E.D.F., Mars 2004

Systèmes de particules en interaction, et modèles d'arbres généalogiques, en estimation non linéaire

Pierre DEL MORAL
LSP-C.N.R.S, Univ. Paul Sabatier, Toulouse.

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

$\supset$ *Collaborations* : D. Crisan, D. Dawson, A. Doucet, A. Guionnet, J. Jacod, M. Ledoux, T. Lyons, L. Miclo, Ph. Protter, S. Tindel, F. Viens, T. Zajic,...

Plan

1. Introduction $+$ qq domaines d'applications.

2. Description semi-groupes FKS
   ($\in$ esp. de chemins, $\sim$ EDP Boltzmann, Zakai, Kushner-Stratonovitch).

3. Interprétations probabilistes de McKean.

4. QQ exemples $\subset$ Ingéniérie Math.
   (analyse évènements rares, filtrage/lissage, Filtre K-B en interaction, dualité filtrage/régulation, algo. de Metropolis/recuit en interaction)

5. Comportement en temps long, $\leq$ de contraction.

6. Systèmes de particules en interaction $+$ arbres généalogiques.

7. QQ résultats d'analyse asymptotique (temps/espace).

Formules de Feynman-Kac/Interprétations particulaires

Physique $\longleftrightarrow$ Biologie $\longleftrightarrow$ Sci. Ingénieur

- **Physique**:
  - $FKS \in$ Eq. intégro-différentielles non linéaires ($\sim$ Boltzmann généralisée).
  - Evolution de particules $\in$ milieux absorbants.
  - Analyse spectrale de semigroupes de FKS.
  - Résolution de problèmes de Dirichlet (conditions aux bords).
    $\Rightarrow$ FKS $\in$ espaces d'excursions.

- **Biologie**:
  - Processus de branchements/Algorithmes d'évolution de type génétique.
  - Processus de coalescence. Analyse d'arbres généalogiques. $\longleftarrow$
  - Analyse macromoléculaire/polymères dirigés (ex: chaînes auto-évitantes).
- **Traitement du signal/filtrage non linéaire** $\longleftarrow$
  - Prédiction/filtrage/lissage optimal.
  - Filtres de Kalman-Bucy en interaction.
  - Modèles d'approximations particulaires (branchement/interaction).

- **Évènements rares**: $\longleftarrow$
  - Algorithmes de branchements Multi-niveaux (méth. RESTART). Arbres généalogique$\simeq$Loi du processus $\in$ regime rare.
  - $\neq$ Changement de Probab. de référence (méth. "importance sampling").
  - Estimation de temps de sortie/hauteurs d'arbres.
- **Régulation/contrôle/optimisation/simulation**: $\longleftarrow$
  - Simulation de chaînes de Markov à conditions terminales fixées.
  - Simulation de lois cibles, algorithmes de Metropolis en interaction.
  - Algorithmes évolutionnaires de recherche aléatoire.

$\sim$ *Algorithmes stochastiques heuristiques*: multi-level splitting (Khan-Harris 51), prune enrichment (Rosenbluth 1955), switching algo. (Magill 65), genetic models (Holland 75), matrix reconfiguration (Hetherington 84), restart (Villen-Altamirano 91), go-with-the-winner (Vazirani-Aldous 94), SIR-particle filters (Rigal-Salut-DM 92, Gordon-Salmon-Smith 93, Kitagawa 96), Sequential Monte Carlo methods...

Notations $(E, \mathcal{E})$ espace mesurable

$$\begin{array}{rcl} \mathcal{P}(E) & = & \{\mu \text{ probabilité sur E}\} \\ \mathcal{B}_b(E) & = & \{f : E \to \mathbb{R},\ \mathcal{E}\text{-mesurable + bornée}\} \end{array}$$

- $(\mu, f) \in (\mathcal{P}(E) \times \mathcal{B}_b(E)) \quad \longrightarrow \quad \mu(f) = \int f(x)\mu(dx)$

- $M(x, dx')$ transition Markov sur E

$$\mu M(dx') = \int \mu(dx) M(x, dx') \quad \text{et} \quad M(f)(x) = \int M(x, dx') f(x')$$

- $M, L,\ (M_n)_{n \geq 0}$ Markov

$$(ML)(x, dz) = \int M(x, dy) L(y, dz) \quad (\text{s.g.: } M^n = MM^{n-1},\ M_{p,n} = M_{p+1} M_{p+1,n})$$

Formules de Feynman-Kac

Déf.: [Chaîne de Markov X_n sur E_n, potentiel $G_n : E_n \to [0,\infty)$]

Flot de mesures $\eta_n \in \mathcal{P}(E_n)$, définies pour tout $f_n \in \mathcal{B}_b(E_n)$

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1)$$

avec

$$\gamma_n(f_n) \;=\; \mathbb{E}(f_n(X_n) \prod_{0\le p<n} G_p(X_p))$$

Formules trajectorielles $\rightsquigarrow$ idem $\in$ espaces de chemins

$$X_n = (X'_0,\ldots,X'_n) \in E_n = (E'_0 \times \ldots \times E'_n)$$
$$\Downarrow$$
$$\gamma_n(f_n) = \mathbb{E}(f(X'_0,\ldots,X'_n)\ \prod_{0\le p<n} G_p(X'_0,\ldots,X'_p))$$

Approximations numériques :

Tech. Monte Carlo classique $[X_n^i \text{ iid} \sim X_n]$

$$\eta_n \simeq \sum_{i=1}^{N} \frac{\prod_{0\leq p<n} G_p(X_p^i)}{\sum_{j=1}^{N} \prod_{0\leq p<n} G_p(X_p^j)} \; \delta_{(X_0^i,\ldots,X_n^i)} \rightarrow \text{poids+exploration dégénérés}$$

$$\Downarrow$$

Systèmes de particules en interaction
$\in$ {Maillage adaptatif stochastique, tech. simulation}

$$\eta_n \simeq \frac{1}{N} \sum_{i=1}^{N} \; \delta_{(\xi_{0,n}^i,\xi_{1,n}^i,\ldots,\xi_{n,n}^i)} \longrightarrow \textbf{arbre généalogique}$$

Rmq.:

- S.g. FKS renormalisé *non linéaire* : (X_n transitions M_n)

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) = \Psi_n(\eta_n) M_{n+1} \quad \text{avec} \quad \Psi_n(\eta_n)(dx_n) \propto G_n(x_n)\ \eta_n(dx_n)$$

- Normalisées $\rightsquigarrow$ non-normalisées

$$\begin{aligned} \gamma_n(f_n) &= \mathbb{E}(f_n(X_n) \prod_{0 \leq p < n} G_p(X_p)) \\ &= \mathbb{E}(f_n(X'_0, \ldots, X'_n) \prod_{0 \leq p < n} G_p(X'_0, \ldots, X'_p)) = \eta_n(f_n) \prod_{0 \leq p < n} \eta_p(G_p) \end{aligned}$$

- Approx. arbre généalogique *(sans biais)*

$$\gamma_n(f_n) \simeq \gamma_n^N(f_n) = \eta_n^N(f_n) \prod_{0 \leq p < n} \eta_p^N(G_p)$$

Semi-groupe non linéaire de FKS

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) =_{\text{déf.}} \Psi_n(\eta_n) M_{n+1} \quad (\text{avec} \quad \Psi_n(\eta_n)(dx_n) \propto G_n(x_n)\eta_n(dx_n))$$

$$\downarrow$$

<u>*Interprétation Probabiliste de McKean :*</u> $\exists ?\overline{X}_n$ Markov : Loi$(\overline{X}_n) = \eta_n$

$\Rightarrow$ Mesure de McKean

$$\mathbb{P}((\overline{X}_0, \ldots, \overline{X}_n) \in d(x_0, \ldots, x_n)) = \eta_0(dx_0) K_{\eta_0,1}(x_0, dx_1) \ldots K_{\eta_{n-1},n}(x_{n-1}, dx_n)$$

$$\text{avec} \quad \eta_p = \eta_{p-1} K_{\eta_{p-1},p}$$

$$\Downarrow$$

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

S.g. de FKS $\rightarrow$ Interprétation de McKean
[$X_n \in E_n,\ \epsilon_n$ t.q. $\epsilon_n G_n \in [0,1]$]

$$\eta_{n+1} = \eta_n K_{\eta_n,n+1} \quad \text{avec} \quad K_{\eta_n,n+1} = S_{\eta_n,n} M_{n+1}$$

i.e.

$$\eta_n \xrightarrow{\text{correction}} \widehat{\eta}_n = \eta_n S_{\eta_n,n} = \Psi_n(\eta_n) \xrightarrow{\text{prediction}} \eta_{n+1} = \widehat{\eta}_n M_{n+1}$$

où

$$S_{\eta,n}(x,dy) = \epsilon_n G_n(x)\ \delta_x(dy) + (1 - \epsilon_n G_n(x))\ \Psi_n(\eta)(dy)$$

Rmq :

$X_n = ((X'_0, \ldots, X'_{n-1}), X'_n) = (X_{n-1}, X'_n) \Rightarrow$ "même modèle" $\in$ espace de chemins

Markov $\in$ Milieux absorbants $G(x) = e^{-V(x)} \in [0,1]$, $X_n^c \in E^c = E \cup \{c\}$

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

Absorption

$$G^c(x, dy) = G(x)\ \delta_x(dy) + (1 - G(x))\ \delta_c(dy)$$

$$\Downarrow$$

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

$$\Downarrow$$

$$\gamma_n = \text{Loi}(X_n^c\ ;\ T \geq n) \quad \text{et} \quad \eta_n = \text{Loi}(X_n^c \mid T \geq n)$$

Rmq.: interprétation McKean $\neq$ interprétation par absorption

Analyse dévènements rares ($\neq$ "importance sampling" $\rightsquigarrow$ interp. FK évidente)
($E = A \cup A^c$), Y_n Markov, $Y_0 \in A_0 (\subset A)$ $\rightsquigarrow$ $A^c = (B \cup C)$, $C =$ ens. absorbant.
Décomposition par niveaux

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

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

Excursions entre niveaux ou absorption : $T_n = \inf\{T_{n-1} \leq p \ : \ Y_p \in B_n \cup C\}$

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

$$\downarrow$$

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

Filtrage (non linéaire) de signaux (Markov $X_n \in E_n$)

Obs. $\quad Y_n = H_n(X_n, V_n) \quad$ avec $\quad \mathbb{P}(H_n(x_n, V_n) \in dy_n) = g_n(x_n, y_n)\lambda_n(dy_n)$

$\Downarrow$ FK avec $[X_n$ et $G_n(x_n) = g_n(x_n, y_n)]$

FK=Prédiction/filtrage/lissage:

$$\begin{aligned} \eta_n &= \text{Loi}(X_n \mid Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \\ &= \text{Loi}((X'_0, \ldots, X'_n) \mid Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \end{aligned}$$

Filtrage linéaire/Gaussien partiel (Markov $X_n^1 \in E_n$)

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

Trajectoire $X^1 = x$ fixée $\rightarrow$ *Prédicteur de Kalman-Bucy*

$$\begin{array}{rcl} \widehat{X}^{2\ -}_{x,n+1} & = & \mathbb{E}(X^2_{n+1} \mid Y_0, \ldots, Y_n,\ X^1 = x) \\ P^-_{x,n+1} & = & \mathbb{E}([X^2_{n+1} - \widehat{X}^{2\ -}_{x,n+1}][X^2_{n+1} - \widehat{X}^{2\ -}_{x,n+1}]') \end{array}$$

$$\downarrow$$

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

Représentation de Feynman-Kac : $\eta_n \;\sim\; (\mathbf{X}_n, \mathbf{G}_n)$ t.q.

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

$$\Downarrow \quad [\textit{capteur virtuel :} \quad Y_n = \{C_n(X_n^1)\ \widehat{X}^{2\;-}_{X^1,n} + c_n(X_n^1)\} + \widehat{V}_{X^1,n}\ ]$$

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

Rmq.:

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

Dualité Filtrage/Régulation

Ex. Filtrage scalaire/bruits Gaussien$\leftrightarrow$Régulation quadratique:

$$\left.\begin{array}{rcll} X_n & = & F_n(X_{n-1},W_n) & \text{(signal)} \\ Y_n & = & H_n(X_n)+V_n & \text{(observation)} \end{array}\right\} \longleftrightarrow \left\{\begin{array}{rcll} X_n^w & = & F_n(X_{n-1}^w,w_n) & \text{(entrée)} \\ Y_n^w & = & H_n(X_n^w) & \text{(sortie)} \end{array}\right.$$

$$\text{Loi}((W_0,\ldots,W_T)\mid Y_0=y_0,\ldots,Y_T=y_T) \leftrightarrow \text{Arg}\min_{w_0,\ldots,w_n}\sum_{p=0}^{T}\{|w_p|^2+|y_p-Y_p^w|^2\}$$

$$\updownarrow$$

$$\mathbb{P}(W_{[0,T]}\in dw_{[0,T]}\mid Y_0=y_0,\ldots,Y_T=y_T) \propto \exp -\frac{\beta}{2}\underbrace{\sum_{p=0}^{T}\{|w_p|^2+|y_p-H_p(x_p^w)|^2\}}_{\sim\text{masse de proba. conditionnelle/y}} \quad dw_{[0,T]}$$

Rmq.: ($\beta = 1$/température) (vraisemblance $\leftrightarrow$ optimalité)

Simulation $\sim$ lois conditionnelles $\leftrightarrow$ arbre de décisions/contrôles

S.g. de Feynman-Kac-Metropolis
(Doucet,DM Séminaire Probab. 37 (2003))

Potentiel de Metropolis [π mesure "cible"]+[K, L Markov]

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

Ex. π **mesure de Gibbs:**

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

-($K = L\quad \lambda -$ réversible)

-($\lambda L = \lambda,$ et $K(y_1, dy_2) = \lambda(dy_2)\frac{dL(y_2, \bullet)}{d\lambda}(y_1)$) $\qquad \Rightarrow G(y_1, y_2) = e^{(V(y_1)-V(y_2))}$

Ex. $V(w_0, \ldots, w_T) \sum_{p=0}^{T} \{|w_p|^2 + |y_p - Y_p^w|^2\}$

Rmq.: $\rightsquigarrow$ *Algo. recuit simulés en interaction* $\in$ espace de chemins,

$\sim$ gel asympt.=schéma température sous-lin.$\neq$ logarithmiques.

Notation $\mathbb{E}^M_\nu(\,.\,)$=Espérance Markov [transition M, cond. initiale ν]

$$\Downarrow$$

Th. [**Form. inversion du temps**]

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

$\oplus$ FK-Metropolis n-marginale : $\lim_{n\to\infty}\eta_n = \pi$ (vitesse $\perp\pi$)

$$\Updownarrow \quad X_n = (Y_n, Y_{n+1}) \in E^2 = \text{espace des transitions}$$

Algo. de Metropolis/recuit simulé non linéaire :

$$\eta_{n+1} = \eta_n \quad \underbrace{S_{\eta_n,n}}_{\text{(accept./rejet)}} \quad \underbrace{M_{n+1}}_{\text{(proposition)}} \longrightarrow \pi\times K \quad \text{vitesse } \perp\pi!!$$

Objectifs/Problèmes

Analyser, estimer, simuler les flots de mesures de FK η_n

- *Étude de la stabilité des semigroupes non linéaires.*
  - Stabilité asympto. du filtre optimal (oubli cond. initiales).
  - Convergence d'algo. de Metropolis non linéaire.
  - Robustesse des flots de mesures $\to$ linéarisation stochastique locale.

- *Mise en forme de schémas particulaires*
  - $\in$ espace de chemins $\rightsquigarrow$ arbres généalogiques.
  - Choix du couple (potentiel/transition Markov)

$$\downarrow$$

Lissage signaux, décisions optimales, lois de processus dans un régime rare, stratégies de sorties de poches d'obstacles, polymérisations,...

- Filtres particulaires/estimation/contrôle/simulation de type génétique (accept./rejet).
- Filtres de Kalman-Bucy en interaction.
- Algo. de Metropolis en interaction.
- .../...

• *Analyse du comportement asymptotique spatio-temporel*

- Estimations $\mathbb{L}_p$, déviations exponentielles, PGD, TCL, $\leq$ Berry-Esseen,...
- Propagations du chaos, degré d'indépendence/adéquation.
- Concentration des mesures limites (physique/optimisation).
- Estimations uniformes par rapport au temps (qualité schémas).

Prop. Stabilité asympt. Semi-groupe FKS non linéaire $\Phi_{p,n}(\eta_p) = \eta_n$

$$\downarrow$$

Th. [A. Guionnet, P. DM, Ann. IHP (2001)]

$$(H) \qquad M_{p,p+m}(x_p, \centerdot) \geq \rho \ M_{p,p+m}(y_p, \centerdot) \quad \text{et} \quad r = \sup_{p,x_p,y_p} (G_p(x_p)/G_p(y_p))$$

$$\Downarrow$$

$$\|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\eta_p)\|_{\mathrm{tv}} \leq 2r^m \rho^{-1} \ (1 - \rho^2/r^{m-1})^{\lfloor (n-p)/m \rfloor} \ \|\mu_p - \eta_p\|_{\mathrm{tv}}$$

et

$$\|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\eta_p)\|_{\mathrm{tv}} \leq (1 - \rho^2/r^{m-1})^{\lfloor (n-p)/m \rfloor}$$

Extensions :

a) Sg FK-Metropolis (π, K, L), [$\beta(M)$=coef. contraction Dobrushin de M Markov]

$$\|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\eta_p)\|_{\mathsf{tv}} \leq c(\pi)\ \beta(L^{n-p})$$

b) $\leq$ fonctionnelles/h-entropies relatives
(ex.: Boltzmann, Havrda-Charvat, Kakutani-Hellinger).

$$\mathrm{Ent}(\Phi_{p,n}(\mu_p) \mid \Phi_{p,n}(\eta_p)) \leq r^m \rho^{-1}\ (1 - \rho^2/r^{m-1})^{\lfloor (n-p)/m \rfloor}\ \mathrm{Ent}(\mu_p \mid \eta_p)$$

c) FKS temps continu + fonctionnelles $\times$, $Z_{r,t} = Z_{r,s} Z_{s,t}$.

d) Cas réversible : État fondamental op. Schrödinger=point fixe s.g FKS

e) $G(x) = e^{-\beta V(x)} \rightarrow$ concentration points fixes.

Interprétation particulaire : $\xi_n \in E_n^N$ **Markov t.q.** $(N \to \infty)$

$$\eta_n^N = m(\xi_n) = \frac{1}{N}\sum_{j=1}^{N} \delta_{\xi_n^j} \longrightarrow \eta_n$$

Syst. de N particules en interaction $\sim$ [interp. McKean $(K_{n,\eta})$]

$$\mathbb{P}(\xi_n \in d(x_n^1,\ldots,x_n^N) \mid \xi_{n-1}) = \prod_{i=1}^{N} K_{m(\xi_{n-1}),n}(\xi_{n-1}^i, dx_n^i)$$

$$\Updownarrow$$

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

- $\xi_n^i \to \widehat{\xi}_n^i \quad \sim \quad S_{m(\xi_n),n}(\xi_n^i, \cdot\,) = \epsilon_n G_n(\xi_n^i)\ \delta_{\xi_n^i} + (1 - \epsilon_n G_n(\xi_n^i))\ \Psi_n(m(\xi_n))$
- $\widehat{\xi}_n^i \to \xi_{n+1}^i \ \sim \ M_{n+1}(\widehat{\xi}_n^i, \cdot\,)$
  Rmq.: $\epsilon_n = 0 \rightsquigarrow$ algo. génétique simple

Arbres généalogiques

FKS $\in$ Espace de chemins : $\quad X_n = (X'_0, \ldots, X'_n) \quad$ et $\quad G_n(X_n) = G'_n(X'_n)$

$$\downarrow$$

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

- Sélection de trajectoires.

- Mutation=extensions élémentaires de trajectoires

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

Mesures d'occupations
particule $\in$ {individu, ligne ancestrale complète ou $\vee$ partielle}

Mesures FKS normalisés/non-normalisés :

$$\begin{aligned}
\mathbb{K}_n^N &= \frac{1}{N}\sum_{i=1}^N \delta_{(\xi_0^i,\ldots,\xi_n^i)} \longrightarrow \mathbb{K}_n = \eta_0 \times K_{\eta_0,1} \times \ldots \times K_{\eta_{n-1},n} \\
\eta_n^N &= \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i} \longrightarrow \eta_n \\
\gamma_n^N &= \{\prod_{0\le p<n} \eta_p^N(G_p)\}\ \eta_n^N \longrightarrow \gamma_n = \{\prod_{0\le p<n} \eta_p(G_p)\}\ \eta_n \qquad \text{(sans biais)}
\end{aligned}$$

Arbres généalogiques complet/partiel:

$$\begin{aligned}
\mathbb{K}_n^N &= \frac{1}{N}\sum_{i=1}^N \delta_{[\xi_0^i,(\xi_{0,1}^i,\xi_{1,1}^i),\ldots,(\xi_{0,n}^i,\ldots,\xi_{n,n}^i)]} \longrightarrow \mathbb{K}_n \\
\eta_n^N &= \frac{1}{N}\sum_{i=1}^N \delta_{(\xi_{0,n}^i,\ldots,\xi_{n,n}^i)} \longrightarrow \eta_n \in \mathcal{P}(E_0' \times \ldots \times E_n')
\end{aligned}$$

Analyse asymptotique $(n, N) \to \infty$ (temps long $\sim$ (H))

- *Propagation chaos* (analyse faible, ex. $\mathbb{E}(\eta_n^N(f)) = \mathbb{E}(f(\xi_n^1)) \simeq \eta_n + c(n)/N$).

$$(H) \Longrightarrow \mathsf{Ent}(\mathsf{Loi}(\xi_0^i, \ldots, \xi_n^i)_{1 \le i \le q} \mid \mathbb{K}_n^{\otimes q}) \le c\ q\ n\ /N$$

- *Estimations* $\mathbb{L}_p$, $\le$ Burkhölder-D.G., Th. Glivenko-Cantelli

Ex.: $\quad (H) \Longrightarrow \sqrt{N} \sup_{n \ge 0} \mathbb{E}(\|\eta_n^N - \eta_n\|_{\mathcal{F}}^p)^{1/p} \le b(p)\ I(\mathcal{F})\ c(\rho, m) \quad (b(2p) = 2^{-p}\ (2p)!/p!)$

$\rightsquigarrow$ Th. Limite Centrale (Donsker) $\rightsquigarrow$ $\le$ Berry-Esseen.

- *Déviations exponentielles*, $\le$ Talagrand.

Ex. ($\mathrm{osc}(f) \le 1$): $\quad (H) \Longrightarrow \sup_{n \ge 0} \mathbb{P}(|\eta_n^N(f) - \eta_n(f)| \ge \varepsilon) \le (1 + \varepsilon \sqrt{N/2}) e^{-N\varepsilon^2/2c(\rho, m)}$

$\rightsquigarrow$ PGD $\in$ espaces de distributions (topo. faible ou forte).

Propagations croissantes et uniformes du chaos

Déf. : $\mathbb{P}_n^{(N,q)} = \text{Loi}[(\xi_0^i, \ldots, \xi_n^i)_{1\leq i\leq q}] \longrightarrow$ n-marginale $\mathbb{P}_{[n]}^{(N,q)} = \text{Loi}[(\xi_n^i)_{1\leq i\leq q}]$

Temps continu *sauts en interaction* $\sim$ *éq. collision Boltzmann*

-[(couplage arbres) Graham-Méléard Ann. Proba. (1997)]

-[(FKS-s.g.) Miclo-DM préprint (2000)]

$$\|\mathbb{P}_t^{(N,q)} - \mathbb{K}_t\|_{\text{tv}} \leq c(t)\ q^2/N$$

Problème $\rightsquigarrow c(t) = e^{c\ t}$

1 Solution $\rightsquigarrow$ tech. s.g. $\supset$ *Stabilité des s.g. non linéaires limites.*

Temps discret ($\epsilon_n = 0 \Leftrightarrow$ algo. génétique)

Technique FKS-s.g. q-tensoriel ($\eta_n^{\otimes q}$, $\mathbb{K}_n^{\otimes q}$)

Ex. : $\quad \eta_n^{\otimes q} = \Phi_{p,n}^{(q)}(\eta_p^{\otimes q}) \Rightarrow [(\eta_n^N)^{\otimes q} - \eta_n^{\otimes q}] = \sum_{p=0}^{n} [\Phi_{p,n}^{(q)}((\eta_p^N)^{\otimes q}) - \Phi_{p,n}^{(q)}(\Phi_p^{(q)}(\eta_{p-1}^N)^{\otimes q}))]$

$+$

Prop. chaos mes. de Boltzmann-Gibbs [$m(X) = N^{-1}\sum_{i=1}^N \delta_{X^i}$, (X^i) i.i.d. $\sim \mu$] :

$$\mathbb{E}[\Psi_n(m(X))^{\otimes q}] = \mathbb{E}[\Psi_n^{(q)}(m(X)^{\otimes q})] \simeq \Psi_n^{(q)}(\mu^{\otimes q}) + cte.q^2/N$$

$\longleftrightarrow$ inégalités de Burkhölder précises

Ex. (X^i) i.i.d. $\sim \mu$: $\qquad N^p\ \mathbb{E}([m(X)(f) - \mu(f)]^{2p}) \ \leq\ 2^{-p}\ \mathrm{osc}(f)^{2p}\ (2p)!/p!$

Th. *[version asympt. + (H)]*

$$q^2(N) = o(N) \implies \limsup_{N\to\infty} \frac{N}{q^2(N)} \|\mathbb{P}_n^{(N,q(N))} - \mathbb{K}_n^{\otimes q(N)}\|_{\mathsf{tv}} \leq c\ n^3\ r^{2m}/\rho^2$$

$$n(N)\ q^2(N) = o(N) \implies \limsup_{N\to\infty} \frac{N}{n(N)\ q^2(N)} \|\mathbb{P}_{[n(N)]}^{(N,q(N))} - \eta_{n(N)}^{\otimes q(N)}\|_{\mathsf{tv}} \leq c\ r^{2m}/\rho^2$$

Est. unif. :

$$\|\mathbb{P}_{[n]}^{(N,q)} - \eta_n^{\otimes q}\|_{\mathsf{tv}} \leq c\ [n \wedge (m(r^m/\rho)^q)]\ e_{m,r}(q^2/N)\ q^2/N \qquad (e_{m,r}(0) < \infty)$$

$\Downarrow$ [filtrage $\leftrightarrow$ régulation]

$$\mathbb{P}(\wedge_{i\leq q(N)} J_n(W_{0,n}^i,\ldots,W_{n,n}^i) \geq \inf J_n + \delta \mid Y_0^n) \leq c(n)\ \frac{q^2(N)}{N} + [1 - \mathbb{P}(J_n < \inf J_n + \delta \mid Y_0^n)]^{q(N)}$$