Computing the stationary measure of McKean-Vlasov SDEs

J-F Chassagneux (ENSAE-CREST & IP Paris)
based on a joint work with
Gilles Pagès (Sorbonne Université & LPSM)

International seminar on SDEs and related Topics
Jan 17 2025, online

The question

A solution

Why it works

Stationary distribution of MKV SDEs

Let $(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{F}) := (\mathcal{F}_t)_{t \geqslant 0})$ be a filtered probability space, with $\mathbb{F}$ satisfying the usual conditions. W is $\mathbb{F}$-Brownian motion (independent of $\mathcal{F}_0$). Consider

- a McKean-Vlasov SDE: ($[\xi]$ is the law of the random variable ξ)

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

- As you know: generally obtained as limit of particles systems weakly interacting. Here we are interested in the long time behavior - convergence to equilibrium of this equation $\rightsquigarrow$ stationary distribution.
- A stationary distribution $\nu^\star$ is such that $[X_t] = \nu^\star$ for all $t \geqslant 0$
- Starting with $X_0^\star \sim \nu^\star$ we have simply

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

$\hookrightarrow$ A 'classical' diffusion...

Goal: Find a way to compute $\nu^\star$ (provided it exists!)

Computing the stationary distribution of classical SDE

For a classical SDEs $\mathrm{d}X_t = B(X_t)\,\mathrm{d}t + \Sigma(X_t)\,\mathrm{d}W_t$ with stationary distribution $\hat{\mu}$.

- *First approach*. Use Euler scheme with step h for the SDE:

$$X_{(n+1)h} = X_{nh} + hB(X_{nh}) + \Sigma(X_{nh})(W_{(n+1)h} - W_{nh})$$

  Simulate M samples $(X^m_{nh})_{1 \leqslant m \leqslant M}$ with (M, n) large to obtain $\frac{1}{M}\sum_{m=1}^M \delta_{X^m_{nh}} \simeq \hat{\mu}$

- *Second approach*. From the fact $\frac{1}{t}\int_0^t \delta_{X_s}\,\mathrm{d}s \to \hat{\mu}$ (a.k.a. ergodic behavior under some good conditions)
  1. First idea: Use the scheme (X_{nh}) and let it run forever to obtain $\frac{1}{n}\sum_{k=0}^{n-1} \delta_{X_{kh}} \to \hat{\mu}$...+bias.
  2. Better idea: Use adaptative time stepping to kill the bias asymptotically.

Computing the stationary distribution of SDE

- Let $(\gamma_n)_{n\geqslant 1}$ be a non-increasing sequence of positive steps satisfying

$$\gamma_n \to 0 \quad \text{and} \quad \Gamma_n := \sum_{k=1}^{n} \gamma_k = +\infty \quad \text{as} \quad n \to +\infty$$

$\hookrightarrow$ Define then $\bar{X}$ the Euler scheme with decreasing step size:

$$\bar{X}_{\Gamma_n} = \bar{X}_{\Gamma_{n-1}} + \gamma_n B(\bar{X}_{\Gamma_{n-1}}) + \sqrt{\gamma_n}\Sigma(\bar{X}_{\Gamma_{n-1}})Z_n, \quad \text{with} \;\; Z_n = \frac{W_{\Gamma_n} - W_{\Gamma_{n-1}}}{\sqrt{\gamma_n}}$$

and then set $\bar{\nu}_{\Gamma_n} = \frac{1}{\Gamma_n}\sum_{k=1}^{n} \gamma_k \delta_{\bar{X}_{k-1}}, \quad n \geqslant 1$

- One expects $\bar{\nu}_{\Gamma_n} \to \hat{\mu}$ when n goes to infinity.
- Lamberton-Pagès [LP02, LP03] give a complete picture of the convergence results and how to chose the step optimally in various contexts. (In particular, $\gamma_n = \gamma_1 n^{-\frac{1}{3}}$)
- How to adapt this to our framework?

Computing the stationary distribution (MKV)

- Obstruction for MKV: $(B(x), \Sigma(x)) = (b(x, \nu^\star), \sigma(x, \nu^\star))$ and $\nu^\star$ is what we want to compute...
- Solution: Replace $\nu^\star$ by the empirical measure in the coefficients. Introduce first (the self-interacting diffusion)

$$\mathrm{d}\mathcal{X}_t = b(\mathcal{X}_t, \nu_t^{\mathcal{X}})\,\mathrm{d}t + \sigma(\mathcal{X}_t, \nu_t^{\mathcal{X}})\,\mathrm{d}W_t \text{ with } \nu_t^{\mathcal{X}} := \frac{1}{t}\int_0^t \delta_{\mathcal{X}_s}\,\mathrm{d}s.$$

Then consider its Euler Scheme $(\bar{\mathcal{X}}_{\Gamma_n})_{n\geqslant 0}$ the scheme, which is defined by

$$\bar{\mathcal{X}}_{\Gamma_n} = \bar{\mathcal{X}}_{\Gamma_{n-1}} + \gamma_n b(\bar{\mathcal{X}}_{\Gamma_{n-1}}, \bar{\nu}_{\Gamma_{n-1}}) + \sqrt{\gamma_n}\sigma(\bar{\mathcal{X}}_{\Gamma_{n-1}}, \bar{\nu}_{\Gamma_{n-1}})Z_n,$$
$$\text{with } \bar{\nu}_{\Gamma_n} := \frac{1}{\Gamma_n}\sum_{k=1}^n \gamma_k \delta_{\bar{\mathcal{X}}_{\Gamma_{k-1}}},\ Z_n := \frac{W_{\Gamma_n} - W_{\Gamma_{n-1}}}{\sqrt{\gamma_n}}.$$

- 0ne expects $\bar{\nu}_{\Gamma_n} \to \nu^\star$.
- Observe that this is in sharp contrast with the basic idea of simulating a particles system and letting time run forever $\rightsquigarrow$ only one particle is simulated here!

Related literature and contributions

The idea of using self-interacting diffusion to approximate stationary measure of MKV SDEs is not new.

- The paper [ABRS19] from CEMRACS 2017 mentions this approach.
- The paper [DJL23] studies the convergence with a rate for $\mathbb{E}\big[\mathcal{W}_2^2(\nu^\star, \bar{\nu}_{\Gamma_n})\big]$ (very inspiring) Their setting is the closest to ours.
- The paper [KK$^+$12] (and references therein) give some almost sure rate of convergence: application to physics.
- The paper [DRSW23] uses exponentially weighted empirical measure and combines this with an annealing method to obtain convergence.

Our main contributions:

1. We focus on the implemented scheme.
2. We obtain rate of convergence for $\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n})$ both in the L^2 and almost sure case where

$$\mathcal{W}_2(\mu,\nu) := \inf_{(X,Y) s.t. X\sim\nu, Y\sim\mu} \mathbb{E}\Big[|X-Y|^2\Big]^{\frac{1}{2}}$$

Main setting

1. Lipschitz coefficient: this is because we consider the error for the Euler scheme.

$$|b(x,\mu)-b(y,\nu)| + \|\sigma(x,\mu)-\sigma(y,\nu)\|_F \leqslant L\left(|x-y| + \mathcal{W}_2(\mu,\nu)\right).$$

2. Confluence $(HC)_{p,\alpha,\beta}$: for every x, $y \in \mathbb{R}^d$ and every μ, $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, (b,σ) satisfies:

$$2\big(b(x,\mu)-b(y,\nu)\,|\,x-y\big) + (2p-1)\|\sigma(x,\mu)-\sigma(y,\nu)\|_F^2 \leqslant -\alpha|x-y|^2 + \beta\mathcal{W}_2^2(\mu,\nu).$$

with $\alpha > \beta \geqslant 0$.

$\hookrightarrow$ $(H^\star)$=(HL) & $(HC)_{1,\alpha,\beta}$ hold and we set $\vartheta^\star := 1-\frac{\beta}{\alpha}$.

3. Mean-reversion: to obtain some integrability $(HMV)_{p,K',\alpha',\beta'}$: For every x, $y \in \mathbb{R}^d$ and every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$2\big(b(x,\mu)\,|\,x\big) + (2p-1)\|\sigma(x,\mu)\|_F^2 \leqslant K' - \alpha'|x|^2 + \beta'\mathcal{W}_2^2(\mu,\delta_0).$$

with $\alpha' > \beta' \geqslant 0$

4. σ is uniformly elliptic, to obtain the almost sure rate of convergence only.

(There are links between assumptions 1,2 & 3)

On the stationary process $X^\star$

- For $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, denote $\Pi(\mu)$ the set of invariant measure for

$$\mathrm{d}X_t^\mu = b(X_t^\mu, \mu)\,\mathrm{d}t + \sigma(X_t^\mu, \mu)\,\mathrm{d}W_t \text{ and } X_0^\mu \sim \mu$$

  - Under the confluence assumption, $\Pi(\mu)$ is single valued and

$$\mathcal{W}_2(\Pi(\mu), \Pi(\nu)) \leqslant \sqrt{\frac{\beta}{\alpha}}\mathcal{W}_2(\mu, \nu)$$

  - Work in the case where $\alpha > \beta \geqslant 0$: $\nu^\star$ is the fixed point of Π then $(X^\star = X^{\nu^\star})$.

- The mean-reversion assumption allows then to "control" the level of integrability of $\nu^\star$ (minimal case is $\mathcal{L}^{2p^\star}$, for some $p^\star > 1$)

- Example "OU like process": $\mathrm{d}X_t = (\mathfrak{b}\mathbb{E}[X_t] - X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$. Set $m_t = \mathbb{E}[X_t]$, so that $\mathrm{d}m_t = (\mathfrak{b} - 1)m_t$ and $m_t = \mathbb{E}[X_0]\, e^{(\mathfrak{b}-1)t}\ \forall t$

  1. if $|\mathfrak{b}| < 1$, $(H^\star)$ holds: $\mathrm{d}X_t^\star = -X_t^\star\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$, O.U. process $\nu^\star = \mathcal{N}(0,1)$.
  2. if $\mathfrak{b} = 1$: $\mathrm{d}X_t^\star = (m_0 - X_t^\star)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$ Stationary distributions $\nu^\star = \mathcal{N}(m_0, 1)$ parametrised by initial mean m_0...
  3. if $\mathfrak{b} < -1$: $(H^\star)$ does not hold but stationary distribution $\nu^\star = \mathcal{N}(0,1)$

Some computations for the contraction property

- $i = 1, 2$, $\mu^i \in \mathcal{P}^2(\mathbb{R}^d)$, $\mathrm{d}X_t^i = b(X_t^i, \mu^i)\,\mathrm{d}t + \sigma(X_t^i, \mu^i)\,\mathrm{d}W_t$, $X^i \sim \Pi[\mu^i]$ (stationary)
  - Apply Ito's formula to $(e^{\alpha t}|X_t^1 - X_t^2|^2)_{t \geqslant 0}$:

$$e^{\alpha t}|X_t^1 - X_t^2|^2 = |X_0^1 - X_0^2|^2 + \alpha \int_0^t e^{\alpha s}|X_s^1 - X_s^2|^2\,\mathrm{d}s + M_t(loc.mart.)$$

$$+ \int_0^t e^{\alpha s}\Big(\underbrace{2\big(b(X_s^1, \mu^1) - b(X_s^2, \mu^2)|X_s^1 - X_s^2\big) + \|\sigma(X_s^1, \mu^1) - \sigma(X_s^2, \mu^2)\|_F^2}_{\text{via } (HC) \leqslant -\alpha|X_s^1 - X_s^2|^2 + \beta\mathcal{W}_2^2(\mu^1, \mu^2)} \Big)\,\mathrm{d}s$$

  - After localization if need be:

$$\underbrace{\mathbb{E}\Big[|X_t^1 - X_t^2|^2\Big]}_{\geqslant \mathcal{W}_2^2(\Pi[\mu^1], \Pi[\mu^2])} \leqslant e^{-\alpha t}\mathbb{E}\Big[|X_0^1 - X_0^2|^2\Big] + \beta\mathcal{W}_2^2(\mu^1, \mu^2)\int_0^t e^{\alpha(s-t)}\,\mathrm{d}s$$

  - Integrating and letting $t \to +\infty$, we do obtain

$$\mathcal{W}_2^2(\Pi[\mu^1], \Pi[\mu^2]) \leqslant \frac{\beta}{\alpha}\mathcal{W}_2^2(\mu^1, \mu^2)$$

Convergence results, see [CP24].

- We obtain rates of convergence for $\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n})$ in the L2 and a.s. sense where:

$\hookrightarrow$ $\nu^\star$ is the unique stationary distribution of the MKV SDE (distribution of $X^\star$)

$\hookrightarrow$ $\bar{\nu}_{\Gamma_n} := \frac{1}{\Gamma_n}\sum_{k=1}^n \gamma_k \delta_{\bar{\mathcal{X}}_{\Gamma_{k-1}}}$ with $\bar{\mathcal{X}}$ Euler scheme with stepsize (γ_n) for the SID.

- *Results in the L2 sense*. $(H^\star)$ holds. Denote $r^\star := \frac{\vartheta^\star}{1+\vartheta^\star} \in (0,1)$ with $\vartheta^\star = 1 - \frac{\beta}{\alpha}$.

Set $\gamma_n = \gamma_1 n^{-r^\star}$, $\gamma_1 > 0$.

i) for any small $\eta > 0$, set $\zeta_{p^\star} := \frac{2p^\star - 1}{2(2(d+2)+(2p^\star-1)(d+3))} \underset{p^\star\to\infty}{\longrightarrow} \frac{1}{2(d+3)}$,

$$\mathbb{E}\Big[\mathcal{W}_2^2(\nu^\star, \bar{\nu}_{\Gamma_n})\Big]^{\frac{1}{2}} = \underbrace{O_\eta\left(n^{-(1-r^\star)\zeta_{p^\star}}\right)}_{\text{convergence of } \nu^{X^\star} \text{ to } \nu^\star} + \underbrace{(1+\mathcal{W}_2(\nu^\star, [\mathcal{X}_0]))\, o_\eta\left(n^{-\frac{r^\star}{2}+\eta}\right)}_{\text{rate for the Euler scheme to the SID}}$$

ii) If moreover, σ is bounded then

$$\mathbb{E}\Big[\mathcal{W}_2^2(\nu^\star, \bar{\nu}_{\Gamma_n})\Big]^{\frac{1}{2}} = \underbrace{O\left(n^{-\frac{1-r^\star}{2(d+3)}} \log(n)^{\frac{d+2}{2(d+3)}}\right)}_{\text{improved rate}} + \left(1+\mathcal{W}_2^2(\nu^\star, [\mathcal{X}_0])\right) o_\eta\left(n^{-\frac{r^\star}{2}+\eta}\right).$$

Note: If one has $\mathbb{E}\Big[\mathcal{W}_2^2(\nu_t^{X^\star}, \nu^\star)\Big]^{\frac{1}{2}} = O(t^{-\zeta})$ then it can be used above!

Convergence results in the a.s. sense

- $(H^\star) + (HMV)_2 + \sigma$ unif. elliptic: Set $\gamma_n = \gamma_1 n^{-(r^\star \wedge \frac{1}{3})}$

i) for any small $\eta', \eta > 0$, set $\hat{\zeta}_{p^\star} := \frac{(2p^\star - 1)^2}{2(2p^\star+1)\{2(d+2)+(d+3)(d+2p^\star-1)\}}$

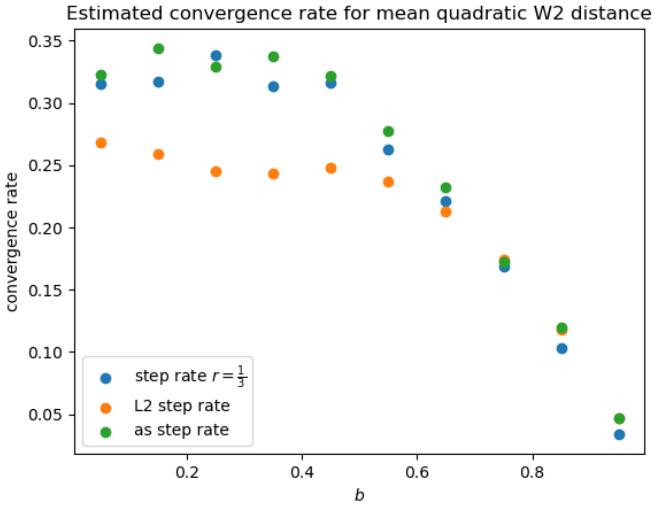
$$\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n}) = \underbrace{o_{\eta'}\left(n^{-(1-(r^\star \wedge \frac{1}{3}))\hat{\zeta}_{p^\star}} \log(n)^{\frac{1}{2}+\eta'}\right)}_{\text{convergence of } \nu^{X^\star} \text{ to } \nu^\star} + \underbrace{(1 + |X_0^\star - \mathcal{X}_0|)\, o_\eta\left(n^{-(\frac{r^\star}{2} \wedge \frac{1}{6})+\eta}\right)}_{\text{rate for the Euler scheme to the SID}}$$

ii) If moreover, σ is bounded then, for any small $\eta', \eta > 0$,

$$\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n}) = \underbrace{o_{\eta'}\left(n^{-\frac{1-(r^\star \wedge \frac{1}{3})}{2(d+3)}} \log(n)^{\frac{1}{2}+\eta'}\right)}_{\text{improved rate}} + (1 + |X_0^\star - \mathcal{X}_0|)\, o_\eta\left(n^{-(\frac{r^\star}{2} \wedge \frac{1}{6})+\eta}\right).$$

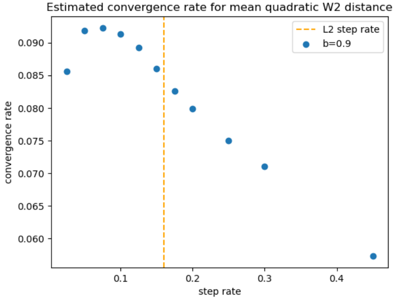
Numerical illustration

- MKV Example: $\mathrm{d}X_t = (\mathfrak{b}\mathbb{E}[X_t] - X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$ $(b \in \mathbb{R})$
  - Stationary version: $\mathrm{d}X_t^\star = -X_t^\star\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$ (OU process)
  - Stationary distribution $\mathfrak{b} < 1$: $\nu^\star = \mathcal{L}(X_t^\star) = \mathcal{N}(0,1)$
- Scheme for this equation is easy to implement:
  $\bar{\mathcal{X}}_{\Gamma_n} = \bar{\mathcal{X}}_{\Gamma_{n-1}} + \gamma_n(b\bar{m}_{n-1} - \mathcal{X}_{\Gamma_{n-1}}) + \sqrt{2\gamma_n}Z_n$ with $\bar{m}_{n-1} = \frac{1}{\Gamma_{n-1}}\sum_{k=1}^{n-1}\gamma_k X_{k-1}$
  (observe that $m_n = \frac{\gamma_n}{\Gamma_n}X_{n-1} + (1 - \frac{\gamma_n}{\Gamma_n})m_{n-1}$)
- Question: how to choose the step rate r in $\gamma_n = \gamma_1 n^{-r}$?
  $\hookrightarrow$ Previous theoretical results indicates $r^\star := \frac{\vartheta^\star}{1+\vartheta^\star} \in (0,1)$ with $\vartheta^\star = 1 - \frac{\beta}{\alpha}$.
  An upper bound for $\vartheta^\star$ here is $1 - b^2 \longrightarrow$ "L2 step rate" on the graph.
  $\hookrightarrow$ we also consider $r^\star \wedge \frac{1}{3}$: "as step rate" and $r = \frac{1}{3}$ (classical SDE case).
- We estimate convergence rate for $\mathbb{E}\big[\mathcal{W}_2^2(\bar{\nu}_{\Gamma_n}, \nu^\star)\big]^{\frac{1}{2}}$ (emp. mean on 500 samples, n up to $N = 100000$):



Empirical estimation of the convergence rate as a function of $\mathfrak{b}$. Tests realised for three different specifications of the algorithm input step rate, $N = 100000$ and $M = 500$.

Is there an optimal step rate?



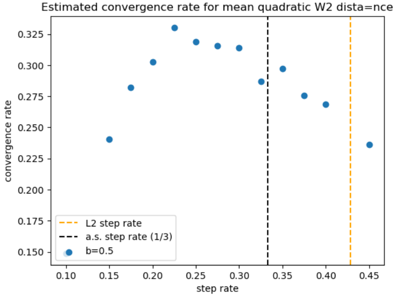
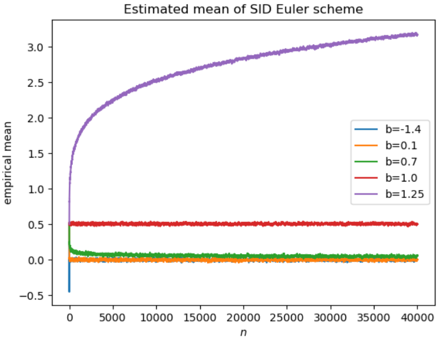
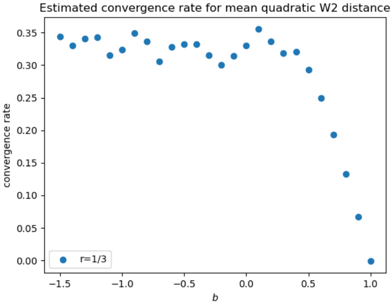


Figure: Comparison of optimal rates for $\mathfrak{b} = 0.9$ (left) and $\mathfrak{b} = 0.5$ (right).

Empirical convergence for $r = \frac{1}{3}$



(a) Empirical mean of $(\bar{\mathcal{X}}_{\Gamma_n})$ (M=10000) for various values of $\mathfrak{b}$



(b) Convergence rate as a function of the parameter $\mathfrak{b}$. $(M = 500)$

Global approach to the problem

• We decompose the main error as follows:

$$\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n}) \leqslant \mathcal{W}_2(\nu^\star, \nu^{\mathcal{X}}_{\Gamma_n}) + \mathcal{W}_2(\nu^{\mathcal{X}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})$$

▸ The term $\mathcal{W}_2(\nu^\star, \nu^{\mathcal{X}}_{\Gamma_n})$ reveals the ergodic behavior of $\mathcal{X}$ at the limit. It is further controled by

$$\mathcal{W}_2(\nu^\star, \nu^{X^\star}_{\Gamma_n}) + \mathcal{W}_2(\nu^{X^\star}_{\Gamma_n}, \nu^{\mathcal{X}}_{\Gamma_n})$$

▸ The term $\mathcal{W}_2(\nu^{\mathcal{X}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})$ is linked to discretisation errors. It is further controled by

$$\mathcal{W}_2(\nu^{\mathcal{X}}_{\Gamma_n}, \nu^{\bar{\mathcal{X}}}_{\Gamma_n}) + \mathcal{W}_2(\nu^{\bar{\mathcal{X}}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})$$

↪ Notation: $\bar{\mathcal{X}}$ is the continuous Euler scheme:

$$\bar{\mathcal{X}}_t = \bar{\mathcal{X}}_0 + \int_0^t b(\bar{\mathcal{X}}_{\underline{s}}, \bar{\nu}_{\underline{s}})\,\mathrm{d}s + \int_0^t \sigma(\bar{\mathcal{X}}_{\underline{s}}, \bar{\nu}_{\underline{s}})\,\mathrm{d}W_s \text{ with } \underline{s} := \Gamma_n \text{ if } s \in [\Gamma_n, \Gamma_{n+1})$$

▸ For the blue terms (distance between empirical measures):

$$\mathcal{W}_2^2(\nu^Y_t, \nu^Z_t) \leqslant \frac{1}{t}\int_0^t |Y_s - Z_s|^2\,\mathrm{d}s, \text{ using the transport plan } \pi(\mathrm{d}y, \mathrm{d}z) = \frac{1}{t}\int_0^t \delta_{(Y_s, Z_s)}(\mathrm{d}y, \mathrm{d}z)\,\mathrm{d}s.$$

They both involve the self-interacting diffusion.

Stability for self-interacting diffusion (L^2-case)

• Under our Lipschitz assumption, it is not too difficult to obtain existence and uniqueness for the SID.

• However, we need to for some $p \geqslant 1$ (using mean reversion assumption)

$$\sup_{t \geqslant 0} \mathbb{E}\Big[|\mathcal{X}_t|^{2p}\Big] < +\infty$$

• **Stability:** Consider, for some perturbation η^b, η^σ

$$\tilde{\mathcal{X}}_t = \tilde{\mathcal{X}}_0 + \int_0^t \left(b(\tilde{\mathcal{X}}_s, \nu_s^{\tilde{\mathcal{X}}}) + \eta_s^b \right) \mathrm{d}s + \int_0^t \left(\sigma(\tilde{\mathcal{X}}_s, \nu_s^{\tilde{\mathcal{X}}}) + \eta_s^\sigma \right) \mathrm{d}W_s.$$

Then, for any $\vartheta \in (0, \vartheta^\star)$, the following holds

$$\frac{1}{t}\int_0^t \mathbb{E}\Big[|\tilde{\mathcal{X}}_s - \mathcal{X}_s|^2\Big]\,\mathrm{d}s \leqslant Ct^{-\vartheta}\left(\mathbb{E}\Big[|\tilde{\mathcal{X}}_0 - \mathcal{X}_0|^2\Big] + C_{\vartheta^\star-\vartheta}\int_0^t s^{\vartheta-1}\mathbb{E}[|\eta_s^b|^2 + |\eta_s^\sigma|^2]\,\mathrm{d}s\right).$$

• Example of application: say perturbation $\mathbb{E}[|\eta_s^b|^2 + |\eta_s^\sigma|^2] = O(s^{-a})$,
then overall error is: $O(t^{-\vartheta}\mathbb{E}\Big[|\tilde{\mathcal{X}}_0 - \mathcal{X}_0|^2\Big]) + O(t^{-a})$

Some Computations for stability

- SID: $\mathrm{d}\mathcal{X}_t = b(\mathcal{X}_t, \nu_t^{\mathcal{X}})\,\mathrm{d}t + \mathrm{d}W_t$, perturbation $\tilde{\mathcal{X}}$.
  - Applying Ito's formula to $(e^{\alpha t}|\mathcal{X}_t - \tilde{\mathcal{X}}_t|^2)_{t\geqslant 0}$:

$$e^{\alpha t}|\mathcal{X}_t - \tilde{\mathcal{X}}_t|^2 = |\mathcal{X}_0 - \tilde{\mathcal{X}}_0|^2 + \alpha \int_0^t e^{\alpha s}|\mathcal{X}_s - \tilde{\mathcal{X}}_s|^2 \,\mathrm{d}s + M_t(loc.mart.)$$

$$+ \int_0^t e^{\alpha s}\Big(\underbrace{2\big(b(\mathcal{X}_s, \nu_s^{\mathcal{X}}) - b(\tilde{\mathcal{X}}_s, \nu_s^{\tilde{\mathcal{X}}})|\mathcal{X}_s - \tilde{\mathcal{X}}_s\big)}_{\text{via } (HC) \leqslant -\alpha|\mathcal{X}_s - \tilde{\mathcal{X}}_s|^2 + \beta\mathcal{W}_2^2(\nu_s^{\mathcal{X}}, \nu_s^{\tilde{\mathcal{X}}})} + perturbation\Big)\,\mathrm{d}s$$

  - Since $\mathcal{W}_2^2(\nu_t^{\mathcal{X}}, \nu_t^{\tilde{\mathcal{X}}}) \leqslant \frac{1}{t}\int_0^t |\mathcal{X}_s - \tilde{\mathcal{X}}_s|^2 \,\mathrm{d}s$, taking expectation (localizing if need be)

$$\mathbb{E}\Big[|\mathcal{X}_t - \tilde{\mathcal{X}}_t|^2\Big] \leqslant e^{-\alpha t}\mathbb{E}\Big[|\mathcal{X}_0 - \tilde{\mathcal{X}}_0|^2\Big] + \beta \int_0^t e^{-\alpha(s-t)} \frac{1}{s}\int_0^s \mathbb{E}\Big[|\mathcal{X}_u - \tilde{\mathcal{X}}_u|^2\Big]\,\mathrm{d}u\,\mathrm{d}s + \dots$$

  - Set $g(t) := \frac{1}{t}\int_0^t \mathbb{E}\Big[|\mathcal{X}_s - \tilde{\mathcal{X}}_s|^2\Big]\,\mathrm{d}s$, integrate the previous inequality + some Fubini:

$$g(t) \leqslant \frac{\beta}{\alpha}\frac{1}{t}\int_0^t g(s)\,\mathrm{d}s + \dots$$

say equality holds with $\dots = 0$, "$t\,g(t) = \frac{\beta}{\alpha}\int_0^t g(s)\,\mathrm{d}s$", $g(t) \sim t^{\frac{\beta}{\alpha}-1}$ $(\frac{\beta}{\alpha} - 1 < 0)$

Convergence of $\mathbb{E}\left[\mathcal{W}_2^2(\nu^\star, \nu_t^{\mathcal{X}})\right]$

(*have in mind:* $\nu_t^Y := \frac{1}{t}\int_0^t \delta_{Y_s}\,\mathrm{d}s$ for a process Y)

- Classically, we observe: $X^\star$ is a perturbed SID...

$$X_t^\star = X_0^\star + \int_0^t \left(b(X_s^\star, \nu_s^{X^\star}) + \eta_s^b\right)\,\mathrm{d}s + \int_0^t \left(\sigma(X_s^\star, \nu_s^{X^\star}) + \eta_s^\sigma\right)\,\mathrm{d}W_s,$$

forcing $\eta_s^b = b(X_s^\star, \nu^\star) - b(X_s^\star, \nu_s^{X^\star})$, $\eta_s^\sigma = \sigma(X_s^\star, \nu^\star) - \sigma(X_s^\star, \nu_s^{X^\star})$.

- The Lipschitz assumption leads to

$$\mathbb{E}[|\eta_t^b|^2 + |\eta_t^\sigma|^2] \leqslant C\mathbb{E}\left[\mathcal{W}_2^2(\nu^\star, \nu_t^{X^\star})\right] = O(t^{-\zeta}) \text{ (not obvious!)}.$$

- Then stability yields, for any $\epsilon > 0$,

$$\mathbb{E}\left[\mathcal{W}_2^2(\nu_t^{X^\star}, \nu_t^{\mathcal{X}})\right] \leqslant C_\epsilon\left(t^{-\vartheta^\star+\epsilon}\mathbb{E}\left[|X_0^\star - \mathcal{X}_0|^2\right] + t^{-\zeta}\right).$$

- and finally

$$\mathbb{E}\left[\mathcal{W}_2^2(\nu^\star, \nu_t^{\mathcal{X}})\right] \leqslant 2\mathbb{E}\left[\mathcal{W}_2^2(\nu^\star, \nu_t^{\mathcal{X}}) + \mathcal{W}_2^2(\nu^\star, \nu_t^{\mathcal{X}})\right] \leqslant C_\epsilon\left(t^{-\vartheta^\star+\epsilon}\mathbb{E}\left[|X_0^\star - \mathcal{X}_0|^2\right] + t^{-\zeta}\right)$$

Convergence of $\mathbb{E}\left[\mathcal{W}_2^2(\nu_{\Gamma_n}^{\mathcal{X}}, \bar{\nu}_{\Gamma_n})\right]$

(*have in mind:* (γ_n) decreasing time step for the scheme, $\Gamma_n = \sum_{k=1}^n \gamma_k$)

- We have $\mathcal{W}_2(\nu_{\Gamma_n}^{\mathcal{X}}, \bar{\nu}_{\Gamma_n}) \leqslant \mathcal{W}_2(\nu_{\Gamma_n}^{\mathcal{X}}, \nu_{\Gamma_n}^{\bar{\mathcal{X}}}) + \mathcal{W}_2(\nu_{\Gamma_n}^{\bar{\mathcal{X}}}, \bar{\nu}_{\Gamma_n})$

▸ Second term is a time discretisation error for the integral namely

$$\mathbb{E}\left[\mathcal{W}_2^2(\nu_{\Gamma_n}^{\bar{\mathcal{X}}}, \bar{\nu}_{\Gamma_n})\right] \leqslant \frac{1}{\Gamma_n}\int_0^{\Gamma_n} \mathbb{E}\left[|\bar{\mathcal{X}}_s - \bar{\mathcal{X}}_{\underline{s}}|^2\right] \mathrm{d}s$$

▸ Two key properties of $\bar{\mathcal{X}}$ (under $(HMV)_{p,K',\alpha',\beta'}$):

1. $\sup_{t\geqslant 0} \mathbb{E}\left[|\bar{\mathcal{X}}_t|^{2p}\right] \leqslant C$ (from tedious computations)
2. $\mathbb{E}\left[|\bar{\mathcal{X}}_t - \bar{\mathcal{X}}_{\underline{t}}|^{2p}\right]^{\frac{1}{p}} \leqslant C(t - \underline{t})$ (comes from the previous point and Lipschitz assumption)

$\hookrightarrow$ And then: $\mathcal{W}_2^2(\nu_{\Gamma_n}^{\bar{\mathcal{X}}}, \bar{\nu}_{\Gamma_n}) = O\left(\frac{1}{\Gamma_n^{1-\varrho}} \sum_{k=1}^n \frac{\gamma_k^2}{\Gamma_k^{\varrho}}\right)$

▸ For $\mathcal{W}_2(\nu_{\Gamma_n}^{\mathcal{X}}, \nu_{\Gamma_n}^{\bar{\mathcal{X}}})$: See $\bar{\mathcal{X}}$ (continuous Euler scheme) as a perturbed SID and use *stability property*

$$\mathbb{E}\left[\mathcal{W}_2(\nu_{\Gamma_n}^{\mathcal{X}}, \nu_{\Gamma_n}^{\bar{\mathcal{X}}})\right] = O_{\varrho}\left((\Gamma_n)^{\varrho-1}\sum_{k=1}^{n-1}\frac{\gamma_k^2}{\Gamma_k^{\varrho}}\right)$$

Global convergence (L^2-case)

- Finally, we obtain $\mathbb{E}\left[\mathcal{W}_2^2(\nu^{\mathcal{X}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})\right] = O_\varrho\left((\Gamma_n)^{\varrho-1}\sum_{k=1}^{n-1}\frac{\gamma_k^2}{\Gamma_k^\varrho}\right)$.
- Set $\gamma_n = cn^{-r}$ for $r \in (0,1)$ and $r^\star = \frac{\vartheta^\star}{1+\vartheta^\star}$ then

$$\mathbb{E}\left[\mathcal{W}_2^2(\nu^{\mathcal{X}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})\right] \leqslant C_\eta \begin{cases} \gamma_n & \text{if } r < r^\star \\ \gamma_n^{(\frac{1}{r}-1)(1-\frac{\beta^\star}{\alpha^\star})-\eta} & \text{if } r \geqslant r^\star \text{ for every (small) } \eta > 0 \end{cases}$$

$\hookrightarrow$ the 'optimal' rate is $r^\star$.

- Since $\mathcal{W}_2(\nu^\star, \bar{\nu}_{\Gamma_n}) \leqslant \mathcal{W}_2(\nu^\star, \nu^{\mathcal{X}}_{\Gamma_n}) + \mathcal{W}_2(\nu^{\mathcal{X}}_{\Gamma_n}, \bar{\nu}_{\Gamma_n})$, one gets

$$\mathbb{E}\left[\mathcal{W}_2^2(\nu^\star, \bar{\nu}_{\Gamma_n})\right] \leqslant Cn^{-\zeta} + C_\eta(1 + \mathcal{W}_2(\nu^\star, [\mathcal{X}_0]))\, n^{-r^\star+\eta}$$

for any (small) η.

Conclusion

- We show that the Euler scheme of the self-interacting diffusion converges to the stationary measure of the associated MKV SDE
- The convergence is obtained with a rate for $\mathcal{W}_2$-distance in the L^2 and almost sure case.
- All the rates are suboptimal...
- What happens when there is more than one stationary measure? (in some special cases, recall Ornstein-Uhlenbeck example, [KK$^+$12]) shows convergence to a random limit for the self-interacting diffusion.)

References I

Houssam AlRachid, Mireille Bossy, Cristiano Ricci, and Lukasz Szpruch, *New particle representations for ergodic mckean-vlasov sdes*, ESAIM: Proceedings and Surveys **65** (2019), 68–83.

Jean-François Chassagneux and Gilles Pagès, *Computing the invariant distribution of mckean-vlasov sdes by ergodic simulation*, arXiv preprint arXiv:2406.13370 (2024).

Kai Du, Yifan Jiang, and Jinfeng Li, *Empirical approximation to invariant measures for mckean–vlasov processes: Mean-field interaction vs self-interaction*, Bernoulli **29** (2023), no. 3, 2492–2518.

Kai Du, Zhenjie Ren, Florin Suciu, and Songbo Wang, *Self-interacting approximation to mckean-vlasov long-time limit: a markov chain monte carlo method*, arXiv preprint arXiv:2311.11428 (2023).

Victor Kleptsyn, Aline Kurtzmann, et al., *Ergodicity of self-attracting motion*, Electronic Journal of Probability **17** (2012).

References II

Damien Lamberton and Gilles Pagès, *Recursive computation of the invariant distribution of a diffusion*, Bernoulli **8** (2002), no. 3, 367–405.

Damien Lamberton and Gilles Pagès, *Recursive computation of the invariant distribution of a diffusion: The case of a weakly mean reverting drift*, Stoch. Dyn. **3** (2003), no. 4, 435–451 (English).