上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

Mean field error estimate of the random batch method for large interacting particle system

Zhenyu Huang
PhD candidate, Institute of Natural Sciences, SJTU

Joint work with Shi Jin and Lei Li

Gradient flows face-to-face 5, Granada

Interacting particle system

As a class of fundamental microscopic models, interacting particle systems (or many-body systems) play important roles in the fields ranging from physics, biology to social sciences and data sciences etc.

We focus on first order system of N particles:

$$dX^{i,N} = b\left(X^{i,N}\right)dt + \frac{1}{N-1}\sum_{j:j\neq i} K\left(X^{i,N} - X^{j,N}\right)dt + \sqrt{2\sigma}dW^{i,N}, \quad i = 1, 2, \ldots, N. \tag{1}$$

One expects that as the number N of particles goes to infinity the system (1) will converge to the following Fokker-Plank equation:

$$\partial_t \rho = -\nabla \cdot ((b + K * \rho)\rho) + \sigma \Delta \rho. \tag{2}$$

Random batch method

If one numerically discretizes (1) directly, the computational cost per time step is $O(N^2)$, which is prohibitively expensive for large N. The Random Batch Method[1] is a simple and generic random algorithm to reduce the computation cost per time step from $O(N^2)$ to $O(N)$.

Algorithm 1 The Random Batch Method (RBM)

1: **for** $k = 1 : [T/\tau]$ **do**
2: Divide $\{1, 2, \dots, N\}$ into $n = N/p$ batches randomly.
3: **for** each batch ξ_k **do**
4: Update $\bar{X}^{i,N}$'s $(i \in \xi_k)$ by solving the following stochastic differential equation (SDE) with $t \in [t_{k-1}, t_k)$:

$$d\bar{X}^{i,N} = b\left(\bar{X}^{i,N}\right) dt + \frac{1}{p-1} \sum_{j \in \xi_k, j \neq i} K\left(\bar{X}^{i,N} - \bar{X}^{j,N}\right) dt + \sqrt{2\sigma} dW^i. \qquad (1.3)$$

5: **end for**
6: **end for**

[1] Shi Jin, Lei Li, and Jian-Guo Liu. "Random batch methods (RBM) for interacting particle systems". In: *Journal of Computational Physics* 400 (2020), p. 108877.

Due to the simplicity and scalability, RBM already has a variety of applications:

- Efficient particle methods for homogeneous Landau equation[2] in plasma physics;
- Random batch Monte Carlo method[3] for sampling from Gibbs distributions of interacting particle systems with singular kernels;
- Random batch Ewald method[4] for molecular dynamics simulations of particle systems with long-range Coulomb interactions.
- Reduce the computational cost of calculating the weighted average in the consensus-based optimization method[5] .

[2] José Antonio Carrillo, Shi Jin, and Yijia Tang. "Random batch particle methods for the homogeneous Landau equation". In: *Communications in Computational Physics* 31 (2021).

[3] Lei Li, Zhenli Xu, and Yue Zhao. "A Random-Batch Monte Carlo Method for Many-Body Systems with Singular Kernels". In: *SIAM Journal on Scientific Computing* 42.3 (2020), A1486–A1509.

[4] Shi Jin, Lei Li, Zhenli Xu, et al. "A Random Batch Ewald Method for Particle Systems with Coulomb Interactions". In: *SIAM Journal on Scientific Computing* 43.4 (2021), B937–B960.

[5] José A Carrillo et al. "A consensus-based global optimization method for high dimensional machine learning problems". In: *ESAIM: Control, Optimisation and Calculus of Variations* 27 (2021), S5.

Theoretical result of RBM

One marginal of N particle system

Nonlinear Fokker-Planck

$$
\begin{tikzcd}
\mu_N^{(1)} \arrow[r, "N \to \infty"] & \mu = \varrho\, dx \quad \mu(\cdot, t) = \mathcal{S}(t)(\mu_0) \\
\tilde{\mu}_N^{(1)} \arrow[u, "\Delta t \to 0"] \arrow[r, "N \to \infty"'] & \tilde{\mu} \arrow[u, "\Delta t \to 0"']
\end{tikzcd}
$$

One marginal of RBM

Mean field limit of RBM

Goal: Mean field error estimate

Denote

$$\tilde{\rho}_t^N = \mathrm{Law}(\tilde{X}_t^{1,N}, \cdots, \tilde{X}_t^{N,N}),$$

and

$$\rho_t^N(x_1, \ldots, x_N) = \rho_t^{\otimes N}.$$

Propagation of chaos: the $k-$marginal distribution of the particle system converges to the tensor product of the limit law $\rho_t^{\otimes k}$ as N goes to infinity, given for instance the i.i.d. initial data:

$$\lim_{N\to\infty} \tilde{\rho}_t^{N,k} = \rho_t^{\otimes k}.$$

Or equivalently (for exchangeable particles), the mean field limit:

$$\tilde{\mu}_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{\tilde{X}_t^{i,N}} \to \rho_t.$$

Assumption

(a) The field b is Lipschitz:

$$|b(x_1) - b(x_2)| \leq r\,|x_1 - x_2|\,.$$

Moreover, the field b is twice differentiable and its Hessian have at most polynomial growth:

$$\left|\nabla^2 b(x)\right| \leq C(1+|x|)^q.$$

(b) The field b is strongly confining in the sense that there exists two constants α and β such that for any $x_1 \neq x_2$, :

$$(x_1 - x_2)\cdot(b(x_1) - b(x_2)) \leq \alpha - \beta|x_1 - x_2|^2$$

for some constant $\beta > 0$.

(c) The interaction kernel K is bounded, and Lipschitz:

$$|K(x) - K(y)| \leq L|x - y|.$$

Moreover, the interaction kernel K is twice differentiable and their Hessians have at most polynomial growth:

$$\left|\nabla^2 K(x)\right| \leq \tilde{C}(1+|x|)^q.$$

Assumption

There exists a constant $C_{LS} > 0$ such that for any nonnegative smooth functions f, one has

$$\mathrm{Ent}_{\rho_t}(f) := \int f \log f d\rho_t - \left(\int f d\rho_t\right) \log \left(\int f d\rho_t\right) \leq C_{LS} \int \frac{|\nabla f|^2}{f} d\rho_t. \tag{3}$$

Such LSI assumption is a common ingredient in the proof of uniform-in-time propagation of chaos. One crucial property of the LSI is the tensorization, i.e. if ρ_t satisfies a LSI then $\rho_t^{\otimes N}$ satisfies the same inequality with the same constant.

Uniform-in-time relative entropy bound[6]

Under the previous assumptions, we have

$$\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) \leq e^{c_1 t}\mathcal{H}_N\left(\tilde{\rho}_0^N \mid \rho_0^{\otimes N}\right) + c_2(T)\left(\tau^2 + \frac{1}{N}\right), \tag{4}$$

where the constants c_1 and $c_2(T)$ are independent of N and τ. Here,

$$\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^N\right) = \frac{1}{N}\int_{\mathbb{R}^{Nd}} \tilde{\rho}_t^N\left(\mathrm{x}^N\right)\log\frac{\tilde{\rho}_t^N\left(\mathrm{x}^N\right)}{\rho_t^N\left(\mathrm{x}^N\right)}d\mathrm{x}^N,$$

is the rescaled relative entropy and $\mathrm{x}^N = (x_1, \cdots, x_N) \in \mathbb{R}^{Nd}$. Moreover, if $\beta > 2L$ and $\|K\|_{L^\infty}^2 \leq \frac{\sigma}{8e^2 C_{LS}}$, then $c_1 < 0$ and c_2 can be taken to be independent of T so the above bound is uniform-in-time.

[6] Zhenyu Huang, Shi Jin, and Lei Li. "Mean field error estimate of the random batch method for large interacting particle system". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 59.1 (2025), pp. 265–289.

Main results

Propagation of chaos

By Csiszár-Kullback-Pinsker inequality and transport inequality, we have

$$\|\tilde{\rho}_t^{N,k} - \rho_t^{\otimes k}\|_{L^1} + W_2\left(\tilde{\rho}_t^{N,k}, \rho_t^{\otimes k}\right) \leq C_1\tau + \frac{C_2}{\sqrt{N}}. \tag{5}$$

Here we define $\tilde{\rho}_t^{N,k}$ to be the density of the law of the k marginals of the random batch N particle system,

$$\tilde{\rho}_t^{N,k}(x_1, \cdots, x_k) = \int_{\mathbb{R}^{Nd}} \tilde{\rho}_t^N(x_1, \cdots, x_N)dx_{k+1}\cdots dx_N.$$

And we define the usual Wasserstein-2 distance by

$$W_2(\mu,\nu) = \left(\inf_{\gamma\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 d\gamma\right)^{1/2}.$$

上海交通大学 SHANGHAI JIAO TONG UNIVERSITY

An analogue of the Liouville equation

Denote by $\tilde{\varrho}_t^{N,\boldsymbol{\xi}}$ the probability density function of $\tilde{\mathrm{X}}_t^N = \left(\tilde{X}_t^{1,N}, \cdots, \tilde{X}_t^{N,N}\right)$ defined in (7) for $t \in [T_k, T_{k+1})$. Then the following Liouville equation holds:

$$\partial_t \tilde{\varrho}_t^{N,\boldsymbol{\xi}} + \sum_{i=1}^N \operatorname{div}_{x_i} \left(\tilde{\varrho}_t^{N,\boldsymbol{\xi}} \left(\hat{b}_t^{\boldsymbol{\xi},i}\left(\mathrm{x}^N\right) + \hat{K}_t^{\boldsymbol{\xi},i}\left(\mathrm{x}^N\right)\right)\right) = \sum_{i=1}^N \sigma \Delta_{x_i} \tilde{\varrho}_t^{N,\boldsymbol{\xi}}, \tag{8}$$

where

$$\hat{b}_t^{\boldsymbol{\xi},i}\left(\mathrm{x}^N\right) = \mathbb{E}\left[b\left(\tilde{X}_{T_k}^{i,N}\right) \mid \tilde{\mathrm{X}}_t^N = \mathrm{x}^N, \boldsymbol{\xi}\right], \quad t \in [T_k, T_{k+1}), \tag{9}$$

and

$$\hat{K}_t^{\boldsymbol{\xi},i}\left(\mathrm{x}^N\right) := \mathbb{E}\left[K^{\xi_k}\left(\tilde{X}_{T_k}^{i,N}\right) \mid \tilde{\mathrm{X}}_t^N = \mathrm{x}^N, \boldsymbol{\xi}\right], \quad t \in [T_k, T_{k+1}). \tag{10}$$

Here, $\boldsymbol{\xi} := (\xi_0, \xi_1, \cdots, \xi_k, \cdots)$ is a given sequence of batches.

On each time interval, for ξ_k given, by Markov property, we can define:

$$\tilde{\rho}_t^{N,\xi_k} := \mathbb{E}\left[\tilde{\varrho}_t^{N,\boldsymbol{\xi}} \mid \xi_i, i \geq k\right] = \mathcal{S}_{T_k,t}^{N,\xi_k} \tilde{\rho}_{T_k}^N, \quad t \in [T_k, T_{k+1}), \tag{11}$$

and we have

$$\partial_t \tilde{\rho}_t^{N,\xi_k} + \sum_{i=1}^N \operatorname{div}_{x_i}\left(\tilde{\rho}_t^{N,\xi_k}\left(\tilde{b}_t^{\xi_k,i}\left(\mathrm{x}^N\right) + \tilde{K}_t^{\xi_k,i}\left(\mathrm{x}^N\right)\right)\right) = \sum_{i=1}^N \sigma \Delta_{x_i} \tilde{\rho}_t^{N,\xi_k}, \quad \tilde{\rho}_{T_k}^{N,\xi_k} = \tilde{\rho}_{T_k}^N, \tag{12}$$

where

$$\tilde{b}_t^{\xi_k,i}\left(\mathrm{x}^N\right) := \mathbb{E}\left[b\left(\tilde{X}_{T_k}^{i,N}\right) \mid \tilde{\mathrm{X}}_t^N = \mathrm{x}^N, \xi_k\right], \quad t \in [T_k, T_{k+1}),$$

and

$$\tilde{K}_t^{\xi_k,i}\left(\mathrm{x}^N\right) := \mathbb{E}\left[K^{\xi_k}\left(\tilde{X}_{T_k}^{i,N}\right) \mid \tilde{\mathrm{X}}_t^N = \mathrm{x}^N, \xi_k\right], \quad t \in [T_k, T_{k+1}).$$

Time evolution of the relative entropy

$$
\begin{aligned}
\frac{d}{dt}\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) &= \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^{Nd}} \mathbb{E}_{\xi_k}\left(\rho_t^{N,\xi_k}\left(\tilde{b}_t^{\xi_k,i}(\mathrm{x}^N) - b(x_i)\right)\right)\cdot\nabla_{x_i}\log\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}d\mathrm{x}^N \\
&+ \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^{Nd}} \mathbb{E}_{\xi_k}\left(\tilde{\rho}_t^{N,\xi_k}\tilde{K}_t^{\xi_k,i}(\mathrm{x}^N) - \tilde{\rho}_t^{N,\xi_k}K^{\xi_k}(x_i)\right)\cdot\nabla_{x_i}\log\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}d\mathrm{x}^N \\
&+ \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^{Nd}} \mathbb{E}_{\xi_k}\left(\tilde{\rho}_t^{N,\xi_k}K^{\xi_k}(x_i) - \tilde{\rho}_t^{N}F^{N}(x_i)\right)\cdot\nabla_{x_i}\log\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}d\mathrm{x}^N \\
&+ \frac{1}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^{Nd}} \left(F^N(x_i) - K*\rho_t(x_i)\right)\tilde{\rho}_t^N\cdot\nabla_{x_i}\log\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}d\mathrm{x}^N \\
&- \frac{\sigma}{N}\sum_{i=1}^{N}\int_{\mathbb{R}^{Nd}} \tilde{\rho}_t^N\left|\nabla_{x_i}\log\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}\right|^2 d\mathrm{x}^N := \frac{1}{N}\sum_{i=1}^{N}(J_1^i + J_2^i + J_3^i + J_4^i + J_5^i).
\end{aligned}
\tag{15}
$$

Introduce another copy of RBM

Intuitively, the term $\left|\tilde{\rho}_t^{N,\xi_k} K^{\xi_k}(x_i) - \tilde{\rho}_t^N F^N(x_i)\right|$ in J_3^i is of $O(1)$, since $|K^{\xi_k} - F^N(x_i)| = O(1)$, which is not small.

We introduce another copy of RBM $\hat{\mathrm{X}}^N$ that depends on another batch $\tilde{\xi}_k$ such that:

- $\hat{\mathrm{X}}_{T_k}^N = \tilde{\mathrm{X}}_{T_k}^N$;
- the Brownian motion are the same in $[T_k, T_{k+1})$;
- the batch $\tilde{\xi}_k$ on $[T_k, T_{k+1})$ is independent of ξ_k.

Consequently, density of the law $\tilde{\rho}_t^{N,\tilde{\xi}_k}$ for $\hat{\mathrm{X}}^N$ satisfies both (11) and (12). Then

$$
\begin{aligned}
J_3^i &= \int_{\mathbb{R}^{Nd}} \mathbb{E}_{\xi_k}\left[\left(K^{\xi_k}(x_i) - F^N(x_i)\right)\left(\tilde{\rho}_t^{N,\xi_k} - \tilde{\rho}_t^N\right)\right] \cdot \nabla_{x_i} \log \frac{\tilde{\rho}_t^N}{\rho_t^N} d\mathrm{x}^N \\
&= \int_{\mathbb{R}^{Nd}} \mathbb{E}_{\xi_k,\tilde{\xi}_k}\left[\left(K^{\xi_k}(x_i) - F^N(x_i)\right)\left(\tilde{\rho}_t^{N,\xi_k} - \tilde{\rho}_t^{N,\tilde{\xi}_k}\right)\right] \cdot \nabla_{x_i} \log \frac{\tilde{\rho}_t^N}{\rho_t^N} d\mathrm{x}^N .
\end{aligned} \tag{16}
$$

New local error brought by RBM

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

Note that

$$\int_{\mathbb{R}^{Nd}} \frac{\left|\tilde{\rho}_t^{N,\tilde{\xi}_k} - \tilde{\rho}_t^{N,\xi_k}\right|^2}{\tilde{\rho}_t^{N,\tilde{\xi}_k}} d\mathrm{x}^N = \int_{\mathbb{R}^{Nd}} \left|\frac{\tilde{\rho}_t^{N,\xi_k}}{\tilde{\rho}_t^{N,\tilde{\xi}_k}} - 1\right|^2 \tilde{\rho}_t^{N,\tilde{\xi}_k} d\mathrm{x}^N.$$

Making use of the Girsanov transform in the path space:

$$\frac{\tilde{\rho}_t^{N,\xi_k}}{\tilde{\rho}_t^{N,\tilde{\xi}_k}}(\mathrm{x}^N) = \mathbb{E}\left[\frac{dP_{\tilde{\mathrm{X}}^N}}{dP_{\hat{\mathrm{X}}^N}} \mid \hat{\mathrm{X}}_t^N = \mathrm{x}^N, \xi_k, \tilde{\xi}_k\right]$$

$$= \mathbb{E}\left[\exp\left(\sqrt{\frac{1}{2\sigma}} \int_{T_k}^t \left(\delta \mathrm{K}^N\right)(\mathrm{y}^N) d\mathrm{W}_s - \frac{1}{4\sigma}\int_{T_k}^t \left|\left(\delta \mathrm{K}^N\right)(\mathrm{y}^N)\right|^2 ds\right) \mid \hat{\mathrm{X}}_t^N = \mathrm{x}^N, \xi_k, \tilde{\xi}_k\right].$$

Here, we denote

$$\delta \mathrm{K}^N(\mathrm{y}^N) := \frac{1}{\sqrt{2\sigma}}\left(\mathrm{K}^{N,\tilde{\xi}_k} - \mathrm{K}^{N,\xi_k}\right)(\mathrm{y}^N).$$

Overall estimate

Gathering the previous results, by the Log-Sobolev inequality taking $f = \frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}$, we have

$$\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) = \frac{1}{N}\operatorname{Ent}_{\rho_t^{\otimes N}}\left(\frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}\right) \le \frac{C_{LS}}{N}\sum_{i=1}^N \int_{\mathbb{R}^{Nd}} \tilde{\rho}_t^N \left|\nabla_{x_i} \log \frac{\tilde{\rho}_t^N}{\rho_t^{\otimes N}}\right|^2 \mathrm{dx}^N = C_{LS}\mathcal{I}_N(t).$$

Then we yield the following desired estimate for $t \in [T_k, T_{k+1})$:

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) \le& \left(4e^2\|K\|_{L^\infty}^2 - \frac{\sigma}{2C_{LS}}\right)\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) + c_1\tau^2\left(1 + \frac{1}{N}\mathcal{I}\left(\tilde{\rho}_{T_k}^N\right)\right) + \frac{c_2}{N}\\ \le& C_0\mathcal{H}_N\left(\tilde{\rho}_t^N \mid \rho_t^{\otimes N}\right) + C_1\tau^2 + \frac{C_2}{N},\end{aligned}$$

here the constants C_0, C_1 and C_2 are independent of N, τ and ξ_k, then by Gronwall's inequality, we end the proof. If $\beta > 2L$, the constants C_1 and C_2 can be made independent of t. Moreover, if $\|K\|_{L^\infty}^2 \le \frac{\sigma}{8e^2C_{LS}}$, then the constant C_0 becomes negative. Therefore, we have a uniform-in-time bound for the relative entropy.

Ongoing work

- Random batch method for Biot-Savart Law kernel (vortex method for simulating 2D Navier-Stokes equation);
- Random batch method for homogeneous Landau equation.

Thank you!

Zhenyu Huang
PhD candidate, Institute of Natural Sciences, SJTU

Joint work with Shi Jin and Lei Li · Mean field error estimate of the random batch method for large interacting particle system