

# A new class of interacting Markov Chain Monte Carlo methods

**P. Del Moral, A. Doucet**

INRIA Bordeaux & UBC Vancouver

Workshop on Numerics and Stochastics, Helsinki, August 2008

# Outline

- 1 Introduction
  - Stochastic sampling problems
  - Some stochastic engineering models
  - Interacting sampling methods
- 2 Some classical distribution flows
  - Feynman-Kac models
  - "Static" Boltzmann-Gibbs models
- 3 Interacting stochastic sampling technology
  - Mean field particle methods
  - Interacting Markov chain Monte Carlo models (i-MCMC)
- 4 Functional fluctuation & comparisons
- 5 Some references

## Stochastic sampling problems

- "Nonlinear" distribution flow with  $\uparrow$  level of complexity.

$$\eta_n(dx_n) = \frac{\gamma_n(dx_n)}{\gamma_n(1)} \quad \text{Time index } n \in \mathbb{N} \quad \text{State var. } x_n \in E_n$$

- Two objectives :

- 1  $\sim$  "Sampling independent" random variables w.r.t.  $\eta_n$
- 2 Computation of the normalizing constants  $\gamma_n(1)$   
(=  $\mathcal{Z}_n$  Partition functions).

## Stochastic engineering $\rightsquigarrow$ Conditional & Boltzmann-Gibbs' measures

- **Filtering:** Signal-Observation  $(X_n, Y_n)$  [Radar, Sonar, GPS, ...]

$$\eta_n = \text{Law}(X_n \mid (Y_0, \dots, Y_n))$$

- **Rare events:** [Overflows, ruin processes, epidemic propagations, ...]

$$\eta_n = \text{Law}(X_n \mid n \text{ intermediate events}) \ \& \ \mathcal{Z}_n = \mathbb{P}(\text{Rare event})$$

- **Molecular simulation:** [ground state energies, directed polymers...]

$$\eta_n := \text{Feynman-Kac/Boltzmann-Gibbs}$$

$$\sim \text{Free Markov motion in an absorbing medium}$$

- **Combinatorial counting, Global optimization, HMM**

$$\eta_n = \frac{1}{\mathcal{Z}_n} e^{-\beta_n V(x)} \lambda(dx) \quad \text{or} \quad \eta_n = \frac{1}{\mathcal{Z}_n} 1_{A_n}(x) \lambda(dx)$$

## Two simple ingredients

- Find or Understand the probability mass transformation

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

~ Cooling schemes, temp. variations, constraints sequences, subset restrictions, observation data, conditional events,...

- Natural interacting sampling idea :

Use  $\eta_{n-1}$  or its empirical approx. to sample w.r.t.  $\eta_n$

- Monte-Carlo/ Mean Field models :

$$\eta_n = \text{Law}(\bar{X}_n) \quad \text{with} \quad \text{Markov} : \bar{X}_{n-1} \xrightarrow{\sim \eta_{n-1}} \bar{X}_n$$

- Interacting MCMC models :

$$\left\{ \begin{array}{l} \text{Use the occupation measures} \\ \text{of an MCMC with target } \eta_{n-1} \end{array} \right\} \rightsquigarrow \text{MCMC target } \eta_n$$

## Feynman-Kac distribution flows

- **Weak representation:**  $[f_n$  test funct. on a state space  $E_n]$

$$\eta_n(f_n) = \frac{\gamma_n(f_n)}{\gamma_n(\mathbf{1})} \quad \text{with} \quad \gamma_n(f_n) = \mathbb{E} \left( f_n(X_n) \prod_{0 \leq p < n} G_p(X_p) \right)$$

- **A Key Formula:**  $Z_n = \mathbb{E} \left( \prod_{0 \leq p < n} G_p(X_p) \right) = \prod_{0 \leq p < n} \eta_p(G_p)$
- **Path space models**  $X_n = (X'_0, \dots, X'_n)$

## Examples

- $G_n \in [0, 1] \rightsquigarrow$  particle absorption models.
- $G_n =$  Observation likelihood function  $\rightsquigarrow$  Filtering models.
- $G_n = 1_{A_n} \rightsquigarrow$  Conditional/Restriction models.

## Nonlinear distribution flows

### Evolution equation:

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

### With the only 2 transformations :

- **X-Free Markov transport eq. :**  $[M_n(x_{n-1}, dx_n)$  from  $E_{n-1}$  into  $E_n]$

$$(\eta_{n-1}M_n)(dx_n) := \int_{E_{n-1}} \eta_{n-1}(dx_{n-1}) M_n(x_{n-1}, dx_n)$$

- **Bayes-Boltzmann-Gibbs transformation :**

$$\Psi_{G_n}(\eta_n)(dx_n) := \frac{1}{\eta_n(G_n)} G_n(x_n) \eta_n(dx_n)$$

## Boltzmann-Gibbs distribution flows

- Target distribution flow :  $\eta_n(dx) \propto g_n(x) \lambda(dx)$
- 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 modeling

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

- **Stable equation :**

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

$$\implies \eta_n = \eta_n M_n = \Psi_{G_{n-1}}(\eta_{n-1}) M_n = \text{FK-model}$$

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

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

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

## Mean field interpretation

- **Nonlinear Markov models** : Always  $\exists K_{n,\eta}(x, dy)$  Markov s.t.

$$\eta_n = \Phi_n(\eta_{n-1}) = \eta_{n-1} K_{n,\eta_{n-1}} = \text{Law}(\bar{X}_n)$$

i.e. :

$$\mathbb{P}(\bar{X}_n \in dx_n \mid \bar{X}_{n-1}) = K_{n,\eta_{n-1}}(\bar{X}_{n-1}, dx_n)$$

## Mean field particle interpretation

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

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

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

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

# Discrete generation mean field particle model

Schematic picture :  $\xi_n \in E_n^N \rightsquigarrow \xi_{n+1} \in E_{n+1}^N$

$$\begin{array}{ccc}
 \xi_n^1 & \xrightarrow{K_{n+1, \eta_n^N}} & \xi_{n+1}^1 \\
 \vdots & & \vdots \\
 \xi_n^i & \longrightarrow & \xi_{n+1}^i \\
 \vdots & & \vdots \\
 \xi_n^N & \longrightarrow & \xi_{n+1}^N
 \end{array}$$

Rationale :

$$\begin{aligned}
 \eta_n^N \simeq_{N \uparrow \infty} \eta_n &\implies K_{n+1, \eta_n^N} \simeq_{N \uparrow \infty} K_{n+1, \eta_n} \\
 &\implies \xi_n^i \text{ almost iid copies of } \bar{X}_n
 \end{aligned}$$

## Ex.: Feynman-Kac distribution flows

- **FK-Nonlinear Markov models :**

$\epsilon_n = \epsilon_n(\eta_n) \geq 0$  s.t.  $\eta_n$ -a.e.  $\epsilon_n G_n \in [0, 1]$  ( $\epsilon_n = 0$  not excluded)

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

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

- **Mean field genetic type particle model :**

$$\xi_n^i \in E_n \xrightarrow{\text{accept/reject/selection}} \widehat{\xi}_n^i \in E_n \xrightarrow{\text{proposal/mutation}} \xi_{n+1}^i \in E_{n+1}$$

- **Examples :**

- $G_n = 1_A \rightsquigarrow$  killing with uniform replacement.
- $M_n$ -Metropolis/Gibbs moves  $\rightsquigarrow$   $G_n$ -interaction function (subsets fitting or change of temperatures)

## Mean field genetic type particle model :

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

Accept/Reject/Selection transition :

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

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

**Ex. :**  $G_n = 1_A$ ,  $\epsilon_n = 1 \rightsquigarrow G_n(\xi_n^i) = 1_A(\xi_n^i)$

## 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 :**

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

- **Ex.**  $G_n = 1_A :$

$$\Rightarrow \gamma_n^N(1) = \prod_{0 \leq p < n} (\text{success \% at } p)$$

- **FK-Mean field particle models** = *sequential Monte Carlo, population Monte Carlo, particle filters, pruning, spawning, reconfiguration, quantum Monte carlo, go with the winner...*

## Objective

- Find a series of MCMC models  $X^{(n)} := (X_k^{(n)})_{k \geq 0}$  s.t.

$$\eta_k^{(n)} = \frac{1}{k+1} \sum_{0 \leq l \leq k} \delta_{X_l^{(n)}}$$

$$\simeq_{k \uparrow \infty} \eta_n$$

$\Rightarrow$  Use  $\eta_k^{(n)} \simeq \eta_n$  to define  $X^{(n+1)}$  with target  $\eta_{n+1}$

## Advantages

- Using  $\eta_n$  the sampling  $\eta_{n+1}$  is often easier.
- Improve the proposition step in any Metropolis type model with target  $\eta_{n+1}$  ( $\rightsquigarrow$  enters the stability prop. of the flow  $\eta_n$ )
- Increases the precision at every time step.  
But CLT variance often  $\geq$  CLT variance mean field models.
- Easy to combine with mean field stochastic algorithms.

## Interacting Markov chain Monte Carlo models

- Find  $M_0$  and a collection of transitions  $M_{n,\mu}$  s.t.

$$\eta_0 = \eta_0 M_0 \quad \text{and} \quad \Phi_n(\mu) = \Phi_n(\mu) M_{n,\mu}$$

- $(X_k^{(0)})_{k \geq 0}$  Markov chain  $\sim M_0$ .
- Given  $X^{(n)}$ , we let  $X_k^{(n+1)}$  with Markov transitions  $M_{n+1, \eta_k^{(n)}}$

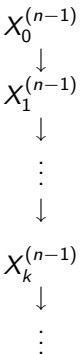
**Rationale :**

$$\begin{aligned} \eta_k^{(n)} \simeq \eta_n &\implies \begin{cases} \Phi_{n+1}(\eta_k^{(n)}) \simeq \Phi_{n+1}(\eta_n) = \eta_{n+1} \\ M_{n+1, \eta_k^{(n)}} \simeq M_{n+1, \eta_n} \quad \text{with fixed point } \eta_{n+1} \end{cases} \\ &\implies \eta_k^{(n+1)} \simeq \eta_{n+1} \end{aligned}$$

**Example :**  $M_{n,\mu}(x, dy) = \Phi_n(\mu)(dy) \rightsquigarrow X_k^{(n+1)}$  r.v.  $\sim \Phi_{n+1}(\eta_k^{(n)})$

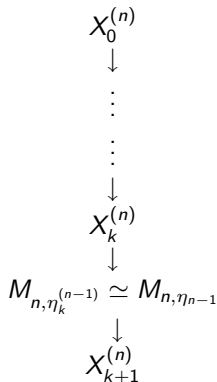


$((n-1)$ -th chain)



$$\xrightarrow{\eta_k^{(n-1)} \simeq \eta_{n-1}}$$

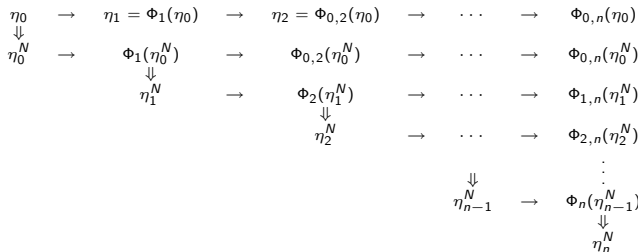
$(n$ -th chain)



[MEAN FIELD PARTICLE MODEL] Nonlinear semigroup  $\longrightarrow \Phi_{p,n}(\eta_p) := \eta_n$

Local fluctuation theorem :  $W_n^N := \sqrt{N} [\eta_n^N - \Phi_n(\eta_{n-1}^N)] \simeq W_n \perp$  Centered Gaussian field

Local transport formulation :



$\rightsquigarrow$  Key decomposition formula :

$$\begin{aligned}
 \eta_n^N - \eta_n &= \sum_{q=0}^n [\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))] \\
 &\simeq \frac{1}{\sqrt{N}} \sum_{q=0}^n W_q^N D_{q,n} \leftrightarrow \text{First order decomp. } \Phi_{p,n}(\eta) - \Phi_{p,n}(\mu) \simeq (\eta - \mu)D_{p,n} + (\eta - \mu)^{\otimes 2} \dots
 \end{aligned}$$

$$\Rightarrow \text{Example Functional CLT : } \sqrt{N} [\eta_n^N - \eta_n] \simeq \sum_{q=0}^n W_q D_{q,n}$$

[i-MCMC] **Nonlinear** sg  $\Phi_{p,n}(\eta_p) = \eta_n$  with a first order decomp. :

$$\Phi_{p,n}(\eta) - \Phi_{p,n}(\mu) \simeq (\eta - \mu)D_{p,n} + (\eta - \mu)^{\otimes 2} \dots$$

↓

**Functional CLT for correlated/interacting MCMC models :**

$$\sqrt{k} [\eta_k^{(n)} - \eta_n] \simeq \sum_{q=0}^n \frac{\sqrt{(2(n-q))!}}{(n-q)!} V_q D_{q,n}$$

with  $(V_q)_{q \geq 0} \perp$  Centered Gaussian field

$$\mathbb{E} \left( V_q(f)^2 \right) = \eta_q \left[ (f - \eta_q(f))^2 \right] + 2 \sum_{m \geq 1} \eta_q \left[ (f - \eta_q(f)) M_{q, \eta_{q-1}}^m (f - \eta_q(f)) \right]$$

"Comparisons" : **[Mean field case]**  $(W_q)_{q \geq 0} \perp$  Centered Gaussian field

$$\mathbb{E} \left( W_q(f)^2 \right) = \eta_{q-1} \left\{ K_{q, \eta_{q-1}} (f - K_{q, \eta_{q-1}}(f))^2 \right\}$$

**Case :**  $K_{q, \eta}(x, dy) = M_{q, \eta}(x, dy) = \Phi_q(\eta)(dy) \implies (V_q = W_q) \implies$  **[Mean field] > [i-MCMC]**

## Some references

### Interacting stochastic simulation algorithms

- **Mean field and Feynman-Kac particle models :**

- Feynman-Kac formulae. Genealogical and interacting particle systems, Springer (2004)  $\oplus$  Refs.
- joint work with L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Processes and their Applications*, Vol. 86, 193-216 (2000).
- 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).

- **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. *Sém. de Proba.* 37 (2003).

## 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).