Concentration analysis of multivariate elliptic diffusions

Mathematisches Seminar – Christian-Albrechts-Universität zu Kiel

Lukas Trottner
joint work with Cathrine Aeckerle-Willems and Claudia Strauch

10 November 2022

Aarhus University University of Mannheim

AARHUS UNIVERSITY

Some known concentration results for Markov processes I

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

$$\mathbb{C}_\nu(f, T, x) \coloneqq \mathbb{P}^\nu\Big(\Big| \frac{1}{T} \int_0^T f(X_t) \, \mathrm{d}t - \mu(f) \Big| > x \Big), \quad f \in \mathbb{L}^2(\mu), x, T > 0.$$

Some known concentration results for Markov processes I

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

$$\mathbb{C}_\nu(f, T, x) \coloneqq \mathbb{P}^\nu\left(\left|\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t - \mu(f)\right| > x\right), \quad f \in \mathbb{L}^2(\mu), x, T > 0.$$

Bounds have been mostly studied with two approaches (*Lyapunov vs. Poincaré* [BCG08]):

1. Functional inequalities:
   - Poincaré inequality (PI):

     $$\mathrm{Var}_\mu(g) \coloneqq \mu(g^2) - \mu(g)^2 \leqslant -C_{\mathrm{P}}\langle Lg, g\rangle_\mu, \quad g \in D(L).$$

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

Some known concentration results for Markov processes I

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

$$\mathbb{C}_\nu(f,T,x) \coloneqq \mathbb{P}^\nu\Big(\Big|\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t - \mu(f)\Big| > x\Big), \quad f\in\mathbb{L}^2(\mu), x, T>0.$$

Bounds have been mostly studied with two approaches (*Lyapunov vs. Poincaré* [BCG08]):

1. Functional inequalities:
   - Poincaré inequality (PI):
   $$\mathrm{Var}_\mu(g) \coloneqq \mu(g^2) - \mu(g)^2 \leqslant -C_{\mathrm{P}}\langle Lg, g\rangle_\mu, \quad g\in D(L).$$
   Implies: $\|P_t f - \mu(f)\|_{\mathbb{L}^2(\mu)} \leqslant \mathrm{e}^{-2t/C_P}\|f-\mu(f)\|_{\mathbb{L}^2(\mu)}$
   - log-Sobolev inequality (LS): $(P_t)_{t\geqslant 0}$ symmetric and
   $$\mathrm{Ent}_\mu(g^2) \coloneqq \mu(g^2\log g^2) - \mu(g^2)\log\mu(g^2) \leqslant 2C_{\mathrm{LS}}\|(-L)^{1/2}g\|_\mu^2, \quad g\in D((-L)^{1/2}),$$

Some known concentration results for Markov processes I

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

$$\mathbb{C}_\nu(f, T, x) \coloneqq \mathbb{P}^\nu\left(\left|\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t - \mu(f)\right| > x\right), \quad f \in \mathbb{L}^2(\mu), x, T > 0.$$

Bounds have been mostly studied with two approaches (*Lyapunov vs. Poincaré* [BCG08]):

1. Functional inequalities:
   - Poincaré inequality (PI):
   $$\mathrm{Var}_\mu(g) \coloneqq \mu(g^2) - \mu(g)^2 \leqslant -C_{\mathrm{P}}\langle Lg, g\rangle_\mu, \quad g \in D(L).$$
   Implies: $\|P_t f - \mu(f)\|_{\mathbb{L}^2(\mu)} \leqslant \mathrm{e}^{-2t/C_P}\|f - \mu(f)\|_{\mathbb{L}^2(\mu)}$
   - log-Sobolev inequality (LS): $(P_t)_{t\geqslant 0}$ symmetric and
   $$\mathrm{Ent}_\mu(g^2) \coloneqq \mu(g^2\log g^2) - \mu(g^2)\log\mu(g^2) \leqslant 2C_{\mathrm{LS}}\|(-L)^{1/2}g\|_\mu^2, \quad g \in D((-L)^{1/2}),$$
2. Mixing assumptions:
$$(\alpha(\nu,\varphi)): \quad \alpha_\nu(t) \coloneqq \sup_{s\geqslant 0}\ \sup_{A\in\sigma(X_u, u\leqslant s), B\in\sigma(X_u, u\geqslant s+t)} |\mathbb{P}^\nu(A\cap B) - \mathbb{P}^\nu(A)\mathbb{P}^\nu(B)| \leqslant \varphi(t) \underset{t\to\infty}{\longrightarrow} 0.$$
   For reasonable ν implied by ergodicity of P_t, i.e., $\|P_t(x,\cdot) - \mu\|_{\mathrm{TV}} \leqslant CV(x)\varphi(t)$

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U\colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying
$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$
on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U\colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$

on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.
   - invariant distribution: $\mu = |\mathcal{O}|^{-1}\mathrm{Leb}$, since $\int_{\mathcal{O}} Lf(x)\,\mathrm{d}x = 0$ for all $f \in H^1(\mathcal{O})$ s.t. $\partial_n f = 0$ on $\partial\mathcal{O}$

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U \colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$

on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.

- invariant distribution: $\mu = |\mathcal{O}|^{-1}\mathrm{Leb}$, since $\int_{\mathcal{O}} Lf(x)\,\mathrm{d}x = 0$ for all $f \in H^1(\mathcal{O})$ s.t. $\partial_n f = 0$ on $\partial\mathcal{O}$
- Poincaré inequality follows from analytical Poincaré inequality for Sobolev functions:

$$\int_{\mathcal{O}} f(x)(-Lf(x))\,\mathrm{d}x = \int_{\mathcal{O}} U(x)|\nabla f(x)|^2\,\mathrm{d}x \geqslant u_{\min}\|\nabla f\|_{\mathbb{L}^2(\mathcal{O})} \geqslant v_{\min} C_{\mathrm{P}}^{-1}\|f - \mu(f)\|_{\mathbb{L}^2(\mathcal{O})}.$$

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U\colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$

on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.

- invariant distribution: $\mu = |\mathcal{O}|^{-1}\mathrm{Leb}$, since $\int_{\mathcal{O}} Lf(x)\,\mathrm{d}x = 0$ for all $f \in H^1(\mathcal{O})$ s.t. $\partial_n f = 0$ on $\partial\mathcal{O}$
- Poincaré inequality follows from analytical Poincaré inequality for Sobolev functions:

$$\int_{\mathcal{O}} f(x)(-Lf(x))\,\mathrm{d}x = \int_{\mathcal{O}} U(x)|\nabla f(x)|^2\,\mathrm{d}x \geqslant u_{\min}\|\nabla f\|_{\mathbb{L}^2(\mathcal{O})} \geqslant v_{\min} C_{\mathrm{P}}^{-1}\|f - \mu(f)\|_{\mathbb{L}^2(\mathcal{O})}.$$

2. Let $U\colon \mathbb{R}^d \to \mathbb{R}$ be strongly convex with Lipschitz gradient ∇U (e.g., $U(x) = x^\top \Sigma x$, Σ p.d.). Let X be the strong solution to the Langevin SDE

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t.$$

Then $L = (-\nabla U)^\top \nabla + \Delta$

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U\colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$

on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.

- invariant distribution: $\mu = |\mathcal{O}|^{-1}\mathrm{Leb}$, since $\int_{\mathcal{O}} Lf(x)\,\mathrm{d}x = 0$ for all $f \in H^1(\mathcal{O})$ s.t. $\partial_n f = 0$ on $\partial\mathcal{O}$
- Poincaré inequality follows from analytical Poincaré inequality for Sobolev functions:

$$\int_{\mathcal{O}} f(x)(-Lf(x))\,\mathrm{d}x = \int_{\mathcal{O}} U(x)|\nabla f(x)|^2\,\mathrm{d}x \geqslant u_{\min}\|\nabla f\|_{\mathbb{L}^2(\mathcal{O})} \geqslant v_{\min} C_{\mathrm{P}}^{-1}\|f - \mu(f)\|_{\mathbb{L}^2(\mathcal{O})}.$$

2. Let $U\colon \mathbb{R}^d \to \mathbb{R}$ be strongly convex with Lipschitz gradient ∇U (e.g., $U(x) = x^\top \Sigma x$, Σ p.d.). Let X be the strong solution to the Langevin SDE

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t.$$

Then $L = (-\nabla U)^\top \nabla + \Delta$

- invariant distribution: $\mu(\mathrm{d}x) = \exp(-U(x))\,\mathrm{d}x$ (log-concave), since

$$\int \Delta f(x) \mathrm{e}^{-U(x)}\,\mathrm{d}x = \int (\nabla U(x))^\top \nabla f(x) \mathrm{e}^{-U(x)}\,\mathrm{d}x.$$

Examples

1. Let $\mathcal{O} \subset \mathbb{R}^d$ be an open, bounded and convex domain and $U\colon \mathcal{O} \to [u_{\min}, \infty) \subset (0, \infty)$ a smooth function. Let X be the reflected diffusion on $\mathcal{O}$ obeying

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2U(X_t)}\,\mathrm{d}W_t,$$

on the interior. Then $L = \nabla \cdot U\nabla$ subject to Neumann boundary conditions.

- invariant distribution: $\mu = |\mathcal{O}|^{-1}\mathrm{Leb}$, since $\int_{\mathcal{O}} Lf(x)\,\mathrm{d}x = 0$ for all $f \in H^1(\mathcal{O})$ s.t. $\partial_n f = 0$ on $\partial\mathcal{O}$
- Poincaré inequality follows from analytical Poincaré inequality for Sobolev functions:

$$\int_{\mathcal{O}} f(x)(-Lf(x))\,\mathrm{d}x = \int_{\mathcal{O}} U(x)|\nabla f(x)|^2\,\mathrm{d}x \geqslant u_{\min}\|\nabla f\|_{\mathbb{L}^2(\mathcal{O})} \geqslant v_{\min} C_{\mathrm{P}}^{-1}\|f - \mu(f)\|_{\mathbb{L}^2(\mathcal{O})}.$$

2. Let $U\colon \mathbb{R}^d \to \mathbb{R}$ be strongly convex with Lipschitz gradient ∇U (e.g., $U(x) = x^\top \Sigma x$, Σ p.d.). Let X be the strong solution to the Langevin SDE

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t.$$

Then $L = (-\nabla U)^\top \nabla + \Delta$

- invariant distribution: $\mu(\mathrm{d}x) = \exp(-U(x))\,\mathrm{d}x$ (log-concave), since

$$\int \Delta f(x) \mathrm{e}^{-U(x)}\,\mathrm{d}x = \int (\nabla U(x))^\top \nabla f(x) \mathrm{e}^{-U(x)}\,\mathrm{d}x.$$

- X satisfies a log-Sobolev inequality

Some known concentration results for Markov processes II

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

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

where $\sigma^2(f) = \lim_{t\to\infty} t^{-1}\mathrm{Var}_{\mathbb{P}^\mu}\left(\int_0^t f(X_s)\,\mathrm{d}s\right)$

Some known concentration results for Markov processes II

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

$$\mathbb{C}_\nu(f, T, x) \leqslant 2\Big\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\Big\|_{\mathbb{L}^2(\mu)} \exp\Big(-\frac{Tx^2}{2(\sigma^2(f) + 2C_{\mathrm{P}}\|f\|_\infty x)}\Big),$$

where $\sigma^2(f) = \lim_{t\to\infty} t^{-1}\mathrm{Var}_{\mathbb{P}^\mu}(\int_0^t f(X_s)\,\mathrm{d}s)$

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

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

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

Some known concentration results for Markov processes II

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

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

where $\sigma^2(f) = \lim_{t\to\infty} t^{-1}\mathrm{Var}_{\mathbb{P}^\mu}(\int_0^t f(X_s)\,\mathrm{d}s)$

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

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

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

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

Martingale approximation for diffusions

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

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t,$$

  $b \in \mathrm{Lip}_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R}^d)$, $\sigma \in \mathrm{Lip}(\mathbb{R}^d;\mathbb{R}^{d\times d})$ and bounded, $a \coloneqq \sigma\sigma^\top$ s.t. $\lambda_- \mathbb{I} \leqslant a(x) \leqslant \lambda_+ \mathbb{I}$, $\forall x$
- Let $L = b^\top \nabla + \sum_{i,j} a_{i,j} \partial_{x_i} \partial_{x_j}$ and suppose that for given $f\colon \mathbb{R}^d \to \mathbb{R}$ the Poisson equation $Lg = f$ has some sufficiently regular solution $L^{-1}[f]$
- By Itō's formula: $L^{-1}[f](X_t) - L^{-1}[f](X_0) = \int_0^t LL^{-1}[f](X_s)\,\mathrm{d}s + \int_0^t (\nabla L^{-1}[f](X_s))^\top \sigma(X_s)\,\mathrm{d}W_s$ and hence

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

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

Statistics based on martingale approximation

[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 $(P_t)_{t \geqslant 0}$
- at most linear drift

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

[NR20] multivariate case $d \geqslant 1$: $\sigma = \mathbb{I}$ 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

Statistics based on martingale approximation

[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 $(P_t)_{t \geqslant 0}$
- at most linear drift

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

[NR20] multivariate case $d \geqslant 1$: $\sigma = \mathbb{I}$ 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 $q \in (-1, 1)$, $\mathfrak{r}, A > 0$,

$$\langle b(x), x/\|x\| \rangle \leqslant -\mathfrak{r}\|x\|^{-q}, \quad \|x\| > A. \qquad (\mathcal{D}(q))$$

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

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

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

$$\langle b(x), x/\|x\| \rangle \leqslant -\mathfrak{r}\|x\|^{-q}, \quad \|x\| > A. \tag{$\mathcal{D}(q)$}$$

[PV01; BRS18] If $\mu(f) = 0$ and $|f(x)| \lesssim 1 + \|x\|^\eta$, then for $L^{-1}[f](x) := -\int_0^\infty P_t f(x)\,\mathrm{d}t$ we have $L^{-1}[f] \in \mathcal{W}^{2,p}_{\mathrm{loc}}(\mathbb{R}^d)$ 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 $\|\nu\|_f := \sup_{|g| \leqslant f} |\nu(g)|$ for some $f \geqslant 1$ and (Ψ_1, Ψ_2) be either pairs of inverse Young functions (i.e., $xy \leqslant \Psi_1^{-1}(x) + \Psi_2^{-1}(y)$) or $(\mathrm{Id}, 1)$ or $(1, \mathrm{Id})$.

Proposition [DFG09; AWST22]

Given ($\mathcal{D}(q)$) we have

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

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

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

Continuous-time concentration result

Theorem [AWST22]

Assume ($\mathcal{D}(q)$), $\|b(x)\| \lesssim 1 + \|x\|^{\kappa}$ and $|f(x)| \leqslant \mathfrak{L}(1 + \|x\|^{\eta})$. Let

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

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

$$\mathbb{C}_{\mu}(f, T, x) \coloneqq \mathbb{P}^{\mu}\left(\left|\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t - \mu(f)\right| > x\right) \leqslant \exp\left(-\mathfrak{c}\left(\frac{x\sqrt{T}}{\mathfrak{L}}\right)^{1/\rho(\eta,\kappa,q)}\right).$$

Continuous-time concentration result

Theorem [AWST22]

Assume $(\mathcal{D}(q))$, $\|b(x)\| \lesssim 1 + \|x\|^\kappa$ and $|f(x)| \leqslant \mathfrak{L}(1 + \|x\|^\eta)$. Let

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

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

$$\mathbb{C}_\mu(f, T, x) \coloneqq \mathbb{P}^\mu\left(\left|\frac{1}{T}\int_0^T f(X_t)\,\mathrm{d}t - \mu(f)\right| > x\right) \leqslant \exp\left(-\mathfrak{c}\left(\frac{x\sqrt{T}}{\mathfrak{L}}\right)^{1/\rho(\eta,\kappa,q)}\right).$$

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

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

Let observations $(X_{k\Delta})_{k=1,\dots,n}$ be given for some $\Delta \leqslant 1$. Define $\mathbb{H}_{n,\Delta}(f) \coloneqq \frac{1}{\sqrt{n\Delta}} \mathbb{G}_{n,\Delta}(f)$, where

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

Then for $\mathbb{G}_t(f) \coloneqq t^{-1/2} \int_0^T f(X_t)\,\mathrm{d}t$, $f = \widetilde{f} - \mu(\widetilde{f})$, $\Phi_k(t) \coloneqq \int_t^{k\Delta} (L\widetilde{f}(X_s) - \mu(L\widetilde{f}))\,\mathrm{d}s$, $\omega_k(t) \coloneqq \int_t^{k\Delta} \nabla \widetilde{f}(X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$,

$$\sqrt{n\Delta}(\mathbb{G}_{n,\Delta}(f) - \mathbb{G}_{n\Delta}(f)) = \mu(L\widetilde{f})\frac{n\Delta^2}{2} + \sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \Phi_k(t)\,\mathrm{d}t + \sum_{k=1}^{n} \int_{(k-1)\Delta}^{k\Delta} \omega_k(t)\,\mathrm{d}t.$$

Discrete-time concentration result

Let observations $(X_{k\Delta})_{k=1,\dots,n}$ be given for some $\Delta \leqslant 1$. Define $\mathbb{H}_{n,\Delta}(f) \coloneqq \frac{1}{\sqrt{n\Delta}}\mathbb{G}_{n,\Delta}(f)$, where

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

Then for $\mathbb{G}_t(f) \coloneqq t^{-1/2}\int_0^T f(X_t)\,\mathrm{d}t$, $f = \widetilde{f} - \mu(\widetilde{f})$, $\Phi_k(t) \coloneqq \int_t^{k\Delta}(L\widetilde{f}(X_s) - \mu(L\widetilde{f}))\,\mathrm{d}s$, $\omega_k(t) \coloneqq \int_t^{k\Delta} \nabla\widetilde{f}(X_s)^\top \sigma(X_s)\,\mathrm{d}W_s$,

$$\sqrt{n\Delta}(\mathbb{G}_{n,\Delta}(f) - \mathbb{G}_{n\Delta}(f)) = \mu(L\widetilde{f})\frac{n\Delta^2}{2} + \sum_{k=1}^{n}\int_{(k-1)\Delta}^{k\Delta} \Phi_k(t)\,\mathrm{d}t + \sum_{k=1}^{n}\int_{(k-1)\Delta}^{k\Delta} \omega_k(t)\,\mathrm{d}t.$$

Theorem [AWST22]

Assume $(\mathcal{D}(q))$, $\|b(x)\| \lesssim 1 + \|x\|^\kappa$, $f \in C^2(\mathbb{R}^d;\mathbb{R})$ 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 $\widetilde{\gamma} > 1 + q$, $r > 1$, s.t. $\widetilde{\gamma} - (1+q) > r(\alpha \vee (1+q)/(r-1))$. Then, for $p \geqslant 2$,

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

and

$$\mathbb{P}^\mu\Big(|\mathbb{H}_{n,\Delta}(f) - \mu(f)| > \mathrm{e}(n\Delta)^{-1/2}\Phi(n,\Delta,x)\Big) \leqslant \mathrm{e}^{-x}, \quad x \geqslant 2.$$

Applications

Lasso for parametrized drifts

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

$$\mathrm{d}X_t = b_{\theta^0}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t, \quad \text{where} \quad b_{\theta^0}(x) = \sum_{i=1}^{N} \theta_i^0 \psi_i(x).$$

Let $\boldsymbol{\psi}(x) = (\psi_1(x), \ldots, \psi_N(x))$, $\boldsymbol{\Psi}(x) \coloneqq (\sigma^{-1}(x)\boldsymbol{\psi}(\boldsymbol{x}))^\top \sigma^{-1}(x)\boldsymbol{\psi}(\boldsymbol{x})$ and $\overline{\boldsymbol{\Psi}}_T \coloneqq T^{-1}\int_0^T \boldsymbol{\Psi}(X_t)\,\mathrm{d}t$.
Then for $b_\theta \coloneqq \boldsymbol{\psi}\theta$, negative log-likelihood given by

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

Goal

Study convergence guarantees of Lasso estimator

$$\widehat{\theta}_T \coloneqq \operatorname*{arg\,min}_{\theta \in \mathbb{R}^N} \left\{ \mathcal{L}_T(\theta) + \lambda \|\theta\|_1 \right\},$$

under sparsity assumptions on θ^0, i.e., $\|\theta^0\|_0 \leqslant s_0$.

Assumptions and examples

We assume

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

Example 1: [GM19; CMP20] 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, θ^0 s.t. for $b_{\theta^0}(\cdot) = A_{\theta^0}\cdot$, A_{θ^0} is symmetric, negative definite $\rightsquigarrow q = -1, \eta = 1$.

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

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

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

Oracle inequality

Theorem [GM19; CMP20; AWST22]

Suppose $\|\theta^0\|_0 \leqslant s_0$ and let $\|\theta\|_{L^2}^2 := \theta^\top \overline{\mathbf{\Psi}}_T \theta = T^{-1}\int_0^T \|\sigma^{-1}(X_t) b_\theta(X_t)\|^2 \, dt$. Suppose that for all $\varepsilon_0 \in (0,1)$ and $\forall T \geqslant T_0(\varepsilon_0, s, c_0, c)$ the restricted eigenvalue property

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

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

$$T_0(\varepsilon_0, s, c_0, c) := \left\{\log\left(21^{2s}\left(d \wedge \left(\frac{ed}{2s}\right)^{2s}\right)\right) - \log \varepsilon_0\right\}^{\rho(q,\eta)} \cdot \frac{18^2\,(c_0 + 2)^2\, e^2 c^2}{\lambda_{\min}(\mathbb{E}[\overline{\mathbf{\Psi}}_1])^2}.$$

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

$$\lambda \geqslant 2\sqrt{\frac{(2\max_{i=1,\dots,N} \mathbb{E}[\overline{\mathbf{\Psi}}_1]_{i,i} + \lambda_{\min}(\mathbb{E}[\overline{\mathbf{\Psi}}_1]))}{T} \cdot \log\left(\frac{6N}{\varepsilon_0}\right)} \quad \text{and} \quad T \geqslant T_0\left(\frac{\varepsilon_0}{3}, s_0, 1 + \frac{2}{\gamma}, c\right),$$

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

$$\|\widehat{\theta}_T - \theta^0\|_{L^2}^2 \leqslant (1 + \gamma) \inf_{\theta \in \mathbb{R}^N : \|\theta\|_0 \leqslant s_0} \left\{ \|\theta - \theta^0\|_{L^2}^2 + \frac{4(2 + \gamma)^2}{\gamma(1 + \gamma)\lambda_{\min}(\mathbb{E}[\overline{\mathbf{\Psi}}_1])} \|\theta\|_0 \lambda^2 \right\}.$$

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

$$T \geqslant \left\{\log\left(21^{2s}\left(d \wedge \left(\tfrac{ed}{2s}\right)^{2s}\right)\right) - \log \varepsilon_0\right\}^{\frac{6\eta+2q+3-q_+}{1-q_+}} \cdot \frac{18^2\,(c_0+2)^2\,e^2 c^2}{\lambda_{\min}(\mathbb{E}[\overline{\boldsymbol{\Psi}}_1])^2},$$

we have

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

MCMC for moderately heavy tailed targets

- Langevin diffusion

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t,$$

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

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

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

$$\langle \nabla U(x), x/\|x\| \rangle \geqslant \mathfrak{r}\|x\|^{-q}, \quad \|x\| > A. \qquad (\mathcal{U}(q))$$

$\rightsquigarrow \exists \lambda > 0 : \int_{\mathbb{R}^d} \exp\left(\lambda \|x\|^{\widetilde{q}}\right)\pi(x)\,\mathrm{d}x < \infty \iff \widetilde{q} \leqslant 1 - q$
$\rightsquigarrow$ prototypical example: $\pi(x) \propto \exp(-\beta\|x\|^{1-q})$ outside some ball around the origin

Convergence guarantees

Proposition [AWST22]

Assume ($\mathcal{U}(q)$) and that ∇U is bounded. Let $f \in C^2(\mathbb{R}^d)$ s.t. $\|D^k f(x)\| \lesssim 1 + \|x\|^{\eta_k}$, $k = 0, 1, 2$, and consider the burn-in estimator

$$\mathbb{H}_{n,m,\Delta}(f) := \mathbb{H}_{n,\Delta}(f) \circ \theta_m = \frac{1}{n} \sum_{k=m+1}^{n+m} f(X_{k\Delta}).$$

Then we have the following approximation guarantees for $\int f(x)\pi(x)\,dx$:

	step length Δ	sample size n	burn-in m
ε-prec. sampling	$\frac{\varepsilon^2}{d(\log(\mathfrak{C}/\varepsilon))^{(1-q)/(1+q)}}$	$\frac{d(\log(\mathfrak{C}/\varepsilon))^{2(1-q)/(1+q)}}{\varepsilon^2}$	—
(ε, δ)-PAC bound	$\frac{(\delta\varepsilon)^2}{d(\log(1/\delta))^{2(\eta_0+(q+3)/2)/(1-q)}}$	$\frac{d\mathfrak{D}^2(\log(1/\delta))^{(4(\eta_0+(q+3)/2))/(1-q)}}{\delta^2\varepsilon^4}$	$\frac{d(\log(1/\delta))^{2(\eta_0+q+2)/(1-q)}}{(\delta\varepsilon)^2}$

Table 2: Order of sufficient sampling frequency Δ, sample size n and burn-in m for (ε, δ)-PAC bounds and sampling within ε-TV margin

Summary

- we provide concentration inequalities for subexponentially ergodic diffusions and polynomially bounded functions given continuous observations
- Concentration inequalities for sampled chains are derived from the continuous observation result
- we demonstrate implications on sufficient sample sizes for parametric high-dimensional drift estimaton and MCMC for moderately heavy tailed targets

Summary

- we provide concentration inequalities for subexponentially ergodic diffusions and polynomially bounded functions given continuous observations
- Concentration inequalities for sampled chains are derived from the continuous observation result
- we demonstrate implications on sufficient sample sizes for parametric high-dimensional drift estimaton and MCMC for moderately heavy tailed targets

Thank you for your attention!

C. Aeckerle-Willems and C. Strauch. "Concentration of scalar ergodic diffusions and some statistical implications". In: *Ann. Inst. Henri Poincaré Probab. Stat.* (2021), to appear.

C. Aeckerle-Willems, C. Strauch, and L. Trottner. *Concentration analysis of multivariate elliptic diffusion processes*. 2022.

D. Bakry, P. Cattiaux, and A. Guillin. "Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré". In: *J. Funct. Anal.* 254.3 (2008), pp. 727–759. ISSN: 0022-1236.

V. I. Bogachev, M. Röckner, and S. V. Shaposhnikov. "The Poisson equation and estimates for distances between stationary distributions of diffusions". In: *J. Math. Sci. (N.Y.)* 232.3, Problems in mathematical analysis. No. 92 (Russian) (2018), pp. 254–282. ISSN: 1072-3374.

P. Cattiaux and A. Guillin. "Deviation bounds for additive functionals of Markov processes". In: *ESAIM Probab. Stat.* 12 (2008), pp. 12–29. ISSN: 1292-8100.

G. Ciołek, D. Marushkevych, and M. Podolskij. "On Dantzig and Lasso estimators of the drift in a high dimensional Ornstein-Uhlenbeck model". In: *Electron. J. Stat.* 14.2 (2020), pp. 4395–4420.

A. S. Dalalyan. "Theoretical guarantees for approximate sampling from smooth and log-concave densities". In: *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 79.3 (2017), pp. 651–676. ISSN: 1369-7412.

R. Douc, G. Fort, and A. Guillin. "Subgeometric rates of convergence of f-ergodic strong Markov processes". In: *Stochastic Process. Appl.* 119.3 (2009), pp. 897–923. ISSN: 0304-4149.

A. S. Dalalyan and A. Karagulyan. "User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient". In: *Stochastic Process. Appl.* 129.12 (2019), pp. 5278–5311. ISSN: 0304-4149.

A. Durmus and E. Moulines. "Nonasymptotic convergence analysis for the unadjusted Langevin algorithm". In: *Ann. Appl. Probab.* 27.3 (2017), pp. 1551–1587. ISSN: 1050-5164.

A. Durmus, S. Majewski, and B. a. Miasojedow. "Analysis of Langevin Monte Carlo via convex optimization". In: *J. Mach. Learn. Res.* 20 (2019), Paper No. 73, 46. ISSN: 1532-4435.

F. Gao, A. Guillin, and L. Wu. "Bernstein-type concentration inequalities for symmetric Markov processes". In: *Theory Probab. Appl.* 58.3 (2014), pp. 358–382. ISSN: 0040-585X.

S. Gaïffas and G. Matulewicz. “Sparse inference of the drift of a high-dimensional Ornstein-Uhlenbeck process”. In: *J. Multivariate Anal.* 169 (2019), pp. 1–20. ISSN: 0047-259X.

L. Galtchouk and S. Pergamenshchikov. “Uniform concentration inequality for ergodic diffusion processes”. In: *Stochastic Process. Appl.* 117.7 (2007), pp. 830–839. ISSN: 0304-4149.

P. Lezaud. “Chernoff and Berry-Esséen inequalities for Markov processes”. In: *ESAIM Probab. Statist.* 5 (2001), pp. 183–201. ISSN: 1292-8100.

R. Nickl and K. Ray. “Nonparametric statistical inference for drift vector fields of multi-dimensional diffusions”. In: *Ann. Statist.* 48.3 (2020), pp. 1383–1408. ISSN: 0090-5364.

E. Pardoux and A. Y. Veretennikov. “On the Poisson equation and diffusion approximation. I”. In: *Ann. Probab.* 29.3 (2001), pp. 1061–1085. ISSN: 0091-1798.