Bias behaviour for mean-field rank based particle approximations of one dimensional viscous scalar conservation law

Oumaima Bencheikh & Benjamin Jourdain

CERMICS – École des Ponts ParisTech

oumaima.bencheikh@enpc.fr

École des Ponts ParisTech

Introduction

Let $\Lambda : [0,1] \to \mathbb{R}$ be a C^1 Lipschitz function with first order derivative λ. For m a probability measure on $\mathbb{R}$ and F_0 the cdf of m, we consider the SDE nonlinear in the sense of McKean:

$$\begin{cases} X_t = X_0 + \sigma W_t + \int_0^t \lambda\left(F(s,X_s)\right) ds, \quad t \in [0,T] \\ F(t,x) = \mathbb{P}\left(X_t \le x\right) \end{cases}$$

where the random variable X_0 is distributed according to m and independent from the Brownian motion $(W_t)_{t\ge 0}$.

For $t > 0$, we denote by μ_t the marginal law of X_t and $p(t,x)$ its probability density function that exists. $F(t,x)$ is solution to the following viscous conservation law: $\partial_t F(t,x) + \partial_x\Big(\Lambda(F(t,x))\Big) = \frac{\sigma^2}{2}\partial_{xx}F(t,x)$ with $F_0(x) = m\left((-\infty,x]\right)$.

We approximate the cdf of X_t by the empirical cdf $F^{N,h}(t,x) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\left\{X_t^{i,N,h}\le x\right\}}$ of N interacting particles $(X_t^{i,N,h})_{i\in[\![1,N]\!]}$ evolving according to:

$$X_t^{i,N,h} = X_0^i + \sigma W_t^i + \int_0^t \lambda^N\left(\sum_{j=1}^N \mathbf{1}_{\left\{X_{\tau_s^h}^{j,N,h} \le X_{\tau_s^h}^{i,N,h}\right\}}\right) ds, \ 1\le i\le N, \ t\in[0,T]$$

with time-step $h\in(0,T]$, $(W^i)_{i\ge1}$ i.i.d. copies of W and $\tau_s^h = \lfloor s/h\rfloor h$. The drift coefficient of the i^{th} particle in the increasing order is defined by $\lambda^N(i) = N\left(\Lambda\left(\frac{i}{N}\right) - \Lambda\left(\frac{i-1}{N}\right)\right)$. We also define the associated empirical measure by $\mu_t^{N,h} = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N,h}}$.

Wasserstein distance

Let us introduce the Wasserstein distance of index 1 defined for two probability measures μ and ν as:

$$\begin{aligned} \mathcal{W}_1(\mu,\nu) &= \inf\left\{\mathbb{E}\left[|X-Y|\right]; \mathrm{Law}(X)=\mu, \mathrm{Law}(Y)=\nu\right\} \\ &= \sup_{\varphi\in\mathcal{L}}\left(\int_{\mathbb{R}^d}\varphi(x)\mu(dx) - \int_{\mathbb{R}^d}\varphi(x)\nu(dx)\right) \end{aligned}$$

where $\mathcal{L}$ denotes the set of all 1-Lipschitz function $\varphi:\mathbb{R}^d\to\mathbb{R}$. Moreover, if $d=1$:

$$\mathcal{W}_1(\mu,\nu) = \int_0^1 \left|F_\mu^{-1}(u) - F_\nu^{-1}(u)\right| du = \int_{\mathbb{R}}\left|F_\mu(x) - F_\nu(x)\right| dx$$

where $F_\eta(x) = \eta\left((-\infty;x]\right)$ and $F_\eta^{-1}(u) = \inf\left\{x\in\mathbb{R} : F_\eta(x)\ge u\right\}$ denote respectively the cdf and the quantile function of a probability measure η on $\mathbb{R}$.

Particle initialization

The initial positions $(X_0^i)_{i\ge1}$ of the particles are either deterministic or random variables.

▶ When choosing a random initialization, we denote by $\hat{F}_0^N(x) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{X_0^i\le x\}}$ the empirical cdf of the N first random variables in the sequence $(X_0^i)_{i\ge1}$ i.i.d. according to m with $\hat{\mu}_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_0^i}$ its empirical measure.

▶ When choosing a deterministic initialization, we seek to construct a family $\left(x_i^N\right)_{1\le i\le N}$ of initial positions such that the piecewise constant function $\tilde{F}_0^N(x) = \frac{1}{N}\sum_{i=1}^N \mathbf{1}_{\{x_i^N\le x\}}$ approximates $F_0^{N,h}(x)$ with a sufficiently high accuracy. So we minimize the application $y\mapsto \int_{\frac{i-1}{N}}^{\frac{i}{N}}\left|F_0^{-1}(u) - y\right| du$ for each $i\in[\![1,N]\!]$. We then choose the optimal deterministic initialization $\left(x_i^N = F_0^{-1}\left(\frac{2i-1}{2N}\right)\right)_{i\ge1}$ that reaches $\inf_{(x_1^N,\dots,x_N^N)\in\mathbb{R}^N}\mathcal{W}_1\left(\tilde{\mu}_0^N, m\right) = \inf_{(x_1^N,\dots,x_N^N)\in\mathbb{R}^N}\sum_{i=1}^N\int_{\frac{i-1}{N}}^{\frac{i}{N}}\left|F_0^{-1}(u) - x_i^N\right| du$. We also denote by $\tilde{\mu}_0^N = \frac{1}{N}\sum_{i=1}^N \delta_{x_i^N}$ its associated empirical measure.

The next proposition gives the assumptions under which the L^1-norm of the difference between F_0 and $\hat{F}_0^N$ or $\tilde{F}_0^N$ is of order $N^{-1/2}$.

Proposition 0.1:

We have the following results concerning the $\mathcal{O}(N^{-1/2})$ behaviour of the errors:

$$\sup_{N\ge1}\sqrt{N}\mathbb{E}\left[\mathcal{W}_1\left(\hat{\mu}_0^N, m\right)\right] < \infty$$

$$\Updownarrow$$

$$\int_{\mathbb{R}}|x|^{2+}m(dx)<\infty \Rightarrow \int_{\mathbb{R}}\sqrt{F_0(x)(1-F_0(x))}\,dx<\infty \Rightarrow \int_{\mathbb{R}}|x|^2 m(dx)<\infty$$

$$\Rightarrow \sup_{x\ge1} x\int_x^{+\infty}\left(F_0(-y) + 1 - F_0(y)\right) dy < \infty \Rightarrow \int_{\mathbb{R}}|x|^{2-}m(dx)<\infty$$

$$\Downarrow$$

$$\sup_{N\ge1}\sqrt{N}\mathcal{W}_1\left(\tilde{\mu}_0^N, m\right) < \infty$$

Moreover, none of the implications is an equivalence and there exists a probability measure m such that $\int_{\mathbb{R}}|x|^{2-}m(dx)<\infty$ and $\lim_{N\to\infty}\sqrt{N}\mathcal{W}_1\left(\tilde{\mu}_0^N, m\right) = \infty$.

Concerning the L^1-weak error between F_0 and $F_0^{N,h}$, since the empirical cdf of i.i.d. samples is unbiased $\mathbb{E}\left[\hat{F}_0^N(x)\right] = F_0(x)$ for all $N\ge1$ and $x\in\mathbb{R}$ then $\int_{\mathbb{R}}\left|\mathbb{E}\left[\hat{F}_0^N(x)\right] - F_0(x)\right| dx = 0$. As for the deterministic initialization, we prove in the next proposition that $\int_{\mathbb{R}}\left|\tilde{F}_0^N(x) - F_0(x)\right| dx$ is of order N^{-1}.

Proposition 0.2:

When m is compactly supported i.e. $\exists -\infty < c \le d < \infty$ such that $m([c,d]) = 1$, then

$$\mathcal{W}_1\left(\tilde{\mu}_0^N, m\right) \le \frac{d-c}{2N}.$$

L^1-strong error

We estimate the mean of the Wasserstein distance of index 1 for the marginal law $\mu_t^{N,h}$ of the Euler discretization with time-step h of a system of N interacting particles and its limit μ_t in the following result.

Theorem 0.1:

Assume either that the initial positions are i.i.d. according to m and $\int_{\mathbb{R}}\sqrt{F_0(x)(1-F_0(x))}\,dx < \infty$ or the initial positions are optimal deterministic and $\sup_{x\ge1} x\int_x^{+\infty}\left(F_0(-y)+1-F_0(y)\right) dy$. Then:

$$\exists C<\infty, \forall N\in\mathbb{N}^*, \quad \sup_{t\le T}\mathbb{E}\left[\mathcal{W}_1\left(\mu_t^{N,0},\mu_t\right)\right] \le \frac{C}{\sqrt{N}}.$$

Moreover, if λ is Lipschitz continuous then:

$$\exists C<\infty, \forall N\in\mathbb{N}^*, \forall h\in(0,T], \quad \sup_{t\le T}\mathbb{E}\left[\mathcal{W}_1\left(\mu_t^{N,h},\mu_t\right)\right] \le C\left(\frac{1}{\sqrt{N}}+h\right).$$

Under additional regularity upon Λ and F_0, and with deterministic initial conditions, M. Bossy proved a stronger result for the L^1 and L^∞ norms.

Theorem 0.2:

Assume that Λ is C^3, F_0 is C^2 bounded with bounded first and second order derivatives in x and $\exists\ M, C, \beta > 0,\ \alpha\ge0$ such that $|\partial_x F_0(x)| \le \alpha\exp\left(-\beta x^2/2\right)$ when $|x|>M$. Moreover, the initial positions are deterministic and given by $x_i^N = F_0^{-1}\left(\frac{i}{N}\right)$ when $i=1,..,N-1$ and $x_N^N = F_0^{-1}\left(1-\frac{1}{2N}\right)$. Then $\forall t,h\in[0,T], N\in\mathbb{N}^*$,

$$\mathbb{E}\left[\mathcal{W}_1\left(\mu_t^{N,h},\mu_t\right)\right] + \sup_{x\in\mathbb{R}}\left(\mathbb{E}\left[\left|F^{N,h}(t,x) - F(t,x)\right|\right]\right) = \mathcal{O}\left(\frac{1}{\sqrt{N}}+h\right).$$

L^1-weak error

Now, we prove that the L^1-weak error between the empirical cumulative distribution function $F^{N,h}$ of the Euler discretization with time-step h of the system of N interacting particles and its limit F is $\mathcal{O}\left(\frac{1}{N}+h\right)$. Let $\varphi\in\mathcal{L}$ and $\mathbb{E}\left[\mu_t^{N,h}\right]$ denote the probability measure on $\mathbb{R}$ s.t. $\mathbb{E}\left[\int_{\mathbb{R}}\varphi(x)\mu_t^{N,h}(dx)\right] = \frac{1}{N}\sum_{i=1}^N\mathbb{E}\left[\varphi\left(X_t^{i,N,h}\right)\right]$.

Theorem 0.3:

Assume that λ is Lipschitz continuous and either that the initial positions are optimal deterministic with m being compactly supported or that the initial positions are i.i.d. according to m and for some $\rho>1$, $\int_{\mathbb{R}}|x|^\rho m(dx)<\infty$. Then:

$$\exists C_b<\infty, \forall N\in\mathbb{N}^*, \forall h\in[0,T], \quad \sup_{t\le T}\mathcal{W}_1\left(\mathbb{E}\left[\mu_t^{N,h}\right],\mu_t\right) \le C_b\left(\frac{1}{N}+h\right).$$

Remark 1 *Using the dual formulation of the Wasserstein distance exposed above, we have the additional result:*

$$\exists C<\infty, \forall N\in\mathbb{N}^*, \forall t,h\in[0,T], \sup_{\varphi\in\mathcal{L}}\left|\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N\varphi\left(X_t^{i,N,h}\right)\right] - \int_{\mathbb{R}}\varphi(x)\mu(dx)\right| \le C\left(\frac{1}{N}+h\right).$$

Existing result

▶ Results of weak error behaviour in $\mathcal{O}(N^{-1})$ for particle approximations of general McKean-Vlasov SDEs have been exposed by Kolokoltsov *(2010)* under high constraints of regularity of the coefficients.

▶ In the context of systems of diffusive particles interacting through moments, B. and Jourdain *(18)*prove under specific constraints of regularity on the coefficients that the weak error behaves in $\mathcal{O}\left(N^{-1}+h\right)$ when h denotes the time step of the Euler discretization.

▶ Under higher constraints of smoothness on the coefficients, Chassagneux, Szpruch and Tse *(19)* prove for particle approximations of more general McKean-Vlasov SDEs that the weak error behaves in terms of $\sum_{j=1}^{k-1}\frac{C_j}{N^j} + \mathcal{O}(N^{-k})$ where $k\in\mathbb{N}^*$ refers to the order of smoothness (coefficients are $(2k+1)$-times differentiable).