On the concentration of interacting particle processes

Part III : Theoretical analysis

P. Del Moral

INRIA & Bordeaux Mathematical Institute

Sino-French Summer Institute in Stochastic Modeling and Applications, Beijing, June 2011
Stability properties
  Basic notation
  Nonlinear semigroups
  Contraction properties

Independent empirical processes
  Orlicz’ norm recipes
  Empirical processes
  Kinchine’s inequalities
  Laplace techniques

Interacting processes
  Mixtures of interacting processes
  Perturbation analysis (marginal models)
  Perturbation analysis (empirical processes)

Feynman-Kac particle models
  First order expansions
  Uniform concentration w.r.t. time
  Particle free energy
Stability properties
  Basic notation
  Nonlinear semigroups
  Contraction properties

Independent empirical processes

Interacting processes

Feynman-Kac particle models
Basic notation

▶ Dobrushin's contraction coefficient \( M \) Markov \( E_1 \rightsquigarrow E_2 \)

\[
\beta(M) := \sup \{ \text{osc}(M(f)) ; \ f \in \text{Osc}(E_2) \} \\
= \sup \{ \|M(x, \cdot) - M(y, \cdot)\|_{tv} ; \ (x, y) \in E_2 \}
\]

▶ Boltzmann-Gibbs transformation (\( G \in (0, 1] \))

\[
\Psi_G(\mu) = \mu S_{\mu,G}
\]

with

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

Properties

\[
\Psi_G(\mu) - \Psi_G(\nu) = \frac{1}{\nu(G)} (\mu - \nu) S_\mu \quad \text{and} \quad \beta(S_{\mu,G}) \leq 1 - \|G\|
\]

Proof:

\[
\Psi_G(\mu) - \Psi_G(\nu) = (\mu - \nu) S_\mu + \nu (S_\mu - S_\nu)
\]

and

\[
\nu(S_\mu - S_\nu) = (1 - \nu(G)) \left[ \Psi_G(\mu) - \Psi_G(\nu) \right]
\]
Nonlinear semigroups

Normalized & Unnormalized semigroups

\[ \Phi_{p,n}(\eta_p) = \eta_n \quad \text{and} \quad \gamma_p Q_{p,n} = \gamma_n \]

Linear integral operators

\[ Q_{p,n}(f_n)(x_p) := \mathbb{E} \left( f_n(X_n) \prod_{p \leq q < n} G_q(X_q) \mid X_p = x_p \right) \]

Simplified notation \( Q_{n-1,n}(x, dy) = Q_n(x, dy) (= G_{n-1}(x) M_n(x, dy)) \)

\[ Q_{p,n} = Q_{p+1} Q_{p+2} \cdots Q_n \]

Nonlinear updating-prediction transformations

\[ \Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \ldots \circ \Phi_{p+1} \]
Lipschitz’s regularity

\[
Q_{p,n}(1)(x) = G_{p,n}(x) \quad \text{and} \quad P_{p,n}(f) = \frac{Q_{p,n}(f)}{Q_{p,n}(1)}
\]

\[\Downarrow\]

\[
\eta_n(f) = \Phi_{p,n}(\eta_p)(f) = \Psi_{G_{p,n}(\eta_p)}P_{p,n}
\]

\[\Downarrow\]

Lipschitz’s regularity

\[
\|\Phi_{p,n}(\eta_p) - \Phi_{p,n}(\eta'_p)\|_{tv} \leq g_{p,n} \beta(P_{p,n}) \|\eta_p - \eta'_p\|_{tv}
\]

with

\[
g_{p,n} := \sup_{x,y} \frac{G_{p,n}(x)}{G_{p,n}(y)}
\]
Contraction properties

Key observation

\[ P_{p,n}(f) = \frac{M_{p+1}(Q_{p+1,n}(f))}{M_{p+1}(Q_{p+1,n}(1))} = \frac{M_{p+1}(G_{p+1,n} P_{p+1,n}(f))}{M_{p+1}(G_{p+1,n})} = R_{p+1}^{(n)} P_{p+1,n}(f) \]

with the triangular array of Markov transitions

\[ R_{p+1}^{(n)}(f) := \frac{M_{p+1}(G_{p+1,n} f)}{M_{p+1}(G_{p+1,n})} \Rightarrow P_{p,n} = R_{p+1}^{(n)} R_{p+2}^{(n)} \ldots R_{n}^{(n)} \]

Strong mixing condition \( M_n(x, dy) \leq \chi M_n(x', dy) \)

\[ R_{p+1}^{(n)}(x, dy) := \frac{M_{p+1}(x, dy) G_{p+1,n}(y)}{M_{p+1}(G_{p+1,n})} \leq \chi^2 R_{p+1}^{(n)}(x', dy) \]

\[ \Rightarrow \beta(R_{p+1}^{(n)}) \leq 1 - \chi^{-2} \Rightarrow \beta(P_{p,n}) \leq (1 - \chi^{-2})^{n-p} \]
Contraction properties

Second key observation: **Mixing condition** \( \oplus G_n(x) \leq gG_n(y) \)

\[
\frac{G_{p,n}(x)}{G_{p,n}(y)} = \frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)} = \frac{G_p(x)}{G_p(y)} \frac{M_{p+1}(G_{p+1,n})(x)}{M_{p+1}(G_{p+1,n})(y)} \leq g \chi
\]

\[\Downarrow\]

**Theorem: Strong contraction property**

\[
\| \Phi_{p,n}(\eta_p) - \Phi_{p,n}(\eta'_p) \|_{tv} \leq g \chi (1 - \chi^{-2})^{n-p} \| \eta_p - \eta'_p \|_{tv}
\]

*Extensions:*
*Weak formulation,* \( M_{p,p+m}(x, dy) \leq \chi_m M_{p,p+m}(x', dy), g\beta(M) < 1, \) etc.
Stability properties

Independent empirical processes
  Orlicz’ norm recipes
  Empirical processes
  KINchine’s inequalities
  Laplace techniques

Interacting processes

Feynman-Kac particle models
Orlicz’ norm and Gaussian moments

\[ \pi_\psi[Y] \text{ Orlicz norm of } Y, \psi(u) = e^{u^2} - 1 \]

\[ \pi_\psi(Y) = \inf \{ a \in (0, \infty) : \mathbb{E}(\psi(|Y|/a)) \leq 1 \} \]

\text{U Gaussian and centered random variable } U, \text{ s.t. } \mathbb{E}(U^2) = 1:\]

\[ \pi_\psi(U) = \sqrt{8/3} \]

\text{and}

\[ \mathbb{E}(U^{2m}) = b(2m)^{2m} := (2m)_m 2^{-m} \]

\[ \mathbb{E}\left(|U|^{2m+1}\right) \leq b(2m + 1)^{2m+1} := \frac{(2m + 1)(m+1)}{\sqrt{m + 1/2}} 2^{-(m+1/2)} \]
Orlicz’ norm properties

5 key properties ((\( Y_i, Y \) positive)):

1. \( Y_1 \leq Y_2 \implies \pi_\psi(Y_1) \leq \pi_\psi(Y_2) \)
2. \( (\forall m \geq 0 \quad \mathbb{E}(Y_1^{2m}) \leq \mathbb{E}(Y_2^{2m})) \implies \pi_\psi(Y_1) \leq \pi_\psi(Y_2) \)
3. \( (\pi_\psi(f(x, Y)) \leq c \quad \text{for } \mathbb{P}-\text{a.e. } x) \implies \pi_\psi(f(X, Y)) \leq c \)
4. \( \mathbb{E}(Y^{2m}) \leq m! \pi_\psi(Y)^{2m} \)
5. \( \mathbb{E}(e^{tY}) \leq \min \left(2 e^{\frac{1}{4} (t\pi_\psi(Y))^2}, (1 + t\pi_\psi(Y)) e^{(t\pi_\psi(Y))^2}\right) \)

\[ \implies \mathbb{P}\left(Y \leq \pi_\psi(Y) \sqrt{x + \log 2}\right) \geq 1 - e^{-x} \]
Empirical processes

\( X^i \) independent \( \sim \mu^i \rightarrow m(X) := \frac{1}{N} \sum_{i=1}^{N} \delta_{X^i} \) and \( \mu := \frac{1}{N} \sum_{i=1}^{N} \mu^i \)

Fluctuation centered random fields

\[ V(X) = \sqrt{N} (m(X) - \mu) \]

\[ \sigma(f)^2 = \mathbb{E} (V(X)(f)^2) = \frac{1}{N} \sum_{i=1}^{N} \mu^i ([f - \mu^i(f)]^2) \]

\( \mathcal{F} \) separable class of functions \( \|f\| \leq 1 \)

\[ \|\mu - \nu\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |\mu(f) - \nu(f)|, \]

\[ \mathcal{N}(\epsilon, \mathcal{F}) = \sup \{ \mathcal{N}(\epsilon, \mathcal{F}, \mathbb{L}_2(\eta)); \eta \in \mathcal{P}(E) \} \]

\[ I(\mathcal{F}) = \int_{0}^{2} \sqrt{\log (1 + \mathcal{N}(\epsilon, \mathcal{F}))} \, d\epsilon \]
Some useful properties \((G(x) \in [0, 1], \, M \text{ Markov})\)

\[
\leadsto \text{Two classes of functions}
\]

\[
G \cdot M(\mathcal{F}) = \{ G \cdot M(f) : f \in \mathcal{F} \}
\]
\[
G \cdot (M - \mu M)(\mathcal{F}) = \{ G \cdot [M(f) - \mu M(f)] : f \in \mathcal{F} \}
\]

\[\Downarrow [\text{Exercice}]\]

\[
\mathcal{N} [G \cdot M(\mathcal{F}), \epsilon] \leq \mathcal{N}(\mathcal{F}, \epsilon)
\]
\[
\mathcal{N} [G \cdot (M - \mu M)(\mathcal{F}), 2\epsilon \beta(M)] \leq \mathcal{N}(\mathcal{F}, \epsilon)
\]
Kinchine’s inequalities \((\text{osc}(f) \leq 1)\)

> Marginal models

\[
\mathbb{E}(|V(X)(f)|^m)^{1/m} \leq b(m) \text{osc}(f)
\]

\[
\downarrow
\]

\[
\pi_\psi(V(X)(f)) \leq \sqrt{3/8}
\]

> Empirical processes

\[
\pi_\psi(\|V(X)\|_{\mathcal{F}}) \leq c \ I(\mathcal{F})
\]
Laplace techniques

Legendre-Fenchel transform

\[ \forall \lambda \geq 0 \quad L^*(\lambda) := \sup_{t \in \text{Dom}(L)} (\lambda t - L(t)) \]

\[ L_A(t) := \log \mathbb{E}(e^{tA}) \rightsquigarrow \text{Cramér-Chernov-Chebychev inequalities} \]

\[ \log \mathbb{P}(A \geq \lambda) \leq -L_A^*(\lambda) \quad \text{and} \quad \mathbb{P}\left( A \geq (L_A^*)^{-1}(x) \right) \leq e^{-x} \]

- **Comparison property**

\[ L_1 \leq L_2 \Rightarrow L_2^* \leq L_1^* \Rightarrow (L_1^*)^{-1} \leq (L_2^*)^{-1} \]

- **J. Bretagnolle & E. Rio’s Lemma**

\[ (L_{A+B}^*)^{-1}(x) \leq (L_A^*)^{-1}(x) + (L_B^*)^{-1}(x) \]
3 examples-exercices

- \( L(t) = t^2/(1 - t), \ t \in [0, 1] \)
  \[
  L^*(\lambda) = \left( \sqrt{\lambda} + 1 - 1 \right)^2 \quad \& \quad (L^*)^{-1}(x) = \left( 1 + \sqrt{x} \right)^2 - 1 = x + 2\sqrt{x}
  \]

- \( L_0(t) := -t - \frac{1}{2} \log (1 - 2t), \ t \in [0, 1/2] \)
  \[
  L_0^*(\lambda) = \frac{1}{2} (\lambda - \log (1 + \lambda)) \quad \& \quad (L_0^*)^{-1}(x) \leq 2(x + \sqrt{x})
  \]

- \( L_1(t) := e^t - 1 - t \)
  \[
  L_1^*(\lambda) = (1 + \lambda) \log (1 + \lambda) - \lambda \quad \& \quad (L_1^*)^{-1}(x) \leq \frac{x}{3} + \sqrt{2x}
  \]
Applications (part 1)

- Centered $A \leq 1$ & $\sigma_A = \mathbb{E}(A^2)^{1/2} \Rightarrow L_A(t) \leq \sigma_A^2 L_1(t)$
  - The probability of the following events is greater than $1 - e^{-x}$
    \[
    A \leq \sigma_A^2 (L_1^*)^{-1} \left( \frac{x}{\sigma_A^2} \right) \leq \frac{x}{3} + \sigma_A \sqrt{2x}
    \]

- $B$ s.t. $\mathbb{E}(|B|^m)^{1/m} \leq b(2m)^2 c \Rightarrow L_B(t) \leq ct + L_0(ct)$
  - The probability of the following events
    \[
    B \leq c \left[ 1 + (L_0^*)^{-1} (x) \right] \leq c \left[ 1 + 2(x + \sqrt{x}) \right]
    \]
    is greater than $1 - e^{-x}$.

- Concentration of $A + B$ using J. Bretagnollle & E. Rio’s Lemma
Applications (part 2)

- $0 < \text{osc}(f) \leq a \Rightarrow L_\sqrt{NV(X)(f)}(t) \leq N \sigma^2(f/a) L_1(at)$

  $\Rightarrow$ the probability of the following events is greater than $1 - e^{-x}$,

\[
V(X)(f) \leq a^{-1}\sigma^2(f)\sqrt{N} (L_1^*)^{-1} \left( \frac{xa^2}{N\sigma^2(f)} \right) \leq \frac{xa}{3\sqrt{N}} + \sqrt{2x\sigma(f)^2}
\]

- Concentration of $F(m(X)(f))$ [marginal or empirical processes] using J. Bretagnolle & E. Rio’s Lemma
Stability properties

Independent empirical processes

Interacting processes
  Mixtures of interacting processes
  Perturbation analysis (marginal models)
  Perturbation analysis (empirical processes)

Feynman-Kac particle models
Interacting processes

Markov $X_n = (X^i_n)_{1 \leq i \leq N} \in E^N_n$, conditionally independent $| \mathcal{G}_{n-1}$.

$\mu^i_n = \text{Law}(X^i_n \mid X_0, \ldots, X_{n-1}) \rightsquigarrow V(X_n) := \sqrt{N} (m(X_n) - \mu_n)$

Definition.: $f_n \in \mathcal{G}_{n-1}$, $\text{osc}(f_n) \leq 1 \rightsquigarrow \mathbb{E} (V(X_n)(f_n)^2 \mid \mathcal{G}_{n-1}) \leq \sigma^2_n$

$\bar{\sigma}^2_n := \sum_{0 \leq p \leq n} \sigma^2_p$ and $a^*_n := \max_{0 \leq p \leq n} a_p$

$\triangleright V_n(X)(f) = \sum_{p=0}^n a_p \ V(X_p)(f_p)$

$L \sqrt{N} V_n(X)(f)(t) \leq N \bar{\sigma}^2_n \ L_1(t a^*_n)$

$\Rightarrow$ the probability of the following events is greater than $1 - e^{-x}$

$V_n(X)(f) \leq \sqrt{N} \ a^*_n \ \bar{\sigma}^2_n \ (L_1^*)^{-1} \left( \frac{x}{N \bar{\sigma}^2_n} \right) \leq a^*_n \left( \frac{x}{3 \sqrt{N}} + \sqrt{2 \bar{\sigma}^2_n \ x} \right)$
Perturbation analysis (marginal models)

\[ W_n(X)(f) = V_n(X)(f) + \frac{1}{\sqrt{N}} R_n(X)(f) \]

with

\[ V_n(X)(f) = \sum_{p=0}^{n} a_p V(X_p)(f_p) \]

& \[ \mathbb{E} \left( |R_n(X)(f)|^m \right)^{1/m} \leq b(2m)^2 r_n \]

Using J. Bretagnolle & E. Rio’s Lemma

⇒ the probability of the following events is greater than \( 1 - e^{-x} \),

\[ \sqrt{N} \ W_n(X)(f) \leq r_n \left( 1 + \left( L_0^* \right)^{-1}(x) \right) + N a_n^* \sigma_n^2 \left( L_1^* \right)^{-1} \left( \frac{x}{N \sigma_n^2} \right) \]
Perturbation analysis (empirical processes)

\[ W_n(X)(f) = V_n(X)(f) + \frac{1}{\sqrt{N}} R_n(X)(f) \]

with

\[ V_n(X)(f) = \sum_{p=0}^{n} a_p V(X_p)(f_p) \quad \& \quad \mathbb{E} \left( \| R_n(X) \|_{\mathcal{F}}^m \right) \leq m! \ r_n^m \]

Using J. Bretagnolle & E. Rio’s Lemma

⇒ the probability of the following events is greater than \( 1 - e^{-x} \),

\[ \| W_n(X) \|_{\mathcal{F}} \leq c \left[ \sum_{p=0}^{n} a_p \right] I(\mathcal{F}) \left( 1 + 2\sqrt{x} \right) + \frac{r_n}{\sqrt{N}} \left( 1 + \left( L_0^* \right)^{-1} \left( \frac{x}{2} \right) \right) \]
Stability properties

Independent empirical processes

Interacting processes

Feynman-Kac particle models
  First order expansions
  Uniform concentration w.r.t. time
  Particle free energy
First order expansions

Key telescoping decomposition

\[ \eta^n_N - \eta_n = \sum_{p=0}^n \left[ \Phi_{p,n}(\eta^N_p) - \Phi_{p,n}(\Phi_p(\eta^N_{p-1})) \right] \]

⊕ First order expansion

\[ \sqrt{N}\Phi_{p,n}(\eta^N_p) - \Phi_{p,n}(\Phi_p(\eta^N_{p-1})) \]

= \sqrt{N}\Phi_{p,n}\left(\Phi_p(\eta^N_{p-1}) + \frac{1}{\sqrt{N}} V^N_p\right) - \Phi_{p,n}(\Phi_p(\eta^N_{p-1}))

≃ V^N_p D_{p,n} + \frac{1}{\sqrt{N}} R^N_{p,n} \]

with \( D_{p,n} \in G_{p-1}\)-first order integral operator ⊕ 2nd-order remainder \( R^N_{p,n} \)
First order expansions

Stochastic perturbation model

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

Under the mixing condition of FK semigroups

\[ \text{osc} (D_{p,n}(f)) \leq c g_{p,n} \beta(P_{p,n}) \leq c(1 - \epsilon)^{n-p} \]

and

\[ \mathbb{E} \left( |R_n^N(f)|^m \right) \leq b(2m)^{2m} c \]

\[ \Downarrow \]

Uniform concentration estimates w.r.t. the time parameter
Particle free energy

Multiplicative formulae

\[ Z_n^N = \prod_{0 \leq p < n} \eta_p^N(G_p) = \gamma_n^N(1) \rightarrow_{N \uparrow \infty} \gamma_n(1) = \prod_{0 \leq p < n} \eta_p(G_p) \]

Taylor first order expansion

\[ \forall x, y > 0 \quad \log y - \log x = \int_0^1 \frac{(y - x)}{x + t(y - x)} \, dt \]

\[ \downarrow \]

\[ \log \left( \frac{\gamma_n^N(1)}{\gamma_n(1)} \right) \]

\[ = \sum_{0 \leq p < n} \left( \log \eta_p^N(G_p) - \log \eta_p(G_p) \right) \]

\[ = \sum_{0 \leq p < n} \left( \log \left( \eta_p(G_p) + \frac{1}{\sqrt{N}} W_{p}^{\eta,N}(G_p) \right) - \log \eta_p(G_p) \right) \]

\[ = \frac{1}{\sqrt{N}} \sum_{0 \leq p < n} \int_0^1 \frac{W_{p}^{\eta,N}(G_p)}{\eta_p(G_p) + \frac{t}{\sqrt{N}} W_{p}^{\eta,N}(G_p)} \, dt \]

\[ \leadsto \text{first order expansion [exo]} \]