the input random variables X may be fixed by the experimenter and can influence the output random variable Y.
- ▶ We want to determine the impact of such tuning of parameters on the realization of the critic event, i.e. to compute the law of the parameters Θ conditionally on the rare event, denoted by $\pi \triangleq \mathsf{Law}(\Theta \ \phi(X) > S)$.







The splitting algorithm : a way to evaluate $\mathbb{P}(\phi(X) > S | \Theta = \theta)$

[Ref: S-K. Au and J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, 2001]

Consider an increasing sequence of thresholds

$$-\infty \triangleq S_0 < S_1 < ... < S_m \triangleq S$$
,

and decompose, using Bayes' formula,

$$\mathbb{P}(\phi(X) > S | \Theta = \theta) = \prod_{p=0}^{m-1} \mathbb{P}(\phi(X) > S_{p+1} | \phi(X) > S_p, \Theta = \theta).$$

$$\begin{bmatrix} X_n(1) \\ \vdots \\ X_n(N_2) \end{bmatrix} \xrightarrow{\begin{array}{c} \text{Selection} \\ 1 \\ \vdots \\ X_{n-1} \\ \end{array}} \begin{bmatrix} \widehat{X}_n(1) \\ \vdots \\ \widehat{X}_n(N_2) \end{bmatrix} \xrightarrow{\begin{array}{c} \text{Mutation} \\ \vdots \\ X_{n-1} \\ \end{array}} \begin{bmatrix} X_{n+1}(1) \\ \vdots \\ \vdots \\ X_{n+1}(N_2) \end{bmatrix}$$





The splitting algorithm : a way to evaluate $\mathbb{P}(\overline{\phi}(X) > S | \Theta = heta)$

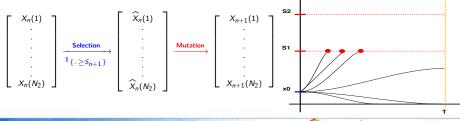
[Ref: S-K. Au and J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, 2001]

Consider an increasing sequence of thresholds

$$-\infty \triangleq S_0 < S_1 < ... < S_m \triangleq S$$
,

and decompose, using Bayes' formula,

$$\mathbb{P}(\phi(X) > S | \Theta = \theta) = \prod_{p=0} \mathbb{P}(\phi(X) > S_{p+1} | \phi(X) > S_p, \Theta = \theta).$$



The splitting algorithm : a way to evaluate $\mathbb{P}(\phi(X) > S | \Theta = \theta)$

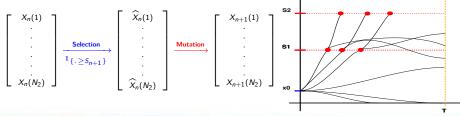
[Ref: S-K. Au and J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, 2001]

Consider an increasing sequence of thresholds

$$-\infty \triangleq S_0 < S_1 < ... < S_m \triangleq S$$
,

and decompose, using Bayes' formula,

$$\mathbb{P}(\phi(X) > S | \Theta = \theta) = \prod_{p=0} \mathbb{P}(\phi(X) > S_{p+1} | \phi(X) > S_p, \Theta = \theta).$$



The splitting algorithm : a way to evaluate $\mathbb{P}(\phi(X) > S | \Theta = \theta)$

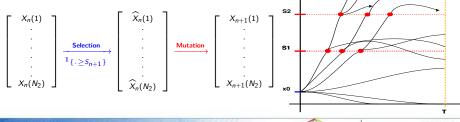
[Ref: S-K. Au and J.L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, 2001]

Consider an increasing sequence of thresholds

$$-\infty \triangleq S_0 < S_1 < ... < S_m \triangleq S$$
,

and decompose, using Bayes' formula,

$$\mathbb{P}(\phi(X) > S | \Theta = \theta) = \prod_{p=0} \mathbb{P}(\phi(X) > S_{p+1} | \phi(X) > S_p, \Theta = \theta).$$



The SMC² algorithm

- ▶ One way to sample from $\pi \triangleq \text{Law}(\Theta|\phi(X) > S)$, is to use a standard SMC algorithm with N_1 particles $\{\theta_n(i)\}_{i=1}^{N_1}$.
- ▶ For that purpose, we create a dynamic introducing intermediary thresholds $S_1 < ... < S_m \triangleq S$, and defining the distributions $\{\pi_n\}_{n \in \mathbb{N}^*}$ by

$$\pi_n(\mathrm{d}\theta) \triangleq \mathsf{Law}(\Theta|\phi(X) > S_n).$$

- Instead of trying to sample directly from π , sample successively from $\pi_1, \ldots, \pi_m \triangleq \pi$.
- ▶ The particles $\{\theta_n(i)\}_{i=1}^{N_1}$ evolve according to usual selection and mutation steps :
 - Selection : Multinomial resampling with weights proportional to : $\{\mathbb{P}(\phi(X) > S_{n+1} | \phi(X) > S_n, \Theta = \theta_n(i))\}_{i=1}^{N_1},$
 - Mutation : An acceptance / rejection step involving the probabilities $\mathbb{P}(\phi(X) > S|\theta_n(i))$, which are not computable.
- For each parameter $\theta_n(i)$, we run a splitting with N_2 particles $\{X_n(i,j)\}_{j=1}^{N_2}$, in order to replace every incalculable quantity by an unbiased estimator. We then have 2 embedded SMC algorithms \hookrightarrow the SMC² algorithm.

[Ref: N. Chopin, P. Jacob and O. Papaspiliopoulos, JRSSB, 2013]

[Ref : C. Vergé, J. Morio, P. Del Moral, preprint]







Analysis of SMC²

Theorem:

The SMC² algorithm converges to the target law π when $N_1 \to \infty$, for any fixed N_2 .

Sketch of proof : The SMC^2 algorithm can be viewed as an SMC algorithm on an extended state space.

Toy case : threshold exceedence for a Gaussian tail. We can compute explicitly $\text{Law}(\Theta|X\geq 5)$ when $X\sim \mathcal{N}(\Theta,1)$ and $\Theta\sim \mathcal{N}(0,1)$.

In this simulation, we use 2000×20 particles for the SMC² algorithm.

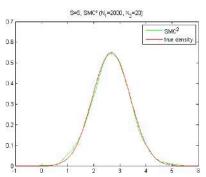


FIGURE: Law($\Theta|X \ge 5$)

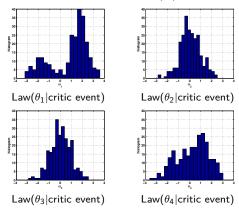






Application of SMC² algorithm to the fallout zone of a launch vehicle

We apply SMC² algorithm where ϕ simulates the distance between the true position of the fallout zone of a stage rocket and its predicted position. X is a Gaussian vector with covariance matrix equal to I_4 and mean $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)^t$ with a Gaussian *prior* i.e. for all $i \in [1, 4], \theta_i \sim \mathcal{N}(0, 1)$. The critic event is when the output distance $\phi(X)$ exceeds 0.72 km.









Influence of the parameters on the probability of interest

Monte Carlo estimates for different sets of parameters :

θ_1	θ_2	θ_3	θ_4	$\hat{\mathbb{P}}\left(\phi(X) > S \theta\right)$
0	0	0	0	$8.5 \ 10^{-4}$
1	0	0	1	$1.05 \ 10^{-2}$
-1	0	0	1	$1.02 \ 10^{-2}$
-1	0	0	-1	$1.14 \ 10^{-2}$

A bad tuning of the parameters can imply a large increase of the probability of the critic event and an underestimation of the associated risk ⇒ security matter.

[Ref : C. Vergé, J. Morio, P. Del Moral, preprint]







Plan

Introduction to island particle models : a way to parallelize sequential Monte Carlo (SMC) algorithms

An example of island particle algorithm for rare event analysis

Conclusion







Conclusions and perspectives

Conclusions:

- Introduction to island particle models,
- ▶ Definition of operations on islands,
- Establishment of a criterion to determine when islands should be considered in parallel or may interact,
- Study of asymptotic properties of island particle models,
- Transposition of an existing island particle model for rare event analysis.

Application:

Reliability analysis of a launch vehicle stage fallout.

Perspectives and future application:

- Application to reliability analysis for collision between a space debris and a satellite.
- ► Study of the SMC² algorithm.







Publications

Book:



P. Del Moral and C. Vergé, *Algorithmes Stochastiques : Modèles et Applications*, Springer Series : Maths & Applications, SMAI, vol.75, 2014, 487 pages, DOI = 10.1007/978-3-642-54616-7 (published).

Book chapters :



Contribution to two chapters of: Estimation of rare event probabilities in complex (and other) systems - a practical approach, J. Morio and M. Balesdent, Elsevier-Woodhead Publishing (August 2015).

- Chapter 5 : Simulation techniques
- Chapter 11: Estimation of collision probability between a space debris and a satellite







Publications

Journal publications :

Published

- C. Vergé, C. Dubarry, P. Del Moral, E. Moulines, On parallel implementation of Sequential Monte Carlo methods: the island particle model, Statistics and Computing, vol. 25, Issue 2, Mars 2015, pp. 243-260, DOI = 10.1007/s11222-013-9429-x.
- J. Morio, M. Balesdent, D. Jacquemart, C. Vergé, A survey of rare event estimation methods for static input-output models, Simulation Modelling Practice and Theory, vol. 49, pp 287-304, 2014.

Submitted

- C. Vergé, P. Del Moral, E. Moulines, J. Olsson, Asymptotic properties of weighted archipelagos with application to particle island methods.
- C. Vergé, J. Morio, P. Del Moral, An island Particle Markov Chain Monte Carlo algorithm for safety analysis.
- 3. C. Vergé, C. Ichard, Introduction to labeled island particle models and their asymptotic properties.

Conference publication :

- P. Del Moral, G. W. Peters, C. Vergé, An introduction to particle integration methods: with applications to risk and insurance, Monte Carlo and Quasi-Monte Carlo Methods 2012, Springer Proceedings in Mathematics & Statistics, volume 65, 2013, p.39-81, DOI = 10.1007/978-3-642-41095-6_3.
- C. Vergé, E. Moulines, J. Olsson, Asymptotic properties of particle island models with application to the double bootstrap filter, Signal Processing Conference (EUSIPCO), Proceedings of the 22nd European, IEEE (submitted, march 2015).
- C. Vergé, E. Moulines, J. Olsson, Fluctuation analysis of island particle models, 18th INFORMS Applied Probability Conference, Istanbul University, Turkey (submitted, march 2015).







Thank you for your attention!





