

# Particle rare event simulation

**P. Del Moral**

INRIA, Centre Bordeaux-Sud Ouest

Credit Risk Workshop, Evry University, June 2008.

# Outline

- 1 Stochastic sampling strategies
- 2 Feynman-Kac models
  - Some notation
  - Feynman-Kac semigroups
  - Mean field particle interpretations
  - Some theoretical convergence results
- 3 Some rare event models
  - Boltzmann-Gibbs/Tail distribution flows
  - Markov processes with fixed terminal values
  - Multi-splitting rare events excursions
  - Fixed time level set entrances
- 4 Some references

## Stochastic modeling

- Rare event = **cascade of intermediate (less) rare events** (increasing energies, critical levels, multilevel gateways).
- $\eta_n = \text{Law}(\text{process} \mid \text{a series of } n \text{ intermediate } \downarrow \text{ events})$   
= **nonlinear distribution flow** with  $\uparrow$  level of complexity.

$$\eta_0 \rightarrow \eta_1 \rightarrow \dots \rightarrow \eta_{n-1} \rightarrow \eta_n(dx) = \frac{1}{\gamma_n(\mathbf{1})} \gamma_n(dx) \rightarrow \dots$$

- Rare event probabilities = normalizing constants  $\gamma_n(\mathbf{1}) = \mathcal{Z}_n$ .

## Interacting stochastic sampling strategy

- **Interacting stoch. algo. = sampling w.r.t. a flow of meas.**
  - **Mean field particle models** = (sequential Monte Carlo, population Monte Carlo, particle filters, pruning, spawning, reconfiguration, quantum Monte carlo, go with the winner).
  - **Interacting MCMC models (new i-MCMC tech.)**  $\rightsquigarrow$  Ref..

## Notation

$E$  measurable space,  $\mathcal{P}(E)$  proba. on  $E$ ,  $\mathcal{B}(E)$  bounded meas. functions.

- $(\mu, f) \in \mathcal{P}(E) \times \mathcal{B}(E) \longrightarrow \mu(f) = \int f(x)\mu(dx)$
- $M(x, dy)$  **integral operator on E**

$$M(f)(x) = \int M(x, dy)f(y)$$

$$[\mu M](dy) = \int \mu(dx)M(x, dy) \iff [\mu M](f) = \mu[M(f)]$$

- **Boltzmann-Gibbs transformation** :  $G : E \rightarrow [0, 1]$  with  $\mu(G) > 0$

$$\Psi_G(\mu)(dx) = \frac{1}{\mu(G)} G(x) \mu(dx)$$

## Updating-prediction transformations

Time parameter  $n \in \mathbb{N}$ ,  $M_n(x, dy)$  Markov transitions and  $G_n : E \rightarrow [0, 1]$

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) := \Psi_{G_n}(\eta_n)M_{n+1} \quad (1)$$

Markov chain  $X_n$  with transitions  $M_n$  and initial condition  $X_0 \simeq \eta_0$  :

$$(1) \iff \eta_n(f) \propto \gamma_n(f) = \mathbb{E} \left( f(X_n) \prod_{0 \leq p < n} G_p(X_p) \right)$$

## Observations

- **Path space models** :  $X_n = (X'_0, \dots, X'_n) \in E_n = (E'_0 \times \dots \times E'_n)$
- **Normalizing constants** :

$$\gamma_n(\mathbf{1}) = \mathbb{E} \left( \prod_{0 \leq p < n} G_p(X_p) \right) = \prod_{0 \leq p < n} \eta_p(G_p)$$

- $G_n = \mathbf{1} \implies \eta_n = \Phi_n(\eta_{n-1}) = \eta_{n-1}M_n = \text{Law}(X_n)$ .

## Nonlinear distribution flows

- **Nonlinear Markov models** :  $\eta_n = \eta_{n-1} K_{n, \eta_{n-1}} = \text{Law}(\bar{X}_n)$

$$K_{n+1, \eta_n}(x, dz) = \int S_{n, \eta_n}(x, dy) M_{n+1}(y, dz)$$

$$S_{n, \eta_n}(x, dy) := \epsilon_n G_n(x) \delta_x(dy) + (1 - \epsilon_n G_n(x)) \Psi_{G_n}(\eta_n)(dy)$$

## Mean field particle interpretation

- **Markov chain**  $\xi_n = (\xi_n^1, \dots, \xi_n^N) \in E_n^N$  s.t.

$$\eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i} \simeq_{N \uparrow \infty} \eta_n$$

- Particle approximation transitions ( $\forall 1 \leq i \leq N$ )

$$\xi_{n-1}^i \rightsquigarrow \xi_n^i \sim K_{n, \eta_{n-1}^N}(\xi_{n-1}^i, dx_n)$$

## Mean field genetic type particle model :

$$\begin{array}{c} \xi_n^1 \\ \vdots \\ \xi_n^i \\ \vdots \\ \xi_n^N \end{array} \Bigg] \xrightarrow{S_{n,\eta_n^N}} \begin{array}{c} \widehat{\xi}_n^1 \\ \vdots \\ \widehat{\xi}_n^i \\ \vdots \\ \widehat{\xi}_n^N \end{array} \begin{array}{c} \xrightarrow{M_{n+1}} \\ \longrightarrow \\ \longrightarrow \end{array} \begin{array}{c} \xi_{n+1}^1 \\ \vdots \\ \xi_{n+1}^i \\ \vdots \\ \xi_{n+1}^N \end{array} \Bigg]$$

Accept/Reject/Selection transition :

$$S_{n,\eta_n^N}(\xi_n^i, dx)$$

$$:= \epsilon_n G_n(\xi_n^i) \delta_{\xi_n^i}(dx) + (1 - \epsilon_n G_n(\xi_n^i)) \sum_{j=1}^N \frac{G_n(\xi_n^j)}{\sum_{k=1}^N G_n(\xi_n^k)} \delta_{\xi_n^j}(dx)$$

- $\epsilon_n = 1 / \sup_j G_n(\xi_n^j) \rightsquigarrow$  better fitted individuals always accepted.
- Path space models  $\rightsquigarrow$  genealogical tree evolution.

## Path space models

- $X_n = (X'_0, \dots, X'_n) \rightsquigarrow$  genealogical tree/ancestral lines

$$\eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i} = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{(\xi_{0,n}^i, \xi_{1,n}^i, \dots, \xi_{n,n}^i)} \simeq_{N \uparrow \infty} \eta_n$$

- **Unbias particle approximations :**

$$\gamma_n^N(\mathbf{1}) = \prod_{0 \leq p < n} \eta_p^N(G_p) \simeq_{N \uparrow \infty} \gamma_n(\mathbf{1}) = \prod_{0 \leq p < n} \eta_p(G_p)$$

## Summary : Feynman-Kac particle sampling recipes

Nonlinear Feynman-Kac type flow  $\sim (G_n, M_n)$

$$\eta_n = \Phi_n(\eta_{n-1}) = \Psi_{G_{n-1}}(\eta_{n-1})M_n$$



- Interacting stochastic sampling algorithm

acceptance/rejection/selection/branching  $\rightsquigarrow G_n$

exploration/proposition/mutation/prediction  $\rightsquigarrow M_n$

- Normalizing constants  $\rightsquigarrow$  key multiplicative formula.
- Path space models  $\rightsquigarrow$  path-particles=ancestral lines

**Occupation meas. of genealogical trees**  $\simeq \eta_n \in$  path-space

## Some Theoretical results : TCL, PGD, PDM, ... (n, N) :

- McKean particle measure

$$\frac{1}{N} \sum_{i=1}^N \delta_{(\xi_0^i, \dots, \xi_n^i)} \simeq_N \text{Law}(\bar{X}_0, \dots, \bar{X}_n) \ \& \ \eta_n^N = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i} \simeq_N \eta_n$$

- Empirical Processes :  $\sup_{n \geq 0} \sup_{N \geq 1} \sqrt{N} \mathbb{E}(\|\eta_n^N - \eta_n\|_{\mathcal{F}_n}^p) < \infty$
- Uniform concentration inequalities :

$$\sup_{n \geq 0} \mathbb{P}(|\eta_n^N(f_n) - \eta_n(f_n)| > \epsilon) \leq c \exp\{- (N\epsilon^2)/(2\sigma^2)\}$$

- Propagations of chaos :  $\mathbb{P}_{n,q}^N := \text{Law}(\xi_n^1, \dots, \xi_n^q)$

$$\mathbb{P}_{n,q}^N \simeq \eta_n^{\otimes q} + \frac{1}{N} \partial^1 \mathbb{P}_{n,q} + \dots + \frac{1}{N^k} \partial^k \mathbb{P}_{n,q} + \frac{1}{N^{k+1}} \partial^{k+1} \mathbb{P}_{n,q}^N$$

with  $\sup_{N \geq 1} \|\partial^{k+1} \mathbb{P}_{n,q}^N\|_{\text{tv}} < \infty$  &  $\sup_{n \geq 0} \|\partial^1 \mathbb{P}_{n,q}\|_{\text{tv}} \leq c q^2$ .

# Boltzmann-Gibbs/Tail distribution flows

## Boltzmann-Gibbs measures

- $X$  r.v.  $\in (E, \mathcal{E})$  with  $\mu = \text{Law}(X)$
- Target measures associated with  $g_n : E \rightarrow \mathbb{R}_+$

$$\eta_n(dx) := \Psi_{g_n}(\mu)(dx) = \frac{1}{\mu(g_n)} g_n(x) \mu(dx)$$

Running examples :

$$\begin{aligned} g_n = 1_{A_n} &\Rightarrow \eta_n(dx) \propto 1_{A_n}(x) \mu(dx) \\ g_n = e^{-\beta_n V} &\Rightarrow \eta_n(dx) \propto e^{-\beta_n V(x)} \mu(dx) \end{aligned}$$

**Applications :** global optimization pb., tails distributions, hidden Markov chain models, etc.

# Boltzmann-Gibbs distribution flows

## Boltzmann-Gibbs distribution flows

- Target distribution flow :  $\eta_n(dx) \propto g_n(x) \mu(dx)$
- Product hypothesis :

$$g_n = g_{n-1} \times G_{n-1} \implies \eta_n = \Psi_{G_{n-1}}(\eta_{n-1})$$

Running Ex.:

$$\begin{aligned} g_n &= 1_{A_n} & \text{with } A_n \downarrow & \implies G_{n-1} = 1_{A_n} \\ g_n &= e^{-\beta_n V} & \text{with } \beta_n \uparrow & \implies G_{n-1} = e^{-(\beta_n - \beta_{n-1})V} \end{aligned}$$

- **Problem** :  $\eta_n = \Psi_{G_{n-1}}(\eta_{n-1}) = \text{unstable equation.}$

## Feynman-Kac distribution flows

## FK-stabilization

- Choose  $M_n(x, dy)$  s.t. local fixed point eq.  $\rightarrow \eta_n = \eta_n M_n$   
(Metropolis, Gibbs,...)

- Stable equation :

$$\begin{aligned} g_n = g_{n-1} \times G_{n-1} &\implies \eta_n = \Psi_{G_{n-1}}(\eta_{n-1}) \\ &\implies \eta_n = \eta_n M_n = \Psi_{G_{n-1}}(\eta_{n-1}) M_n \end{aligned}$$

- Feynman-Kac "dynamical" formulation ( $X_n$  Markov  $M_n$ )

$$\int f(x) g_n(x) \mu(dx) \propto \mathbb{E} \left( f(X_n) \prod_{0 \leq p < n} G_p(X_p) \right)$$

- $\rightsquigarrow$  Interacting Metropolis/Gibbs/... stochastic algorithms.

## Objectives - Markov processes with fixed terminal values

- $X_n$  Markov with transitions  $L(x, dy)$  on  $E$
- $\text{Law}(X_0) = \pi$  only known up to a normalizing constant.
- For a given fixed **terminal value**  $x$  solve/sample inductively the following flow of measures

$$n \mapsto \text{Law}_\pi((X_0, \dots, X_n) \mid X_n = x)$$

## FK-formulation - Markov processes with fixed terminal values

- $\pi$  "target type" measure +  $(K, L)$  pair Markov transitions

$$\text{Metropolis potential } G(x_1, x_2) = \frac{\pi(dx_2)L(x_2, dx_1)}{\pi(dx_1)K(x_1, dx_2)}$$

- Theorem [Time reversal formula] :

$$\begin{aligned} & \mathbb{E}_\pi^L(f_n(X_n, X_{n-1}, \dots, X_0) | X_n = x) \\ &= \frac{\mathbb{E}_x^K(f_n(X_0, X_1, \dots, X_n) \{\prod_{0 \leq p < n} G(X_p, X_{p+1})\})}{\mathbb{E}_x^K(\{\prod_{0 \leq p < n} G(X_p, X_{p+1})\})} \end{aligned}$$

- $\rightsquigarrow$  time reversal genealogical tree simulation

## Rare event excursions

- $(E = A \cup A^c)$ ,  $Y_n$  Markov,  $C \subset A^c$  absorbing set

$$Y_0 \in A_0(\subset A) \rightsquigarrow A^c = (B \cup C)$$

- Objectives :

$$\mathbb{P}(Y \text{ hits } B \text{ before } C) \quad \text{and} \quad \text{Law}(Y \mid Y \text{ hits } B \text{ before } C)$$

## Multi-splitting rare events

- *Multi-level decomposition*  $B_0 \supset B_1 \supset \dots \supset B_m = B$   
( $A_0 = B_1 - B_0$ ,  $B_0 \cap C = \emptyset$ )
- *Inter-level excursions* :  $T_n = \inf \{p \geq T_{n-1} : Y_p \in B_n \cup C\}$

$$X_n = (Y_p ; T_{n-1} \leq p \leq T_n) \quad \text{and} \quad G_n(X_n) = 1_{B_n}(Y_{T_n})$$

Feynman-Kac formulations :

$$\mathbb{P}(Y \text{ hits } B \text{ before } C) = \mathbb{E}\left(\prod_{1 \leq p \leq m} G_p(X_p)\right)$$

$$\mathbb{E}(f(Y_0, \dots, Y_{T_m}) 1_{B_m}(Y_{T_m})) = \mathbb{E}(f(X_0, \dots, X_m) \prod_{1 \leq p \leq m} G_p(X_p))$$

↪ genealogical tree in excursion space.

## Fixed time level set entrances

### Fixed time level set entrances

- $X_n$  Markov  $\in E_n$ ,  $V_n : E_n \rightarrow \mathbb{R}_+$ ,  $a \in \mathbb{R}$
- Objectives :

$$\mathbb{P}(V_n(X_n) \geq a) \quad \text{and} \quad \text{Law}((X_0, \dots, X_n) \mid V_n(X_n) \geq a)$$

# Large deviation analysis

## Large deviation analysis

$$\begin{aligned} \mathbb{P}(V_n(X_n) \geq a) &\stackrel{\forall \lambda}{=} \mathbb{E} \left( \mathbf{1}_{V_n(X_n) \geq a} e^{\lambda V_n(X_n)} e^{-\lambda V_n(X_n)} \right) \\ &\leq e^{-(\lambda a - \Lambda_n(\lambda))} \text{ with } \Lambda_n(\lambda) = \log \mathbb{E}(e^{\lambda V_n(X_n)}) \end{aligned}$$

$$\text{Ex.: } V_n(X_n) = X_n \quad \text{and} \quad \Delta X_n = N(0, 1) \implies \lambda^* = a/n$$

Twisted measures :

$$\eta_n(dx_n) \propto e^{\lambda V_n(x_n)} \mathbb{P}(X_n \in dx_n) := \gamma_n(dx_n)$$

$$\rightsquigarrow \gamma_n(1) = \mathbb{E}(e^{\lambda V_n(X_n)})$$

$$\implies \mathbb{P}(V_n(X_n) \geq a) = \eta_n(\mathbf{1}_{V_n \geq a} e^{-\lambda V_n}) \times \gamma_n(1)$$

## Feynman-Kac representation formula

Feynman-Kac twisted measures ( $V_{-1} = 0$ )

$$\mathbb{E}(f_n(X_n) e^{\lambda V_n(X_n)}) = \mathbb{E} \left( f_n(X_n) \prod_{0 \leq p \leq n} e^{\lambda(V_p(X_p) - V_{p-1}(X_{p-1}))} \right)$$

and

$$\begin{aligned} & \mathbb{E}(f_n(X_0, \dots, X_n) \mid V_n(X_n) \geq a) \\ & \propto \\ & \mathbb{E} \left( T_n(f_n)(X_0, \dots, X_n) \prod_{0 \leq p \leq n} e^{\lambda(V_p(X_p) - V_{p-1}(X_{p-1}))} \right) \end{aligned}$$

with

$$T_n(f_n)(X_0, \dots, X_n) = f_n(X_0, \dots, X_n) e^{-\lambda V_n(X_n)} \mathbf{1}_{V_n(X_n) \geq a}$$

↔ genealogical tree interpretation model in transition space

## Some references

### Interacting stochastic simulation algorithms

- **Mean field and Feynman-Kac particle models :**
  - joint work with F. Patras and S. Rubenthaler Coalescent tree based functional representations for some Feynman-Kac particle models, Preprint Univ. Nice-Sophia Antipolis (2006).
  - Feynman-Kac formulae. Genealogical and interacting particle systems, Springer (2004)  $\oplus$  Refs.
  - joint work with L. Miclo. Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae. *Séminaire de Probabilités XXXIV, Lecture Notes in Mathematics, Springer-Verlag Berlin, Vol. 1729, 1-145 (2000).*

## Some references

### Interacting stochastic simulation algorithms

- **Sequential Monte Carlo models :**
  - joint work with Doucet A., Jasra A. Sequential Monte Carlo Samplers. JRSS B (2006).
  - joint work with A. Doucet. On a class of genealogical and interacting Metropolis models. Sémin. de Proba. 37 (2003).

## Some references

### Particle rare event simulation algorithms

- **Twisted Feynman-Kac measures**

- joint work with J. Garnier. Genealogical Particle Analysis of Rare events. *Annals of Applied Probab.*, 15-4 (2005).
- joint work with J. Garnier. Simulations of rare events in fiber optics by interacting particle systems. *Optics Communications*, Vol. 267 (2006).

- **Multi splitting excursion models**

- joint work with P. Lezaud. Branching and interacting particle interpretation of rare event proba.. *Stochastic Hybrid Systems : Theory and Safety Critical Applications*, eds. H. Blom and J. Lygeros. Springer (2006).
- joint work with F. Cerou, Le Gland F., Lezaud P. *Genealogical Models in Entrance Times Rare Event Analysis, Alea*, Vol. I, (2006).

## Some references

### Interacting stochastic simulation algorithms

- **i-MCMC algorithms :**

- joint work with A. Doucet. Interacting Markov Chain Monte Carlo Methods For Solving Nonlinear Measure-Valued Eq., HAL-INRIA RR-6435, (Feb. 2008).
- joint work with B. Bercu and A. Doucet. Fluctuations of Interacting Markov Chain Monte Carlo Models. HAL-INRIA RR-6438, (Feb. 2008).
- joint work with C. Andrieu, A. Jasra, A. Doucet. *Non-Linear Markov chain Monte Carlo via self-interacting approximations*. Tech. report, Dept of Math., Bristol Univ. (2007).
- joint work with A. Brockwell and A. Doucet. *Sequentially interacting Markov chain Monte Carlo*. Tech. report, Dept. of Statistics, Univ. of British Columbia (2007).