

# Concentration analysis of multivariate elliptic diffusions

Statistics for Stochastic processes – Luxembourg 2022

---

Lukas Trottner

joint work with Cathrine Aeckerle-Willems and Claudia Strauch

Aarhus University

University of Mannheim



AARHUS UNIVERSITY

## Some known concentration results for Markov processes

Let  $X$  be a nice ergodic Markov processes with semigroup  $(P_t)_{t \geq 0}$ , invariant distribution  $\mu$  and generator  $L$  on  $\mathbb{L}^2(\mu)$  (endowed with inner product  $\langle f, g \rangle_\mu = \int fg d\mu$ ) and denote

$$C_\nu(f, T, x) := \mathbb{P}^\nu \left( \left| \frac{1}{T} \int_0^T f(X_t) dt - \mu(f) \right| > x \right), \quad f \in \mathbb{L}^2(\mu), x, T > 0.$$

Bounds have been mostly studied with two approaches (*Lyapunov vs. Poincaré* [Bak08]):

1. **Functional inequalities:** If  $(P_t)_{t \geq 0}$  is symmetric let  $\mathcal{E}$  be the Dirichlet form associated to  $-L$  (then  $D(\mathcal{E}) = D(\sqrt{-L})$  and  $\mathcal{E}(f, g) = -\langle Lf, g \rangle_\mu$  for  $f \in D(L), g \in D(\mathcal{E})$ )

- **Poincaré inequality (PI):**

$$\text{Var}_\mu(g) := \mu(g^2) - \mu(g)^2 \leq -C_P \langle Lg, g \rangle_\mu, \quad g \in D(L).$$

Implies:  $\|P_t f - \mu(f)\|_{\mathbb{L}^2(\mu)} \leq e^{-2t/C_P} \|f - \mu(f)\|_{\mathbb{L}^2(\mu)}$

- **log-Sobolev inequality (LS):**  $(P_t)_{t \geq 0}$  symmetric and

$$\text{Ent}_\mu(g^2) := \mu(g^2 \log g^2) - \mu(g^2) \log \mu(g^2) \leq 2C_{LS} \mathcal{E}(g, g), \quad g \in D(\mathcal{E}),$$

2. **Mixing assumptions:**

$$(\alpha(\nu, \varphi)) : \quad \alpha_\nu(t) := \sup_{s \geq 0} \sup_{A \in \sigma(X_u, u \leq s), B \in \sigma(X_u, u \geq s+t)} |\mathbb{P}^\nu(A \cap B) - \mathbb{P}^\nu(A)\mathbb{P}^\nu(B)| \leq \varphi(t) \xrightarrow[t \rightarrow \infty]{} 0.$$

For reasonable  $\nu$  implied by **ergodicity** of  $P_t$ , i.e.,  $\|P_t(x, \cdot) - \mu\|_{TV} \leq CV(x)\varphi(t)$

## Some known concentration results for Markov processes

- [Lez01] Suppose  $\nu \ll \mu$ ,  $d\nu/d\mu \in \mathbb{L}^2(\mu)$  and  $\|f\|_\infty < \infty$ . If  $\mu$  satisfies (PI) then we have the **Bernstein inequality (BI)**

$$\mathbb{C}_\nu(f, T, x) \leq 2 \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}^2(\mu)} \exp\left(-\frac{Tx^2}{2(\sigma^2(f) + 2C_P\|f\|_\infty x)}\right),$$

where  $\sigma^2(f) = \lim_{t \rightarrow \infty} t^{-1} \text{Var}_{\mathbb{P}^\mu} \left( \int_0^t f(X_s) ds \right)$

- [GGW14] If  $\mu$  satisfies (LS),  $\mu(f) = 0$  and  $\mu(\exp(\lambda_\pm f^\pm)) < \infty$  for some  $\lambda_\pm > 0$  then we have the (BI)

$$\mathbb{C}_\nu(f, T, x) \leq 2 \left\| \frac{d\nu}{d\mu} \right\|_{\mathbb{L}^2(\mu)} \exp\left(-\frac{Tx^2}{2(\sigma^2(f) + C_P(\Lambda^*)^{-1}(2C_{LS}/C_P)x)}\right),$$

w.  $\Lambda^* = \Lambda_+^* \vee \Lambda_-^*$  and  $\Lambda_\pm^*$  Legendre transf. of  $[0, \lambda_\pm] \ni s \mapsto \Lambda_\pm(s) := \log \mu(\exp(s(\pm f)))$ .

- [CG08] If  $(\alpha(\mu, \varphi))$  with  $\varphi(t) = c \exp(-t^{\frac{1-q}{1+q}})$ ,  $q \in [0, 1)$  [ $q = 0$ : **exponential mixing**,  $q \in (0, 1)$ : **subexponential mixing**] and  $\|f\|_\infty < \infty$ , then for any  $x \geq C(c, q)/\sqrt{T}$ , it holds

$$\mathbb{C}_\mu(f, T, x) \leq 2 \exp\left(-c(q) \left(\frac{x\sqrt{T}}{\|f\|_\infty}\right)^{1-q}\right).$$

- Let  $X$  be a (weak) solution to the SDE

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

$b \in \text{Lip}_{\text{loc}}(\mathbb{R}^d; \mathbb{R}^d)$ ,  $\sigma \in \text{Lip}(\mathbb{R}^d; \mathbb{R}^{d \times d})$  and bounded,  $a := \sigma \sigma^\top$  s.t.  $\lambda_- \mathbb{I} \leq a(x) \leq \lambda_+ \mathbb{I}$ ,  $\forall x$

- Let  $L = b^\top \nabla + \sum_{i,j} a_{i,j} \partial_{x_i} \partial_{x_j}$  and suppose that for given  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  the **Poisson equation**  $Lg = f$  has some sufficiently regular solution  $L^{-1}[f]$
- By Itô's formula:  $L^{-1}[f](X_t) - L^{-1}[f](X_0) = \int_0^t LL^{-1}[f](X_s) ds + \int_0^t (\nabla L^{-1}[f](X_s))^\top \sigma(X_s) dW_s$  and hence

$$\int_0^t f(X_s) ds = \underbrace{\int_0^t (-\nabla L^{-1}[f](X_s))^\top \sigma(X_s) dW_s}_{(\text{loc.}) \text{ martingale}} + \underbrace{L^{-1}[f](X_t) - L^{-1}[f](X_0)}_{\text{remainder}}$$

- $\rightsquigarrow$  If we have some control on  $L^{-1}[f]$ ,  $\nabla L^{-1}[f]$  we can use martingale approximation for derivation of concentration bounds

[AWS21; GP07] scalar case  $d = 1$ : explicit formula for  $\nabla L^{-1}[f]$  available  $\rightsquigarrow$  careful bounds on  $\nabla L^{-1}[K_h(x - \cdot)b]$  given

- exponential ergodicity of  $(P_t)_{t \geq 0}$
- at most linear drift

provide uniform concentration results that are tight enough for proving minimax estimation rates for drift estimation

[NR20] multivariate case  $d \geq 1$ :  $\sigma = \mathbb{I}$  and  $b$  periodic  $\rightsquigarrow$  for periodic  $f$ ,  $\nabla L^{-1}[f]$  is bounded. This yields sub-Gaussian concentration for such  $f$ , which is used for minimax Bayesian drift inference

### Question

What kind of concentration rates can be achieved in a multivariate setting under relaxed stability assumptions and for unbounded  $f$ ?

## Poisson equation under subexponential drift assumptions

Assume  $\|b(x)\| \lesssim 1 + \|x\|^\kappa$  and for some  $q \in (-1, 1)$ ,  $\tau, A > 0$ ,

$$\langle b(x), x/\|x\| \rangle \leq -\tau \|x\|^{-q}, \quad \|x\| > A. \quad (D_q)$$

[PV01; BRS18] If  $\mu(f) = 0$  and  $\|f(x)\| \lesssim 1 + \|x\|^\eta$ , then for  $L^{-1}[f](x) := -\int_0^\infty P_t f(x) dt$  we have  $L^{-1}[f] \in \mathcal{W}_{\text{loc}}^{2,p}(\mathbb{R}^d)$  for any  $p > 1$ ,  $L^{-1}[f]$  solves the Poisson equation and

$$|L^{-1}[f](x)| \lesssim 1 + \|x\|^{\eta+1+q}, \quad |\nabla L^{-1}[f](x)| \lesssim 1 + \|x\|^{\eta+\kappa+1+q}.$$

Let  $\|v\|_f := \sup_{|g| \leq f} |v(g)|$  for some  $f \geq 1$  and  $(\Psi_1, \Psi_2)$  be either pairs of inverse Young functions (i.e.,  $xy \leq \Psi_1^{-1}(x) + \Psi_2^{-1}(y)$ ) or  $(\text{Id}, 1)$  or  $(1, \text{Id})$ .

### Proposition [DFG09; AWST22]

We have

$$\|P_t f(x) - \mu\|_{\text{TV}} \leq C(q_+) \exp(\iota \|x\|^{1-q_+}) \exp\left(-\iota' t^{\frac{1-q_+}{1+q_+}}\right) \quad \text{and} \quad \int_{\mathbb{R}^d} \exp(\iota \|x\|^{1-q_+}) \mu(dx) < \infty.$$

Moreover, for  $\gamma \geq 1 + q$ ,  $r_{\gamma,q}(t) \sim (1+t)^{(\gamma-(1+q))/(1+q)}$ ,  $f_{\gamma,q}(x) \sim 1 + \|x\|^{\gamma-(1+q)}$ ,

$$(\Psi_1(r_{\gamma,q}(t)) \vee 1) \|P_t(x, \cdot) - \mu\|_{1 \vee \Psi_2 \circ f_{\gamma,q}} \leq C(\Psi)(1 + \|x\|^\gamma).$$

## Theorem [AWST22]

Assume  $(D_q)$ ,  $\|b(x)\| \lesssim 1 + \|x\|^\kappa$  and  $|f(x)| \leq \mathfrak{L}(1 + \|x\|^\eta)$ . Let

$$\rho(\eta, \kappa, q) := \begin{cases} 1/(1 - q_+), & \eta = 0 \\ \frac{1}{2} + \frac{\eta + \kappa + 1 + q}{1 - q_+}, & \eta > 0. \end{cases}$$

Then, there exists a constant  $c > 0$  s.t. for any  $x \geq 2/\sqrt{T}$ ,

$$\mathbb{C}_\mu(f, T, x) := \mathbb{E}^\mu \left[ \left| \frac{1}{T} \int_0^T f(X_t) dt - \mu(f) \right| > x \right] \leq \exp \left( -c \left( \frac{x\sqrt{T}}{\mathfrak{L}} \right)^{1/\rho(\eta, \kappa, q)} \right).$$

Poincaré, $\eta = 0$	log-Sobolev, $\eta \leq 2$	subexponential, $\eta > 0$
$\frac{\log(1/\delta)}{\varepsilon}$	$\frac{\log(1/\delta)}{\varepsilon}$	$\frac{(c^{-2} \log(1/\delta))^{2\rho(\eta, \kappa, q)}}{\varepsilon^2}$

**Table 1:** Order of sufficient sample length  $\Psi(\varepsilon, \delta)$  s.t.  $(\varepsilon, \delta)$ -PAC-bound  $\mathbb{P}^\mu(|\mu_T(f) - \mu(f)| \leq \varepsilon) \geq 1 - \delta$  holds for  $T \geq \Psi(\varepsilon, \delta)$

## Discrete-time concentration result

Let observations  $(X_{k\Delta})_{k=1,\dots,n}$  be given for some  $\Delta \leq 1$ . Define the mean estimator

$\mathbb{H}_{n,\Delta}(f) := \frac{1}{\sqrt{n\Delta}} \mathbb{G}_{n,\Delta}(f)$ , where

$$\mathbb{G}_{n,\Delta}(f) := \frac{1}{\sqrt{n\Delta}} \sum_{k=1}^n f(X_{k\Delta})\Delta.$$

### Theorem [AWST22]

Assume  $(D_q)$ ,  $\|b(x)\| \lesssim 1 + \|x\|^\kappa$ ,  $f \in C^2(\mathbb{R}^d; \mathbb{R})$  s.t.  $\|D^k f(x)\| \lesssim 1 + \|x\|^{\eta_k}$ ,  $k = 0, 1, 2$ . Define  $\alpha =: (\kappa + \eta_1) \vee \eta_2$ , and let  $\tilde{\gamma} > 1 + q$ ,  $r > 1$ , s.t.  $\tilde{\gamma} - (1 + q) > r(\alpha \vee (1 + q))/(r - 1)$ . Then, for some constant  $\mathfrak{D}$  independent of  $n, p, \Delta$ , for any  $p \geq 2$  we have

$$\|\mathbb{G}_{n,\Delta}(f - \mu(f))\|_{L^p(\mathbb{P}^\mu)} \leq \mathfrak{D} \left( \sqrt{n\Delta}^{3/2} + \Delta p^{\frac{\max\{(\tilde{\gamma} + 2\alpha + 1 - q_+)/2, \eta_1 + 1 - q_+\}}{1 - q_+}} + p^{\frac{1}{2} + \frac{\eta + \kappa + 1 + q}{1 - q_+}} \right) := \Phi(n, \Delta, p),$$

and

$$\mathbb{P}^\mu \left( |\mathbb{H}_{n,\Delta}(f) - \mu(f)| > e(n\Delta)^{-1/2} \Phi(n, \Delta, x) \right) \leq e^{-x}, \quad x \geq 2.$$



# Applications

---

## Lasso for parametrized drifts

For a given **dictionary**  $\{\psi_1, \dots, \psi_N\}$  of Lipschitz functions  $\psi_i: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , let  $X$  be the strong solution to

$$dX_t = b_{\theta^0}(X_t) dt + \sigma(X_t) dW_t, \quad \text{where} \quad b_{\theta^0}(x) = \sum_{i=1}^N \theta_i^0 \psi_i(x).$$

Let  $\boldsymbol{\psi}(x) = (\psi_1(x), \dots, \psi_N(x))$ ,  $\boldsymbol{\Psi}(x) := (\sigma^{-1}(x)\boldsymbol{\psi}(x))^\top \sigma^{-1}(x)\boldsymbol{\psi}(x)$  and  $\bar{\boldsymbol{\Psi}}_T := T^{-1} \int_0^T \boldsymbol{\Psi}(X_t) dt$ .

Then for  $b_\theta := \boldsymbol{\psi}\theta$ , negative **log-likelihood** given by

$$\mathcal{L}_T(\theta) = \mathcal{L}_T(b_\theta) = \theta^\top \bar{\boldsymbol{\Psi}}_T \theta - 2\theta^\top \frac{1}{T} \int_0^T \boldsymbol{\psi}(X_t)^\top a^{-1}(X_t) dX_t.$$

### Goal

Study convergence guarantees of **Lasso estimator**

$$\hat{\theta}_T := \arg \min_{\theta \in \mathbb{R}^N} \{ \mathcal{L}_T(\theta) + \lambda \|\theta\|_1 \},$$

under sparsity assumptions on  $\theta_0$ , i.e.,  $\|\theta_0\|_0 \leq s_0$ .

## Assumptions and examples

We assume

1.  $\exists A, \tau > 0, q \in [-1, 1) : \langle b_{\theta^0}(x), x/\|x\| \rangle \leq -\tau \|x\|^{-q}, \quad \|x\| > A;$
2.  $\lambda_{\max}(\Psi(x)) \lesssim 1 + \|x\|^{2\eta};$
3.  $\bar{\Psi}_T$  is positive definite  $\mathbb{P}_{\theta^0}$ -a.s.

Example 1: **OU process:**  $N = d^2$ ,  $\psi_i(x) = E_i x$ , where  $E_i$  is the matrix with  $i$ -th entry (counted row-wise, say) equal to 1 and all other entries equal to 0,  $\theta^0$  s.t. for  $b_{\theta^0}(\cdot) = A_{\theta^0} \cdot$ ,  $A_{\theta^0}$  is symmetric, negative definite  $\rightsquigarrow q = -1, \eta = 1$ .  
[GM19; CMP20]

Example 2:  $N = 2d^2$ ,  $\psi_i = E_i x$  for  $i \leq d^2$  and  $\psi_i = E_{i-d^2} x (\alpha + \|x\|)^{-(1+\tilde{q})}$ , for  $d^2 + 1 \leq i \leq 2d^2$  and some  $\tilde{q} \in (-1, 1), \alpha > 0$ . Let  $\theta^0$  s.t. for

$$b_{\theta^0}(x) = A_{\theta^0} x + B_{\theta^0} x (\alpha + \|x\|)^{-(1+\tilde{q})},$$

$A_{\theta^0}$  is singular and negative semi-definite and  $B_{\theta^0}$  is symmetric negative definite  $\rightsquigarrow q = \tilde{q}, \eta = 1$

## Theorem [GM19; CMP20; AWST22]

Suppose  $\|\theta^0\|_0 \leq s_0$  and let  $\|\theta\|_{L^2}^2 := \theta^\top \bar{\Psi}_T \theta$ . **Suppose** that for all  $\varepsilon_0 \in (0, 1)$  and  $\forall T \geq T_0(\varepsilon_0, s, c_0, c)$  the **restricted eigenvalue property**

$$\mathbb{P} \left( \inf_{\theta \in \mathcal{S}_1(s), \theta' \in \mathcal{S}_2(s, c_0, \theta)} \frac{\|\theta - \theta'\|_{L^2}^2}{\|\theta - \theta'\|^2} \geq \frac{\lambda_{\min}(\mathbb{E}[\bar{\Psi}_T])}{2} \right) \geq 1 - \varepsilon_0,$$

holds, where  $\mathcal{S}_2(s, c_0, \theta) = \{\theta' \in \mathbb{R}^N : \|\theta - \theta'\|_1 \leq (1 + c_0) \|(\theta - \theta')\|_{\mathcal{J}_s(\theta - \theta')}\|_1\}$ ,  $\mathcal{S}_1(s) = \{\theta \in \mathbb{R}^N : \|\theta\|_0 = s\}$ , and for some  $\rho(q, \eta) > 0$ ,

$$T_0(\varepsilon_0, s, c_0, c) := \left\{ \log \left( 21^{2s} \left( d \wedge \left( \frac{ed}{2s} \right)^{2s} \right) - \log \varepsilon_0 \right)^{\rho(q, \eta)} \cdot \frac{18^2 (c_0 + 2)^2 e^2 c^2}{\lambda_{\min}(\mathbb{E}[\bar{\Psi}_T])^2} \right\}.$$

Fix  $\gamma > 0$  and  $\varepsilon_0 \in (0, 1)$ . Then, for

$$\lambda \geq 2 \sqrt{\frac{(2\mathfrak{D}_\infty + \lambda_{\min}(\mathbb{E}[\bar{\Psi}_T]))}{T} \cdot \log \left( \frac{6N}{\varepsilon_0} \right)} \quad \text{and} \quad T \geq T_0 \left( \frac{\varepsilon_0}{3}, s_0, 1 + \frac{2}{\gamma}, c \right),$$

with probability at least  $1 - \varepsilon_0$ , we have

$$\|\hat{\theta}_T - \theta_0\|_{L^2}^2 \leq (1 + \gamma) \inf_{\theta \in \mathbb{R}^N : \|\theta\|_0 \leq s_0} \left\{ \|\theta - \theta_0\|_{L^2}^2 + \frac{4(2 + \gamma)^2}{\gamma(1 + \gamma)\lambda_{\min}(\mathbb{E}[\bar{\Psi}_T])} \|\theta\|_0 \lambda^2 \right\}.$$

### Proposition [AWST22]

The restricted eigenvalue property holds for  $\rho(q, \eta) = \frac{6\eta+2q+3-q_+}{1-q_+}$ , i.e., for any  $\varepsilon_0 \in (0, 1)$  and

$$T \geq \left\{ \log \left( 21^{2s} \left( d \wedge \left( \frac{ed}{2s} \right)^{2s} \right) \right) - \log \varepsilon_0 \right\}^{\frac{6\eta+2q+3-q_+}{1-q_+}} \cdot \frac{18^2 (c_0 + 2)^2 e^2 c^2}{\lambda_{\min}(\mathbb{E}[\bar{\Psi}_T])^2},$$

we have

$$\mathbb{P} \left( \inf_{\theta \in \mathcal{S}_1(s), \theta' \in \mathcal{S}_2(s, \theta)} \frac{(\theta - \theta')^\top \bar{\Psi}_T (\theta - \theta')}{\|\theta - \theta'\|^2} \geq \frac{\lambda_{\min}(\mathbb{E}[\bar{\Psi}_T])}{2} \right) \geq 1 - \varepsilon_0,$$

- Langevin diffusion

$$dX_t = -\nabla U(X_t) dt + \sqrt{2} dW_t,$$

has invariant density  $\pi(x) \propto \exp(-U(x)) \rightsquigarrow$  sampling from  $\pi$  by numerical approximation of  $X$ , e.g., Euler scheme

$$\vartheta_{n+1}^{(\Delta)} = \vartheta_n^{(\Delta)} - \Delta \nabla U(\vartheta_n^{(\Delta)}) + \sqrt{2\Delta} \xi_{n+1}, \quad \vartheta_0^{(\Delta)} \sim X_0, \quad (\xi_n) \underset{\text{iid}}{\sim} \mathcal{N}(0, 1)$$

- abundant literature on sampling precision in TV or Wasserstein distance for  $U$  **strongly convex** or modifications thereof [Dal17; DK19; DM17; DMM19]  $\rightsquigarrow \pi(x) dx$  **sub-Gaussian**
- Assume instead that for some  $q \in (0, 1)$

$$\langle \nabla U(x), x/\|x\| \rangle \geq \tau \|x\|^{-q}, \quad \|x\| > A.$$

$\rightsquigarrow \exists \lambda > 0 : \int_{\mathbb{R}^d} \exp(\lambda \|x\|^{\tilde{q}}) \pi(x) dx < \infty \iff \tilde{q} \leq 1 - q$

$\rightsquigarrow$  prototypical example:  $\pi(x) \propto \exp(-\beta \|x\|^{1-q})$  outside some ball around the origin