Non-asymptotic Gaussian Estimates for the Recursive Approximation of the Invariant Measure of a Diffusion

PDE & Probability Methods for Interactions (Sophia)

I. Honoré, S. Menozzi & G. Pagès

Univ. d'Évry & LPMA–UPMC

Steady regime of diffusion

▷ Let $(X_t)_{t \geqslant 0}$ be the unique strong solution to the stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t.$$

where $b : \mathbf{R}^d \to \mathbf{R}^d$ and $\sigma : \mathbf{R}^d \to \mathcal{M}(d, q)$ ar Lipschitz continuous. Let L be its infinitesimal generator defined on twice differentiable functions $g : \mathbf{R}^d \to \mathbf{R}$ by

$$Lg = (b|\nabla g) + \frac{1}{2}\mathrm{Tr}\big(\sigma^* D^2 g\, \sigma\big).$$

▷ *Existence of a stationary regime (Mean-reversion)*:
If there exists a $\mathcal{C}^2$ *Lyapunov* function $V : \mathbf{R}^d \to \mathbf{R}_+$ such that

$$\lim_{|x| \to +\infty} V(x) = +\infty \quad \text{and} \quad \overline{\lim_{|x| \to +\infty}}\, LV(x) = -\infty.$$

then there exists a distribution ν such that $(X_t)_{t \geqslant 0}$ is $\mathbf{P}_\nu$-stationary i.e

$$X_0 \stackrel{d}{=} \nu \text{ and } \quad \forall\, \theta > 0, \quad (X_{t_1}, \ldots, X_{t_p}) \stackrel{d}{=} (X_{t_1+\theta}, \ldots, X_{t_p+\theta}).$$

Connection and examples

- Connection with stationary Fokker-Planck equation:

$$\nu \text{ invariant } \iff \forall \in \mathcal{C}^2_K(\mathbf{R}^d, \mathbf{R}), \quad \nu(Ag) = 0$$

(made rigorous through Echeverria-Weiss theorem) i.e. if $\nu(dx) = p(x)dx$, if p solution to

$$A^* p = 0.$$

- Setting 2 (Stat. mechanics): If V is $\mathcal{C}^2$ and $\lim_{|x|\to+\infty} V(x) = +\infty$

$$dX_t = -\nabla V(X_t)dt + \sigma dW_t \quad \sigma \in (0, +\infty) \rightsquigarrow \nu_\sigma(dx) = C_V e^{-\frac{V(x)}{2\sigma^2}} dx.$$

Then

$$\nu_\sigma \overset{weakly}{\longrightarrow} \nu = \mathrm{Unif}(\mathrm{argmin}\, V) \quad \text{as} \quad \sigma \to 0.$$

Classical starting result of the theory of simulated annealing ...

Other applications

- Stationary stochastic volatility models: Heston model $\rightsquigarrow \gamma$-distribution, multi-asset Heston models for the pricing and hedging of path dependent options.
- Pricing of swing and spark options, real options for gas plants, gas storages.
- Ergodic stochastic control problems (long term investments, etc).

Ergodicity, stability

▷ **Ergodicity**: If ν is an extremal invariant measure, then X is $\mathbf{P}_\nu$-ergodic:

$$\nu(dx)\text{-}a.s.,\ \mathbf{P}_x\text{-}a.s. \quad \nu_t(\omega, x, d\xi) = \frac{1}{t}\int_0^t \delta_{X_s^x(\omega)} ds \stackrel{(weakly)}{\longrightarrow} \nu \ \text{ as } t \to +\infty$$

▷ **Stability**: Let $\mathcal{I}_{SDE} = \big\{\nu : \mathbf{P}_\nu \ \textit{stationary}\big\}$.

Theorem (Stability result)

If

$$LV \leqslant \beta - \alpha V^\rho,\ \alpha > 0,\ \rho \in (0,1],$$

then $\mathcal{I}_{SDE} = \{\nu,\ \textit{invariant distribution}\} \neq \emptyset$, *convex, weakly compact and*

$$\forall\, x \in \mathbf{R}^d,\ \mathbf{P}_x\text{-}a.s. \quad d_{weak}\big(\nu_t(\omega, x, d\xi), \mathcal{I}_{SDE}\big) \stackrel{weakly}{\longrightarrow} 0.$$

If $\mathcal{I}_{SDE} = \{\nu\}$, *for every continuous f with $f = o(V^p)$ at infinity,*

$$\forall\, x \in \mathbf{R}^d,\ \mathbf{P}_x\text{-}a.s. \quad \frac{1}{t}\int_0^t f(X_s^x) ds \longrightarrow \nu(f)$$

Uniqueness of ν

- $\mathcal{L}(X_t^x) = p_t(x,y)\mu(dy)$ with $(\forall\, x > 0,\ p_t(x,y) > 0)\,\mu(dy)$-*a.s.*
  $\Big[\Leftarrow$ hypo-ellipticity and controlability$\Big]$

  or

- (b,σ) confluence:

$$(b(x)-b(y)|x-y)+\frac{1}{2}\mathrm{Tr}\big((\sigma(x)-\sigma(y))(\sigma(x)-\sigma(y))^*\big)<0$$

then

ν is unique.

Weak rate of convergence

Still assume $\mathcal{I}_{SDE} = \{\nu\}$.

Theorem (Weak rate of convergence, Bhattacharya's CLT)

If $f - \nu(f) = L\varphi$, $\varphi \in \mathcal{C}_b^2$, then

$$\sqrt{t}\big(\nu_t(\omega, x)(f) - \nu(f)\big) \stackrel{(weakly)}{\longrightarrow} \mathcal{N}\big(0, \sigma(f)^2\big) \quad \text{as} \quad t \to +\infty.$$

with

$$\sigma(f)^2 = \int_{\mathbf{R}^d} |\sigma^* \varphi|^2 d\nu.$$

Langevin MC: mimicking the ergodic theorem

▷ Aim: Computing by Langevin Monte Carlo simulation $\nu(f)$.

Langevin/ergodic MC simulation = 1 path of length n.

The constant step approach (Talay,'96)

▷ Euler scheme with constant step $\gamma > 0$:

$$\bar{X}^{\gamma}_{n+1} = \bar{X}^{\gamma}_{n} + \gamma b(\bar{X}^{\gamma}_{n}) + \sigma(\bar{X}^{\gamma}_{n})\big(W_{n\gamma} - W_{(n-1)\gamma}\big),\ n \geqslant 0,\ \bar{X}^{\gamma}_{0} = x.$$

▷ Markov chain $(\bar{X}^{\gamma}_{n})_{n\geqslant 0}$ shares the properties of the diffusions for a small enough step γ.

- (Unique) invariant distribution ν^{γ}.
- stability/positive recurrence of the chain.

▷ "Regular"/uniform empirical measures:

$$\mu^{\gamma}_{n}(\omega, dx) = \frac{1}{n\not{\gamma}} \sum_{k=1}^{n} \not{\gamma}\delta_{\bar{X}^{\gamma}_{k-1}(\omega)} \overset{\mathbf{R}^d}{\Longrightarrow} \nu^{\gamma} \quad \text{as} \quad n \to +\infty \mathbf{P}\text{-}a.s..$$

and (Talay 96): $\nu^{\gamma} \overset{\mathbf{R}^d}{\Longrightarrow} \nu$ at rate $O(\gamma)$.

Unbiased estimation

(Lamberton-P.'00, Lemaire '05, Panloup '06, etc.)

▷ Switch to the Euler scheme with decreasing step:

$$\bar{X}_{n+1} = \bar{X}_n + \gamma_{n+1} b(\bar{X}_n) + \sqrt{\gamma_n}\sigma(\bar{X}_n) U_{n+1}, \; n \geqslant 0, \; \bar{X}_0 = x$$

where $(U_n)_{n\geqslant 1}$ is a(n L^2) white noise and

$$\gamma_n > 0, \qquad \gamma_n \downarrow 0 \qquad \text{and} \qquad \Gamma_n := \gamma_1 + \cdots + \gamma_n \to +\infty.$$

▷ For numerics $U_n = \frac{W_{\Gamma_{n+1}} - W_{\Gamma_n}}{\sqrt{\gamma_n}} \sim \mathcal{N}(0; I_q)$ or Bernoulli$(1/2)^{\otimes q}$.

▷ Weighted empirical measures by mimicking the ergodic continuous time empirical measure:

$$\nu_n(\omega, dx) = \frac{1}{\Gamma_n} \sum_{k=1}^{n} \gamma_k \delta_{\bar{X}_{k-1}(\omega)} \simeq \frac{1}{\Gamma_n} \int_0^{\Gamma_n} \delta_{\bar{X}_{N(s)}(\omega)} ds.$$

Theorem (Lamberton-P. ('02) & ('03))

Mean-reversion: $\rho = 1$ & $\mathcal{I}_{inv} = \{\nu\}$ & $\gamma_n = \gamma_1 n^{-\theta}$, $0 < \theta \leqslant 1$.

(*a*) **Convergence:** $\forall\, x \in \mathbf{R}^d$, $\mathbf{P}_x(d\omega)$-*a.s.* $\nu_n^{\gamma}(\omega, dx) \overset{(weakly)}{\longrightarrow} \nu$.

(*b*) **Rate of convergence:** *Assume* $f - \nu(f) = -L\varphi$, $\varphi \in \mathcal{C}_b^3$, $\mathbf{E}\, U_1^{\otimes 3} = 0$:

- *If* $\gamma_n = \gamma_1 n^{-\theta}$, $\theta \in (\frac{1}{3}, 1)$ *(fast decreasing step): Bhattacharya's CLT for diffusion holds:*

$$n^{\frac{1-\theta}{2}}\left(\nu_n(f) - \nu(f)\right) \overset{(weakly)}{\longrightarrow} c_{\gamma_1,\theta}.\mathcal{N}\left(0, \sigma_f^2\right).$$

- *If* $\gamma_n = \gamma_1 n^{-\theta}$, $\theta = \frac{1}{3}$ *(optimal rate): a bias* m_f *appears:*

$$n^{\frac{1}{3}}\left(\nu_n(f) - \nu(f)\right) \overset{(weakly)}{\longrightarrow} c_{\gamma_1,1/3}.\mathcal{N}\left(\widetilde{\gamma}\, m_f; \sigma_f^2\right).$$

- *If* $\gamma_n = \gamma_1 n^{-\theta}$, $\theta \in (0, \frac{1}{3})$ *(slowly decreasing step): the discretization effect slows down the convergence.*

$$c_{\gamma_1,\theta}\, n^{\theta}\left(\nu_n(f) - \nu(f)\right) \overset{\mathbf{P}_x/a.s.}{\longrightarrow} c_{\gamma_1,\theta}.m_f.$$

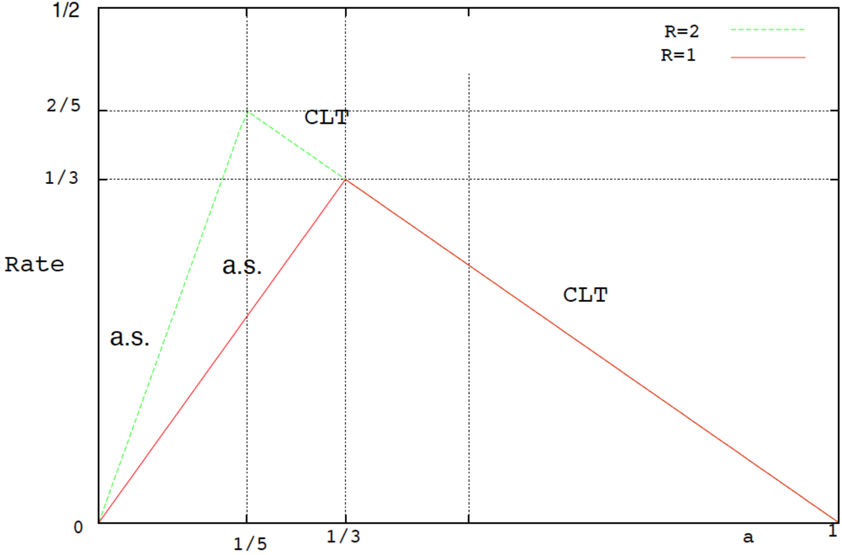

Figure: Rate of convergence depending on a when $\gamma_n \sim \gamma_1 n^{-a}$.

Looking for non-asymptotic confidence intervals

- How to devise non-asymptotic confidence interval in presence of a CLT ?
  Let $\gamma_n \asymp n^{-\theta}$, $\theta \in [\frac{1}{3}, 1]$.

$$\mathbf{P}\Big(n^{\frac{1-\theta}{2}}\big(\nu_n(f) - \nu(f)\big) \geqslant a\Big) \leqslant Ce^{-ca^2}?$$

  - Malrieu-Talay (*MCQMC*^{2} Proc., 2006), for the constant step Euler scheme by a direct approach the Euler semi-group (log-Sob, Poincaré) implemented.
  - Frikha-Menozzi (*ECP*, 2012) by a martingale approach based on the Gaussian deviation inequalities.
  - For every Lipschitz continuous function $f : \mathbf{R}^r \to \mathbf{R}$ and every $\lambda > 0$.

$$\mathbf{E}\big[\exp(\lambda f(U_1))\big] \leq \exp\left(\lambda \mathbf{E}[f(U_1)] + \frac{\lambda^2 [f]_{\rm Lip}^2}{2}\right).$$

Remark: $\|\sigma\|^2 = \mathrm{Tr}\big(\sigma\sigma^*\big)$ denotes the Fröbenius norm.

Theorem (Fast Decreasing Step: $\gamma_k = \gamma_1 k^{-\theta}, \frac{1}{3} < \theta \leqslant 1$)

Assume $U_1 \in L^4$. *Let* $f : \mathbf{R}^d \to \mathbf{R}$ *be a coboundary such that*

$$f - \nu(f) = -L\varphi,\ \varphi = O(\sqrt{V}), \quad \varphi \in \mathcal{C}^2_{b,Lip}(\mathbf{R}^d, R)$$

For every $a > 0$,

$$\mathbf{P}\left[\sqrt{\Gamma_n}|\nu_n(f) - \nu(f))| \geq a\right] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|^2_\infty \|\nabla\varphi\|^2_\infty}\right)$$

where the explicit sequences $C_n \downarrow 1$ *and* $c_n \uparrow 1$.

Can this result be improved?

- We have $\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2$ instead of $\displaystyle\int |\sigma^*\varphi|^2 d\nu$.
- $\|\nabla\varphi\|_\infty^2$ looks somewhat *intrinsic*.
- What about $\|\sigma\|_\infty^2$? In particular, what happens if we also assume that

$$\|\sigma\|^2 - \nu(\|\sigma\|^2) \text{ is a coboundary}\ldots$$

If $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is a coboundary...

Theorem

Still with $\gamma_n = \gamma_1 n^{-\theta}$, $\theta \in (1/3, 1)$. *If furthermore*

$$\|\sigma\|^2 - \nu(\|\sigma\|^2) = -L\varsigma,\ \varsigma \in \mathcal{C}^3_{\rm Lip}(\mathbf{R}^d, \mathbf{R})\ [\dots]$$

Then, for every $a \in \left(0, \varepsilon_0 \frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}}\right)$ *[note that* $\frac{\sqrt{\Gamma_n}}{\Gamma_n^{(2)}} \to +\infty$*].*

$$\mathbf{P}\left[\left|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f))\right| \geq a\right] \leq 2\widetilde{C}_n \exp\left(-\widetilde{c}_n \frac{a^2}{2\,\nu(\|\sigma\|^2)\|\nabla\varphi\|_\infty^2}\right)$$

where $\widetilde{c}_n \uparrow 1$ *and* $\widetilde{C}_n \downarrow 1$ *for large enough* n *respectively.*

Critical case (with bias!), $\gamma_n = \frac{\gamma_1}{n^\theta}$, $\theta = \frac{1}{3}$

Theorem (Critical Decreasing step (case $\theta = \frac{1}{3}$))

- *Let f be s.t. $f - \nu(f) = -L\varphi$, $\varphi \in \mathcal{C}^4(\mathbf{R}^d, \mathbf{R})$. Then $\forall\, n \geq 1$, $\forall\, a > 0$:*

$$\mathbf{P}\Big[\Big|\sqrt{\Gamma_n}\big(\nu_n(f) - \nu(f)\big) - \underbrace{\widetilde{\gamma}\, m}_{bias} + E_n\Big| \geq a\Big] \leq 2C_n \exp\left(-c_n \frac{a^2}{2\|\sigma\|_\infty^2 \|\nabla\varphi\|_\infty^2}\right)$$

where $E_n \overset{n\to+\infty}{\longrightarrow} 0$ and $c_n \uparrow 1$ and $C_n \downarrow 1$, for large enough n(free of a).

- *Moreover, if $\|D^2\varphi(x)\| \leqslant 1/(1+|x|)$, then $\big(E_n\big)_{n\geqslant 1}$ is square exponentially tight:*

$$\exists \lambda_0 > 0,\ \forall \lambda \leqslant \lambda_0, \quad \sup_{n\geqslant 1} \mathbf{E}\Big[\exp(\lambda |E_n|^2)\Big] < +\infty$$

If $\gamma_n = \gamma_1 n^{-1/3}$, $\theta = \frac{1}{3}$ and $\|\sigma\|^2 - \nu(\|\sigma\|^2)$ is a coboundary...

Theorem

If $\gamma_n = \gamma_1 n^{-1/3}$ *and furthermore*

$$\|\sigma\|^2 - \nu(\|\sigma\|^2) = -L\varsigma,\ \varsigma \in \mathcal{C}^3_{\rm Lip}(\mathbf{R}^d, \mathbf{R})\ [\dots]$$

Then, for every $a > 0$,

$$\mathbf{P}\Big[\big|\sqrt{\Gamma_n}(\nu_n(f) - \nu(f)) - \widetilde{\gamma} m + E_n\big| \geq a\Big] \leq 2\widetilde{C}_n \exp\left(-\widetilde{c}_n \frac{\Phi_n(a)}{2\,\nu(\|\sigma\|^2)\|\nabla\varphi\|^2_\infty}\right)$$

where $\widetilde{c}_n \uparrow 1$ *and* $\widetilde{C}_n \downarrow 1$ *for large enough* n *respectively and*

$$\Phi_n(a) = \alpha_n a^2 - \beta_n a^4,\ \alpha_n \to 1,\ \beta_n \geqslant 0,\ \beta_n \to 0 \quad \text{as } n \to \infty.$$