Stochastic particle models and methods in risk analysis

P. Del Moral (INRIA research team ALEA)

INRIA & Bordeaux Mathematical Institute & X CMAP

CSIRO, Sydney, February 20-th, 2012

Some Refs & links

- **Feynman-Kac formulae, Genealogical & Interacting Particle Systems with appl., Springer (2004)**
- **Sequential Monte Carlo Samplers for Rare Events. (joint work with Johansen, Doucet) (2006)**
- **Branching and Interacting Particle Interpretations of Rare Event Probabilities Springer (2006). [+ Réfs]**
- **Plus de références http://www.math.u-bordeaux1.fr/∼delmoral/index.html [+ Links]**

Some examples

- Sanitary risks: pandemic propagation, tumor development, cardio-vascular risks, medical treatment efficiency, food risk analysis, bacterial propagations.
- System reliability : production chains, offshore platforms security, networks overflows, computer security.
- Air traffic control : control system failures, airport and flight collision risks.
- Nuclear plant security : radioactivity storage, nuclear tank cracks.
- Financial and economical risks : payment defaults, financial breakdowns and economic crisis.
- Environmental risks: invasive species, climatic fluctuations, floods and inondations, earthquakes
- .../...

On the notion of risk

- Poorly documented events :
  - Only a few observations, missing statistical data.
  - Poorly formalised empirical models.
- Unpredictable and complex systems :
  - Multilevel and multiple scales interactions.
  - Chaotic and unstable systems.
  - High dimensional random models.

Mathematical tools

- Calibration of models :
  1. Direct statistical estimation :
     - Empirical estimates based on very few real data.
     - Extreme values theory.
  2. Calibration of formalised systems or numerical codes *w.r.t. partial observations and critical events* :
     - Uncertainty and statistical variability.
     - Kinetic & statistical parameters.

  $\rightsquigarrow$ Stochastic methods :
  Filtering, bayesian inference, sequential Monte Carlo techniques, particle methods, stochastic style gradients.

Mathematical tools (continued)

- Uncertainty propagations in formalised models.
  - Rare event simulation :
    - Rare event probabilities.
    - Default random tree simulation.
  - Sensitivity measures estimation :
    - Stability and robustness of the estimation.
    - Sensitivity w.r.t. the initial conditions.
    - Sensitivity w.r.t. kinetic/statistical parameters.

$\rightsquigarrow$ Stochastic methods :

Recycling style acceptance-rejection techniques, interacting particle systems, default genealogy sampling, multi-level sampling techniques, free energy gradient computation, conditional path sampling, non intrusive importance sampling methods, interacting island models,...

Mathematical models

- X = random variable, stochastic process, random excursions
- W = model uncertainties, random noises, stochastic perturbations.
- Θ = kinetic parameters, initial conditions, control variables, statistical parameters.
- A = A critical event ($\mathbb{P}(X \in A) \simeq 10^{-6} - 10^{-12}$)

Mathematical models

- X = random variable, stochastic process, random excursions
- W = model uncertainties, random noises, stochastic perturbations.
- Θ = kinetic parameters, initial conditions, control variables, statistical parameters.
- A = A critical event ($\mathbb{P}(X \in A) \simeq 10^{-6} - 10^{-12}$)

$$\boxed{\begin{array}{c} \text{Objectives} \qquad \mathbb{P}(X \in A) \quad \text{and} \quad \mathrm{Law}\,(X \mid X \in A) \quad ? \\ \Uparrow \\ \mathbb{P}((\Theta, W) \text{ s.t. } X \in A) \quad \text{and} \quad \mathrm{Law}\,((\Theta, W) \mid X \in A) \quad ? \end{array}}$$

Mathematical models

- X = random variable, stochastic process, random excursions
- W = model uncertainties, random noises, stochastic perturbations.
- Θ = kinetic parameters, initial conditions, control variables, statistical parameters.
- A = A critical event ($\mathbb{P}(X \in A) \simeq 10^{-6} - 10^{-12}$)

Objectives $\quad \mathbb{P}(X \in A) \quad \text{and} \quad \text{Law}\,(X \mid X \in A) \quad ?$

$$\Uparrow$$

$$\mathbb{P}((\Theta, W) \text{ s.t. } X \in A) \quad \text{and} \quad \text{Law}\,((\Theta, W) \mid X \in A) \quad ?$$

Sensitivity measures :

$$\frac{\partial}{\partial \theta} \log \mathbb{P}(X \in A) \quad \text{and} \quad \frac{\partial}{\partial \theta} \mathbb{E}\,(f(X) \mid X \in A) \quad ?$$

Importance sampling

$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = 10^{-10} \rightsquigarrow$ Find a twisted $\mathbb{P}_Y$ s.t. $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A) \simeq 1$

$\rightsquigarrow$ Classical Monte Carlo simulation Y^i i.i.d. $\mathbb{P}_Y$

$$\int \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(y)\, 1_A(y)\, \mathbb{P}_Y(dy) = \mathbb{P}_X(A) \simeq \mathbb{P}_X^N(A) := \frac{1}{N} \sum_{1 \leq i \leq N} \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(Y^i)\, 1_A(Y^i)$$

$$\Downarrow$$

$$\text{Variance} \quad = \quad \int \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(x)\, 1_A(x)\, \mathbb{P}_X(dx) - \mathbb{P}_X(A)$$

Importance sampling

$\mathbb{P}(X \in A) = \mathbb{P}_X(A) = 10^{-10} \rightsquigarrow$ Find a twisted $\mathbb{P}_Y$ s.t. $\mathbb{P}_Y(A) = \mathbb{P}(Y \in A) \simeq 1$

$\rightsquigarrow$ Classical Monte Carlo simulation Y^i i.i.d. $\mathbb{P}_Y$

$$\int \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(y)\, 1_A(y)\, \mathbb{P}_Y(dy) = \mathbb{P}_X(A) \simeq \mathbb{P}_X^N(A) := \frac{1}{N} \sum_{1 \leq i \leq N} \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(Y^i)\, 1_A(Y^i)$$

$$\Downarrow$$

$$\text{Variance} \quad = \quad \int \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}(x)\, 1_A(x)\, \mathbb{P}_X(dx) - \mathbb{P}_X(A)$$

Optimal twisted measure

$$\text{Variance} = 0 \Longleftrightarrow \mathbb{P}_Y(dx) = \frac{1_A(x)}{\mathbb{P}_X(A)}\, \mathbb{P}_X(dx) = \mathbb{P}\left(X \in dx \mid X \in A\right)$$

Importance sampling (continued)

The drawbacks

- Complex systems $\Rightarrow$ Difficult choice of the twisted measures.
- Robustness troubles : huge variance for wrong choice of Y.
  Ex.: Stochastic process $X = (X_0, \ldots, X_n)$

$$\frac{d\mathbb{P}_n}{d\mathbb{Q}_n}(X) := \prod_{k=0}^{n} \frac{p_k(X_k|X_{k-1})}{q_k(X_k|X_{k-1})} \quad \text{martingale} \geq 0 \rightarrow \text{degenerated products w.r.t. } n$$

- **Intrusive methods : we sample Y, and not the real X.**

Stochastic particle methods

$\boxed{\textbf{A critical event = A cascade of (much less) rare events}}$

Examples :
increasing energy levels, "physical" gateways, critical excursion sequences, numerical codes inputs levels associated with critical outputs, . . .

Stochastic particle methods

A critical event = A cascade of (much less) rare events

Examples :
increasing energy levels, "physical" gateways, critical excursion sequences, numerical codes inputs levels associated with critical outputs, . . .

$$\Downarrow$$

Conditional distributions = Optimal twisted measures

$$n \rightarrow \eta_n = \textbf{Law(r.v., path, excursion} \mid n \textbf{ intermediate events)}$$

$$\bigoplus$$

$$\begin{aligned} \textbf{Rare event probability} &= \textbf{Normalized constants } \mathcal{Z}_n \\ &= \mathbb{P}\left(n \textbf{ intermediate events}\right) \end{aligned}$$

Genetic type algorithm N individuals $(\xi_n^i)_{1\leq i\leq N}$

- (local) explorations/propositions/predictions of random states (noises, model uncertainties, unknown parameters)
- Branching-Selection-Duplication-Recycling predicted evolutions & "individuals" entering in higher critical events.

Genetic type algorithm N individuals $(\xi_n^i)_{1\leq i\leq N}$

- (local) explorations/propositions/predictions of random states (noises, model uncertainties, unknown parameters)
- Branching-Selection-Duplication-Recycling predicted evolutions & "individuals" entering in higher critical events.

Particle estimations :

$$\begin{aligned} \eta_n &= \text{Law(random states or paths} \mid n \text{ events)} \\ &\simeq_{N\uparrow\infty} \eta_n^N := \text{Occupation measure of a genealogical tree} \end{aligned}$$

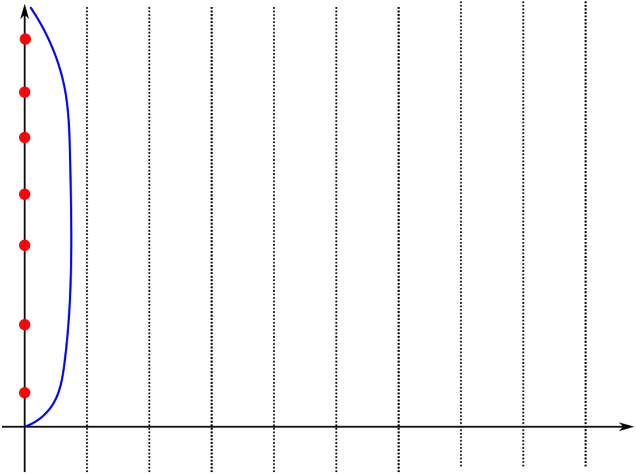
Unbias particle estimate :

$$\mathcal{Z}_n = \mathbb{P}(n \text{ successive events})$$

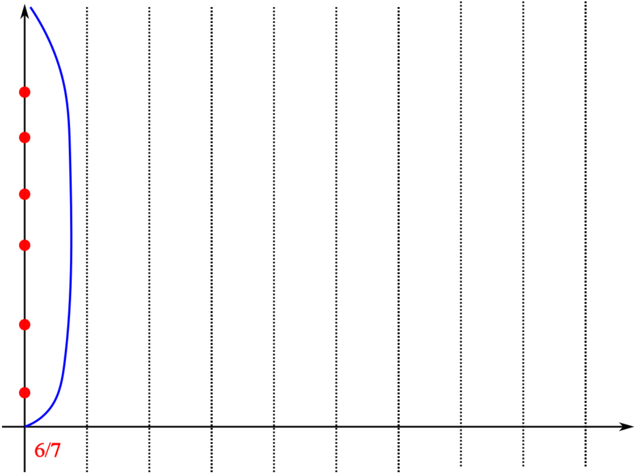
$$= \mathbb{P}(n\text{-th} \mid (n-1) \text{ events}) \times \ldots \times \mathbb{P}(\text{2nd} \mid \text{the first}) \times \mathbb{P}(\text{the first})$$

$$\simeq_{N\uparrow\infty} \mathcal{Z}_n^N := [\% \text{ success } (n-1) \rightsquigarrow n\ ] \times \ldots \times [\% \text{ success } 1 \rightsquigarrow 2\ ] \times [\% \text{ in } 1\ ]$$

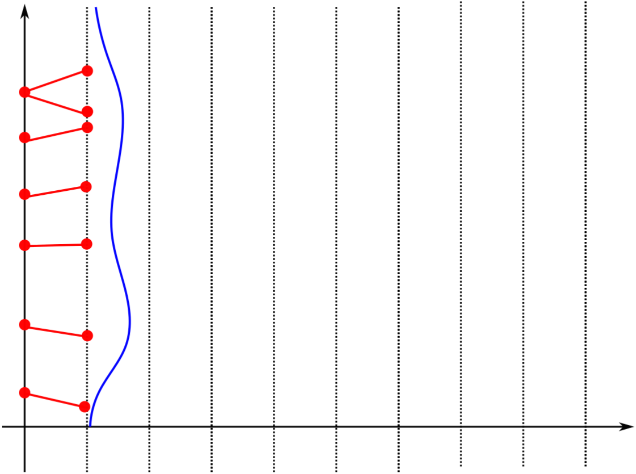
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \le i \le N} \delta_{\xi_n^i}$



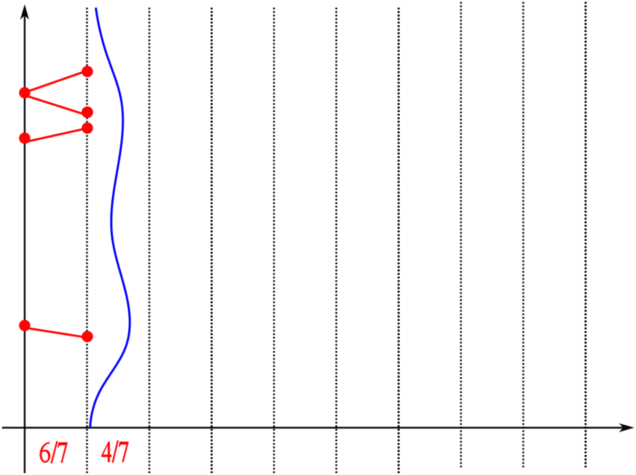
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N}\sum_{1\leq i\leq N}\delta_{\xi_n^i}$



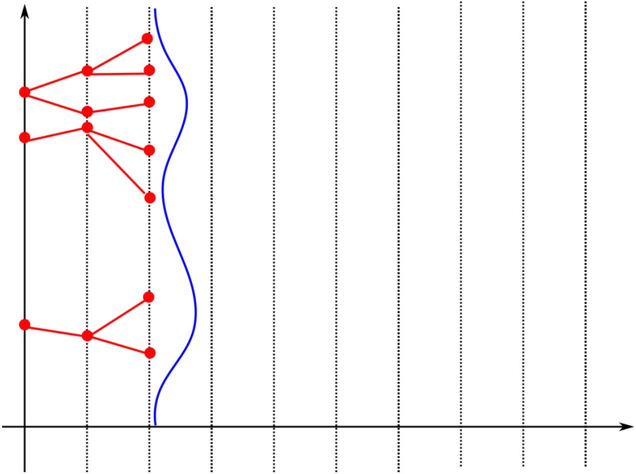
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



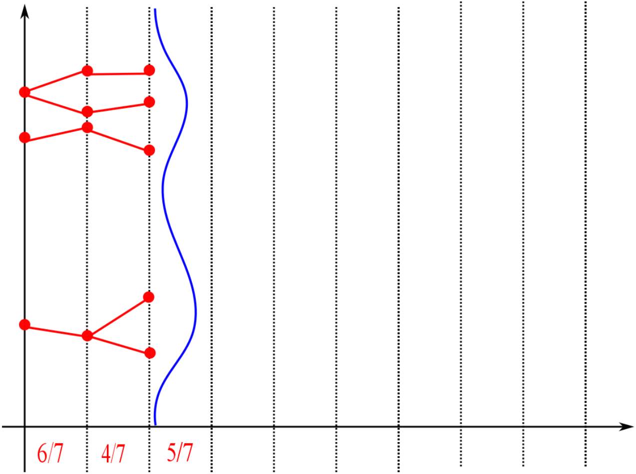
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



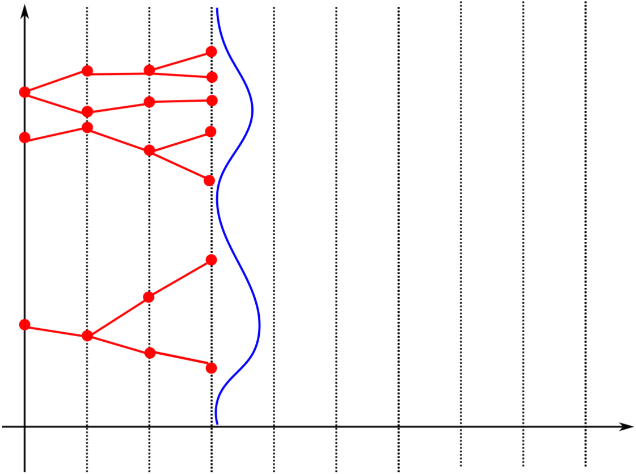
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



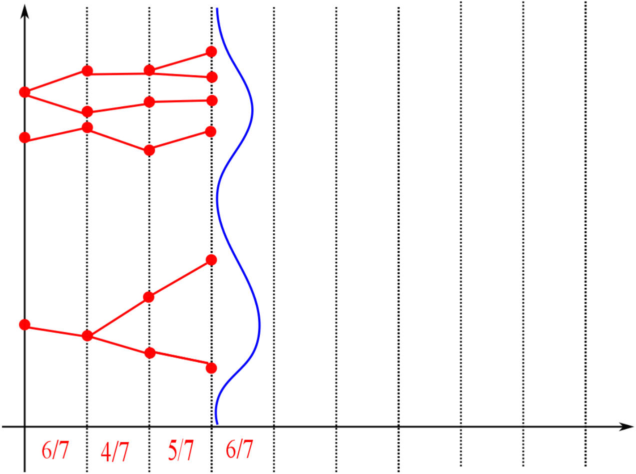
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



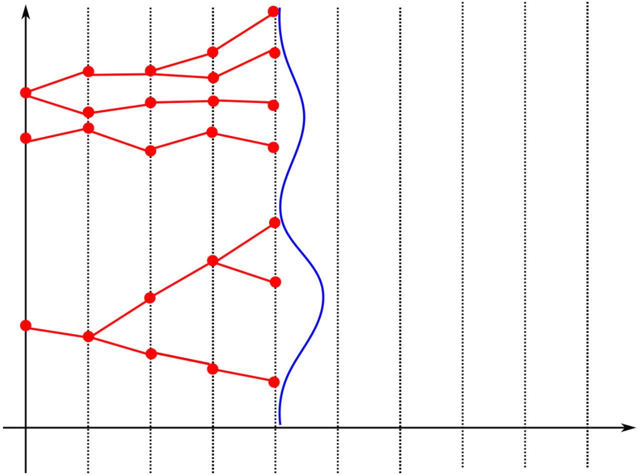
Graphical illustration : $\eta_n \simeq \ \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



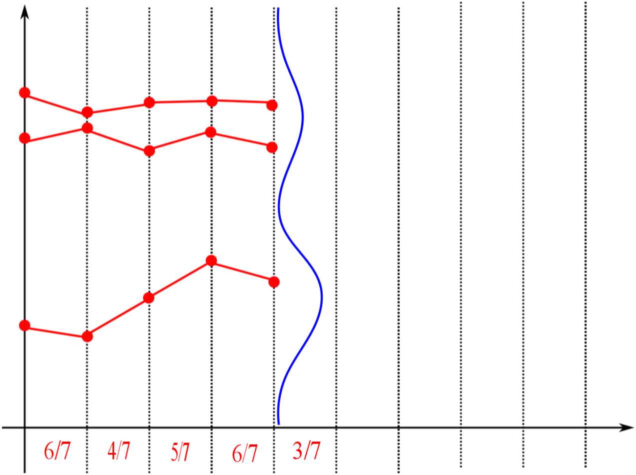
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



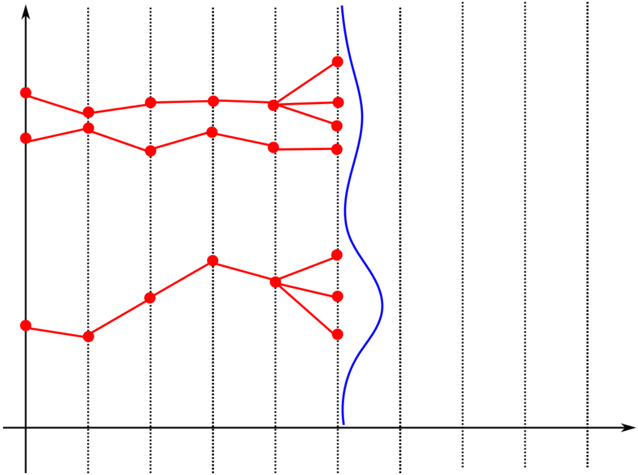
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N}\sum_{1\le i\le N} \delta_{\xi_n^i}$



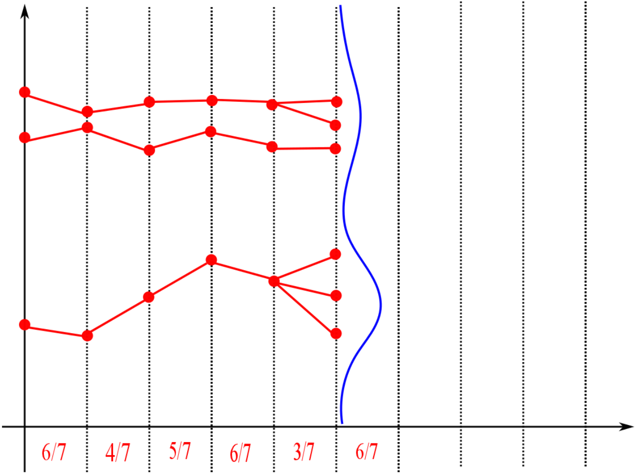
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



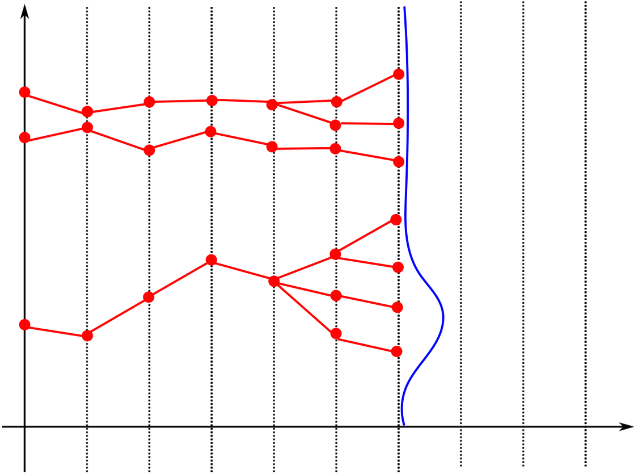
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N}\sum_{1\leq i\leq N} \delta_{\xi_n^i}$



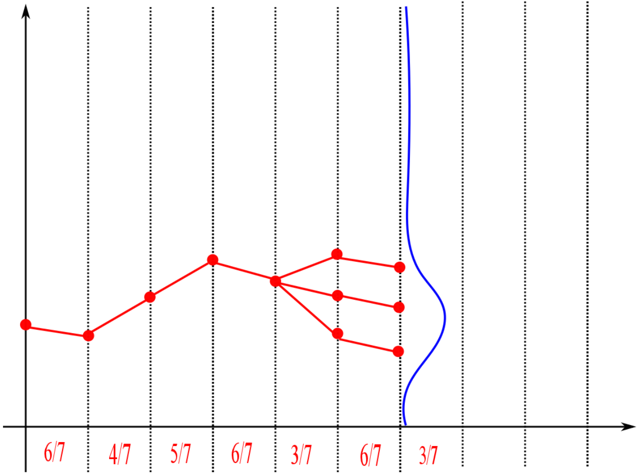
Graphical illustration : $\eta_n \simeq \ \eta_n^N := \frac{1}{N}\sum_{1\leq i\leq N}\delta_{\xi_n^i}$



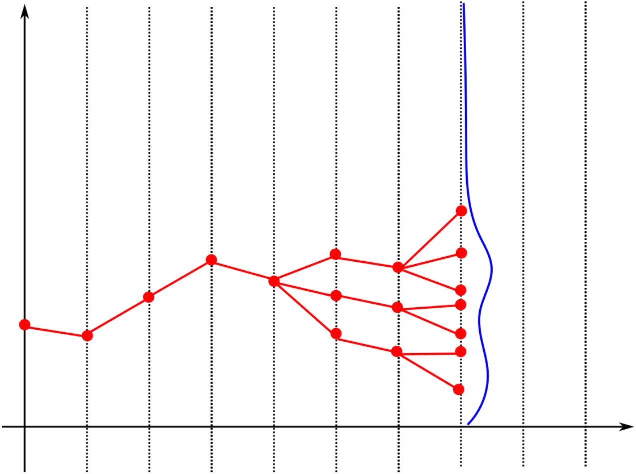
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



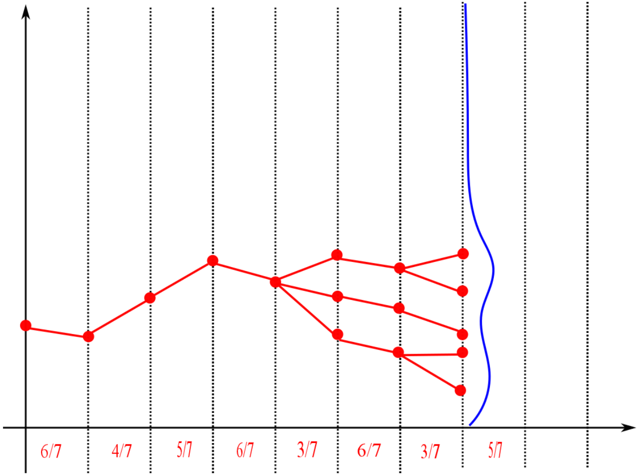
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



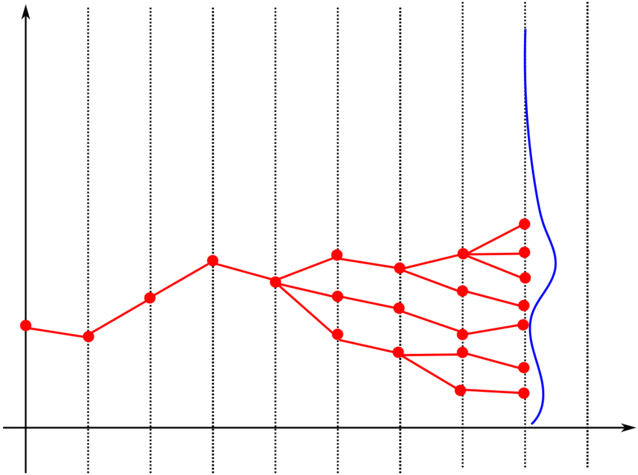
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



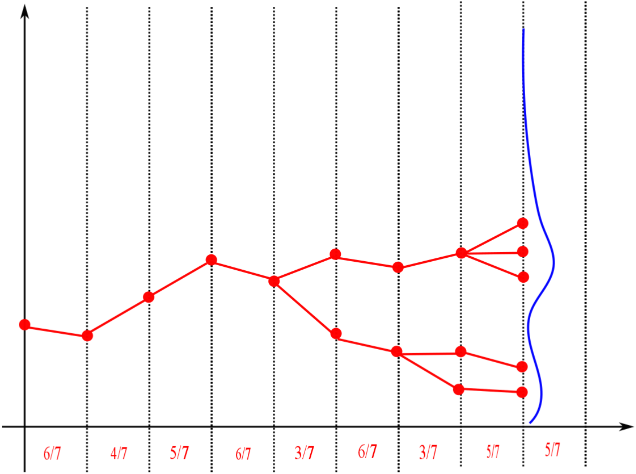
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



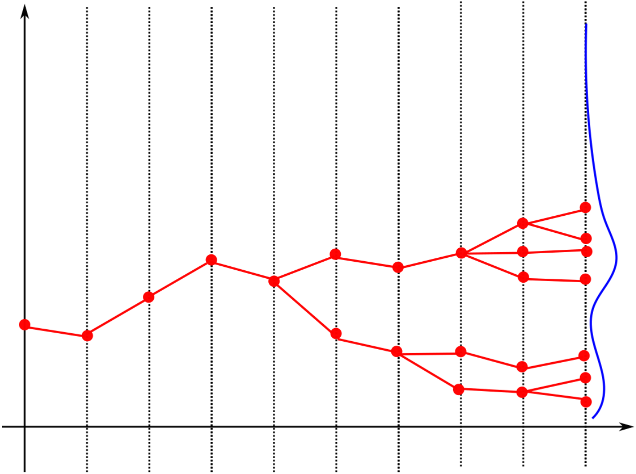
Graphical illustration : $\eta_n \simeq \eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i}$



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



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



Numerical codes: mechan. and hydrodyn.

(contract ALEA-IFREMER 2010)

- **Inputs $\mathcal{I}$:** waves heights, forecasting data.

  temperatures graphs, uncertainties and kinetic parameters.

  (statistical model : $\mathcal{I} = 2000$ Gaussian r.v.)

- **Ouputs $\mathcal{O}$:** force mappings $F_t(\mathcal{I})$ on the offshore structure (oil rigs, gas carriers).

- **Critical event :** hit a critical level a within the time horizon T.

$$A = B(a) = \left\{ \mathcal{I} \text{ s.t. } \sup_{0 \leq t \leq T} F_t(\mathcal{I}) \geq a \right\} \rightsquigarrow \textbf{Law}\left(\textbf{Inputs } \mathcal{I} \mid \mathcal{I} \in A\right)$$

Numerical codes: mechan. and hydrodyn.

(contract ALEA-IFREMER 2010)

- **Inputs $\mathcal{I}$:** waves heights, forecasting data.

  temperatures graphs, uncertainties and kinetic parameters.

  (statistical model : $\mathcal{I} = 2000$ Gaussian r.v.)

- **Ouputs $\mathcal{O}$:** force mappings $F_t(\mathcal{I})$ on the offshore structure (oil rigs, gas carriers).

- **Critical event :** hit a critical level a within the time horizon T.

$$A = B(a) = \left\{ \mathcal{I} \text{ s.t. } \sup_{0 \leq t \leq T} F_t(\mathcal{I}) \geq a \right\} \rightsquigarrow \textbf{Law} \left(\textbf{Inputs } \mathcal{I} \mid \mathcal{I} \in A \right)$$

Multi-level particle branching algorithm w.r.t. $A_n = B(a_n)$ with $a_n \uparrow$

- MCMC exploration of the level sets A_n.
- Selection of the individual $\in A_{n+1}$ ($\oplus$ adaptive choice of A_{n+1})

Nuclear plant security

(Contract CIFRE, ALEA-EDF R&D Industrial Risks)

- **Inputs $\mathcal{I}$:** material specif./defaults, injection temperatures, model uncertainties and unknown parameters.

  (ex. stat. model : $\mathcal{I} = 20$ Gaussian r.v.)

- **Outputs = neutronics/mechanics/hydro. codes : $\mathcal{O} = \mathcal{F}(\mathcal{I})$**

- **Critical event :**

  $\mathcal{G}(\mathcal{O}) \in \mathbb{R}$ risk measure function ($\mathcal{G}^{-1}\{0\}$=default hyper-surface)

$$A = \{\mathcal{I} \text{ s.t. } \mathcal{G} \circ \mathcal{F}(\mathcal{I}) < 0\} \rightsquigarrow \mathrm{Law}\,(\mathcal{I} \mid \mathcal{I} \in A)$$

Nuclear plant security

(Contract CIFRE, ALEA-EDF R&D Industrial Risks)

- **Inputs $\mathcal{I}$:** material specif./defaults, injection temperatures, model uncertainties and unknown parameters.

  (ex. stat. model : $\mathcal{I} = 20$ Gaussian r.v.)

- **Outputs = neutronics/mechanics/hydro. codes :** $\mathcal{O} = \mathcal{F}(\mathcal{I})$

- **Critical event :**

  $\mathcal{G}(\mathcal{O}) \in \mathbb{R}$ risk measure function ($\mathcal{G}^{-1}\{0\}$=default hyper-surface)

$$A = \{\mathcal{I} \text{ s.t. } \mathcal{G} \circ \mathcal{F}(\mathcal{I}) < 0\} \rightsquigarrow \mathrm{Law}\left(\mathcal{I} \mid \mathcal{I} \in A\right)$$

The same particle algorithm as before with $A_n \downarrow$
$\oplus$ **Sensitivity analysis (r.v. correlations, parameters estimations)**

Watermarking & Digital fingerprinting (Illegal copy detection systems)

(ANR Nebbiano 06-09, INRIA ALEA-ASPI-TEMICS)

- **Random variables :** hidden code (fingerprint)

  Formalisation : bit series $X = (X^1, \ldots, X^d) \in \{0, 1\}$ or Gauss. r.v.

- **Rare event :** Innocent prosecution (illegal copy).

Watermarking & Digital fingerprinting (Illegal copy detection systems)

(ANR Nebbiano 06-09, INRIA ALEA-ASPI-TEMICS)

- **Random variables :** hidden code (fingerprint)

  Formalisation : bit series $X = (X^1, \ldots, X^d) \in \{0,1\}$ or Gauss. r.v.

- **Rare event :** Innocent prosecution (illegal copy).

We find a document with some code $y = (y^1, \ldots, y^d)$.
Wrong accusation/Mistaken degree $\sim$ level passing of a function:

$$\begin{aligned}
\mathcal{F}(X) &= \sum_{1 \leq i \leq d} y^i \, f(X^i) \geq a \quad \text{Tardos Codes} \\
\mathcal{F}(X) &= |\langle X, u \rangle| \,/ \|X\| \geq a \quad \text{Zero-bit watermarking } (\|u\| = 1)
\end{aligned}$$

$$A = \{X \text{ such that } \mathcal{F}(X) \geq a\} \rightsquigarrow \text{Law}\,(X \mid \mathcal{F}(X) \geq a)$$

Watermarking & Digital fingerprinting (Illegal copy detection systems)

(ANR Nebbiano 06-09, INRIA ALEA-ASPI-TEMICS)

- **Random variables :** hidden code (fingerprint)

  Formalisation : bit series $X = (X^1, \ldots, X^d) \in \{0,1\}$ or Gauss. r.v.

- **Rare event :** Innocent prosecution (illegal copy).

We find a document with some code $y = (y^1, \ldots, y^d)$.
Wrong accusation/Mistaken degree $\sim$ level passing of a function:

$$\begin{aligned}
\mathcal{F}(X) &= \sum_{1\leq i\leq d} y^i \; f(X^i) \geq a \quad \text{Tardos Codes} \\
\mathcal{F}(X) &= |\langle X, u\rangle| \,/ \|X\| \geq a \quad \text{Zero-bit watermarking } (\|u\| = 1)
\end{aligned}$$

$$A = \{X \text{ such that } \mathcal{F}(X) \geq a\} \rightsquigarrow \text{Law}\left(X \mid \mathcal{F}(X) \geq a\right)$$

$\Rightarrow$ **Same multi-level branching algorithm as before $A_n \downarrow$**
$\oplus$ **Sensitivity analysis (r.v. correlations, parameters estimations)**

Fiber optics communication

- DM + J. Garnier. Simulations of rare events in fiber optics by interacting particle systems. Optics Communications (2006).
- DM + J. Garnier. Genealogical particle analysis of rare events. Ann. Appl. Probab. (2005).

Stochastic model :

- **Elementary pulses X_k in each k-th fiber section (virtual) :** soliton type pulse profiles.
- **Random var. :** ω_k = randomness/perturbations in each section. (speed dispersion fluctuations)

$$X_k = \underbrace{F_k(X_{k-1}, \omega_k)}_{\text{nonlinear Schrödinger equation}}$$

Rare event : critical level of a characteristic value $V(X_T)$ (pulse profile width, modal dispersion, loss of amplitude power).

$$\mathbb{P}(V(X_T) \geq a) \quad \& \quad \textbf{Law}\left((X_t)_{0\leq t\leq T} \mid V(X_T) \geq a\right)$$

Fiber optics communication (continued)

Particle algorithm = N particles $\xi_t^i = (X_t^i, X_{t+1}^i)$

- Free transition space exploration $X_{t-1} \rightsquigarrow X_t = F_t(X_{t-1}, \omega_t)$.
- Selection-Recycling $\uparrow$ levels (criteria/weight $e^{\alpha(V(X_t)-V(X_{t-1}))}$)

Fiber optics communication (continued)

Particle algorithm $=$ N particles $\xi_t^i = (X_t^i, X_{t+1}^i)$

- Free transition space exploration $X_{t-1} \rightsquigarrow X_t = F_t(X_{t-1}, \omega_t)$.
- Selection-Recycling $\uparrow$ levels (criteria/weight $e^{\alpha(V(X_t)-V(X_{t-1}))}$)

$$\Downarrow$$

Unbias estimate of the normalizing constants

$$\mathcal{Z}_T = \mathbb{P}(V_T(X_T) \geq a)$$

$$= \mathbb{E}\left(1_{V_T(X_T)\geq a}\ e^{\alpha V(X_T)}\ \left\{\prod_{0\leq t\leq T} e^{\alpha(V(X_t)-V(X_{t-1}))}\right\}\right)$$

$$\simeq_{N\uparrow} \left(\frac{1}{N}\sum_{1\leq i\leq N} 1_{V_T(X_T^i)\geq a}\ e^{\alpha V(X_T^i)}\right) \times \prod_{0\leq t\leq T} \frac{1}{N}\sum_{1\leq i\leq N} e^{\alpha\left(V(X_t^i)-V(X_{t-1}^i)\right)}$$

$$\oplus$$

Genealogical tree $\simeq_{N\uparrow}$ Default tree model

Food risk & Epidemic propagations

- `ANR VIROSCOPY 08-11: Epidemic propagations analysis INRIA-ENST.`
- `ARC EPS INRA-INRIA : Eco-microbiologie previsonnelle (09-10).`
- CNRS Project : ENS Paris & Bordeaux Mathematical Institute (2011-2013).

Kinetic models $\sim$ unknown parameters Θ :

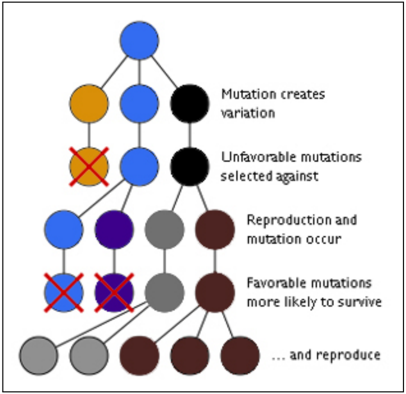
$$X_k = F_k(X_{k-1}, \omega_k, \Theta)$$

- Calibration : partial and noisy observations (web, mesures stat.)
  $\rightsquigarrow$ **Particle filters and sequential Monte Carlo techniques.**
- Risk analysis : level crossing type models $V(X) > a$, critical excursions,...
  $\rightsquigarrow$ **Same type of particle models (branching-selection)**

Genetic type particle models

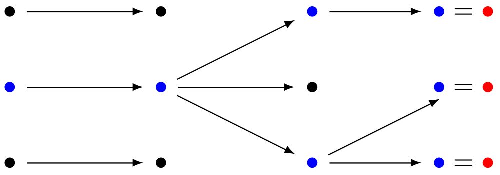
Interacting sampling+ adaptative + universal

- Mutation-Propositions : Markov transitions $X_{n-1} \rightsquigarrow X_n \in E_n$.
- Selection-Accept/Rejet-Recycling : local criteria $\rightsquigarrow$ function G_n.



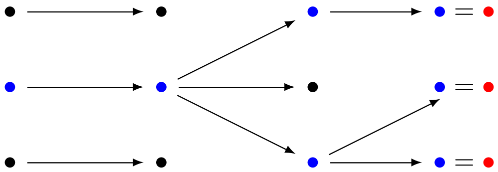
Genealogical tree models

(Population size, Time horizon)=$(N, n) = (3, 3)$



Genealogical tree models

(Population size, Time horizon)=$(N, n) = (3, 3)$



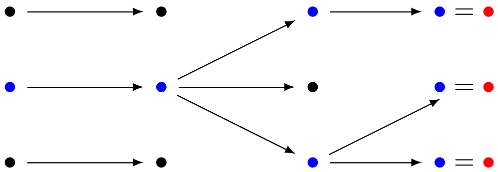
Theoreme (96-98'):

Ancestral lines $\simeq$ i.i.d. $\sim =$ Feynman-Kac measures

$$\mathbb{Q}_n := \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0 \leq p < n} G_p(X_p) \right\} \mathbb{P}_n \quad \text{avec} \quad \mathbb{P}_n := \mathrm{Law}\left(X_0, \ldots, X_n\right)$$

Genealogical tree models

(Population size, Time horizon)=$(N, n) = (3, 3)$



Theoreme (96-98'):

Ancestral lines $\simeq$ i.i.d. $\sim$ = Feynman-Kac measures

$$\mathbb{Q}_n := \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0 \leq p < n} G_p(X_p) \right\} \mathbb{P}_n \quad \text{avec} \quad \mathbb{P}_n := \mathrm{Law}\,(X_0, \ldots, X_n)$$

Example

$$G_n = 1_{A_n} \to \mathbb{Q}_n = \mathrm{Law}((X_0, \ldots, X_n) \mid X_p \in A_p,\ p < n)$$

Sensitivity measures $\theta \in \mathbb{R}^{d=1} \rightsquigarrow \left(X_n^\theta, G_n^\theta, \mathbb{Q}_n^\theta\right)$

Rare event probabilities = Normalizing constants

$$\mathcal{Z}_n(\theta) = \mathbb{P}(\text{critical event} \sim \theta) = \mathbb{E}\left(\prod_{0 \leq p < n} G_p^\theta(X_p^\theta)\right)$$

Sensitivity measures $\theta \in \mathbb{R}^{d=1} \rightsquigarrow \left(X_n^\theta, G_n^\theta, \mathbb{Q}_n^\theta\right)$

Rare event probabilities = Normalizing constants

$$\mathcal{Z}_n(\theta) = \mathbb{P}(\text{critical event} \sim \theta) = \mathbb{E}\left(\prod_{0\leq p<n} G_p^\theta(X_p^\theta)\right)$$

$$\frac{\partial}{\partial\theta}\log \mathcal{Z}_n(\theta) = \mathbb{Q}_n\left(\mathbf{f_n}\right) = \int_{\mathbf{x}=\text{paths}} \mathbb{Q}_n^\theta(\mathbf{dx})\ \ \mathbf{f_n}(\mathbf{x})$$

with the additive functional :

$$\mathbf{f_n}(x_0,\ldots,x_n) = \frac{1}{n}\sum_{0\leq p<n}\frac{\partial}{\partial\theta}\log\left(p_\theta(x_{p+1}|x_p)\ G_p^\theta(x_p)\right)$$

Sensitivity measures $\theta \in \mathbb{R}^{d=1} \rightsquigarrow \left(X_n^\theta, G_n^\theta, \mathbb{Q}_n^\theta\right)$

Rare event probabilities = Normalizing constants

$$\mathcal{Z}_n(\theta) = \mathbb{P}(\text{critical event} \sim \theta) = \mathbb{E}\left(\prod_{0 \leq p < n} G_p^\theta(X_p^\theta)\right)$$

$$\frac{\partial}{\partial\theta} \log \mathcal{Z}_n(\theta) = \mathbb{Q}_n\left(\mathbf{f_n}\right) = \int_{\mathbf{x}=\text{paths}} \quad \mathbb{Q}_n^\theta(\mathbf{dx}) \;\; \mathbf{f_n}(\mathbf{x})$$

with the additive functional :

$$\mathbf{f_n}(x_0, \ldots, x_n) = \frac{1}{n} \sum_{0 \leq p < n} \frac{\partial}{\partial\theta} \log\left(p_\theta(x_{p+1}|x_p) \;\; G_p^\theta(x_p)\right)$$

$$\downarrow$$

Particle approximations

$$\frac{\partial}{\partial\theta} \log \mathcal{Z}_n(\theta) = \mathbb{Q}_n^\theta\left(\mathbf{f_n}\right) \simeq_{N\uparrow} \frac{1}{N} \sum_{1 \leq i \leq N} \mathbf{f_n}(\text{ancestral line}_n(i))$$

In summary

⊕ *Genealogical tree*

$$\begin{aligned}\mathbb{A}_n^N(f_n) &= \frac{1}{N}\sum_{1\leq i\leq N} f_n\left(\text{ancestral line}_n(i)\right)\\ &\simeq_{N\to\infty} \mathbb{Q}_n(f_n) = \mathbb{E}\left(f_n(X_0,\ldots,X_n)\mid \text{critical event}\right)\end{aligned}$$

In summary

- *Genealogical tree*

$$\begin{aligned}\mathbb{A}_n^N(f_n) &= \frac{1}{N}\sum_{1\leq i\leq N} f_n\left(\text{ancestral line}_n(i)\right)\\ &\simeq_{N\to\infty} \mathbb{Q}_n(f_n) = \mathbb{E}\left(f_n(X_0,\ldots,X_n) \mid \text{critical event}\right)\end{aligned}$$

- *Current population*

$$\eta_n^N(f_n) := \frac{1}{N}\sum_{1\leq i\leq N} f_n(\xi_n^i) \longrightarrow_{N\to\infty} \eta_n(f_n) = \mathbb{E}\left(f_n(X_n) \mid \text{critical event}\right)$$

In summary

⊕ *Genealogical tree*

$$\mathbb{A}_n^N(f_n) \quad = \quad \frac{1}{N}\sum_{1\leq i\leq N} f_n\left(\text{ancestral line}_n(i)\right)$$

$$\simeq_{N\to\infty} \quad \mathbb{Q}_n(f_n) = \mathbb{E}\left(f_n(X_0,\ldots,X_n) \mid \text{critical event}\right)$$

⊕ *Current population*

$$\eta_n^N(f_n) := \frac{1}{N}\sum_{1\leq i\leq N} f_n(\xi_n^i) \longrightarrow_{N\to\infty} \eta_n(f_n) = \mathbb{E}\left(f_n(X_n) \mid \text{critical event}\right)$$

⊕ *Unbias particle approximation*

$$\mathcal{Z}_n^N = \prod_{0\leq p<n} \frac{1}{N}\sum_{1\leq i\leq N} G_p(\xi_p^i) \longrightarrow_{N\to\infty} \mathcal{Z}_n = \mathbb{P}\left(\text{critical event}\right)$$

In summary

⊕ *Genealogical tree*

$$\mathbb{A}_n^N(f_n) \quad = \quad \frac{1}{N}\sum_{1\leq i\leq N} f_n\left(\text{ancestral line}_n(i)\right)$$

$$\simeq_{N\to\infty} \quad \mathbb{Q}_n(f_n) = \mathbb{E}\left(f_n(X_0,\ldots,X_n) \mid \text{critical event}\right)$$

⊕ *Current population*

$$\eta_n^N(f_n) := \frac{1}{N}\sum_{1\leq i\leq N} f_n(\xi_n^i) \longrightarrow_{N\to\infty} \eta_n(f_n) = \mathbb{E}\left(f_n(X_n) \mid \text{critical event}\right)$$

⊕ *Unbias particle approximation*

$$\mathcal{Z}_n^N = \prod_{0\leq p<n} \frac{1}{N}\sum_{1\leq i\leq N} G_p(\xi_p^i) \longrightarrow_{N\to\infty} \mathcal{Z}_n = \mathbb{P}\left(\text{critical event}\right)$$

Ex.: $G_n = 1_{A_n} \rightsquigarrow \mathcal{Z}_n^N = \prod \%\ \textit{success} \longrightarrow \mathbb{P}(X_p \in A_p,\ p < n)$

$\oplus$ Compelete genealogical tree $= \left[(\xi_p^i)_{1\leq i\leq N}\ ,\ 0\leq p\leq n\right]$

Backward Markovian model

$$\mathbb{Q}_n^N(\xi_0^{i_0},\xi_1^{i_1},\ldots,\xi_n^{i_n}) := \frac{1}{N}\times \mathbb{M}_n^N(i_n,i_{n-1})\times\ldots\times\mathbb{M}_2^N(i_2,i_1)\times\mathbb{M}_1^N(i_1,i_0)$$

with the random transitions : (p_n=transition densities $x_{n-1}\rightsquigarrow x_n$)

$$\mathbb{M}_{n+1}^N(i_{n+1},i_n) = \frac{p_{n+1}(\xi_{n+1}^{i_{n+1}}|\xi_n^{i_n})G_n(\xi_n^{i_n})}{\sum_{1\leq k\leq N} p_{n+1}(\xi_{n+1}^{i_{n+1}}|\xi_n^k)G_n(\xi_n^k)}$$

$\oplus$ Compelete genealogical tree $= \left[(\xi_p^i)_{1\leq i\leq N}\ ,\ 0\leq p\leq n\right]$

Backward Markovian model

$$\mathbb{Q}_n^N(\xi_0^{i_0},\xi_1^{i_1},\ldots,\xi_n^{i_n}) := \frac{1}{N}\times\mathbb{M}_n^N(i_n,i_{n-1})\times\ldots\times\mathbb{M}_2^N(i_2,i_1)\times\mathbb{M}_1^N(i_1,i_0)$$

with the random transitions : (p_n=transition densities $x_{n-1}\rightsquigarrow x_n$)

$$\mathbb{M}_{n+1}^N(i_{n+1},i_n) = \frac{p_{n+1}(\xi_{n+1}^{i_{n+1}}|\xi_n^{i_n})G_n(\xi_n^{i_n})}{\sum_{1\leq k\leq N}p_{n+1}(\xi_{n+1}^{i_{n+1}}|\xi_n^k)G_n(\xi_n^k)}$$

Example: Additive functional integration

$$\mathbf{f_n}(x_0,\ldots,x_n) = \frac{1}{n+1}\sum_{0\leq p\leq n}f_p(x_p)$$

$$\Downarrow$$

$$\mathbb{Q}_n(\mathbf{f_n})\simeq_{N\uparrow}\mathbb{Q}_n^N(\mathbf{f_n}) := \frac{1}{n+1}\sum_{0\leq p\leq n}\left[\frac{1}{N},\ldots,\frac{1}{N}\right]\mathbb{M}_n^N\mathbb{M}_{n-1}^N\ldots\mathbb{M}_{p+1}^N\begin{bmatrix}f_p(\xi_p^1)\\ \vdots\\ f_p(\xi_p^N)\end{bmatrix}$$

Concentration inequalities

Constants (c_1, c_2) related to (biais,variance) $\perp$ n, c universal ct, $\|f_n\| \leq 1$, et $\mathbf{F_n}$ normalised additive functional

$\forall$ $(x \geq 0, n \geq 0, N \geq 1, \epsilon \in \{+1, -1\})$, the probability of the following inequalities is larger than $1 - e^{-x}$:

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

$$\left[\mathbb{A}_n^N - \mathbb{Q}_n\right](f_n) \leq c_1\ \frac{n+1}{N}\ \left(1 + x + \sqrt{x}\right) + c_2\ \sqrt{\frac{(n+1)}{N}}\ \sqrt{x}$$

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

$$\frac{\epsilon}{n} \log \frac{\mathcal{Z}_n^N}{\mathcal{Z}_n} \leq \frac{c_1}{N}\ \left(1 + x + \sqrt{x}\right) + \frac{c_2}{\sqrt{N}}\ \sqrt{x}$$

Ongoing research projects

- Genealogical trees = Defaults trees = Typical critical trajectories
- Computation of the influent parameters w.r.t. critical event

$$\Downarrow$$

- $\rightsquigarrow$ Control of processes evolving in critical regimes
- $\rightsquigarrow$ Increase the reliability of systems
- $\rightsquigarrow$ Prediction of events w.r.t. observations