Particle Methods,
World Congress of the Bernoulli Society, Barcelona 2004

Genealogical tree models and Nonlinear estimation

Pierre DEL MORAL
LSP-C.N.R.S, Univ. Paul Sabatier, Toulouse. ⇝ Univ. Nice-Sophia Antipolis

↪ *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*, *Springer NY. Series: Probability and Applications, (2004)*

↪ (delmoral@cict.fr), ↪ [preprints+info.] http://www.lsp.ups-tlse.fr/Delmoral/

Evolutionary type models

Simple Genetic Branching Algo.	*Mutation*	*Selection/Branching*
Metropolis-Hastings Algo.	*Proposal*	*Acceptance/Rejection*
Sequential Monte Carlo methods	*Sampling*	*Resampling (SIR)*
Filtering/Smoothing	*Prediction*	*Updating/Correction*
Particle $\in$ Absorbing Medium	*Evolution*	*Killing/Creation/Anhiling*

Other Botanical Names: multi-level splitting (Khan-Harris 51), prune enrichment (Rosenbluth 1955), switching algo. (Magill 65), matrix reconfiguration (Hetherington 84), restart (Villen-Altamirano 91), particle filters (Rigal-Salut-DM 92), SIR filters (Gordon-Salmon-Smith 93, Kitagawa 96), go-with-the-winner (Vazirani-Aldous 94), ensemble Kalman-filters (Evensen 1994), quantum Monte Carlo methods (Melik-Nightingale 1999), spawning filters (Fisher-Maybeck 2002), SIR Pilot Exploration Resampling (Liu-Zhang 2002),...

$\Longleftrightarrow$ Particle Interpretations of Feynman-Kac models

Since R. Feynman's phD. on path integrals 1942

Physics $\longleftrightarrow$ Biology $\longleftrightarrow$ Engineering Sciences $\longleftrightarrow$ Probability/Statistics

- **Physics** :
  - $FKS \in$ nonlinear integro-diff. éq. ($\sim$ 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.
- **Biology**:
  - Self-avoiding walks, macromolecular polymerizations.
  - Branching and genetic population models.
  - Coalescent and Genealogical evolutions.

- **Rare events analysis**:
  - Multisplitting and branching particle models (Restart).
  - Importance sampling and twisted probability measures.
  - Genealogical tree based simulation methods.
- **Advanced Signal processing**:
  - Optimal filtering/smoothing/regulation, open loop optimal control.
  - Interacting Kalman-Bucy filters.
  - Stochastic and adaptative grid approximation-models
- **Statistics/Probability**:
  - Restricted Markov chains (w.r.t terminal values, visiting regions,...)
  - Analysis of Boltzmann-Gibbs type distributions (simulation, partition functions,...).
  - Random search evolutionary algorithms, interacting Metropolis/simulated annealing algo.

Simple Genetic evolution/simulation models $\longrightarrow$ only 2 ingredients!!

(Discrete time parameter $n \in \mathbb{N} = \{0, 1, 2, ...\}$, state spaces E_n ($\in \{\mathbb{Z}^d, \mathbb{R}^d, \underbrace{\mathbb{R}^d \times \ldots \times \mathbb{R}^d}_{(n+1)-times}, ...\}$)

- *Mutation/exploration/prediction/proposal* :

  $\rightarrow$ Markov transitions $M_n(x_{n-1}, dx_n)$ from E_{n-1} into E_n.

- *Selection/absorption/updating/acceptance* :

  $\rightarrow$ Potential functions G_n from E_n into $[0, 1]$.

A Genetic Evolution Model $\Rightarrow$ Markov chain $\xi_n = (\xi_n^1, \ldots, \xi_n^N) \in E_n^N = \underbrace{E_n \times \ldots \times E_n}_{N-times}$

$$\xi_n \in E_n^N \xrightarrow{\text{selection}} \widehat{\xi}_n \in E_n^N \xrightarrow{\text{mutation}} \xi_{n+1} \in E_{n+1}^N$$

- **Selection transition** ($\exists \neq$ types $\rightarrow$ Ex.: accept/reject)

$$\xi_n^i \rightsquigarrow \widehat{\xi}_n^i = \xi_n^i \quad \text{with proba. } G_n(\xi_n^i) \quad \textbf{[Acceptance]}$$

Otherwise we select a better fitted individual in the current configuration

$$\widehat{\xi}_n^i = \xi_n^j \quad \text{with proba. } G_n(\xi_n^j)/\sum_{k=1}^N G_n(\xi_n^k) \quad \textbf{[Rejection + Selection]}$$

- **Mutation transition**

$$\widehat{\xi}_n^i \rightsquigarrow \xi_{n+1}^i \sim M_{n+1}(\widehat{\xi}_n^i, \centerdot)$$

A Genealogical tree model

Important observation [Historical process]

$$X'_n \in E'_n \quad \text{Markov chain}$$

$$\Downarrow$$

$$X_n = (X'_0, \ldots, X'_n) \in E_n = (E'_0 \times \ldots \times E'_n) \quad \text{Markov chain} \in \text{path spaces}$$

$\rightarrow$ *Markov transitions* $M_n(x_{n-1}, dx_n)$ [elementary extensions]

$$X_{n+1} = ((X'_0, \ldots, X'_n), X'_{n+1}) = (X_n, X'_{n+1})$$

Genetic Evolution Model on Path Spaces=Genealogical tree model

$X_n = (X'_0, \ldots, X'_n)$ Markov transitions M_n and $G_n(X_n) = G'_n(X'_n)$

$$\downarrow$$

Genetic path-valued particle Model

$$\begin{cases} \xi^i_n &= (\xi^i_{0,n}, \xi^i_{1,n}, \ldots, \xi^i_{n,n}) \\ \widehat{\xi}^i_n &= (\widehat{\xi}^i_{0,n}, \widehat{\xi}^i_{1,n}, \ldots, \widehat{\xi}^i_{n,n}) \in E_n = (E'_0 \times \ldots \times E'_n) \end{cases}$$

- Path acceptance/(rejection+selection).
- Path mutation = path elementary extensions.

Occupation/Empirical measures ($\forall f_n$ test function on E_n)

$$\eta_n^N(f_n) = \frac{1}{N}\sum_{i=1}^N f_n(\xi_n^i) = \frac{1}{N}\sum_{i=1}^N f_n \underbrace{(\xi_{0,n}^i, \xi_{1,n}^i, \ldots, \xi_{n,n}^i)}_{i\text{-th ancestral lines}}$$

$$\downarrow$$

Unbias-particle measures & Unnormalized Feynman-Kac measures :

$$\gamma_n^N(f_n) \quad = \quad \eta_n^N(f_n) \times \prod_{0\leq p<n} \eta_p^N(G_p) \longrightarrow_{N\to\infty} \gamma_n(f_n) = \mathbb{E}(f_n(X_n) \prod_{0\leq p<n} G_p(X_p))$$

Notes:

- $f_n = 1 \Rightarrow \gamma_n^N(1) = \prod_{0\leq p<n} \eta_p^N(G_p) \longrightarrow_{N\to\infty} \gamma_n(1) = \mathbb{E}(\prod_{0\leq p<n} G_p(X_p))$
- *Path-space models*

$$[\ X_n = (X'_0, \ldots, X'_n) \text{ and } G_n(X_n) = G'_n(X'_n)\ ] \Rightarrow \gamma_n(f_n) = \mathbb{E}(f_n(X'_0, \ldots, X'_n) \prod_{0\leq p<n} G'_p(X'_p))$$

$\Longrightarrow$ *Occupation measure & Normalized Feynman-Kac measures:*

$$\eta_n^N(f_n) \quad = \quad \frac{1}{N}\sum_{i=1}^N f_n(\xi_n^i) = \gamma_n^N(f_n)/\gamma_n^N(1) \longrightarrow_{N\to\infty} \eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1)$$

Path-space models

$$[\ X_n = (X'_0,\ldots,X'_n) \text{ and } G_n(X_n) = G'_n(X'_n)\ ]$$

$$\Downarrow$$

$$\eta_n(f_n) = \frac{\mathbb{E}(f_n(X'_0,\ldots,X'_n)\prod_{0\le p<n} G'_p(X'_p))}{\mathbb{E}(\prod_{0\le p<n} G'_p(X'_p))}$$

Note:

$$\gamma_n(f_n) = \eta_n(f_n) \times \prod_{0\le p<n} \eta_p(G_p) \quad (\longleftarrow \gamma_n^N(f_n) = \eta_n^N(f_n) \times \prod_{0\le p<n} \eta_p^N(G_p))$$

Complete asymptotic theory $(n, N) \to \infty$ (weak+strong LLN, CLT, LDP,...)

$\hookrightarrow$ *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications, Springer NY. Series: Probability and Applications (2004)*

Some examples:

- *Weak convergence* [$p \geq 1$ + $\mathcal{F}_n$ not too large + regular mutations]

$$\sup_{n\geq 0} \mathbb{E}(\sup_{f_n\in\mathcal{F}_n} |\eta_n^N(f_n) - \eta_n(f_n)|^p)^{1/p} \leq c(p)/\sqrt{N}$$

Ex.:

$$E_n = \mathbb{R}, \quad \mathcal{F}_n = \{1_{]-\infty,x]} \ ; \ x \in \mathbb{R}\} \Rightarrow \sup_{n\geq 0} \mathbb{E}(\sup_{x\in\mathbb{R}} |\eta_n^N(1_{]-\infty,x]}) - \eta_n(1_{]-\infty,x]})|^p)^{1/p} \leq c(p)/\sqrt{N}$$

- *Propagation-of-chaos estimates* [$q \leq N$ finite block size, regular mutations]

$$\text{Law}(\xi_n^1, \ldots, \xi_n^q) \simeq \eta_n^{\otimes q} + \frac{1}{N} \mathcal{M}_n^{(q)} \quad \text{with} \quad \mathcal{M}_n^{(q)} \quad \text{signed meas. s.t.} \sup_{n\geq 0} \sup_{\|F\|\leq 1} |\mathcal{M}_n^{(q)}(F)| \leq c\ q^2$$

Particle physics: Markov $X_n \in$ **Absorbing medium** $G(x) = e^{-V(x)} \in [0,1]$

$$X_n^c \in E^c = E \cup \{c\} \xrightarrow{absorption} \widehat{X}_n^c \xrightarrow{exploration} X_{n+1}^c$$

Absorption/killing: $\longrightarrow \widehat{X}_n^c = X_n^c$, with proba $G(X_n^c)$; otherwise the particle is killed and $\widehat{X}_n^c = c$.

$$\Downarrow$$

$$\begin{aligned} A &= \{x \; : \; G(x) = 0\} \longrightarrow \text{Hard obstacles} \\ T &= \inf \{n \geq 0 \; ; \; \widehat{X}_n^c = c\} \longrightarrow \text{Absorption time} \;\; X_{T+n}^c = \widehat{X}_{T+n}^c = c \end{aligned}$$

$\Longrightarrow$ **Feynman-Kac models** (G, X_n) **:** $\quad \gamma_n = \text{Law}(X_n^c \; ; \; T \geq n) \quad$ and $\quad \gamma_n(1) = \text{Proba}(T \geq n)$

$$\Downarrow$$

$$\eta_n = \text{Law}(X_n^c \mid T \geq n) = \text{Law}((X_0'^c, \ldots, X_n'^c) \mid T \geq n)$$

Biology: Macromolecules and Directed Polymers

- *Self avoiding walks* $X'_n \in \mathbb{Z}^d$

$$X_n = (X'_0, \ldots, X'_n) \quad \text{and} \quad G_n(X_n) = 1_{\not\in \{X'_0, \ldots, X'_{n-1}\}}(X'_n)$$

$$\gamma_n(1) = \text{Proba}(\forall 0 \leq p \neq q \leq n,\ \ X'_p \neq X'_q) \quad \text{and} \quad \eta_n = \text{Law}(X'_0, \ldots, X'_n \mid \forall 0 \leq p \neq q \leq n,\ \ X'_p \neq X'_q)$$

- *Edwards' model*

$$X_n = (X'_0, \ldots, X'_n) \quad \text{and} \quad G_n(X_n) = \exp\{-\beta \sum_{0 \leq p < n} 1_{X'_p}(X'_n)\}$$

Statistics: Sequential MCMC and Feynman-Kac-Metropolis models

Metropolis potential [π target measure]+[(K,L) pair Markov transitions]

$$G(y_1,y_2)=\frac{\pi(dy_2)L(y_2,dy_1)}{\pi(dy_1)K(y_1,dy_2)}$$

Ex. π **Gibbs measure:**

$$\pi(dy)\propto e^{-V(y)}\ \lambda(dy)\Rightarrow G(y_1,y_2)=e^{(V(y_1)-V(y_2))}\ \frac{\lambda(dy_2)L(y_2,dy_1)}{\lambda(dy_1)K(y_1,dy_2)}$$

Note: $(K=L\quad\lambda-\text{reversible})$ or $(\lambda K=\lambda$ and $L(y_2,dy_1)=\lambda(dy_1)\frac{dK(y_1,\,\centerdot\,)}{d\lambda}(y_2))$

$$\Downarrow$$

$$G(y_1,y_2)=\exp\left(V(y_1)-V(y_2)\right)$$

Notation $\mathbb{E}_{\nu}^{M}(\,\centerdot\,)$=Expectation w.r.t. Markov [transition M, initial condition ν]

<u>**Theorem:**</u> (**Time reversal formula**), [*A. Doucet, P.DM; (Séminaire Probab. 2003)*]

$$\mathbb{E}_{\pi}^{L}(f_n(Y_n, Y_{n-1} \ldots, Y_0)|Y_n = y) = \frac{\mathbb{E}_{y}^{K}(f_n(Y_0, Y_1, \ldots, Y_n)\ \{\prod_{0\leq p<n} G(Y_p, Y_{p+1})\})}{\mathbb{E}_{y}^{K}(\{\prod_{0\leq p<n} G(Y_p, Y_{p+1})\})}$$

In addition :

$\oplus$ *FK-Metropolis n-marginal:* $\lim_{n\to\infty} \eta_n = \pi$ (cv. decays $\perp \pi$)

$\oplus$ *Nonhomogeneous models:* (π_n, L_n, K_n)

$\pi_n(dy) \propto e^{-\beta_n V(y)}\ \lambda(dy)$, cooling schedule $\beta_n \uparrow \infty$, mutation s.t. $\pi_n = \pi_n K_n$, and Law$(X_0) = \pi_0$

$$\Downarrow$$

$$G_n(y_1, y_2) = \exp\left[-(\beta_{n+1} - \beta_n)V(y_1)\right] \Longrightarrow \eta_n = \pi_n$$

Rare events analysis

- **Importance sampling and Twisted Feynman-Kac measures**

$$\mathbb{P}(V_n(X_n) \geq a) \quad = \quad \mathbb{E}(1_{V_n(X_n)\geq a}\ e^{-\beta_n V_n(X_n)}\ e^{+\beta_n V_n(X_n)})$$

$$\Downarrow$$

Importance potentials/measures:

$$G_n(X_n, X_{n+1}) = e^{\beta_n(V_{n+1}(X_{n+1})-V_n(X_n))} \Longrightarrow \mathbb{P}(V_n(X_n) \geq a) = \gamma_n(1_{V_n\geq a}e^{-\beta_n V_n})$$

In addition:

$$\mathbb{E}(f_n(X_n) \mid V_n(X_n) \geq a) = \eta_n(f_n\ 1_{V_n\geq a}e^{-\beta_n V_n})/\eta_n(1_{V_n\geq a}e^{-\beta_n V_n})$$

$\oplus$ *Path-space models $\Rightarrow$ weighted genealogies*

$$X_n = (X'_0, \ldots, X'_n) \text{ and } V_n(X_n) = V'_n(X'^{n})$$

$$\Downarrow$$

$$\mathbb{E}(f_n(X'_0, \ldots, X'_n) \mid V'_n(X'_n) \geq a) = \eta_n(f_n\ 1_{V_n\geq a}e^{-\beta_n V_n})/\eta_n(1_{V_n\geq a}e^{-\beta_n V_n})$$

- **Multi-splitting Feynman-Kac models** ($\neq$ importance sampling)

  $(E = A \cup A^c)$, Y_n Markov, $Y_0 \in A_0 (\subset A)$ $\rightsquigarrow$ $A^c = (B \cup C)$, $C =$ absorbing set/hard obstacle

  Multi-level decomposition $B = B_m \subset \ldots \subset B_1 \subset B_0$ $\qquad (A_0 = B_1 - B_0,\ B_0 \cap C = \emptyset)$

$$\Downarrow$$

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

  Inter-level excursions : $T_n = \inf\{p \geq T_{n-1} \ : \ Y_p \in B_n \cup C\}$

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

$$\Downarrow$$

  FK interpretation

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

Advanced signal processing $\rightarrow$ filtering/hidden Markov chains/Bayesian methodology

Signal process $\quad X_n =$ Markov chain $\in E_n$

Observation/Sensor eq. $\quad Y_n = H_n(X_n, V_n) \in F_n$ with $\mathbb{P}(H_n(x_n, V_n) \in dy_n) = g_n(x_n, y_n)\ \lambda_n(dy_n)$

Example: $Y_n = h_n(X_n) + V_n \in F_n = \mathbb{R}$, with Gaussian noise $V_n = \mathcal{N}(0,1)$

$$\Downarrow$$

$$\mathbb{P}(h_n(x_n) + V_n \in dy_n) = (2\pi)^{-1/2} e^{-\frac{1}{2}(y_n - h_n(x_n))^2}\ dy_n = \underbrace{\exp\left[h_n(x_n) y_n - h_n^2(x_n)/2\right]}_{g_n(x_n, y_n)}\ \underbrace{\mathcal{N}(0,1)(dy_n)}_{\lambda_n(dy_n)}$$

Prediction/filtering/smoothing $\rightarrow$ **Feynman-Kac representation** $G_n(x_n) = g_n(x_n, y_n)$

$$\eta_n = \mathrm{Law}(X_n \mid Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) = \mathrm{Law}(X'_0, \ldots, X'_n \mid Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1})$$

Partially linear/Gaussian models

$$X_n^1 \;=\; \text{Markov} \quad \in E_n \quad + \quad \left\{ \begin{array}{rcl} X_n^2 & = & A_n(X_n^1)\; X_{n-1}^2 + a_n(X_n^1) + B_n(X_n^1)\; W_n \in \mathbb{R}^d \\ Y_n & = & C_n(X_n^1)\; X_n^2 + c_n(X_n^1) + D_n(X_n^1)\; V_n \in \mathbb{R}^{d'} \end{array} \right.$$

Given a realization $X^1 = x \rightarrow$ *Kalman-Bucy optimal one step predictor*

$$\widehat{X}^{2\;-}_{x,n+1} \;=\; \mathbb{E}(X^2_{n+1} \mid Y_0, \ldots, Y_n,\; X^1 = x) \quad \text{and} \quad P^-_{x,n+1} = \mathbb{E}([X^2_{n+1} - \widehat{X}^{2\;-}_{x,n+1}][X^2_{n+1} - \widehat{X}^{2\;-}_{x,n+1}]')$$

$$\Downarrow$$

Quenched Kalman-Bucy recursion: $(\widehat{X}^{2\;-}_{x,n+1}, P^-_{x,n+1}) = \mathcal{B}_{n+1}[(x_n, x_{n+1}), (\widehat{X}^{2\;-}_{x,n}, P^-_{x,n})]$

Feynman-Kac representation: $\eta_n \sim (\mathbf{X}_n, \mathbf{G}_n)$ s.t.

$$\begin{aligned}\mathbf{X}_n &= (X_n^1, (\widehat{X}^{2\ -}_{X^1,n+1}, P^-_{X^1,n+1})) \text{ Markov chain} \in \mathbf{E}_n = (E_n \times \mathbb{R}^d \times \mathbb{R}^{d\times d}) \\ \mathbf{G}_n(x,m,P) &= \frac{d\mathcal{N}(C_n(x)\ m + c_n(x), C_n(x)\ P\ C_n(x)' + D_n(x) R_n^v D_n(x)')}{d\mathcal{N}(0, D_n(x) R_n^v D_n(x)')}(y_n)\end{aligned}$$

$\Downarrow$ [*virtual sensor :* $Y_n = \{C_n(X_n^1)\ \widehat{X}^{2\ -}_{X^1,n} + c_n(X_n^1)\} + \widehat{V}_{X^1,n}$]

$$\begin{aligned}F_n(x,m,P) = f_n(x) &\implies \eta_n(F_n) = \mathbb{E}(f_n(X_n^1) \mid Y_0, \ldots, Y_{n-1}) \\ F_n(x,m,P) = \mathcal{N}(m,P)(f_n) &\implies \eta_n(F_n) = \mathbb{E}(f_n(X_n^2) \mid Y_0, \ldots, Y_{n-1})\end{aligned}$$

Note: $\rightsquigarrow$ Interacting Kalman-Bucy filters and for path-space models we have

$$X_n^1 = (X_0^{1\ \prime}, \ldots, X_n^{1\ \prime}) \rightsquigarrow \text{Law}((X_0^{1\ \prime}, \ldots, X_n^{1\ \prime}) \mid Y_0, \ldots, Y_{n-1})$$