Genetic type particle methods: An introduction with applications

P. Del Moral

Centre INRIA Bordeaux-Sud Ouest

EVOLVE 2011, Luxembourg May 25-27 2011

↪ Feynman-Kac formulae. Genealogical and interacting particle systems, Springer (2004), + Ref.
↪ DM, Doucet, Jasra. SMC Samplers. *JRSS B (2006).*
↪ DM, N. G. Hadjiconstantinou. An introduction to probabilistic methods with applications + *Ref.*.
↪ DM, A. Doucet. Particle Methods: An introduction with applications. *HAL-INRIA RR-6991(09), 2008 MLSS, Springer (2011?).*

Summary

Particle Interpretation models

- **Mathematical physics and molecular chemistry** ($\geq 1950's$) : Particle/microscopic interpretation models, particle absorption, macro-molecular chains, quantum and diffusion Monte Carlo.
- **Environmental studies and biology** ($\geq 1950's$): Population, gene evolutions, species genealogies, branching/birth and death models.
- **Evolutionary mathematics and engineering sciences** ($\geq 1970's$): Adaptive stochastic search method, evolutionary learning models, interacting stochastic grids approximations, genetic algorithms.
- **Applied Probability and Bayesian Statistics** ($\geq 1990's$): Approximating simulation technique (recursive acceptance-rejection model), Sequential Monte Carlo, `http-ref : interacting Monte Carlo Markov chains (Andrieu, Bercu, DM, Doucet, Jasra).`
- **Pure mathematics ($\geq 1960's$ for fluid models, $\geq 1990's$ for discrete time and interacting jump models)**: Stochastic linearization tech., mean field particle interpretations of nonlinear PDE and measure valued equations.

- **Central idea of particle/SMC in stochastic engineering :**

$$\left\{ \begin{array}{l} \textbf{Physical and Biological intuitions} \\ \textit{[learning, adaptation, optimization,...]} \end{array} \right\} \in \textbf{Engineering problems}$$

Sequential Monte Carlo	Sampling	Resampling
Particle Filters	Prediction	Updating
Genetic Algorithms	Mutation	Selection
Evolutionary Population	Exploration	Branching
Diffusion Monte Carlo	Free evolutions	Absorption
Quantum Monte Carlo	Walkers motions	Reconfiguration
Sampling Algorithms	Transition proposals	Acceptance-rejection

More botanical names : spawning, cloning, pruning, enrichment, go with the winner, replenish, and many others.

- **Pure mathematical point of view :**
  = Mean field particle interpretation of Feynman-Kac measures

Some application areas of Feynman-Kac formulae

- **Physics** :
  - Feynman-Kac-Schroedinger semigroups $\in$ nonlinear integro-differential equations ($\sim$ generalized Boltzmann models).
  - Spectral analysis of Schrödinger operators and large matrices with nonnegative entries.
  - Particle evolutions in disordered/absorbing media.
  - Multiplicative Dirichlet problems with boundary conditions.
  - Microscopic and macroscopic interacting particle interpretations.
- **Chemistry and Biology**:
  - Self-avoiding walks, macromolecular simulation, directed polymers.
  - Spatial branching and evolutionary population models.
  - Coalescent and Genealogical tree based evolutions.

Some application areas of Feynman-Kac formulae

- **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).
- **Advanced Signal processing**:
  - Optimal filtering, prediction, smoothing.
  - Open loop optimal control, optimal regulation.
  - Interacting Kalman-Bucy filters.
  - Stochastic and adaptative grid approximation-models
- **Statistics/Probability**:
  - Restricted Markov chains (w.r.t terminal values, visiting regions, constraints simulation problems,...)
  - Analysis of Boltzmann-Gibbs type distributions (simulation, partition functions, localization models...).
  - Random search evolutionary algorithms, interacting Metropolis/simulated annealing algo, combinatorial counting.

The filtering problem $\subset$ Bayesian statistics

- X_t**:=Signal=Stochastic process**

Engineering/physics/biology/economics :

- Non cooperative targets (defense : missile, boat, plane,...).
- Physics (Fluids : twisters, cyclones, ocean models, pressure/temperature/diffusion coefficients,...).
- Finance (assets, portfolios, volatilities, default indexes,...).
- Signal (speech, codes, informations transmissions, waves,...).

Dynamics and sources of randomness :

- Physical evolution equations (example : $\sum_i u_i \vec{F}_i = \vec{A}$)
- Perturbations and random sources:
  - Model uncertainties $\oplus$ External perturbations.
  - **Unknown controls and related model parameters.**

$\rightsquigarrow$ **A Priori Law/Knowledge** (unknown quantities=random samples.)

The filtering model

- Y_t=**Partial and Noisy observations of the signal** X_t **:**

Engineering/physics/biology/economics :

- Engineering : Radar, Sonar, GPS, ...
- Physics (sensors : pressure/temperature/...).
- Finance (assets, portfolios,...).
- Statistics (real data: medecine, pharmacology, politics, economics,...).

Dynamics and sources of randomness :

- Partial observations : complex mixtures, partial coordinates.
- Perturbations et random sources :
  - Noisy sensor measures (thermal noise).
  - External/environmental perturbations.
  - Model uncertainties.

Objectives

Compute/Sample/Estimate inductively the flow of measures

$$t \in \mathbb{R}_+ \quad \text{or} \quad t = n \in \mathbb{N} \longrightarrow \eta_t = \mathrm{Law}(X_t \mid Y_0, \ldots, Y_t)$$

Note

- Filtering the trajectories : $X_t = (X'_0, \ldots, X'_t) \in E_t$

$$\Updownarrow \textbf{[State space enlargement]}$$

$$\eta_t = \mathrm{Law}((X'_0, \ldots, X'_t) \mid (Y_0, \ldots, Y_t)) = \mathrm{Law}(X_t \mid Y_0, \ldots, Y_t)$$

Equivalent terminologies :

- Data Assimilation (forecasting, fluids/ocean models).
- Hidden Markov Chains Models (HMM).
- A Posteriori Law=Law($X|Y$) (A Priori=$\mathrm{Law}(X)$).

Heuristic particle filters

Sample a population of N "individuals"/particles" s.t. at **any time**

$$(\widehat{\xi}_t^1, \ldots, \widehat{\xi}_t^N) \in E_t^N \rightsquigarrow \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{\widehat{\xi}_t^i} = \mathrm{Law}(X_t \mid (Y_0, \ldots, Y_t))$$

Heuristic learning/filtering scheme :

- Prediction/Exploration $\rightsquigarrow$ sampling N local transitions of the signal.
- Updating/Correction $\rightsquigarrow$ birth and death process $=$ branching particle algo (fixed size N).
  - Kill/stop individuals/proposal with poor likelihood value.
  - Multiply/increase individuals with high likelihood value.

$\rightsquigarrow$ Path space models : $X_t = (X'_0, \ldots, X'_t)$

$\Rightarrow$ **Genealogical tree based learning algorithm :**

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{\text{i-th ancestral line(t)}} = \mathrm{Law}((X'_0, \ldots, X'_t) \mid (Y_0, \ldots, Y_t))$$

Some typical rare events

- Physical/biological/economical stochastic process : atomic/molecular configurations fluctuations, queueing evolutions, communication network, portfolio and financial assets, ...
- Potential function-Event restrictions : Energy/Hamiltonian potential functions, overflows levels, critical thresholds, epidemic propagations, radiation dispersion, ruin levels.

Objectives

Rare event probabilities & the law of the process $\in$ critical regime

Particle heuristic model

Default tree model $=$ Branching particle genealogical tree model (Branching on "more likely" gateways to critical regimes)

Event restrictions and confinements

- **Non intersecting simple random walks on $\mathbb{Z}^d$**

$$\mathbb{P}\left(\forall p < q \leq n,\ X_p \neq X_q\right) = \frac{1}{(2d)^n} \times \#\{\text{not} \cap \text{ walks length n}\}$$
$$\simeq \exp\left(c\ n\right)$$

$$\text{Law}((X_0, \ldots, X_n) \mid \forall p < q \leq n \quad X_p \neq X_q)$$

- **Confinement model/Lyap. exp. and top eigenval.**

$$\mathbb{P}(\forall 0 \leq p \leq n\ \ X_p \in A) \simeq \exp\left(-\lambda(A)\ n\right)$$

$$\text{Law}((X_0, \ldots, X_n) \mid \forall 0 \leq p \leq n \quad X_p \in A)$$

- **Tube confinement :** as above with $(X_p \in A) \rightsquigarrow (X_p \in A_p)$

Heuristic particle model :

$\rightsquigarrow$ Accept-Reject interacting X-motions

Terminal levels conditioning and excursion models

1. **Terminal level set conditioning :**

$$\mathbb{P}(V_n(X_n) \geq a) \quad \& \quad \mathrm{Law}((X_0, \ldots, X_n) \mid V_n(X_n) \geq a)$$

2. **Fixed terminal value :** $\mathrm{Law}_{\pi,K}((X_0, \ldots, X_n) \mid X_n = x_n)$.
3. **Critical excursion behavior :**
   $\mathbb{P}(X \text{ hits } B \text{ before } C) \quad \& \quad \mathrm{Law}(X \mid X \text{ hits } B \text{ before } C)$

Heuristic particle models :

1. Interacting X-transitions increasing the potential V_n.
2. Interacting M-transitions increasing the Metropolis type potential ratio $\frac{\pi(dx_2)K(x_2,dx_1)}{\pi(dx_1)M(x_1,dx_2)}$
3. Interacting X-excursions on gateways levels $\rightsquigarrow B$.

A pair of target Boltzmann-Gibbs measures

1. $\eta_n(dx) \propto e^{-\beta_n V(x)}\ \lambda(dx)$ with $\beta_n \uparrow$
2. $\eta_n(dx) \propto 1_{A_n}(x)\ \lambda(dx)$ with $A_n \downarrow$
3. Normalizing constants $\lambda(e^{-\beta_n V})$ and $\lambda(A_n)$

Heuristic particle models :

1. $e^{-(\beta_{n+1}-\beta_n)V}$-interacting MCMC moves with local targets η_n
2. A_{n+1}-interacting MCMC moves with local targets η_n
3. Time product of the empirical interaction potential functions.

Summary

Previous heuristic type models $\subset$ A single Feynman-Kac formulation:

$$d\mathbb{Q}_n = \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0 \leq p < n} G_p(X_p) \right\} d\mathbb{P}_n^X$$

$$\stackrel{G_n = 1_{A_n}}{=} \text{Law}((X_0, \ldots, X_n) \mid X_p \in A_p,\ 0 \leq p < n)$$

and $\mathcal{Z}_n = \mathbb{P}(X_p \in A_p,\ 0 \leq p < n)$

Note : $\eta_n :=$ The n-th time marginal $\stackrel{G_n = 1_{A_n}}{=}$ $\text{Law}(X_n \mid X_p \in A_p,\ 0 \leq p < n)$
$\eta_n =$ "nonlinear" transformation of the proba. meas. η_{n-1}

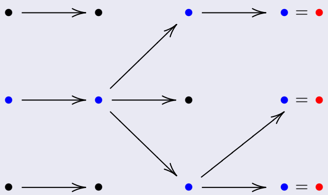
$$\left\{ \prod_{0 \leq p \leq n} G_p(X_p) \right\} = \left\{ \prod_{0 \leq p \leq (n-1)} G_p(X_p) \right\} \times G_n(X_n)$$

Same heuristic $\sim$ multiplicative structure :

$\rightsquigarrow$ (Accept-Reject) G-interacting X-motions [and inversely!]

Interaction/branch. process $\hookrightarrow$ 4 types of occupation measures

$(N=3)$



- Current population $\hookrightarrow \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_n^i \leftarrow i\text{-th individual at time } n} \simeq \eta_n$

- Genealogical tree $\hookrightarrow \frac{1}{N}\sum_{i=1}^{N} \delta_{(\xi_{0,n}^i,\xi_{1,n}^i,\ldots,\xi_{n,n}^i) \leftarrow i\text{-th ancestral line}} \simeq \mathbb{Q}_n$

- **Normalizing constants** [Unbias estimates]

$$\mathcal{Z}_n^N := \prod_{0\leq p<n} \frac{1}{N}\sum_{1\leq i\leq N} G_p(\xi_p^i) \simeq \mathcal{Z}_n$$

Forward particle approximation $\sim$ complete genealogical tree

Mutation transition with a density

$$M_n(x, dx') = H_n(x, x')\ \lambda(dx')$$

$$\Downarrow$$

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

with the backward Markov transitions

$$\mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^i, dx_{n-1}) = \sum_{1 \leq j \leq N} \frac{G_{n-1}(\xi_{n-1}^j) H_n(\xi_{n-1}^j, \xi_n^i)}{\sum_{1 \leq k \leq N} G_{n-1}(\xi_{n-1}^k) H_n(\xi_{n-1}^k, \xi_n^i)}\ \delta_{\xi_{n-1}^j}(dx_{n-1})$$

Some convergence results

Notation (Monte Carlo approximations)

$$\mu(f) := \int f(x)\ \mu(dx) = \mathbb{E}(f(X)) \simeq \mu^N(f) = \frac{1}{N}\sum_{1\leq i\leq N} f(X^i)$$

Proba to have any of the following estimates is greater than $1-2e^{-\lambda}$, $\forall\lambda$

- Current population & time marginal models

$$\left|\eta_n^N(f)-\eta_n(f)\right| \leq 2(1+\sqrt{2\lambda})\ c/\sqrt{N} \qquad \textbf{(Uniform w.r.t. time)}$$

- Genealogical trees & Path-space measures

$$\left|\mathbb{Q}_n^N(F)-\mathbb{Q}_n(F)\right| \leq \frac{\mathbf{n}}{N}\ c_1\left(1+2(\lambda+\sqrt{\lambda})\right)+\sqrt{\frac{\mathbf{n}\lambda c_2}{N}}$$

- Normalizing constants

$$\left|\frac{\mathcal{Z}_n^N}{\mathcal{Z}_n}-1\right| \leq \mathbf{n}\left(\frac{c_1}{N}\ \left(1+2(\lambda+\sqrt{\lambda})\right)+\sqrt{\frac{\lambda c_2}{N}}\right)$$

Backward particle models

Test functions = Additive functionals

$$\overline{F}_n(x_0,\ldots,x_n) = \frac{1}{n+1}\sum_{p=0}^{n} f_p(x_p)$$

Proba to have the following estimates is greater than $1-2e^{-\lambda}$, $\forall\lambda$

$$\left|\mathbb{Q}_n^N(\overline{F}_n) - \mathbb{Q}_n(\overline{F}_n)\right| \leq c\ \left(\frac{1}{N}\ \left(1+2(\lambda+\sqrt{\lambda})\right) + \sqrt{\frac{\lambda}{2N(n+1)}}\right)$$

http - references & Web links resources

- Master lecture notes on Stochastic engineering with scilab programs (in french)
- A pedagogical book on simulation and stochastic algorithms (in french)
- A series of selected research articles on Feynman-Kac models and particle algorithms : convergence, performance analysis, fluctuations, large deviations, propagations of chaos properties, exponential estimates. (see also more recent articles)
- Some web-links to Feynman-Kac and Interacting particle application model areas : particle filtering, robotics, image processing, audio signal, tracking, GPS, fluid mechanics, financial math, biology, chemistry, rare event, optics, hybrid systems,...