Concentration inequalities for mean field particle models

P. Del Moral & **E. Rio**

INRIA Centre Bordeaux-Sud Ouest & Versailles University

Joint Seminar Feb. 2011 : *ILIAS Seminar CSC Research Unit & Proba. Seminar of the Mathematics Research Unit*

↪ DM, Rio E., Concentration Inequalities for Mean Field Particle Models, HAL-INRIA 6901 (2009), to appear in AAP 2011.
↪ Rio E., *Local invariance principles and their applications to density estimation*. Probability Theory and related Fields, vol. 98, 21-45 (1994).
↪ `Feynman-Kac formulae Genealogical and interacting particle systems, Springer (2004)`
↪ `DM, Doucet, Jasra. SMC Samplers.` JRSS B (2006).

Outline

Outline

Notation

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

- $(\mu, f) \in \mathcal{P}(E) \times \mathcal{B}(E) \quad \longrightarrow \quad \mu(f) = \int f(x)\mu(dx)$
- $M(x, dy)$ **integral operator on E**

$$\begin{aligned} M(f)(x) &= \int M(x, dy) f(y) \\ [\mu M](dy) &= \int \mu(dx) M(x, dy) \Longleftrightarrow [\mu M](f) = \mu[M(f)] \end{aligned}$$

- **Boltzmann-Gibbs transformation :** $G : E \rightarrow [0, 1]$ with $\mu(G) > 0$

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

Linear Markov models [Time $n \in \mathbb{N}$, $K_n(x, dy)$ Markov transitions]

Markov chain states X_n with transitions K_n :

$$\eta_n = \eta_{n-1} K_n = \mathrm{Law}(X_n)$$

- Law of large numbers $\rightsquigarrow$ iid copies $(X_n^i)_{i \geq 1}$

$$\eta_n^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{X_n^i} \simeq_{N \uparrow \infty} \eta_n$$

- Time homogeneous models $K_n = K$ $\rightsquigarrow$ Occupation mesures

$$\widehat{\eta}_n := \frac{1}{n} \sum_{p=1}^{n} \delta_{X_p} \simeq_{n \uparrow \infty} \eta_\infty = \eta_\infty K$$

$\rightsquigarrow$ Concentration inequalities for iid sequences or for Markov models

Nonlinear/nonhomogeneous Markov models

$K_{n,\eta}(x, dy)$ **collection of Markov transitions** $\sim \eta$ **probability meas.**

$$\rightsquigarrow \quad \eta_n = \eta_{n-1} K_{n,\eta_{n-1}} = \mathrm{Law}(\overline{X}_n) \in \mathcal{P}(E_n)$$

i.e. :

$$\mathbb{P}(\overline{X}_n \in dx_n \mid \overline{X}_{n-1}) = K_{n,\eta_{n-1}}(\overline{X}_{n-1}, dx_n) \quad \text{with} \quad \mathrm{Law}(\overline{X}_{n-1}) = \eta_{n-1}$$

McKean measures :

$$\mathbb{P}\left((\overline{X}_0, \overline{X}_1, \ldots, \overline{X}_n) \in d\left(x_0, x_1, \ldots, x_n\right)\right)$$

$$= \eta_0(dx_0) K_{1,\eta_0}(x_0, dx_1) \ldots K_{n,\eta_{n-1}}(x_{n-1}, dx_n)$$

Mean field particle interpretation

- **Objective :** Markov chain $\xi_n = (\xi_n^1, \ldots, \xi_n^N) \in E_n^N$ s.t.

$$\eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i} \simeq_{N \uparrow \infty} \eta_n$$

- $\rightsquigarrow$ **Particle approximation transitions** $(\forall 1 \leq i \leq N)$

$$\xi_{n-1}^i \rightsquigarrow \xi_n^i \sim K_{n,\eta_{n-1}^N}(\xi_{n-1}^i, dx_n)$$

Schematic picture : $\xi_n \in E_n^N \rightsquigarrow \xi_{n+1} \in E_{n+1}^N$

$$\begin{array}{ccc}
\xi_n^1 & \xrightarrow{K_{n+1,\eta_n^N}} & \xi_{n+1}^1 \\
\vdots & & \vdots \\
\xi_n^i & \longrightarrow & \xi_{n+1}^i \\
\vdots & & \vdots \\
\xi_n^N & \longrightarrow & \xi_{n+1}^N
\end{array}$$

Rationale :

$$\eta_n^N := \frac{1}{N} \sum_{1 \leq i \leq N} \delta_{\xi_n^i} \simeq_{N\uparrow\infty} \eta_n \quad \Longrightarrow \quad K_{n+1,\eta_n^N} \simeq_{N\uparrow\infty} K_{n+1,\eta_n}$$

$$\Longrightarrow \quad \xi_n^i \text{ almost iid copies of } \overline{X}_n$$

Concentration/fluctuations properties : $\eta_n^N \simeq_{N\uparrow\infty} \eta_n$??

Outline

Nonlinear models with Gaussian transitions

McKean diffusion type model

$$\eta_{n+1} := \eta_n K_{n+1,\eta_n}$$

with

$$K_{n,\eta}(x, dy)$$

$$= \frac{1}{\sqrt{(2\pi)^d \det(Q_n)}} \exp\left\{-\frac{1}{2}\left(y - d_n(x,\eta)\right)' \; Q_n^{-1}\left(y - d_n(x,\eta)\right)\right\} dy ,$$

$$\Updownarrow$$

$$\overline{X}_{n+1} := d_{n+1}\left(\overline{X}_n, \eta_n\right) + W_{n+1} \quad \text{with} \quad \eta_n = \mathrm{Law}(\overline{X}_n)$$

A McKean model of gases

Collision-jump type model

$$\eta_{n+1} := \eta_n K_{n+1,\eta_n}$$

with

$$K_{n+1,\eta}(x,dy) = \int \nu_n(ds)\ \eta(du)\ a_n(s,u)\ M_{n+1}((s,x),dy)$$

Example :

ν counting on $\{-1,+1\}$, $a_n(s,u) = 1_s(u)$, $M_{n+1}((s,x),dy) = \delta_{sx}(dy)$

$$\Downarrow$$

$$K_{n+1,\eta}(x,dy) = \eta(1)\ \delta_x(dy) + \eta(-1)\ \delta_{-x}(dy)$$

Updating-prediction transformations

$M_n(x, dy)$ Markov transitions and $G_n : E \to [0,1]$

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

Markov chain X_n with transitions M_n and initial condition $X_0 \simeq \eta_0$:

$$(1) \Longleftrightarrow \eta_n(f) \propto \mathbb{E}\left(f(X_n) \prod_{0 \le p < n} G_p(X_p) \right)$$

$\rightsquigarrow$ Nonlinear flow $\eta_{n+1} := \eta_n K_{n+1,\eta_n}$

- **Nonlinear Markov models :**

$$K_{n+1,\eta_n}(x, dz) \quad = \quad \int S_{n,\eta_n}(x, dy)\ M_{n+1}(y, dz)$$

$$S_{n,\eta_n}(x, dy) := G_n(x)\ \delta_x(dy) + (1 - G_n(x))\ \Psi_{G_n}(\eta_n)(dy)$$

Mean field genetic type particle model :

$$
\left.\begin{matrix} \xi_n^1 \\ \vdots \\ \xi_n^i \\ \vdots \\ \xi_n^N \end{matrix}\right] \xrightarrow{S_{n,\eta_n^N}} \left[\begin{matrix} \widehat{\xi}_n^1 & \xrightarrow{M_{n+1}} & \xi_{n+1}^1 \\ \vdots & & \vdots \\ \widehat{\xi}_n^i & \longrightarrow & \xi_{n+1}^i \\ \vdots & & \vdots \\ \widehat{\xi}_n^N & \longrightarrow & \xi_{n+1}^N \end{matrix}\right]
$$

Accept/Reject/Selection transition :

$$
S_{n,\eta_n^N}(\xi_n^i, dx)
$$

$$
:= \epsilon_n G_n(\xi_n^i)\ \delta_{\xi_n^i}(dx) + \left(1 - \epsilon_n G_n(\xi_n^i)\right)\ \sum_{j=1}^N \frac{G_n(\xi_n^j)}{\sum_{k=1}^N G_n(\xi_n^k)} \delta_{\xi_n^j}(dx)
$$

Ex. : $G_n = 1_A,\ \epsilon_n = 1 \rightsquigarrow G_n(\xi_n^i) = 1_A(\xi_n^i)$

$\hookrightarrow$ **FK particle models** $\supset$ *sequential Monte Carlo, population Monte Carlo, genetic algorithms, particle filters, pruning, spawning, reconfiguration, quantum-diffusion Monte carlo, go with the winner, etc.*

Nonlinear filtering

Filtering model

$$\mathbb{P}\left((X_n, Y_n) \in d(x', y') | (X_{n-1}, Y_{n-1}) = (x, y)\right) := M_n(x, dx')\; g_n(x', y')\; \lambda_n(dy')$$

- Given the observation sequence $Y = y$ with $G_n(x_n) = g_n(x_n, y_n)$

$$\eta_n = \mathrm{Law}(X_n \mid \forall 0 \leq p < n \quad Y_p = y_p) \quad \text{and} \quad \mathcal{Z}_{n+1} \propto p_n(y_0, \ldots, y_n)$$

- In path space settings

$$X_n = (X'_0, \ldots, X'_n) \quad \& \quad G_n(X'_0, \ldots, X'_n) = g_n(X'_n, y_n)$$

$$\Downarrow$$

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

Confinement, optimization, combinatorial pb, rare events

1. $\eta_n = \mathrm{Loi}((X_0, \ldots, X_n) \mid \forall 0 \leq p \leq n \quad X_p \in A_p)$
2. $\eta_n(dx) \propto e^{-\beta_n V(x)}\ \lambda(dx)$ with $\beta_n \uparrow$
3. $\eta_n(dx) \propto 1_{A_n}(x)\ \lambda(dx)$ with $A_n \downarrow$
4. $\eta_n = \mathrm{Loi}_{\pi}^{K}((X_0, \ldots, X_n) \mid X_n = x_n)$.
5. $\eta_n = \mathrm{Loi}(\ X \text{ hits } B_n \mid X \text{ hits } B_n \text{ before } A)$

Stochastic particle algorithms

1. M_n-local moves $\oplus$ individual selections $\in A_n$ i.e. $\sim G_n = 1_{A_n}$
2. MCMC local moves $\eta_n = \eta_n M_n$ $\oplus$ individual selections $\propto G_n = e^{-(\beta_{n+1}-\beta_n)V}$
3. MCMC local moves $\eta_n = \eta_n M_n$ $\oplus$ individual selections $\propto G_n = 1_{A_{n+1}}$
4. M-local moves $\oplus$ Selection $G(x_1, x_2) = \frac{\pi(dx_2)K(x_2,dx_1)}{\pi(dx_1)M(x_1,dx_2)}$
5. M_n-local moves $\oplus$ Selection-branching on upper/lower levels B_n.

Sub-Markov $\rightsquigarrow$ Markov

- X_n Markov $\in (E_n, \mathcal{E}_n)$ with transitions M_n, and $G_n : E_n \to [0,1]$

$$Q_n(x,dy) = G_{n-1}(x)\ M_n(x,dy) \quad \text{sub-Markov operator}$$

- $\rightsquigarrow E_n^c = E_n \cup \{c\}$.

$$X_n^c \in E_n^c \xrightarrow{\quad absorption\ \sim G_n \quad} \widehat{X}_n^c \xrightarrow{\quad exploration\ \sim M_n \quad} X_{n+1}^c$$

- Absorption: $\widehat{X}_n^c = X_n^c$, with proba $G(X_n^c)$; otherwise $\widehat{X}_n^c = c$.
- Exploration: like $X_n \rightsquigarrow X_{n+1}$

Feynman-Kac integral model

- $T = \inf \{n : \widehat{X}^c_n = c\}$ **absorption time :**

$$\begin{aligned}
\mathbb{P}(T \geq n) &= \gamma_n(1) := \mathbb{E}\left(\prod_{0\leq p<n} G(X_p)\right) \\
\mathbb{E}(f_n(X^c_n)\ ;\ (T \geq n)) &= \gamma_n(f_n) := \mathbb{E}\left(f_n(X_n) \prod_{0\leq p<n} G_p(X_p)\right)
\end{aligned}$$

- **Continuous time models :** $\Delta =$ time step

$$(M, G) = (Id + \Delta\ L, e^{-V\Delta})$$

$\rightsquigarrow$ *L-motions* $\oplus$ *expo. clocks rate V* $\oplus$ *Uniform selection*.

Outline

1. Introduction
2. Some applications
3. Stochastic perturbation analysis
   - Local sampling gaussian fields
   - A local transport formulation
   - First order analysis
4. Fluctuations and Concentration properties

A stochastic perturbation formulation

Nonlinear transport equation

$$\eta_{n+1} = \Phi_{n+1}(\eta_n) := \eta_n K_{n+1,\eta_n}$$

$\rightsquigarrow$ **Mean field particle model**

$$\begin{aligned} W_n^N &:= \sqrt{N}\left[\eta_n^N - \Phi_{n+1}(\eta_n^N)\right] := \frac{1}{\sqrt{N}} \sum_{1\leq i\leq N} \left[\delta_{\xi_n^i} - K_{n+1,\eta_n^N}(\xi_{n-1}^i, \cdot)\right] \\ &\simeq W_n \perp \text{Gaussian fields} \end{aligned}$$

$$\Updownarrow$$

Stochastic perturbation formulation

$$\eta_n^N = \Phi_{n+1}(\eta_n^N) + \frac{1}{\sqrt{N}}\, W_n^N$$

A local transport formulation $\Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \ldots \circ \Phi_{p+1}$

$$
\begin{array}{ccccccccc}
\eta_0 & \longrightarrow & \eta_1 = \Phi_1(\eta_0) & \longrightarrow & \eta_2 = \Phi_{0,2}(\eta_0) & \longrightarrow & \cdots & \longrightarrow & \Phi_{0,n}(\eta_0) \\
\Downarrow & & & & & & & & \\
\eta_0^N & \longrightarrow & \Phi_1(\eta_0^N) & \longrightarrow & \Phi_{0,2}(\eta_0^N) & \longrightarrow & \cdots & \longrightarrow & \Phi_{0,n}(\eta_0^N) \\
 & & \Downarrow & & & & & & \\
 & & \eta_1^N & \longrightarrow & \Phi_2(\eta_1^N) & \longrightarrow & \cdots & \longrightarrow & \Phi_{1,n}(\eta_1^N) \\
 & & & & \Downarrow & & & & \\
 & & & & \eta_2^N & \longrightarrow & \cdots & \longrightarrow & \Phi_{2,n}(\eta_2^N) \\
 & & & & & & \Downarrow & & \vdots \\
 & & & & & & \eta_{n-1}^N & \longrightarrow & \Phi_n(\eta_{n-1}^N) \\
 & & & & & & & & \Downarrow \\
 & & & & & & & & \eta_n^N
\end{array}
$$

$\rightsquigarrow$**Key decomposition formula**

$$\eta_n^N - \eta_n = \sum_{q=0}^{n} \left[\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))\right]$$

(H)
$$\Phi_n(\mu) - \Phi_n(\eta) = (\mu - \eta) \underbrace{D_{n,\eta}\Phi_n}_{\text{Integral operator}} + \underbrace{\mathcal{R}^{\Phi_n}(\mu,\eta)}_{\text{second order measure}}$$

$$\Downarrow$$

$$\Phi_{p,n}(\eta) - \Phi_{p,n}(\mu) = [\eta - \mu] D_\mu \Phi_{p,n} + \mathcal{R}^{\Phi_{p,n}}(\eta,\mu)$$

$$\Downarrow$$

$$\begin{aligned} V_n^N &:= \sqrt{N}\,[\eta_n^N - \eta_n] \\ &= \sum_{q=0}^{n} [\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))] \simeq \sum_{q=0}^{n} W_p^N \mathcal{D}_{p,n} \end{aligned}$$

for some integral operator $\mathcal{D}_{p,n}$ that enters the stability properties of $\Phi_{p,n}$

$$\rightsquigarrow \operatorname{osc}(\mathcal{D}_{p,n}(f)) \leq \underbrace{\beta(\mathcal{D}_{p,n})}_{\textbf{Dobrushin contraction coef.}} \operatorname{osc}(f) \leq e^{-\lambda(n-p)} \operatorname{osc}(f)$$

Outline

A functional Central Limit Theorem

We have the convergence in law of the random fields on $\mathcal{B}(E_n)$ (as $N \uparrow \infty$)

$$V_n^N = \sqrt{N}\,[\eta_n^N - \eta_n] \longrightarrow V_n := \sum_{q=0}^{n} W_p \mathcal{D}_{p,n}$$

with the independent centred gaussian fields W_n s.t.

$$\mathbb{E}(W_n(f)W_n(g)) = \eta_{n-1} K_{n,\eta_{n-1}}([f - K_{n,\eta_{n-1}}(f)][g - K_{n,\eta_{n-1}}(g)])\,.$$

$\rightsquigarrow$ 2nd notation :

$$\sigma_n^2 \text{ uniform local variance parameter } f = g \in \mathcal{B}(E) \text{ s.t. } \mathrm{osc}(f) \leq 1$$

Concentration inequalities

- Feynman-Kac models :
  - (DM, FK Springer 2004) Using Kintchine's type $\mathbb{L}_p$-estimates

$$\mathbb{P}\left(|\eta_n^N(f)-\eta_n(f)|>\epsilon\right)\leq(1+c_0\epsilon\sqrt{N})\ e^{-N\epsilon^2/c_1(n)}$$

  - (DM, A. Doucet, A. Jasra, Hal-INRIA pub. nb. 6700 (2008))

$$\mathbb{P}\left(|\left[\eta_n^N-\eta_n\right](f)|\geq\epsilon\right)\leq 6\ e^{-N\epsilon^2/c_1(n)}$$

with $\epsilon\in[0,1/2]$ and $c_1(n)$ related to the Dobrushin contraction coefficient of

$$P_{p,n}(f):=\frac{Q_{p,n}(f)}{Q_{p,n}(1)}\quad\text{with}\quad Q_{p,n}(f)(x_p):=\mathbb{E}_{p,x_p}\left(f(X_n)\prod_{p\leq p<n}G_q(X_p)\right)$$

In addition : $\Phi_{p,n}$ "stable" $\Rightarrow$ $c_1:=\sup_{n\geq 0}c_1(n)<\infty$.

Some observations

- Rather crude uniform concentration inequalities $\sim \mathbb{L}_p$-bounds
- Refinements $\rightsquigarrow$ Laplace estimates for the first and second order terms

$$\begin{aligned}
\sqrt{N}\,[\eta_n^N - \eta_n] &= \sum_{q=0}^{n} [\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))] \\
&\simeq \sum_{q=0}^{n} \underbrace{W_p^N}_{\sim\, W_p \text{ independent}} \quad \mathcal{D}_{p,n} + \text{second order measures}
\end{aligned}$$

Proof idea :

- First order term $\rightsquigarrow$ Bennett or Hoeffding's inequalities.
- Second order-Bias type term $\rightsquigarrow$ "sharp" $\mathbb{L}_{2p}$-Kintchine's bounds.
- First $\oplus$ Second order terms
  $\rightsquigarrow$ Rio's lemma on the inverse of the Legendre transform of sums.

Concentration for general models [DM, Rio Hal-INRIA, 6901, (2009)]

Notation : (σ_p uniform local variance, $\mathcal{D}_{p,n}$ first order integral operator)

$$\overline{\sigma}_n^2 = \sum_{p=0}^{n} \sigma_p^2 \; \beta(\mathcal{D}_{p,n})^2 \leq \beta_n^2 = \sum_{p=0}^{n} \beta(\mathcal{D}_{p,n})^2 \quad \text{and} \quad b_n^\star = \sup_{0\leq p\leq n} \beta(\mathcal{D}_{p,n})\,.$$

Theorem : $\forall x \geq 0$ the probability of each of the following pair of events

$$[\eta_n^N - \eta_n](f_n) \leq \frac{r_n}{N}\; \left(1+\epsilon_0^{-1}(x)\right) + \underbrace{\overline{\sigma}_n^2 \; b_n^\star \; \epsilon_1^{-1}\left(\frac{x}{N\overline{\sigma}_n^2}\right)}_{\text{Bennett term}}$$

and

$$[\eta_n^N - \eta_n](f_n) \leq \frac{r_n}{N}\; \left(1+\epsilon_0^{-1}(x)\right) + \underbrace{\sqrt{\frac{2x}{N}}\;\beta_n}_{\text{Hoeffding term}}$$

is greater than $1-e^{-x}$, with $r_n \sim$ second order remainder measures and

$$\epsilon_0(\lambda) = \frac{1}{2}\left(\lambda - \log\left(1+\lambda\right)\right)\,, \quad \epsilon_1(\lambda) = (1+\lambda)\log\left(1+\lambda\right) - \lambda$$

Some "direct" consequences

Notice that $(\epsilon_0, \epsilon_1) = (\alpha_0^\star, \alpha_1^\star)$ with

$$\alpha_0(t) := -t - \frac{1}{2}\log(1-2t) \quad \text{and} \quad \alpha_1(t) := e^t - 1 - t$$

$$\Downarrow$$

Corollary 1 [Bernstein type inequalities] :

$$-\frac{1}{N}\ \log \mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \frac{r_n}{N} + \lambda\right)$$

$$\geq\ \frac{\lambda^2}{2}\ \left(\left(b_n^\star\ \overline{\sigma}_n + \frac{\sqrt{2} r_n}{\sqrt{N}}\right)^2 + \lambda\left(2r_n + \frac{b_n^\star}{3}\right)\right)^{-1}$$

and

$$-\frac{1}{N}\ \log \mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \frac{r_n}{N} + \lambda\right) \quad \geq \quad \frac{\lambda^2}{2}\ \left(\left(\beta_n + \frac{\sqrt{2} r_n}{\sqrt{N}}\right)^2 + 2r_n\lambda\right)^{-1}.$$

Feymnan-Kac models s.t. Φ_n "stable"

Time homogeneous models s.t. $\sup_{x,y} G(x)/G(y) < \infty$ and

$$(M)_m \quad \exists m \geq 1 \quad \exists \epsilon_m > 0 \quad \text{s.t.} \quad \forall (x,y) \in E^2 \qquad M^m(x,\cdot) \geq \epsilon_m \; M^m(y,\cdot).$$

Notation :

$$\delta_m := \sup \prod_{0 \leq p < m} (G(x_p)/G(y_p)) \quad \text{and} \quad \varpi_{k,l}(m) \leq m \; \delta_{m-1} \; \delta_m^k / \epsilon_m^{k+2}$$

$$\Downarrow$$

$$r_n \leq 4 \; \varpi_{3,1}(m) \quad \text{and} \quad b_n^\star \leq 2 \; \delta_m/\epsilon_m$$

as well as

$$\overline{\sigma}_n^2 \leq 4 \; \varpi_{2,2}(m) \; \sigma^2 \quad \text{and} \quad \beta_n^2 \leq 4 \; \varpi_{2,2}(m)$$

Corollary 2 [Uniform estimates w.r.t. time horizon] :

For any $n \geq 0$, and any $x \geq 0$ the probability of each of the following pair of events

$$[\eta_n^N - \eta_n](f_n)$$

$$\leq \frac{4}{N}\varpi_{3,1}(m)\left(1+\epsilon_0^{-1}(x)\right) + \frac{8\delta_m}{\epsilon_m}\ \varpi_{2,2}(m)\ \sigma^2\ \ \epsilon_1^{-1}\left(\frac{x}{4\sigma^2\varpi_{2,2}(m)\ N}\right)$$

and

$$[\eta_n^N - \eta_n](f_n) \leq \frac{4}{N}\varpi_{3,1}(m)\left(1+\epsilon_0^{-1}(x)\right) + 2\sqrt{\frac{2\varpi_{2,2}(m)x}{N}}$$

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