

# Concentration inequalities for Sequential Monte Carlo models

P. Del Moral

INRIA Bordeaux & IMB & CMAP X

(Recent) Advances in SMC 2012, Warwick Univ. Sept. 2012

## Some hyper-refs

- ▶ On the concentration of interacting processes. *Foundations & Trends in Machine Learning* [170p.] (2012). (joint work with Peng Hu & Liming Wu) [+ Refs]
- ▶ A Backward Particle Interpretation of Feynman-Kac Formulae *M2AN* (2010). (joint work with A. Doucet & S.S. Singh)
- ▶ Feynman-Kac formulae, *Genealogical & Interacting Particle Systems with appl.*, Springer (2004).
- ▶ Sequential Monte Carlo Samplers *JRSS B.* (2006). (joint work with A. Doucet & A. Jasra)
- ▶ Concentration Inequalities for Mean Field Particle Models. *Annals of Applied Probability* (2011) (joint work with Emmanuel Rio).
- ▶ More references on the website <http://www.math.u-bordeaux1.fr/~delmoral/index.html> [+ Links]

## Feynman-Kac models

Some basic notation

Description of the models

Some examples

Interacting particle interpretations

Concentration inequalities

## Basic notation

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

- ▶  $(\mu, f) \in \mathcal{P}(E) \times \mathcal{B}(E) \longrightarrow \mu(f) = \int \mu(dx) f(x)$
- ▶  $Q(x_1, dx_2)$  **integral operators**  $x_1 \in E_1 \rightsquigarrow x_2 \in E_2$

$$Q(f)(x_1) = \int Q(x_1, dx_2) f(x_2)$$

$$[\mu Q](dx_2) = \int \mu(dx_1) Q(x_1, dx_2) \quad (\implies [\mu Q](f) = \mu[Q(f)] )$$

# Basic notation

- ▶ **Boltzmann-Gibbs transformation**  $G \geq 0$

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

When  $G \leq 1$  we have the (non unique) Markov transport equation

$$\Psi_G(\mu) = \mu S_{\mu,G}$$

with

$$S_{\mu,G}(x, dy) = G(x) \delta_x(dy) + (1 - G(x)) \Psi_G(\mu)(dy)$$

# Feynman-Kac models

Markov chain  $X_n \in E_n$  transition  $M_n$  and a potential function  $G_n \geq 0$

$$d\mathbb{Q}_n := \frac{1}{Z_n} \left\{ \prod_{0 \leq p < n} G_p(X_p) \right\} d\mathbb{P}_n \quad \text{with} \quad \mathbb{P}_n = \text{Law}(X_0, \dots, X_n)$$

- ▶ Equivalent recursive formulation

$$\mathbb{Q}_n(d(x_0, \dots, x_n)) \propto \mathbb{Q}_{n-1}(d(x_0, \dots, x_{n-1})) \times \underbrace{Q_n(x_{n-1}, dx_n)}_{=G_{n-1}(x_{n-1})M_n(x_{n-1}, dx_n)}$$

- ▶ Transition/Excursions/Path spaces

$$X_n = (X'_n, X'_{n+1}) \quad X_n = X'_{[T_n, T_{n+1}]} \quad X_n = (X'_0, \dots, X'_n)$$

- ▶  $\supset$  Continuous time models

$$X_n = X'_{[t_n, t_{n+1}]} \quad \& \quad G_n(X_n) = \exp \int_{t_n}^{t_{n+1}} V_t(X'_t) dt$$

## Some examples

- ▶ **Confinements:**  $X_n \in \mathbb{Z}^d \supset A$  &  $G_n := 1_A$ .

$$\mathbb{Q}_n = \text{Law}((X_0, \dots, X_n) \mid X_p \in A, \forall 0 \leq p < n)$$

and

$$\mathcal{Z}_n = \text{Proba}(X_p \in A, \forall 0 \leq p < n)$$

- ▶ **SAW :**  $X_n = (X'_p)_{0 \leq p \leq n}$  &  $G_n(X_n) = 1_{X'_n \notin \{X'_0, \dots, X'_{n-1}\}}$

$$\mathbb{Q}_n = \text{Law}((X'_0, \dots, X'_n) \mid X'_p \neq X'_q, \forall 0 \leq p < q < n)$$

and

$$\mathcal{Z}_n = \text{Proba}(X'_p \neq X'_q, \forall 0 \leq p < q < n)$$

## Some examples

- ▶ **Filtering** :  $Y_n = H_n(X_n, V_n)$  &  $G_n(x_n) := p_{Y_n|X_n}(y_n|x_n)$ .

$$\mathbb{Q}_n = \text{Law}((X_0, \dots, X_n) \mid Y_p = y_p, \forall 0 \leq p < n)$$

and

$$\mathcal{Z}_n = p_{Y_0, \dots, Y_n}(y_0, \dots, y_n)$$

- ▶ **Multilevel splitting** :  $A_n \downarrow$ , with  $B$  non critical recurrent subset.

$$T_n := \inf \{t \geq T_{n-1} : X'_t \in (A_n \cup B)\}$$

$$X_n = (X'_t)_{t \in [T_n, T_{n+1}]} \quad \& \quad G_n(X_n) = 1_{A_{n+1}}(X'_{T_{n+1}})$$

$$\mathbb{Q}_n = \text{Law}(X'_{[T_0, T_n]} \mid X' \text{ hits } A_{n-1} \text{ before } B)$$

and

$$\mathcal{Z}_n = \mathbb{P}(X' \text{ hits } A_{n-1} \text{ before } B)$$

## Some examples

► **Absorption models :**

$$X_n^c \in E_n^c \xrightarrow{\text{absorption} \sim (1-G_n)} \widehat{X}_n^c \xrightarrow{\text{exploration} \sim M_{n+1}} X_{n+1}^c$$

$$\mathbb{Q}_n = \text{Law}((X_0^c, \dots, X_n^c) \mid T^{\text{abs.}} \geq n) \quad \& \quad \mathbb{Z}_n = \text{Proba}(T^{\text{abs.}} \geq n)$$

► **Quasi-invariant measures :  $(G_n, M_n) = (G, M)$  &  $M$   $\mu$ -reversible**

$$\frac{1}{n} \log \mathbb{P}(T^{\text{abs.}} \geq n) \simeq_{n \uparrow \infty} \lambda = \text{top spect. of } Q(x, dy) = G(x)M(x, dy)$$

[Frobenius theo]  $Q(h) = \lambda h = \lambda \times$  eigenfunction (ground state)

$$\mathbb{P}(X_n^c \in dx \mid T^{\text{abs.}} > n) \simeq_{n \uparrow \infty} \frac{1}{\mu(h)} h(x) \mu(dx)$$



## Some examples

- ▶ **Doob  $h$ -processes  $X^h$  :**

$$M^h(x, dy) = \frac{1}{\lambda} h^{-1}(x) Q(x, dy) h(y) = \frac{Q(x, dy) h(y)}{Q(h)(x)} = \frac{M(x, dy) h(y)}{M(h)(x)}$$

↓

$$\mathbb{Q}_n(d(x_0, \dots, x_n)) \propto \text{Proba}((X_0^h, \dots, X_n^h) \in d(x_0, \dots, x_n)) h^{-1}(x_n)$$

## Some examples : **Sequential Monte Carlo**

$$\underbrace{Q_{n+1}(d(x_0, \dots, x_{n+1}))}_{\text{Target at time (n+1)}} \propto \underbrace{Q_n(d(x_0, \dots, x_n))}_{\text{Target at time (n)}} \times \underbrace{Q_{n+1}(x_n, dx_{n+1})}_{\geq 0 \text{ integral oper.}}$$

- Twisted-IS Markov =  $M_{n+1} \rightsquigarrow$  the weight function :

$$\begin{aligned} G_n(x_n, x_{n+1}) &= \frac{\text{Target at time (n+1)}}{\text{Target at time (n)} \times \text{Twisted transition}} \\ &= \frac{dQ_{n+1}(x_n, \cdot)}{dM_{n+1}(x_n, \cdot)}(x_{n+1}) \end{aligned}$$

$\Updownarrow$

$$Q_{n+1}(d(x_0, \dots, x_{n+1})) \propto Q_n(d(x_0, \dots, x_n)) G_n(x_n, x_{n+1}) M_{n+1}(x_n, dx_{n+1})$$

## Feynman-Kac models

### Interacting particle interpretations

Nonlinear evolution equation

Mean field particle models

Graphical illustration

First order expansions

Uniform concentration w.r.t. time

Particle free energy

### Concentration inequalities

# Flow of $n$ -marginals [ $X_n$ Markov with transitions $M_n$ ]

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

$$\Updownarrow (\gamma_n(1) = \mathcal{Z}_n)$$

**Nonlinear evolution equation :**

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

$$\mathcal{Z}_{n+1} = \eta_n(G_n) \times \mathcal{Z}_n$$

↓

**Nonlinear m.v.p. = Law of a Markov  $\bar{X}_n$  (perfect sampler)**

$$\begin{aligned} \eta_{n+1} &= \Phi_{n+1}(\eta_n) \\ &= \eta_n(S_{n,\eta_n} M_{n+1}) = \eta_n \mathbf{K}_{n+1,\eta_n} = \text{Law}(\bar{X}_{n+1}) \end{aligned}$$

## Examples related to product models

$$\eta_n(dx) := \frac{1}{Z_n} \left\{ \prod_{p=0}^n h_p(x) \right\} \lambda(dx) \quad \text{with } h_p \geq 0$$

### 3 illustrations:

$$h_p(x) = e^{-(\beta_{p+1} - \beta_p)V(x)} \quad \beta_p \uparrow \implies \eta_n(dx) = \frac{1}{Z_n} e^{-\beta_n V(x)} \lambda(dx)$$

$$h_p(x) = 1_{A_{p+1}}(x) \quad A_p \downarrow \implies \eta_n(dx) = \frac{1}{Z_n} 1_{A_n}(x) \lambda(dx)$$

$$h_p(\theta) = p(y_p | \theta, y_0, \dots, y_{p-1}) \implies \eta_n(d\theta) = p(\theta | y_0, \dots, y_n)$$

## Examples related to product models

$$\eta_n(dx) := \frac{1}{Z_n} \left\{ \prod_{p=0}^n h_p(x) \right\} \lambda(dx) \quad \text{with } h_p \geq 0$$

### 3 illustrations:

$$h_p(x) = e^{-(\beta_{p+1}-\beta_p)V(x)} \quad \beta_p \uparrow \implies \eta_n(dx) = \frac{1}{Z_n} e^{-\beta_n V(x)} \lambda(dx)$$

$$h_p(x) = 1_{A_{p+1}}(x) \quad A_p \downarrow \implies \eta_n(dx) = \frac{1}{Z_n} 1_{A_n}(x) \lambda(dx)$$

$$h_p(\theta) = p(y_p|\theta, y_0, \dots, y_{p-1}) \implies \eta_n(d\theta) = p(\theta|y_0, \dots, y_n)$$

For any MCMC transitions  $M_n$  with target  $\eta_n$ , we have

$$\eta_{n+1} = \eta_{n+1} M_{n+1} = \Psi_{h_{n+1}}(\eta_n) M_{n+1} \subset \text{Feynman-Kac model}$$

Mean field particle model = Markov  $\xi_n = (\xi_n^i)_{1 \leq i \leq N} \in E_n^N$

$$\mathbb{P}(\xi_{n+1} \in dx \mid \xi_n) = \prod_{1 \leq i \leq N} K_{n+1, \eta_n^N}(\xi_n^i, dx^i) \quad \text{with} \quad \eta_n^N = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$$

and the (unbiased, ref. DP. in MPRF 96) particle normalizing constants

$$\mathcal{Z}_{n+1}^N = \eta_n^N(G_n) \times \mathcal{Z}_n^N = \prod_{0 \leq p \leq n} \eta_p^N(G_p)$$

Mean field particle model = Markov  $\xi_n = (\xi_n^i)_{1 \leq i \leq N} \in E_n^N$

$$\mathbb{P}(\xi_{n+1} \in dx \mid \xi_n) = \prod_{1 \leq i \leq N} K_{n+1, \eta_n^N}(\xi_n^i, dx^i) \quad \text{with} \quad \eta_n^N = \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$$

and the (unbiased, ref. DP. in MPRF 96) particle normalizing constants

$$\mathcal{Z}_{n+1}^N = \eta_n^N(G_n) \times \mathcal{Z}_n^N = \prod_{0 \leq p \leq n} \eta_p^N(G_p)$$

$$\Updownarrow [K_{n+1, \eta_n} = S_{n, \eta_n} M_{n+1}]$$

↪ **Sequential particle simulation technique**

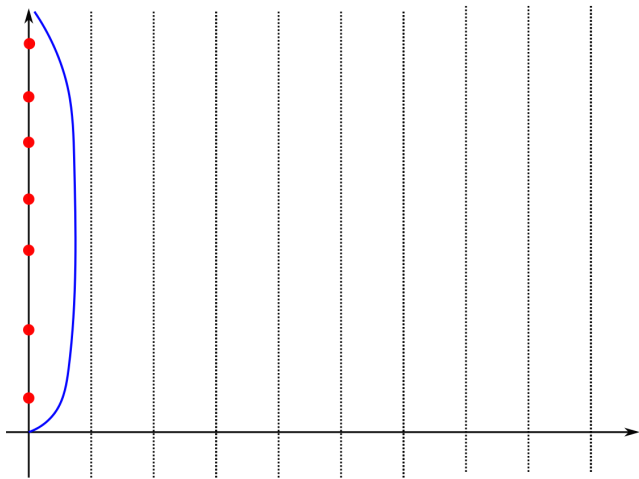
$G_n$ -acceptance-rejection with recycling  $\oplus$   $M_{n+1}$ -propositions

↪ **Genetic type branching particle model**

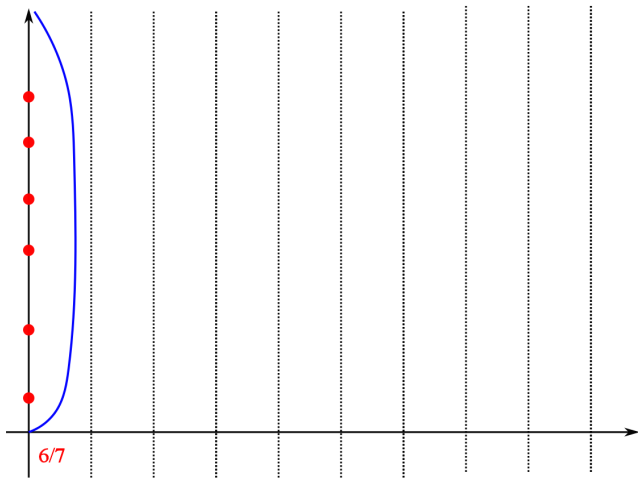
$$\xi_n = (\xi_n^i)_{1 \leq i \leq N} \xrightarrow{G_n\text{-selection}} \hat{\xi}_n = (\hat{\xi}_n^i)_{1 \leq i \leq N} \xrightarrow{M_n\text{-mutation}} \xi_{n+1} = (\xi_{n+1}^i)_{1 \leq i \leq N}$$



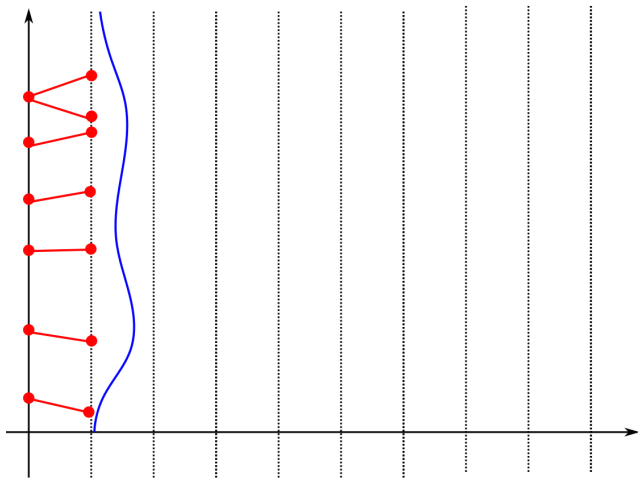
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



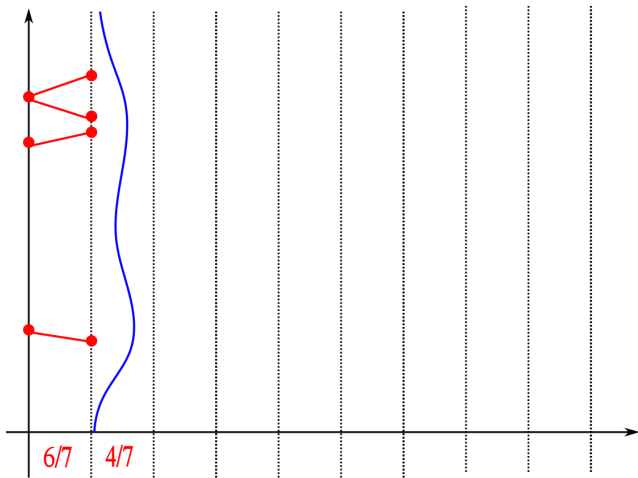
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



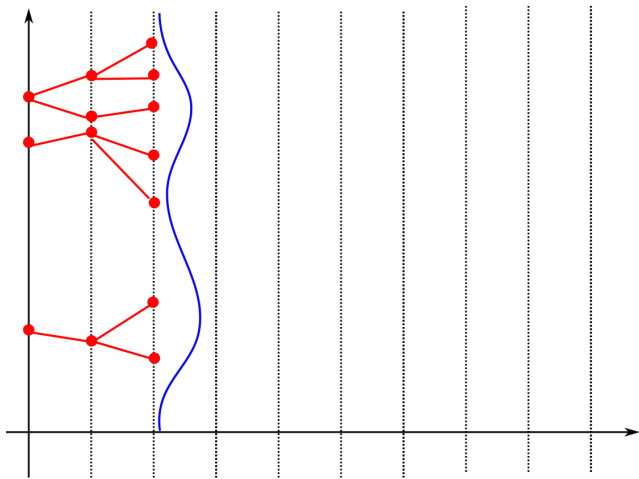
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



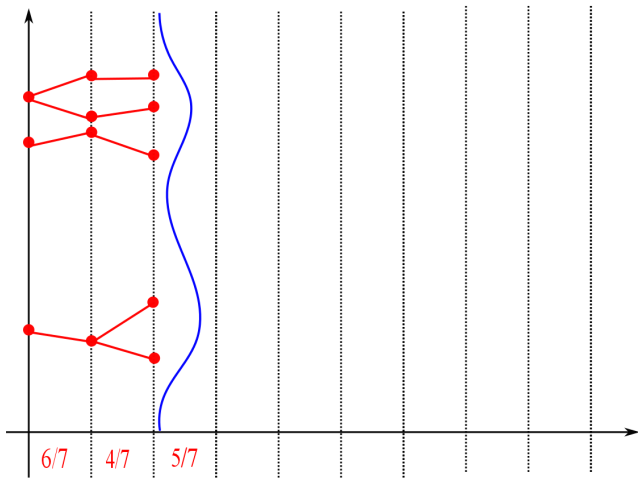
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



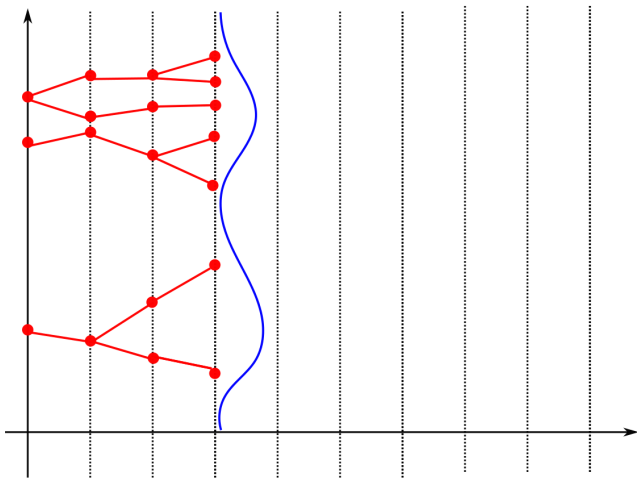
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



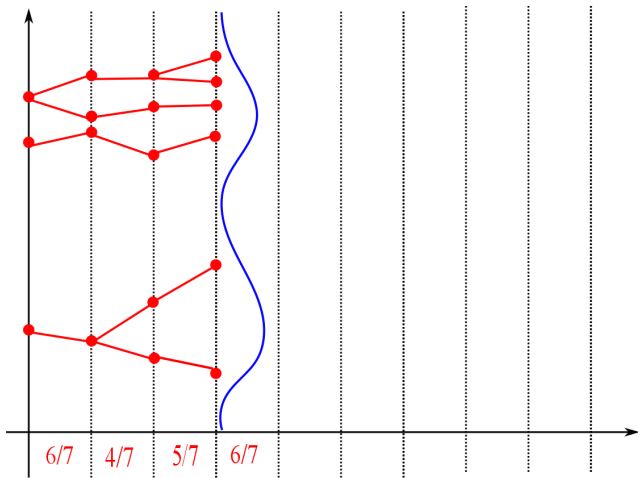
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$

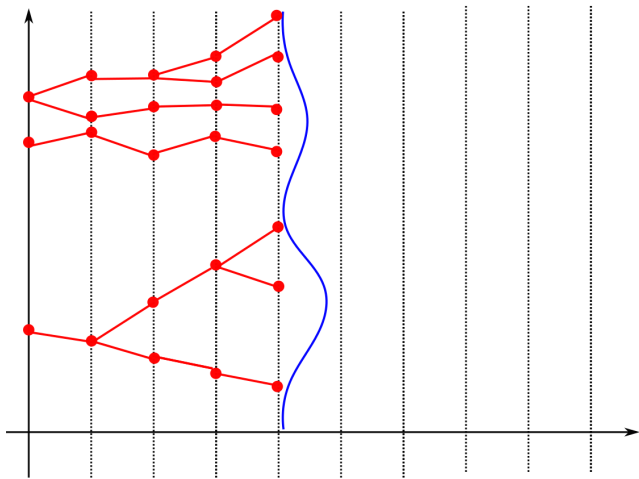


Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$

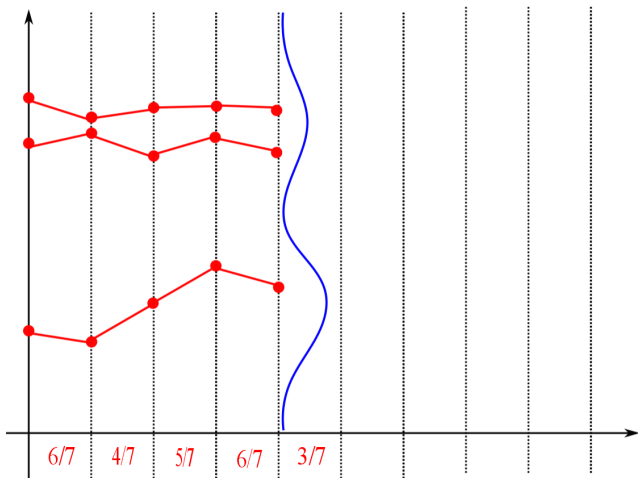




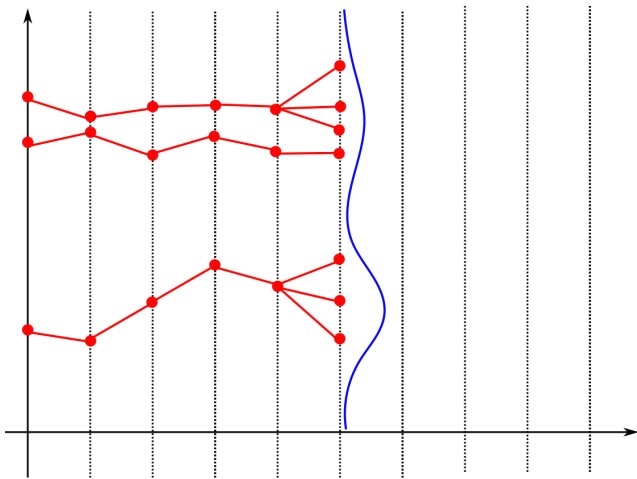
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



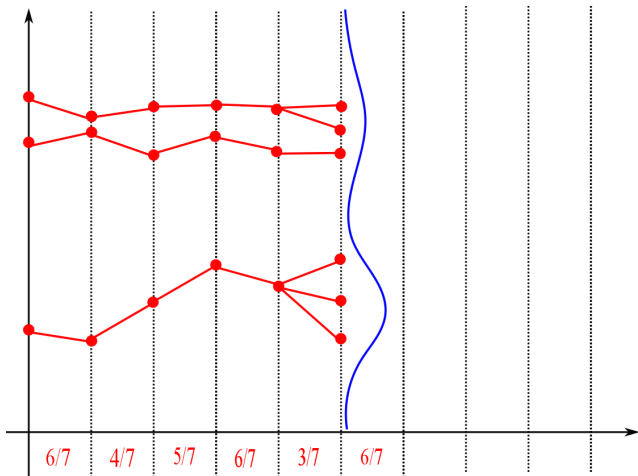
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



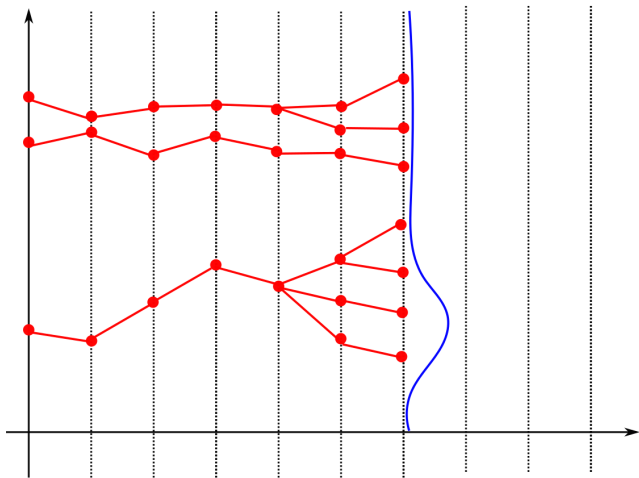
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



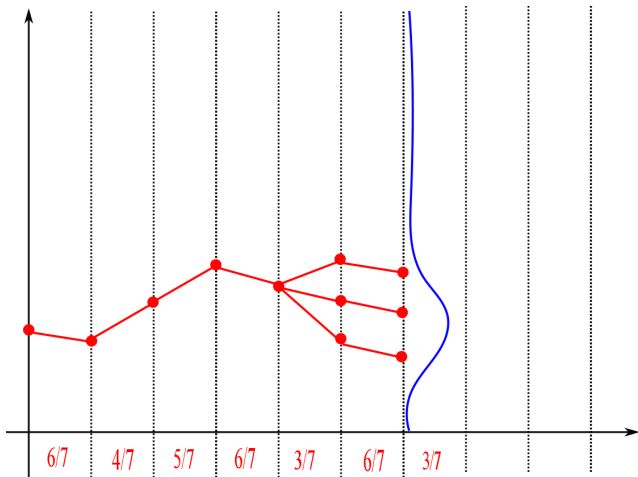
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



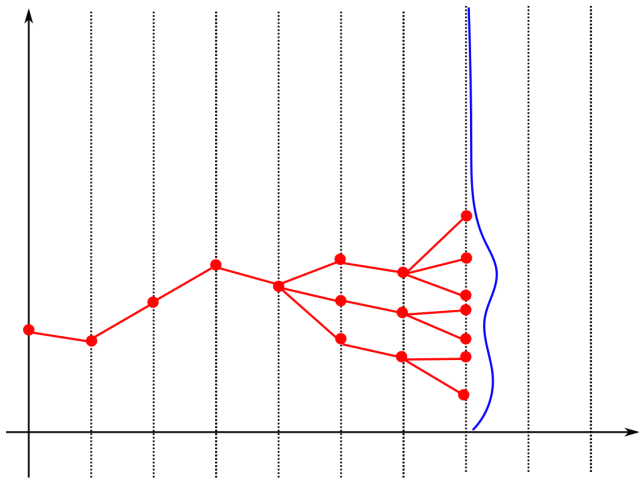
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



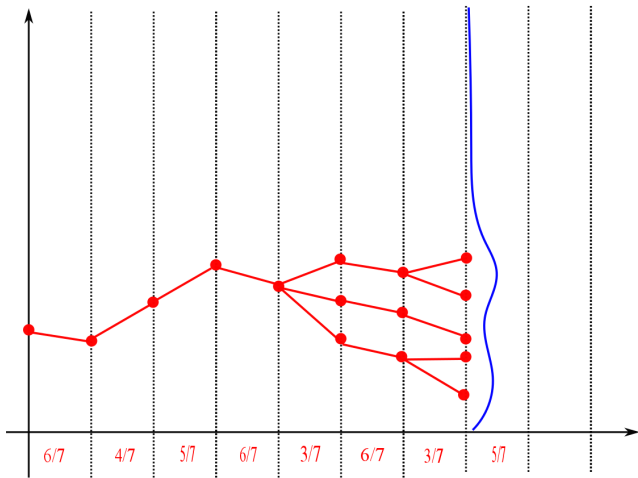
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$

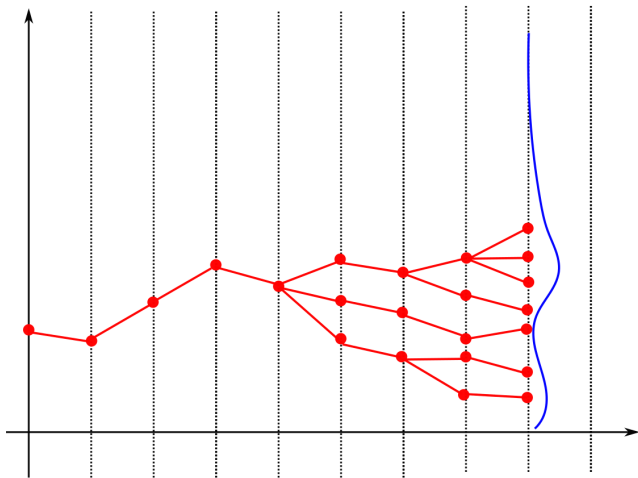


Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$

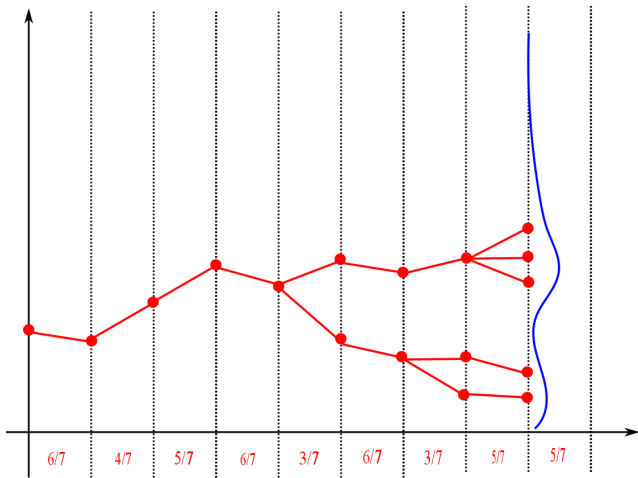




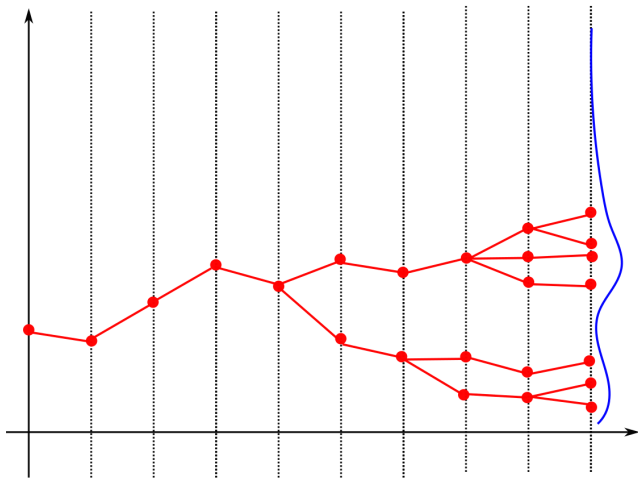
Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



Graphical illustration :  $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



## How to use the full ancestral tree model ?

$$G_{n-1}(x_{n-1})M_n(x_{n-1}, dx_n) \stackrel{\text{hyp}}{=} H_n(x_{n-1}, x_n) \nu_n(dx_n)$$

$$\begin{aligned} \Rightarrow \mathbb{Q}_n(d(x_0, \dots, x_n)) &= \eta_n(dx_n) \underbrace{\mathbb{M}_{n, \eta_{n-1}}(x_n, dx_{n-1}) \dots \mathbb{M}_{1, \eta_0}(x_1, dx_0)}_{\propto \eta_n(dx_n) H_{n+1}(x_n, x_{n+1})} \end{aligned}$$

## How to use the full ancestral tree model ?

$$G_{n-1}(x_{n-1})M_n(x_{n-1}, dx_n) \stackrel{\text{hyp}}{=} H_n(x_{n-1}, x_n) \nu_n(dx_n)$$

$$\Rightarrow \mathbb{Q}_n(d(x_0, \dots, x_n)) = \eta_n(dx_n) \underbrace{\mathbb{M}_{n, \eta_{n-1}}(x_n, dx_{n-1}) \dots \mathbb{M}_{1, \eta_0}(x_1, dx_0)}_{\propto \eta_n(dx_n) H_{n+1}(x_n, x_{n+1})}$$

Particle approximation = Random stochastic matrices

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

## How to use the full ancestral tree model ?

$$G_{n-1}(x_{n-1})M_n(x_{n-1}, dx_n) \stackrel{\text{hyp}}{=} H_n(x_{n-1}, x_n) \nu_n(dx_n)$$

$$\Rightarrow \mathbb{Q}_n(d(x_0, \dots, x_n)) = \eta_n(dx_n) \underbrace{\mathbb{M}_{n, \eta_{n-1}}(x_n, dx_{n-1}) \dots \mathbb{M}_{1, \eta_0}(x_1, dx_0)}_{\propto \eta_n(dx_n) H_{n+1}(x_n, x_{n+1})}$$

Particle approximation = Random stochastic matrices

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

*Ex.: Additive functionals*  $\mathbf{f}_n(x_0, \dots, x_n) = \frac{1}{n+1} \sum_{0 \leq p \leq n} f_p(x_p)$

$$\mathbb{Q}_n^N(\mathbf{f}_n) := \frac{1}{n+1} \sum_{0 \leq p \leq n} \eta_n^N \underbrace{\mathbb{M}_{n, \eta_{n-1}^N} \dots \mathbb{M}_{p+1, \eta_p^N}}_{\text{matrix operations}}(f_p)$$

# The example of the $h$ -process and the absorption model

$$\mathbb{Q}_n(d(x_0, \dots, x_n)) \propto \mathbb{P}((X_0^h, \dots, X_n^h) \in d(x_0, \dots, x_n)) h^{-1}(x_n)$$

- ▶ *Invariant measure  $\mu_h = \mu_h M^h$  & normalized additive functionals*

$$\mathbf{f}_n(x_0, \dots, x_n) = \frac{1}{n+1} \sum_{0 \leq p \leq n} f(x_p) \implies \mathbb{Q}_n(\mathbf{f}_n) \simeq_n \mu_h(f)$$

- ▶ *If  $G = G^\theta$  depends on some  $\theta \in \mathbb{R} \rightsquigarrow f := \frac{\partial}{\partial \theta} \log G^\theta$*

$$\frac{\partial}{\partial \theta} \log \lambda^\theta \simeq_n \frac{1}{n+1} \frac{\partial}{\partial \theta} \log \mathcal{Z}_{n+1}^\theta = \mathbb{Q}_n(\mathbf{f}_n)$$

*NB : Similar expression when  $M^\theta$  depends on some  $\theta \in \mathbb{R}$ .*

## 4 particle estimates

- ▶ **Individuals**  $\xi_n^i$  "almost" iid with law

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

- ▶ **Path space models**  $\rightsquigarrow$  Ancestral lines "almost" iid with law

$$\mathbb{Q}_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\text{ancestral line}_n(i)}$$

- ▶ **Backward particle model**

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

- ▶ **Normalizing constants**

$$\mathcal{Z}_{n+1} = \prod_{0 \leq p \leq n} \eta_p(G_p) \simeq_{N \uparrow \infty} \mathcal{Z}_{n+1}^N = \prod_{0 \leq p \leq n} \eta_p^N(G_p) \quad (\text{Unbiased})$$



On the unbiased property  $\theta \rightsquigarrow (G_{\theta,n}, M_{\theta,n}, \mathcal{Z}_n(\theta), \eta_{\theta,n}^N, \dots)$

$x = (\Theta, \xi) =$  r.v. with distribution  $\lambda(dx) := \nu(d\theta) P(\theta, d\xi)$

$\downarrow$

$h_n(x) = \eta_{\theta,n}^N(G_{n,\theta})$  "easily" computable

$$\Rightarrow \eta_n(dx) := \frac{1}{\mathcal{Z}_n} \left\{ \prod_{p=0}^n h_p(x) \right\} \lambda(dx) \quad \subset \quad \text{Product models examples}$$

$\Downarrow$  [Unbias property  $\Rightarrow$  Target =  $\Theta$ -marginal]

$$\eta_n \circ \Theta^{-1}(d\theta) := \frac{1}{\mathcal{Z}_n} \mathcal{Z}_n(\theta) \nu(d\theta)$$

**Examples :**

$\mathcal{Z}_n(\theta) \in \{p(y_0, \dots, y_n | \theta), \mathbb{P}(\text{cascade of } n \text{ rare events} | \theta), \mathbb{P}(T > n | \theta), \dots\}$

# Unbias $\rightsquigarrow$ : **Quenched** $\oplus$ **Conditional samplers**

**Quenched models :**

$$\mathbb{E} \left( f_n(X_0, \dots, X_n) \prod_{0 \leq p \leq n} G_p(X_p) \right) = \mathbb{E} \left( f_n(X_0^N, \dots, X_n^N) \prod_{0 \leq p \leq n} G_p^N(X_p^N) \right)$$

Ex.:  $\forall$  "doubly random" chains  $X_n^N \sim$  independent random  $Q_n^N$  such that

$$\mathbb{E} (Q_n^N) = Q_n \quad \text{ex.:} \quad G_n^N = G_n \quad \text{and} \quad M_n^N(x, \cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_n^i(x)}$$

# Unbias $\rightsquigarrow$ : **Quenched** $\oplus$ **Conditional samplers**

**Quenched models :**

$$\mathbb{E} \left( f_n(X_0, \dots, X_n) \prod_{0 \leq p \leq n} G_p(X_p) \right) = \mathbb{E} \left( f_n(X_0^N, \dots, X_n^N) \prod_{0 \leq p \leq n} G_p^N(X_p^N) \right)$$

Ex.:  $\forall$  "doubly random" chains  $X_n^N \sim$  independent random  $Q_n^N$  such that

$$\mathbb{E}(Q_n^N) = Q_n \quad \text{ex.:} \quad G_n^N = G_n \quad \text{and} \quad M_n^N(x, \cdot) := \frac{1}{N} \sum_{i=1}^N \delta_{X_n^i(x)}$$

**Conditional models** ( $\oplus$  **Quenched models (ref. DP. in AAP 98)**) :

$$\mathbb{E} \left( f_n(X_0, \dots, X_n) \prod_{0 \leq p \leq n} G_p(X_p) \right) = \mathbb{E} \left( f_n(\hat{X}_0, \dots, \hat{X}_n) \prod_{0 \leq p < n} \hat{G}_p(\hat{X}_p) \right)$$

with the "conditional" Markov transitions and potential functions

$$\hat{M}_n(x_{n-1}, dx_n) := \frac{M_n(x_{n-1}, dx_n) G_n(x_n)}{M_n(G_n)(x_{n-1})} \quad \text{and} \quad \hat{G}_{n-1}(x_{n-1}) := M_n(G_n)(x_{n-1})$$

# How & Why it really works

- ▶ (Computer Sci.) Stochastic adaptive grid approximation.
- ▶ (Stats) Universal acceptance-rejection-recycling sampling schemes.
- ▶ (Probab) Stochastic linearization/perturbation technique.

$$\begin{aligned}\eta_n &= \Phi_n(\eta_{n-1}) \\ \eta_n^N &= \Phi_n(\eta_{n-1}^N) + \frac{1}{\sqrt{N}} V_n^N\end{aligned}$$

**Theorem:**  $(V_n^N)_n \simeq_{N \uparrow \infty} (V_n)_n$  independent centered Gaussian fields.

$\Phi_{p,n}(\eta_p) = \eta_n$  stable sg  $\iff$  No propagation of local errors  
 $\implies$  Uniform control w.r.t. the time horizon

# How & Why it really works

- ▶ (Computer Sci.) Stochastic adaptive grid approximation.
- ▶ (Stats) Universal acceptance-rejection-recycling sampling schemes.
- ▶ (Probab) Stochastic linearization/perturbation technique.

$$\begin{aligned}\eta_n &= \Phi_n(\eta_{n-1}) \\ \eta_n^N &= \Phi_n(\eta_{n-1}^N) + \frac{1}{\sqrt{N}} V_n^N\end{aligned}$$

**Theorem:**  $(V_n^N)_n \simeq_{N \uparrow \infty} (V_n)_n$  **independent centered Gaussian fields.**

$\Phi_{p,n}(\eta_p) = \eta_n$  stable sg  $\iff$  No propagation of local errors  
 $\implies$  **Uniform control w.r.t. the time horizon**

$\rightsquigarrow$  **New concentration inequalities for (general) interacting processes**

# Key idea = First order expansions

## Key telescoping decomposition

$$\eta_n^N - \eta_n = \sum_{p=0}^n [\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N))]$$

## ⊕ First order expansion

$$\begin{aligned} & \sqrt{N} [\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N))] \\ &= \sqrt{N} \left[ \Phi_{p,n} \left( \Phi_p(\eta_{p-1}^N) + \frac{1}{\sqrt{N}} V_p^N \right) - \Phi_{p,n}(\Phi_p(\eta_{p-1}^N)) \right] \\ &\simeq V_p^N D_{p,n} + \frac{1}{\sqrt{N}} R_{p,n}^N \end{aligned}$$

with  $\underbrace{\text{a predictable } D_{p,n} - \text{first order operator}}_{\text{fluctuation term}} \oplus \underbrace{\text{2nd-order measure } R_{p,n}^N}_{\text{bias-term}}$

# Stochastic perturbation model

## Stochastic perturbation model

$$W_n^{\eta, N} := \sqrt{N} [\eta_n^N - \eta_n] = \sum_{0 \leq p \leq n} V_p^N D_{p,n} + \frac{1}{\sqrt{N}} R_n^N$$

Under some mixing condition on the limiting FK semigroups  $\Phi_{p,n}$

$$\text{osc}(D_{p,n}(f)) \leq Cte e^{-(n-p)\alpha}$$

and

$$\mathbb{E} (|R_n^N(f)|^m) \leq Cte 2^{-m}(2m)!/m!$$

↓

Uniform concentration estimates w.r.t. the time parameter

# Particle free energy

## Multiplicative formulae

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

## Taylor first order expansion

$$\begin{aligned} \forall x, y > 0 \quad \log y - \log x &= \int_0^1 \frac{(y-x)}{x+t(y-x)} dt \\ &\Downarrow \\ \log(\gamma_n^N(1)/\gamma_n(1)) & \\ &= \sum_{0 \leq p < n} (\log \eta_p^N(G_p) - \log \eta_p(G_p)) \\ &= \sum_{0 \leq p < n} \left( \log \left( \eta_p(G_p) + \frac{1}{\sqrt{N}} W_p^{\eta, N}(G_p) \right) - \log \eta_p(G_p) \right) \\ &= \frac{1}{\sqrt{N}} \sum_{0 \leq p < n} \int_0^1 \frac{W_p^{\eta, N}(G_p)}{\eta_p(G_p) + \frac{t}{\sqrt{N}} W_p^{\eta, N}(G_p)} dt \end{aligned}$$

↪ first order expansion [exercice]



Feynman-Kac models

Interacting particle interpretations

Concentration inequalities

Current population models

Particle free energy

Genealogical tree models

Backward particle models

# Current population models

Constants  $(c_1, c_2)$  related to (bias, variance),  $c$  universal constant

Test funct.  $\|f_n\| \leq 1, \forall (x \geq 0, n \geq 0, N \geq 1)$ .

- ▶ The probability of the event

$$[\eta_n^N - \eta_n](f) \leq \frac{c_1}{N} (1 + x + \sqrt{x}) + \frac{c_2}{\sqrt{N}} \sqrt{x}$$

is greater than  $1 - e^{-x}$ .

- ▶  $x = (x_i)_{1 \leq i \leq d} \rightsquigarrow (-\infty, x] = \prod_{i=1}^d (-\infty, x_i]$  cells in  $E_n = \mathbb{R}^d$ .

$$F_n(x) = \eta_n(1_{(-\infty, x]}) \quad \text{and} \quad F_n^N(x) = \eta_n^N(1_{(-\infty, x]})$$

The probability of the following event

$$\sqrt{N} \|F_n^N - F_n\| \leq c \sqrt{d(x+1)}$$

is greater than  $1 - e^{-x}$ .

# Particle free energy models

Constants  $(c_1, c_2)$  related to (bias, variance),  $c$  universal constant

$\forall (x \geq 0, n \geq 0, N \geq 1)$

- ▶ Unbias property  $\rightsquigarrow$  Island models (less oscillating potential)

$$\mathbb{E} \left( \eta_n^N(f_n) \prod_{0 \leq p < n} \eta_p^N(G_p) \right) = \mathbb{E} \left( f_n(X_n) \prod_{0 \leq p < n} G_p(X_p) \right)$$

- ▶ For any  $\epsilon \in \{+1, -1\}$ , the probability of the event

$$\frac{\epsilon}{n} \log \frac{Z_n^N}{Z_n} \leq \frac{c_1}{N} (1 + x + \sqrt{x}) + \frac{c_2}{\sqrt{N}} \sqrt{x}$$

is greater than  $1 - e^{-x}$ .

note  $(0 \leq \epsilon \leq 1 \Rightarrow (1 - e^{-\epsilon}) \vee (e^\epsilon - 1) \leq 2\epsilon)$

$$e^{-\epsilon} \leq \frac{z^N}{z} \leq e^\epsilon \Rightarrow \left| \frac{z^N}{z} - 1 \right| \leq 2\epsilon$$

# Genealogical tree models $:= \eta_n^N$ (in path space)

Constants  $(c_1, c_2)$  related to (bias, variance),  $c$  universal constant  
 $\mathbf{f}_n$  test function  $\|\mathbf{f}_n\| \leq 1, \forall (x \geq 0, n \geq 0, N \geq 1)$ .

- ▶ The probability of the event

$$[\eta_n^N - \mathbb{Q}_n](f) \leq c_1 \frac{n+1}{N} (1 + x + \sqrt{x}) + c_2 \sqrt{\frac{(n+1)}{N}} \sqrt{x}$$

is greater than  $1 - e^{-x}$ .

- ▶  $\mathcal{F}_n =$  indicator fct.  $\mathbf{f}_n$  of cells in  $\mathbf{E}_n = (\mathbb{R}^{d_0} \times \dots, \times \mathbb{R}^{d_n})$

The probability of the following event

$$\sup_{\mathbf{f}_n \in \mathcal{F}_n} |\eta_n^N(\mathbf{f}_n) - \mathbb{Q}_n(\mathbf{f}_n)| \leq c (n+1) \sqrt{\frac{\sum_{0 \leq p \leq n} d_p}{N}} (x+1)$$

is greater than  $1 - e^{-x}$ .

# Backward particle models

Constants  $(c_1, c_2)$  related to (bias, variance),  $c$  universal constant.

$\mathbf{f}_n$  normalized additive functional with  $\|f_p\| \leq 1$ ,  $\forall (x \geq 0, n \geq 0, N \geq 1)$

- ▶ The probability of the event

$$[\mathbb{Q}_n^N - \mathbb{Q}_n](\mathbf{f}_n) \leq c_1 \frac{1}{N} (1 + (x + \sqrt{x})) + c_2 \sqrt{\frac{x}{N(n+1)}}$$

- ▶  $\mathbf{f}_{a,n}$  normalized additive functional w.r.t.  $f_p = 1_{(-\infty, a]}$ ,  $a \in \mathbb{R}^d = E_n$

The probability of the following event

$$\sup_{a \in \mathbb{R}^d} |\mathbb{Q}_n^N(\mathbf{f}_{a,n}) - \mathbb{Q}_n(\mathbf{f}_{a,n})| \leq c \sqrt{\frac{d}{N}}(x+1)$$

is greater than  $1 - e^{-x}$ .