Nonparametric estimation of McKean-Vlasov SDEs via deconvolution

Vytautė Pilipauskaitė (Aalborg University, Denmark)

joint work with
Chiara Amorino, Mark Podolskij, Shi-Yuan Zhou
(University of Luxembourg),
Denis Belomestny (University of Duisburg-Essen, Germany)

2024-02-12

Plan of this talk

1. Interacting particle system associated to McKean-Vlasov SDE
2. Assumptions and probabilistic properties
3. Estimation procedure
4. Convergence rates
5. On Fourier transforms

1. Introduction

Consider an interacting particle system: an $\mathbb{R}^N$-valued stochastic process

$$(X_t^1, \dots, X_t^N)_{t \geq 0}$$

defined by

$$dX_t^i = -V'(X_t^i)dt - \frac{1}{N}\sum_{j=1}^{N} W'(X_t^i - X_t^j)dt + dB_t^i, \quad i = 1, \dots, N, \quad t \geq 0,$$

where

- $X_0^1, \dots, X_0^N$ are i.i.d. random variables with a common law $\bar{\Pi}_0$,
- $(B_t^1)_{t \geq 0}, \dots, (B_t^N)_{t \geq 0}$ are independent standard Brownian motions, independent of $X_0^1, \dots, X_0^N$,
- $V, W : \mathbb{R} \to \mathbb{R}$ are called respectively confinement, interaction potentials.

Our objective is to estimate W' from

$$X_T^1, \dots, X_T^N$$

when $N, T \to \infty$.

Let $(\bar{X}_t)_{t\geq 0}$ be the solution of the McKean-Vlasov SDE:

$$d\bar{X}_t = -V'(\bar{X}_t)dt - W' * \bar{\Pi}_t(\bar{X}_t)dt + dB_t, \quad t \geq 0,$$

where

- $\bar{\Pi}_t$ denotes the law of $\bar{X}_t$,
- $*$ denotes convolution:

$$W' * \bar{\Pi}_t(x) = \int_{\mathbb{R}} W'(x-y)\bar{\Pi}_t(dy), \quad x \in \mathbb{R},$$

- $(B_t)_{t\geq 0}$ is a standard Brownian motion, independent of $\bar{X}_0$.

Nonparametric estimation:

- Hoffmann, Liu 2023+, Belomestny, P., Podolskij 2023, Della Maestra, Hoffmann 2022, Lu, Maggioni, Tang 2021, Li, Lu, Maggioni, Tang, Zhang 2021

Recall the McKean-Vlasov SDE

$$d\bar{X}_t = -V'(\bar{X}_t)dt - W' * \bar{\Pi}_t(\bar{X}_t)dt + dB_t, \quad t \geq 0.$$

Assume that

$$V(x) = 0, \quad W(x) = \sum_{j=1}^{J} \alpha_j x^{2j} + \beta(x), \quad x \in \mathbb{R},$$

for some known $J \in \mathbb{N}$ and unknown $\alpha_1 > 0$, $\alpha_j \geq 0$, $1 < j < J$, $\alpha_J > 0$ and unknown even $\beta \in C^2(\mathbb{R})$ with bounded, integrable β' such that

$$\inf_{x \in \mathbb{R}} W''(x) =: \lambda > 0.$$

Proposition (Belomestny, P., Podolskij 2023)

Assume that for some $p > 0$,

$$\limsup_{x\to\infty} x^{2p} \int_{|y|>x} |\beta'(y)|^2 dy < \infty.$$

and let β' be entire and of the first order and type less than $\vartheta > 0$, i.e. $|\beta'(z)| \le A\exp(\vartheta|z|)$, $z \in \mathbb{C}$, for some $A > 0$ and let $\vartheta \le \lambda^{\frac{1}{2}}$. Choose an arbitrary $r_T^N \to 0$ as $N, T \to \infty$ in the estimation procedure. Set

$$N_T^{-1} := N^{-1} + \exp(-\lambda T)$$

Then

$$\mathbb{E}\Big[\int_{\mathbb{R}} |(\beta')_T^N(x) - \beta'(x)|^2 dx\Big] \lesssim (r_T^N)^{-2}(\log N_T)^{-\frac{p}{J}},$$

where $\lesssim$ stands for the inequality up to a constant.

Examples:

$$\beta(x) = \frac{1-\cos(bx)}{x^2}, \quad b > 0, \qquad \beta(x) = \frac{\sin^{2k}(x)}{x^{2k}}, \quad k \in \mathbb{N}.$$

Theorem (Belomestny, P., Podolskij 2023)

Let $X_T^1, \dots, X_T^N$ be i.i.d. from the invariant law $\Pi = \Pi_\beta$ of $(\bar{X}_t)_{t \geq 0}$, $\mathcal{F}$ a specific class of functions $f : \mathbb{R} \to \mathbb{R}$ in accordance with estimation procedure and $p > \frac{1}{2}$. Then there exists $C > 0$ such that

$$\inf_{(\beta')_N} \sup_{\beta \in \mathcal{F}} \Pi_\beta^{\otimes N}(\|(\beta')_N - \beta'\|_2^2 > C(\log N)^{-\frac{p}{J}}) > 0.$$

Remark. The classical deconvolution problem:

$$Y_i = X_i + \varepsilon_i, \quad i = 1, \dots, N,$$

where $(X_i)_{i \in \mathbb{N}}$, $(\varepsilon_i)_{i \in \mathbb{N}}$ are independent sequences of i.i.d. random variables with densities f_X, f_ε respectively. If for $p > \frac{1}{2}$,

$$\int_{\mathbb{R}} |\mathcal{F} f_X(z)|^2 |z|^{2p} dz \leq C, \quad |\mathcal{F} f_\varepsilon(z)| \asymp \exp(-C|z|^{2J}),\ |z| \to \infty,$$

then the minimax rate for the estimation of f_X is $(\log N)^{-\frac{p}{J}}$.

2. Assumptions and probabilistic properties

Assume that

- $V(x) = \frac{\alpha}{2}x^2 + \tilde{V}(x)$, $x \in \mathbb{R}$, where $\alpha > 0$ is unknown, whereas $\tilde{V}$ is known and either

  A1 $\tilde{V} = 0$ or

  A2 $\tilde{V}$ is even with continuous bounded derivatives up to some integer $M \geq 2$,

  moreover, $\inf_{x\in\mathbb{R}} V''(x) =: C_V > 0$,

- $W \in C^2(\mathbb{R})$ is unknown, even with bounded, integrable W' and bounded W'' such that $\inf_{x\in\mathbb{R}} W''(x) =: -C_W < 0$ and

$$\lambda := C_V - 2C_W > 0,$$

- initial law $\bar{\Pi}_0$ admits a density π_0 and for all $a \in \mathbb{R}$,

$$\int_{\mathbb{R}} \exp(ax)\pi_0(x)dx < \infty, \quad \int_{\mathbb{R}} \log(\pi_0(x))\pi_0(x)dx < \infty.$$

Uniform in time propagation of chaos property:

Proposition

Let $p \in \mathbb{N}$. Then for all $N \in \mathbb{N}$, $i = 1, \ldots, N$ and $t \geq 0$,

$$\mathbb{E}[|X_t^i - \bar{X}_t^i|^{2p}] \leq \frac{C}{N^p},$$

where $(\bar{X}_t^i)_{t\geq 0}$ starting at $\bar{X}_0^i = X_0^i$ and driven by $(B_t^i)_{t\geq 0}$, $i = 1, \ldots, N$, are independent copies of $(\bar{X}_t)_{t\geq 0}$.

- A version of results[1,2]

[1] Cattiaux P., Guillin, A., Malrieu, F. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* 140:19–40, 2008.

[2] Malrieu, F. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* 13(2):540–560, 2003.

Combined with exponential convergence to equilibrium for $\bar{\Pi}_t$ as $t \to \infty$[3]:

Theorem

Denote by Π the invariant law of $(\bar{X}_t)_{t\geq 0}$ and

$$\Pi_T^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_T^i}.$$

Then

$$\mathbb{E}[W_1^2(\Pi_T^N, \Pi)] \leq C\Big(\frac{1}{N} + \exp(-\lambda T)\Big) =: \frac{1}{N_T},$$

where W_1 denotes the Wasserstein 1-distance, i.e.

$$W_1(\mu,\nu) := \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}[|X - Y|],$$

where the infimum is taken over all couplings (X, Y) of random variables X and Y with respective laws μ and ν.

[3] Carrillo, J.A., McCann, R.J, Villani, C. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. Rev. Mat. Iberoamericana 19(3):971-1018, 2003.

The invariant law Π admits a density

$$\pi(x) = \frac{1}{Z} \exp(-2V(x) - 2W * \pi(x))$$

with a normalizing constant

$$Z := \int_{\mathbb{R}} \exp(-2V(x) - 2W * \pi(x)) dx < \infty.$$

Lemma

Then, for any $x \in \mathbb{R}$,

$$\exp(-C_0 x^2) \lesssim \pi(x) \lesssim \exp(-C_V x^2)$$

with $C_0 := \sup_{x \in \mathbb{R}} |V''(x)|$. Moreover, let $n \geq 1$ under A1 and $1 \leq n \leq M$ under A2. Then

$$|\pi^{(n)}(x)| \lesssim (1 + |x|)^n \exp(-C_V x^2)$$

3. Estimation procedure

Step 1. We estimate

$$l(x) := (\log \pi(x))' = \frac{\pi'(x)}{\pi(x)} = -2V'(x) - 2W' * \pi(x)$$

by

$$l_T^N(x) := \frac{(\pi')_T^N(x)}{\pi_T^N(x)} \mathbf{1}(\pi_T^N(x) > \delta),$$

with

$$\pi_T^N(x) := (Nh_0)^{-1} \sum_{i=1}^{N} K(h_0^{-1}(x - X_T^i)),$$

$$(\pi')_T^N(x) := (Nh_1^2)^{-1} \sum_{i=1}^{N} K'(h_1^{-1}(x - X_T^i)),$$

where

- K is a smooth kernel of order m, i.e. $\mathfrak{m}_0 = 1$, $\mathfrak{m}_1 = 0$, ..., $\mathfrak{m}_{m-1} = 0$, $\mathfrak{m}_m \neq 0$, where $\mathfrak{m}_k = \int_{\mathbb{R}} x^k K(x) dx$, moreover, $m \geq 2$ under A1 or $2 \leq m \leq M$ under A2.
- $h_0 = h_{T,0}^N \to 0$, $h_1 = h_{T,1}^N \to 0$ and $\delta = \delta_T^N \to 0$ as $N, T \to \infty$.

Step 2. In

$$l(x) = -2\alpha x - 2\tilde{V}'(x) - 2W' * \pi(x)$$

we have that

$$W' * \pi(x) \to 0$$

as $|x| \to \infty$ since W' is integrable.

We define a minimum contrast estimator for α:

$$\alpha_T^N := \arg\min_{\alpha \in \mathbb{R}} \int_{\mathbb{R}} (l_T^N(x) - 2\alpha x - 2\tilde{V}'(x))^2 \frac{1}{U} w\Big(\frac{x}{U}\Big) dx,$$

where

- w is a weight function with support $[\epsilon, 1]$ for some $\epsilon \in (0, 1)$,
- $U = U_T^N \to \infty$ as $N, T \to \infty$.

Step 3. We estimate

$$\Psi(x) = -W' * \pi(x)$$

by

$$\Psi_T^N(x) = \frac{1}{2}(l_T^N(x) + 2\alpha_T^N x + 2\tilde{V}(x))\mathbf{1}(|x| \leq \epsilon U).$$

Step 4[4,5]. We use

$$\mathcal{F}\Psi(z) = -\mathcal{F}W'(z)\mathcal{F}\pi(z), \qquad z \in \mathbb{R} =: \mathcal{L}_0,$$

where

$$\mathcal{F}\phi(z) := \int_{\mathbb{R}} \phi(t) \exp(izt) dt.$$

Choose entire functions ρ_T^N such that for all $z \in \mathcal{L}_0$ it holds $|\mathcal{F}(\Pi_T^N)(z) + \rho_T^N(z)| \geq r_T^N > 0$, where $r_T^N \to 0$ as $N, T \to \infty$. Then, we define the estimator $(W')_T^N$ via

$$\mathcal{F}(W')_T^N(z) := -\frac{\mathcal{F}\Psi_T^N(z)}{\mathcal{F}\Pi_T^N(z) + \rho_T^N(z)}, \quad z \in \mathcal{L}_0,$$

Remark. Actually, we consider deconvolution on $\mathcal{L}_a := \{y + ia : y \in \mathbb{R}\}$ for some $a \geq 0$.

[4] Johannes, J. Deconvolution with unknown error distribution. Ann. Statist. 37(5A) 2301-2323, 2009.

[5] Belomestny, D., Goldenshluger, A. Density deconvolution under general assumptions on the distribution of measurement errors. Ann. Statist. 49(2):615-649, 2021.

4. Convergence rates

Proposition

Let

$$h_0 := N_T^{-\frac{1}{2(m+1)}}, \quad h_1 := N_T^{-\frac{1}{2(m+2)}}, \quad \delta := \frac{c_0}{2}\exp(-C_0U^2),$$

and $U \geq 1$. Then

$$\sup_{|x|\leq U} \mathbb{E}[|l_T^N(x) - l(x)|^2]^{\frac{1}{2}} \lesssim \exp(C_0U^2)(N_T^{-\frac{m}{2(m+2)}} + UN_T^{-\frac{m}{2(m+1)}}).$$

Theorem

In the setting of the above proposition it holds that

$$\mathbb{E}\Big[\int_{\mathbb{R}} |\Psi_T^N(x) - \Psi(x)|^2 dx\Big]^{\frac{1}{2}} \lesssim U^{\frac{1}{2}} \sup_{|x|\leq U} \mathbb{E}[|l_T^N(x) - l(x)|^2]^{\frac{1}{2}} + \frac{\exp(-C_V(\frac{\epsilon U}{2})^2)}{\frac{\epsilon U}{2}} + \Big(\int_{|x|>\frac{\epsilon U}{2}} |W'(x)|^2 dx\Big)^{\frac{1}{2}}.$$

Corollary

In the setting of previous theorem, moreover, assume that for some $C > C_V$,

$$\limsup_{y\to\infty} \exp(2Cy^2) \int_{|x|>y} |W'(x)|^2 dx < \infty.$$

Let $C_0 U^2 := \frac{(1-\gamma)m}{2(m+2)} \log N_T$, where $\gamma := \frac{\epsilon^2 C_V}{\epsilon^2 C_V + 4C_0}$. Then

$$\mathbb{E}\Big[\int_{\mathbb{R}} |\Psi_T^N(x) - \Psi(x)|^2 dx\Big]^{\frac{1}{2}} \lesssim (\log N_T)^{\frac{1}{4}} N_T^{-\frac{\gamma m}{2(m+2)}}.$$

Theorem

In the setting of previous corollary, moreover, assume that

$$\int_{\mathcal{L}_{\pm a}} \Big| \frac{\mathcal{F}W'(z)}{\mathcal{F}\pi(z)} \Big|^2 dz < \infty$$

and let

$$r_T^N := \exp(\frac{a\epsilon}{2} U)(\log N_T)^{\frac{1}{4}} N_T^{-\frac{\gamma m}{4(m+2)}}$$

where, recall, $C_0 U^2 = \frac{(1-\gamma)m}{2(m+2)} \log N_T$ and $\gamma = \frac{\epsilon^2 C_V}{\epsilon^2 C_V + 4C_0}$. Then

$$\mathbb{E}\Big[\int_{\mathbb{R}} |(W')_T^N(x) - W'(x)|^2 dx \Big]^{\frac{1}{2}} \lesssim r_T^N.$$

Remarks:

- polynomial convergence rate,
- $m \leq M$ under A2.

5. On Fourier transforms

Theorem

Denote the order of $\mathcal{F}\pi$ by ρ. Then $\rho \leq 2$. Moreover, denote by ρ_1 the critical exponent of the sequence of zeros $(a_j)_{j\in\mathbb{N}}$ of $\mathcal{F}\pi$. In addition, assume that either $\rho_1 < \rho$ or $\rho_1 = \rho < 2$. Then there exist $C > 0$ and a family of positive numbers $(r_j)_{j\in\mathbb{N}}$ such that, for all $z \in \mathbb{C}$ outside of $\cup_{j\geq 1} B_{r_j}(a_j)$, it holds that

$$|\mathcal{F}\pi(z)| \gtrsim \exp(-C|z|^2).$$

- The proof uses Gaussian tails of π, Hadamard's factorization theorem
- Examples of W': Hermite functions

Example. Under A2, moreover, assume that W, V are infinitely differentiable respectively on $\mathbb{R}$, $\mathbb{R}_0$, $V^{(n)}$, $2 \leq n \leq M+2$, are bounded on $\mathbb{R}_0$ and there exist

$$V^{(M+1)}(0-) \neq V^{(M+1)}(0+).$$

It provides that

$$\pi(x) = \frac{1}{Z} \exp(-2V(x) - 2W * \pi(x))$$

is infinitely differentiable on $\mathbb{R}_0$ and there exist

$$\pi^{(M+1)}(0-) \neq \pi^{(M+1)}(0+).$$

Then, iteratively integrating by parts we obtain

$$\mathcal{F}\pi(z) = \frac{1}{(-iz)^{M+1}} (I_-(z) + I_+(z)),$$

where

$$I_-(z) := \int_{-\infty}^{0} \exp(izx) \pi^{(M+1)}(x) dx, \quad I_+(z) := \int_{0}^{\infty} \exp(izx) \pi^{(M+1)}(x) dx.$$

If we integrate by parts once more, then

$$I_+(z) = \frac{1}{iz}\Big(- \pi^{(M+1)}(0+) - \int_0^{\infty} \exp(izx)\pi^{(M+2)}(x)dx\Big).$$

By Riemann's lemma, the last integral tends to zero and so

$$I_+(z) := -\frac{1}{iz}\pi^{(M+1)}(0+) + o\Big(\frac{1}{z}\Big)$$

as $z \to \infty$. Clearly an analogous reasoning applies to $I_-(z)$. It yields that for all large enough z,

$$|\mathcal{F}\pi(z)| \geq \frac{C}{|z|^{M+2}}.$$

References

- Amorino, C., Belomestny, D., Pilipauskaitė, V., Podolskij, M., Zhou, S.-Y. Polynomial rates via deconvolution for nonparametric estimation in McKean-Vlasov SDEs. Preprint arXiv:2401.04667.
- Belomestny, D., Pilipauskaitė, V., Podolskij, M. Semiparametric estimation of McKean-Vlasov SDEs. *Ann. Inst. H. Poincaré Probab. Statist.* 59(1): 79-96, 2023.

Thank you for your attention