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,

$$\partial_t m_t = L'm_t + Q'_m_t (\lambda m_t m_t) - \lambda m_t m_t.$$ 

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 $(\tilde{X}_t)_{t \geq 0}$ be a Markov process associated to $L$ with $\tilde{X}_0 = X_0$.
- For $F \sim \mathcal{E}(1)$ independent from $(\tilde{X}_t)_{t \geq 0}$, let
  \[ T = \inf \left\{ t > 0, \ F \geq \int_0^t \lambda_{\mu_s}(\tilde{X}_s)ds \right\} \]
- 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 $(\tilde{X}_t)_{t \geq 0}$ be a Markov process associated to $L$ with $\tilde{X}_0 = X_0$.
- For $F \sim \mathcal{E}(1)$ independent from $(\tilde{X}_t)_{t \geq 0}$, let

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

- 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 $m_t = \text{Law}(X_t) := \Psi(t, (\mu_s)_{s \geq 0}, m_0)$ formally solves

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

**Definition**

We say $t \mapsto m_t$ is a solution of the non-linear equation if for all $t \geq 0$

$$m_t = \Psi(t, (m_s)_{s \geq 0}, m_0).$$
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 \((B_t)_{t\geq 0}\) be a 1-d Brownian motion, \(x, y \in \mathbb{R}/\mathbb{Z}\).

\[
X_t = x + B_t \quad \forall t \geq 0
\]

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

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

\[
X_t \sim \delta_x P_t, \quad Y_t \sim \delta_y P_t, \quad \{X_t = Y_t\} = \{t \geq \tau_{\text{merge}}\}.
\]
Mirror coupling for the heat equation

Consider the total variation distance

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

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

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

$$\|\nu P_t - \mu P_t\|_{TV} \leq 2P(X_t \neq Y_t) \leq 2P(X_0 \neq Y_0)P(\tau_{\text{merge}} > t \mid X_0 \neq Y_0) \leq \|\nu - \mu\|_{TV} (1 - \alpha(t))$$

with

$$\alpha(t) = P(\exists s \leq t \text{ s.t. } 2B_s \in \{-1, 1\}) > 0$$
Coupling for the non-linear process

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

**Assumption**

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

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

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

Then

$$\mathbb{P}(X_t = Y_t) \geq \mathbb{P}\left(\tilde{X}_t = \tilde{Y}_t & F > \lambda_* t\right) > \tilde{\alpha}(t)e^{-\lambda_* t}.$$
Coupling for the non-linear process

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

$$\|m_t - h_t\|_{TV} \leq 2\mathbb{P}(X_t \neq Y_t)$$

$$\leq 2\mathbb{P}(X_0 \neq Y_0)(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\mathbb{P}(X_t \neq Y_t \& X_0 = Y_0)$$

$$\leq \|m_0 - h_0\|_{TV}(1 - \tilde{\alpha}(t)e^{-\lambda_* t}) + 2\mathbb{P}(X_t \neq Y_t \mid X_0 = Y_0)$$

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

\[ \frac{dX_t}{dt} = a(X_t, t)dt + dB_t \]
\[ \frac{dY_t}{dt} = b(Y_t, t)dt + dB_t \]

\[ \Rightarrow \frac{d(X_t - Y_t)}{dt} = (a(X_t, t) - b(Y_t, t))dt. \]

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

\[ \mathbb{E} \left( |X_{1,t} - Y_{1,t}|^2 \right) \leq \frac{e^{bt}}{N} \]
Synchronous coupling for jump processes

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

\[ L_{i,t} f(x) = L f(x) + \lambda_{i,t}(x) (Q_{i,t} f(x) - f(x)) , \quad i = 1, 2. \]

Given an initial condition \( x \in E \), define \((X^1_t, X^2_t)_{t \geq 0}\) with \( X^1_0 = X^2_0 = x \) in such a way that

- \( (X^i_t)_{t \geq 0} \) 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 \( \tau_{\text{split}} = \inf \{ t \geq 0, \ X^1_t \neq X^2_t \} \) and

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

Then

\[ \| \delta_x P^1_{0,t} - \delta_x P^2_{0,t} \|_{\text{TV}} \leq 2 \mathbb{P} (\tau_{\text{split}} < t) \leq 2 \left( 1 - e^{-\int_0^t \eta(s) ds} \right). \]
Synchronous coupling for jump processes

For \( i = 1, 2 \), consider \( t \mapsto m^i_t \) two solutions of the non-linear equation.

Assumption

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

\[
\sup_{\|f\|_\infty \leq 1} \left\| \lambda_\nu(Q_\nu f - f) - \lambda_\mu(Q_\mu f - f) \right\|_\infty \leq \theta \|\nu - \mu\|_{TV}
\]

Ex: \( Q_\nu = Q \) for all \( \nu \) and \( \lambda_\nu = \int g d\nu \) for some bounded positive \( g \).

From the synchronous coupling of processes with generator

\[
L_{i,t} = L + \lambda_{m^i_t} (Q_{m^i_t} - Id)
\]

we get

\[
\| m^1_t - m^2_t \|_{TV} \leq \| m^1_0 - m^2_0 \|_{TV} + 2 \left( 1 - e^{-\theta \int_0^t \| m^1_s - m^2_s \|_{TV} ds} \right) 
\]

\[
\leq \| m^1_0 - m^2_0 \|_{TV} + 2 \theta \int_0^t \| m^1_s - m^2_s \|_{TV} ds
\]
Contraction of the TV for non-linear processes

\[ \| m_t^1 - m_t^2 \|_{TV} \leq \| m_0^1 - m_0^2 \|_{TV} \left( 1 - \tilde{\alpha}(t)e^{-\lambda_* t} \right) + 2\mathbb{P}(\tau_{\text{split}} < t | X_0 = Y_0) \]

\[ \leq \| m_0^1 - m_0^2 \|_{TV} \left( 1 - \tilde{\alpha}(t)e^{-\lambda_* t} \right) + 2\theta \int_0^t \| m_s^1 - m_s^2 \|_{TV} ds \]

\[ \leq \| m_0^1 - m_0^2 \|_{TV} \left( 1 - \tilde{\alpha}(t)e^{-\lambda_* t} \right) + 2\theta \int_0^t e^{\theta s} \| m_0^1 - m_0^2 \|_{TV} ds \]

\[ \leq \| m_0^1 - m_0^2 \|_{TV} \left( 1 - \tilde{\alpha}(t)e^{-\lambda_* t} + 2(e^{\theta t} - 1) \right) \]

If there exist \( t_0 \) such that \( \tilde{\alpha}(t_0) > 0 \), given \( \theta \) is small enough, \[
\| m_{t_0}^1 - m_{t_0}^2 \|_{TV} \leq \kappa \| m_0^1 - m_0^2 \|_{TV}
\]

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

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

(Similar result in Cañizo-Yoldaş 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

$$V(x) = U(x) + \frac{\int_E W(x,z) e^{-\beta V(z)} dz}{\int_E e^{-\beta V(z)} dz}.$$  \hfill (1)

Proposition

Suppose that $\beta$ is small enough so that

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

with $\text{osc}(f) = \max f - \min f$. 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

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

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

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

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

- There exists $\mathcal{V}: E \rightarrow [1, \infty)$, $M \geq 0$ and $\gamma \in (0, 1)$ such that
  $$e^{t_0 L} \mathcal{V}(x) \leq \gamma \mathcal{V}(x) + (1 - \gamma)M.$$

- For all $x, y \in E$ with $\mathcal{V}(x) + \mathcal{V}(y) \leq 4M$
  $$\| \delta_x e^{t_0 L} - \delta_y e^{t_0 L} \|_{TV} \leq 2(1 - \alpha).$$

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

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

(Meyn-Tweedie 1993, Hairer-Mattingly 2008)
Non-linear analogous

Assumption

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

\[
m_t(\mathcal{V}) \leq e^{-\rho t} m_0(\mathcal{V}) + (1 - e^{-\rho t}) M
\]

\[
e^{tL} \mathcal{V}(x) \leq e^{\rho^* t} \mathcal{V}(x)
\]

\[
\mathcal{V}(Y) \leq \gamma^* \mathcal{V}(x) \quad \text{almost surely if } Y \sim Q_\mu(x)
\]

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

\[
\mathcal{V}(x) + \mathcal{V}(y) \leq C \quad \Rightarrow \quad \|\delta_x e^{t_0 L} - \delta_y e^{t_0 L}\|_{TV} \leq 2(1 - \alpha).
\]

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

\[
\sup_{\|f\|_{\infty} \leq 1} \|\lambda_\nu(Q_\nu f - f) - \lambda_\mu(Q_\mu f - f)\|_{\infty} \leq \theta \|\nu - \mu\|_{\nu}
\]
Theorem

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

\[
\| m_t - m_\infty \|_{\mathcal{V}} \leq Ce^{-rt} m_0(\mathcal{V}).
\]

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 $(X_{i,t})_{i \in [1, N], t \geq 0}$ by:

- Let $(\tilde{X}_{i,t})_{t \geq 0}$ 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
  \[
  T_i = \inf \left\{ t > 0, \ F_i > \int_0^t \lambda_i(\tilde{X}_s)ds \right\}
  \]
- 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 $\lambda_*, \theta, t_0, \alpha \geq 0$ such that, for all $i \in [1, N]$,

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

\[
\| \lambda_i(x)(Q_i(x) - \delta_{x_i}) - \lambda_i(y)(Q_i(y) - \delta_{y_i}) \|_{\infty} \leq \frac{\theta}{N} \sum_{j=1}^{N} 1_{x_j \neq y_j}.
\]

- For all $x, y \in E$, 

\[
\| \delta_x e^{t_0 L_i} - \delta_y e^{t_0 L_i} \|_{TV} \leq 2(1 - \alpha).
\]

Remark:

$\mathbb{P}$ (the $N$ particles all merge at time $t_0) \geq (\alpha e^{-\lambda_* t_0})^N$. 
Interacting particle system (total variation)

**Theorem**

Under the previous assumption, denote \((P_t^{(N)})_{t \geq 0}\) the semi-group associated to \((X_{i,t})_{i \in \llbracket 1, N \rrbracket, t \geq 0}\), then

\[
\|m_0 P_t^{(N)} - h_0 P_t^{(N)}\|_{TV} \leq Ne^{\theta t_0} \left( e^{\theta t_0} - \alpha e^{-\lambda_* t_0} \right)^{\lfloor t/t_0 \rfloor} \|m_0 - h_0\|_{TV}.
\]

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 \(\llbracket 1, N \rrbracket\) and independent from \(X_t\) and \(Y_t\), then

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

Proof : couple particle independently. Try to keep together those who have already merged with synchronous coupling.
Interacting particle system ($\mathcal{V}$-norm)

**Theorem**

Again,

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

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

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

The analogous of the non-linear theorem would be

$$
\text{lhs} \leq \frac{\theta}{N} \sum_{j=1}^{N} \mathbb{1}_{x_j \neq y_j} (\mathcal{V}_j(x_j) + \mathcal{V}_j(y_j)).
$$
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 $V_i(X_i,t)$ in the interacting particle case
- Propagation of chaos (coupling between interacting particle and independent non-linear processes)

Thank you for your attention!