Couplings for non-linear perturbations of Markov processes via jump mechanisms

Pierre Monmarché

LJLL & LCT, Sorbonne Université

CIRM, september 20, 2018

1 The basic argument

2 Extensions

The non-linear equation

Given

- E a Polish space
- L the generator of a Feller Markov process on E
- For all $\nu \in \mathcal{P}(E)$,
  - $\lambda_\nu : E \to \mathbb{R}_+$ a jump rate
  - $Q_\nu : E \to \mathcal{P}(E)$ a Markov kernel,

and denoting A' the dual on measures of an operator A on functions,

$$(A'\nu)f \ := \ \nu(Af)\,,$$

we are interested in $t \mapsto m_t \in \mathcal{P}(E)$ that solves, in some sense,

$$\partial_t m_t \ = \ L' m_t + Q'_{m_t}\left(\lambda_{m_t} m_t\right) - \lambda_{m_t} m_t\,.$$

Perturbative regime : small non-linearity, long-time behaviour.

The non-linear process

Given $t \mapsto \mu_t \in \mathcal{P}(E)$ measurable and $m_0 \in \mathcal{P}(E)$,

- Draw $X_0 \sim m_0$
- Let $(\tilde{X}_t)_{t \geqslant 0}$ be a Markov process associated to L with $\tilde{X}_0 = X_0$.
- For $F \sim \mathcal{E}(1)$ independent from $(\tilde{X}_t)_{t \geqslant 0}$, let

$$T \;=\; \inf\left\{ t > 0,\; F > \int_0^t \lambda_{\mu_s}(\tilde{X}_s)\mathrm{d}s \right\}$$

- Set $X_t = \tilde{X}_t$ for all $t \in [0, T)$ and draw $X_T \sim Q_{\mu_T}(\tilde{X}_T)$
- Repeat

The non-linear process

Given $t \mapsto \mu_t \in \mathcal{P}(E)$ measurable and $m_0 \in \mathcal{P}(E)$,

- Draw $X_0 \sim m_0$
- Let $(\tilde{X}_t)_{t\geqslant 0}$ be a Markov process associated to L with $\tilde{X}_0 = X_0$.
- For $F \sim \mathcal{E}(1)$ independent from $(\tilde{X}_t)_{t\geqslant 0}$, let

$$T \;=\; \inf\left\{t > 0,\ F > \int_0^t \lambda_{\mu_s}(\tilde{X}_s)\mathrm{d}s\right\}$$

- Set $X_t = \tilde{X}_t$ for all $t \in [0, T)$ and draw $X_T \sim Q_{\mu_T}(\tilde{X}_T)$
- Repeat

Then $m_t = \mathcal{L}aw(X_t) := \Psi(t, (\mu_s)_{s\geqslant 0}, m_0)$ formally solves

$$\partial_t m_t \;=\; L' m_t + Q'_{\mu_t}\left(\lambda_{\mu_t} m_t\right) - \lambda_{\mu_t} m_t\,.$$

Definition

We say $t \mapsto m_t$ is a solution of the non linear equation if for all $t \geqslant 0$ $m_t = \Psi(t, (m_s)_{s\geqslant 0}, m_0)$.

Motivations

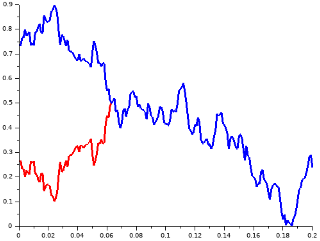
- Biological modelling (neurons, chemotaxy)
- mean-field games
- stochastic algorithms with selection/mutation, adaptive biasing force, Flemming-Viot, etc.
- MC, PDMP or hybrid diffusion/jump sampler for mean-field equation

Mirror coupling for the heat equation

Let $(B_t)_{t\geqslant 0}$ be a 1-d Brownian motion, $x, y \in \mathbb{R}/\mathbb{Z}$.

$$X_t = x + B_t \qquad \forall t \geqslant 0$$

$$Y_t = \begin{cases} y - B_t & \text{for } t \leqslant \tau_{merge} := \inf\{t > 0,\ 2B_t = y - x\} \\ X_t & \text{for } t > \tau_{merge}\,. \end{cases}$$



Denoting $P_t = e^{t\Delta}$ the heat semigroup on $\mathbb{R}/\mathbb{Z}$, then

$$X_t \sim \delta_x P_t\,, \qquad Y_t \sim \delta_y P_t\,, \qquad \{X_t = Y_t\} = \{t \geqslant \tau_{merge}\}.$$

Mirror coupling for the heat equation

Consider the total variation distance

$$\|\nu - \mu\|_{TV} = \sup_{\|f\|_\infty \leqslant 1} |\nu f - \mu f| = 2 \inf_{\Pi(\mu,\nu)} \mathbb{P}(X \neq Y) .$$

with $\Pi(\mu,\nu) := \{(X,Y), X \sim \mu, Y \sim \nu\}$ the set of couplings of μ and ν.

Let (X_t, Y_t) be the mirror coupling starting at an optimal coupling (X_0, Y_0) of $\mu, \nu \in \mathcal{P}(\mathbb{R}/\mathbb{Z})$.

$$\begin{aligned} \|\nu P_t - \mu P_t\|_{TV} &\leqslant 2\mathbb{P}(X_t \neq Y_t) \\ &\leqslant 2\mathbb{P}(X_0 \neq Y_0)\mathbb{P}(\tau_{merge} > t \mid X_0 \neq Y_0) \\ &\leqslant \|\nu - \mu\|_{TV} (1 - \alpha(t)) \end{aligned}$$

with

$$\alpha(t) = \mathbb{P}\left(\exists s \leqslant t \text{ s.t. } 2B_s \in \{-1, 1\}\right) > 0$$

Coupling for the non-linear process

Consider the non-linear process given by L and $\mu \mapsto \lambda_\mu$, Q_μ.

Assumption

There exists $\lambda_* > 0$ such that for all $\mu \in \mathcal{P}(E)$ and $x \in E$, $\lambda_\mu(x) \leqslant \lambda_*$.

For $(\tilde{X}_t, \tilde{Y}_t)_{t\geqslant 0}$ such that $(\tilde{X}_0, \tilde{Y}_0)$ is an optimal coupling of $m_0, h_0 \in \mathcal{P}(E)$ and each marginal is a Markov process with generator L, let

$$\tilde{\tau}_{merge} = \inf\{t > 0,\ \tilde{X}_t = \tilde{Y}_t\}\,.$$

Let $(X_t)_{t\geqslant 0}$ be a non-linear process with $X_t = \tilde{X}_t$ up to time

$$T \ = \ \inf\left\{t > 0,\ F > \int_0^t \lambda_{m_s}(\tilde{X}_s)\mathrm{d}s\right\} \ > \ \frac{1}{\lambda_*}F\,.$$

Then

$$\mathbb{P}\left(X_t = Y_t\right) \ \geqslant \ \mathbb{P}\left(\tilde{X}_t = \tilde{Y}_t \ \&\ F > \lambda_* t\right) \ > \ \tilde{\alpha}(t)e^{-\lambda_* t}\,.$$

Coupling for the non-linear process

If $t \mapsto m_t, h_t$ be two solutions of the non-linear equation,

$$\begin{aligned} \|m_t - h_t\|_{TV} &\leqslant 2\mathbb{P}\left(X_t \neq Y_t\right) \\ &\leqslant 2\mathbb{P}\left(X_0 \neq Y_0\right)\left(1 - \tilde{\alpha}(t)e^{-\lambda_* t}\right) + 2\mathbb{P}\left(X_t \neq Y_t \ \&\ X_0 = Y_0\right) \\ &\leqslant \|m_0 - h_0\|_{TV}(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\mathbb{P}\left(X_t \neq Y_t \mid X_0 = Y_0\right) \end{aligned}$$

Even if $X_t = Y_t$ at some point, still, $m_t \neq h_t$, hence the dynamics of X and Y are different: they may split.

Parallel coupling for diffusions

For diffusions, consider

$$\begin{aligned} \mathrm{d}X_t &= a(X_t, t)\mathrm{d}t + \mathrm{d}B_t \\ \mathrm{d}Y_t &= b(Y_t, t)\mathrm{d}t + \mathrm{d}B_t \\ \Rightarrow \mathrm{d}(X_t - Y_t) &= \left(a(X_t, t) - b(Y_t, t)\right)\mathrm{d}t\,. \end{aligned}$$

For instance, if $X_t = (X_{i,t})_{i\in[\![1,N]\!]}$ are N independent McKean-Vlasov processes and $Y_t = (Y_{i,t})_{i\in[\![1,N]\!]}$ an associated cloud of mean-field interacting particle, we get

$$\mathbb{E}\left(|X_{1,t} - Y_{1,t}|^2\right) \leqslant \frac{e^{bt}}{N}$$

Synchronous coupling for jump processes

M' 2018, Durmus-Guillin-M' 2018, consider two Markov semi-groups

$$L_{i,t}f(x) \quad = \quad Lf(x) + \lambda_{i,t}(x)\left(Q_{i,t}f(x) - f(x)\right), \qquad i = 1,2\,.$$

Given an initial condition $x \in E$, define $(X_t^1, X_2)_{t\geqslant 0}$ with $X_0^1 = X_0^2 = x$ in such a way that

- $(X_t^i)_{t\geqslant 0}$ is a Markov process with generator $L_{i,t}$.
- As much as possible, the processes jump at the same time.
- If they jump at the same time, as much as possible, they jump to the same point.

Set $\tau_{split} = \inf\{t \geqslant 0,\ X_t^1 \neq X_t^2\}$ and

$$\eta(t) \ := \ \sup_{\|f\|_\infty \leqslant 1} \|L_{1,t}f - L_{2,t}f\|_\infty\,.$$

Then

$$\|\delta_x P_{0,t}^1 - \delta_x P_{0,t}^2\|_{TV} \ \leqslant \ 2\mathbb{P}\left(\tau_{split} < t\right) \ \leqslant \ 2\left(1 - e^{-\int_0^t \eta(s)\mathrm{d}s}\right)\,.$$

Synchronous coupling for jump processes

For $i = 1, 2$, consider $t \mapsto m_t^i$ two solutions of the non-linear equation.

Assumption

There exist $\theta > 0$ such that for all $\nu, \mu \in \mathcal{P}(E)$,

$$\sup_{\|f\|_\infty \leqslant 1} \|\lambda_\nu(Q_\nu f - f) - \lambda_\mu(Q_\mu f - f)\|_\infty \leqslant \theta \|\nu - \mu\|_{TV}$$

Ex: $Q_\nu = Q$ for all ν and $\lambda_\nu = \int g \mathrm{d}\nu$ for some bounded positive g.

From the synchronous coupling of processes with generator

$$L_{i,t} = L + \lambda_{m_t^i} \left(Q_{m_t^i} - Id \right) ,$$

we get

$$\begin{aligned} \|m_t^1 - m_t^2\|_{TV} &\leqslant \|m_0^1 - m_0^2\|_{TV} + 2 \left(1 - e^{-\theta \int_0^t \|m_s^1 - m_s^2\|_{TV} \mathrm{d}s} \right) \\ &\leqslant \|m_0^1 - m_0^2\|_{TV} + 2\theta \int_0^t \|m_s^1 - m_s^2\|_{TV} \mathrm{d}s \end{aligned}$$

Contraction of the TV for non-linear processes

$$
\begin{aligned}
\|m_t^1 - m_t^2\|_{TV} &\leqslant \|m_0^1 - m_0^2\|_{TV}(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\mathbb{P}\left(\tau_{split} < t | X_0 = Y_0\right) \\
&\leqslant \|m_0^1 - m_0^2\|_{TV}(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\theta \int_0^t \|m_s^1 - m_s^2\|_{TV} ds \\
&\leqslant \|m_0^1 - m_0^2\|_{TV}(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\theta \int_0^t e^{\theta s}\|m_0^1 - m_0^2\|_{TV} \\
&\leqslant \|m_0^1 - m_0^2\|_{TV}\left(1 - \tilde{\alpha}(t)e^{-\lambda_* t} + 2(e^{\theta t} - 1)\right)
\end{aligned}
$$

If there exist t_0 such that $\tilde{\alpha}(t_0) > 0$, given θ is small enough,

$$\|m_{t_0}^1 - m_{t_0}^2\|_{TV} \leqslant \kappa \|m_0^1 - m_0^2\|_{TV}$$

for some $\kappa \in (0,1)$ and then for all $t \geqslant 0$,

$$\|m_t^1 - m_t^2\|_{TV} \leqslant e^{\theta t_0} \kappa^{\lfloor t/t_0 \rfloor} \|m_0^1 - m_0^2\|_{TV} .$$

(Similar result in Cañizo-Yoldaș 2018 for a neuron model)

Exemple

Let $E=(\mathbb{R}/\mathbb{Z})^d$, $U\in\mathcal{C}(E)$, $W\in\mathcal{C}(E^2)$, $\beta>0$, find V such that

$$V(x) \quad = \quad U(x)+\frac{\int_E W(x,z)e^{-\beta V(z)}\mathrm{d}z}{\int_E e^{-\beta V(z)}\mathrm{d}z}\,. \tag{1}$$

Proposition

Suppose that β is small enough so that

$$\rho \quad := \quad e^{-\beta(osc(U)+osc(W))}-4osc(W)\beta\left(e^{\beta(osc(U)+osc(W))}-1\right) \;>\; 0\,,$$

with $osc(f)=\max f-\min f$. Then (1) admits a unique solution.

Proof: consider a (non-linear) Metropolis-Hasting chain with uniform proposal with target $\propto\exp(-\beta(U+m_t(V)))$. Then

$$\|m_t^1-m_t^2\|_{TV} \;\leqslant\; e^{-\rho t}\|m_0^1-m_0^2\|_{TV}\,,$$

and V solves (1) iff $\exp(-\beta V)$ is an equilibrium of the non-linear chain.

1 The basic argument

2 Extensions

With a Lyapunov function

In the previous argument, we used that there exists $t_0, \alpha > 0$ such that

$$\sup_{x,y \in E} \|\delta_x e^{t_0 L} - \delta_y e^{t_0 L}\|_{TV} \leqslant 2(1-\alpha)\,.$$

Less restrictive assumption (Foster-Lyapunov + local Doeblin) :

- There exists $\mathcal{V} : E \to [1,\infty)$, $M \geqslant 0$ and $\gamma \in (0,1)$ such that

$$e^{t_0 L}\mathcal{V}(x) \leqslant \gamma \mathcal{V}(x) + (1-\gamma) M\,.$$

- For all $x, y \in E$ with $\mathcal{V}(x) + \mathcal{V}(y) \leqslant 4M$

$$\|\delta_x e^{t_0 L} - \delta_y e^{t_0 L}\|_{TV} \leqslant 2(1-\alpha)\,.$$

Then, $e^{t_0 L}$ contracts (a norm equivalent to)

$$\|\mu - \nu\|_{\mathcal{V}} = \int \mathcal{V}(x)|\mu-\nu|(\mathrm{d}x) = \inf_{(X,Y)\in\Pi(\mu,\nu)} \mathbb{E}\left(\mathbb{1}_{X\neq Y}\left(\mathcal{V}(X)+\mathcal{V}(Y)\right)\right)$$

(Meyn-Tweedie 1993, Hairer-Mattingly 2008)

Non-linear analoguous

Assumption

- *Lyapunov-type condition : there exists* $\mathcal{V} : E \to [1,\infty)$, $M,\rho,\rho_*,\gamma_* > 0$ *such that for all* $t \geqslant 0$, $x \in E$, $\mu \in \mathcal{P}(E)$ *and* $t \mapsto m_t$ *solution of the non-linear equation,*

$$\begin{aligned} m_t(\mathcal{V}) &\leqslant e^{-\rho t} m_0(\mathcal{V}) + \left(1 - e^{-\rho t}\right) M \\ e^{tL}\mathcal{V}(x) &\leqslant e^{\rho_* t}\mathcal{V}(x) \\ \mathcal{V}(Y) &\leqslant \gamma_* V(x) \qquad \textit{almost surely if } Y \sim Q_\mu(x) \end{aligned}$$

- *(Linear) local Doeblin condition :* $\forall C > 0$, $\exists \alpha, t_0 > 0$ *such that*

$$\mathcal{V}(x) + \mathcal{V}(y) \leqslant C \qquad \Rightarrow \qquad \|\delta_x e^{t_0 L} - \delta_y e^{t_0 L}\|_{TV} \leqslant 2(1-\alpha)\,.$$

- *Control on the non-linearity : there exists* $\theta > 0$ *such that*

$$\sup_{\|f\|_\infty \leqslant 1} \|\lambda_\nu(Q_\nu f - f) - \lambda_\mu(Q_\mu f - f)\|_\infty \leqslant \theta \|\nu - \mu\|_{\mathcal{V}}$$

Non-linear analoguous

Theorem

Under the previous assumption, if $m_0(\mathcal{V}) < \infty$ then a unique solution $t \mapsto m_t$ of the non-linear equaiont exists. Moreover, there exist $C, r, \theta_* > 0$ such that if $\theta \leqslant \theta_*$ then there exists a unique equilibrium m_∞ and for all $t \geqslant 0$ and $m_0 \in \mathcal{P}(E)$ then

$$\|m_t - m_\infty\|_{\mathcal{V}} \ \leqslant \ Ce^{-rt} m_0(\mathcal{V}) \,.$$

Proof :

- 33% contraction of Hairer-Mattingly
- 33% synchronous coupling
- 33% elbow grease
- 1% faith

Interacting particle system

Let $N \in \mathbb{N}$ and for all $i \in [\![1, N]\!]$ consider L_i the generator of a Feller process on E, $\lambda_i : E^N \to \mathbb{R}_+$ and $Q_i : E^N \to \mathcal{P}(E)$. Define a Markov process $(X_{i,t})_{i\in[\![1,N]\!],t\geqslant 0}$ by:

- Let $(\tilde{X}_{i,t})_{t\geqslant 0}$ be a Markov process associated to a generator L_i
- For $F_i \sim \mathcal{E}(1)$, let $T = \inf\{T_i,\ i \in [\![1, N]\!]\}$ with

$$T_i \ = \ \inf\left\{t > 0,\ F_i > \int_0^t \lambda_i(\tilde{X}_s)\mathrm{d}s\right\}$$

- Set $X_{i,t} = \tilde{X}_{i,t}$ for all $t \in [0, T)$.
- If $T = T_i$, draw $X_{i,T} \sim Q_i(X_T)$. Else, set $X_{i,T} = \tilde{X}_{i,T}$.
- Repeat

Exemple: if non-linear $\mu \mapsto Q_\mu, \lambda_\mu$, then $\lambda_i(x) = \lambda_{\mu_N(x)}(x_i)$ with

$$\mu_N(x) \ := \ \frac{1}{N}\sum_{i=1}^N \delta_{x_i}\,.$$

Interacting particle system (total variation)

Assumption

There exists $\lambda_*, \theta, t_0, \alpha \geqslant 0$ *such that, for all* $i \in [\![1, N]\!]$*,*

- $\lambda_i(x) \leqslant \lambda_*$ *for all* $x \in E^N$ *and* $i \in [\![1, N]\!]$*.*
- *for all* $x, y \in E$ *with* $y_i = x_i$*,*

$$\|\lambda_i(x)\left(Q_i(x) - \delta_{x_i}\right) - \lambda_i(y)\left(Q_i(y) - \delta_{y_i}\right)\|_\infty \quad \leqslant \quad \frac{\theta}{N}\sum_{j=1}^N \mathbb{1}_{x_j \neq y_j}\,.$$

- *for all* $x, y \in E$*,*

$$\|\delta_x e^{t_0 L_i} - \delta_y e^{t_0 L_i}\|_{TV} \quad \leqslant \quad 2(1-\alpha)\,.$$

Remark :

$$\mathbb{P}\left(\text{the } N \text{ particles all merge at time } t_0\right) \geqslant \left(\alpha e^{-\lambda_* t_0}\right)^N.$$

Interacting particle system (total variation)

Theorem

Under the previous assumption, denote $(P_t^{(N)})_{t\geqslant 0}$ *the semi-group associated to* $(X_{i,t})_{i\in[\![1,N]\!],t\geqslant 0}$*, then*

$$\|m_0 P_t^{(N)} - h_0 P_t^{(N)}\|_{TV} \quad \leqslant \quad N e^{\theta t_0}\left(e^{\theta t_0} - \alpha e^{-\lambda_* t_0}\right)^{\lfloor t/t_0 \rfloor} \|m_0 - h_0\|_{TV} \, .$$

Moreover, denoting m'_t *and* h'_t *the respective laws of* $X_{I,t}$ *and* $Y_{I,t}$ *where* $X_t \sim m_0 P_t^{(N)}$, $Y_t \sim h_0 P_t^{(N)}$ *and* I *is uniformly distributed over* $[\![1,N]\!]$ *and independent from* X_t *and* Y_t*, then*

$$\|m'_t - h'_t\|_{TV} \quad \leqslant \quad e^{\theta t_0}\left(e^{\theta t_0} - \alpha e^{-\lambda_* t_0}\right)^{\lfloor t/t_0 \rfloor} \|m'_0 - h'_0\|_{TV} \, .$$

Proof : couple particle independently. Try to keep together those who have already merged with synchronous coupling.

Interacting particle system ($\mathcal{V}$-norm)

Theorem

Again,

- $\lambda_i(x) \leqslant \lambda_*$.
- *Lyapunov-type condition (similar as the non-linear case)*
- *Local Doeblin condition (same)*
- *for all* $x, y \in E$ *with* $y_i = x_i$,

$$\|\lambda_i(x)\left(Q_i(x) - \delta_{x_i}\right) - \lambda_i(y)\left(Q_i(y) - \delta_{y_i}\right)\|_\infty \quad \leqslant \quad \frac{\theta}{N}\sum_{j=1}^N \mathbb{1}_{x_j \neq y_j}\,.$$

Then, nice decay in the $\mathcal{V}$*-norm.*

The analoguous of the non-linear theorem would be

$$lhs \quad \leqslant \quad \frac{\theta}{N}\sum_{j=1}^N \mathbb{1}_{x_j \neq y_j}\left(\mathcal{V}_j(x_j) + \mathcal{V}_j(y_j)\right)\,.$$

To go further

- Self-interaction processes (memory, delay)
- Interactions more general than mean-field
- Link with the sticky coupling for diffusions (via Euler discretization)
- Solve the problem of the $\mathcal{V}_i(X_{i,t})$ in the interacting particle case
- Propagation of chaos (coupling between interacting particle and independent non-linear processes)

To go further

- Self-interaction processes (memory, delay)
- Interactions more general than mean-field
- Link with the sticky coupling for diffusions (via Euler discretization)
- Solve the problem of the $\mathcal{V}_i(X_{i,t})$ in the interacting particle case
- Propagation of chaos (coupling between interacting particle and independent non-linear processes)

Thank you for your attention !