Exponential ergodicity for switching dynamical system

Bertrand CLOEZ - Université Paris-Est Marne-la-Vallée,
Joint work with Martin HAIRER - University of Warwick

Workshop : Piecewise Deterministic Markov Processes
Labex Lebesgue, Rennes, May 16th

1. Model and examples
   - Description of the model
   - Example 1
   - Example 2

2. Limit theorem in the constant case
   - Definition of the Wasserstein distance
   - A limit theorem
   - Idea of the proof

3. Generalisation
   - Hypoelliptic case
   - Fully coupled PDMP
   - General Markov processes with switching

Model

Let us consider

- an irreducible CT Markov chain I, on a finite space F, with an invariant distribution ν,
- for each $i \in F$, a smooth vector field $F^{(i)}$, on $\mathbb{R}^d$, $d \geq 1$.

We consider the process X verifying

$$\forall t \geq 0, \ \partial_t X_t = F^{(I_t)}(X_t).$$

Model

Let us consider

- an irreducible CT Markov chain I, on a finite space F, with an invariant distribution ν,
- for each $i \in F$, a smooth vector field $F^{(i)}$, on $\mathbb{R}^d$, $d \geq 1$.

We consider the process X verifying

$$\forall t \geq 0, \ \partial_t X_t = F^{(I_t)}(X_t).$$

The couple (X, I) is Markovian and is generated by

$$\mathbf{L}f(x,i) = F^{(i)}(x) \cdot \nabla_x f(x,i) + \int_F (f(x,j) - f(x,i)) Q(i,dj).$$

Model

Let us consider

- an irreducible CT Markov chain I, on a finite space F, with an invariant distribution ν,
- for each $i \in F$, a smooth vector field $F^{(i)}$, on $\mathbb{R}^d,\ d \geq 1$.

We consider the process X verifying

$$\forall t \geq 0,\ \partial_t X_t = F^{(I_t)}(X_t).$$

The couple (X, I) is Markovian and is generated by

$$\mathbf{L}f(x,i) = F^{(i)}(x) \cdot \nabla_x f(x,i) + \int_F (f(x,j) - f(x,i)) Q(i, dj).$$

$\rightarrow$ We can also assume that Q depends on the continuous component X.

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)
- Gene Network (Crudu, Debussche, Muller, Radulescu, 2012)

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)
- Gene Network (Crudu, Debussche, Muller, Radulescu, 2012)
- Storage modelling (Boxma, Kaspi, Kella, Perry, 2005)

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)
- Gene Network (Crudu, Debussche, Muller, Radulescu, 2012)
- Storage modelling (Boxma, Kaspi, Kella, Perry, 2005)
- Neuronal activity (Genadot, Pakdaman, Thieullen, Wainrib, 2012)

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)
- Gene Network (Crudu, Debussche, Muller, Radulescu, 2012)
- Storage modelling (Boxma, Kaspi, Kella, Perry, 2005)
- Neuronal activity (Genadot, Pakdaman, Thieullen, Wainrib, 2012)
- Molecular biology (Faggionato, Gabrielli, Crivellari, 2008)

Motivations

- Chemostat (Collet, Martinez, Méléard, San Martin, 2012)
- Gene Network (Crudu, Debussche, Muller, Radulescu, 2012)
- Storage modelling (Boxma, Kaspi, Kella, Perry, 2005)
- Neuronal activity (Genadot, Pakdaman, Thieullen, Wainrib, 2012)
- Molecular biology (Faggionato, Gabrielli, Crivellari, 2008)

$\rightarrow$ **Natural questions** :

- Ergodicity criterion : $\mathcal{L}(X_t) \rightarrow \pi$
- Rate of convergence dist$(\mathcal{L}(X_t), \pi) \leq \varphi(X_0, t)$

An explosive switched vector fields

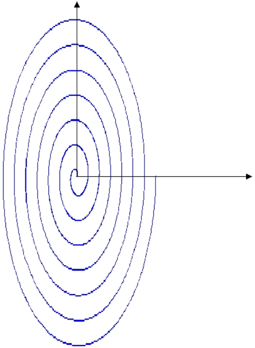

FIGURE: First vector field : $F^{(1)} : x \mapsto A_1 \cdot x$

An explosive switched vector fields

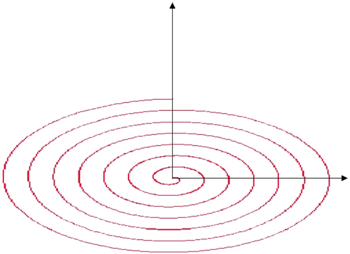

FIGURE: Second vector field : $F^{(2)} : x \mapsto A_2 \cdot x$

Explosive switched vector fields

Let $a > 0$, we consider the following generator :

$$\mathbf{L}f(x,i) = A_i \cdot \nabla_x f(x,i) + a(f(x,1-i) - f(x,i)),$$

where $x \in \mathbb{R}^2$, $i \in \{0,1\}$ and f is smooth.

Explosive switched vector fields

Let $a > 0$, we consider the following generator :

$$\mathbf{L}f(x,i) = A_i \cdot \nabla_x f(x,i) + a(f(x,1-i) - f(x,i)),$$

where $x \in \mathbb{R}^2$, $i \in \{0,1\}$ and f is smooth. If we fix $i \in \{0,1\}$ then the solutions of

$$\forall t \geq 0,\ \partial_t y_t = A_i y_t$$

satisfy

$$\|y_t\| \leq Ce^{-t}\|y_0\|.$$

Explosive switched vector fields

Let $a > 0$, we consider the following generator :

$$\mathbf{L}f(x,i) = A_i \cdot \nabla_x f(x,i) + a(f(x,1-i) - f(x,i)),$$

where $x \in \mathbb{R}^2$, $i \in \{0,1\}$ and f is smooth. If we fix $i \in \{0,1\}$ then the solutions of

$$\forall t \geq 0,\ \partial_t y_t = A_i y_t$$

satisfy

$$\|y_t\| \leq Ce^{-t}\|y_0\|.$$

Nevertheless if a is large enough then

$$\lim_{t\to+\infty} X_t = +\infty$$

Explosive switched vector fields

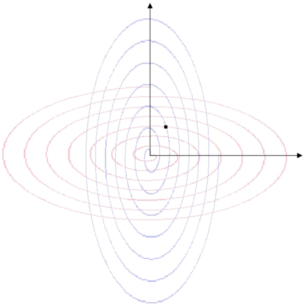

FIGURE: A typical trajectory

Explosive switched vector fields

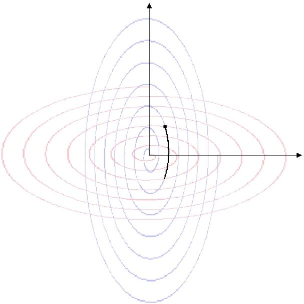

FIGURE: A typical trajectory

Explosive switched vector fields

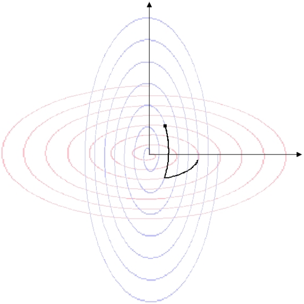

FIGURE: A typical trajectory

Explosive switched vector fields

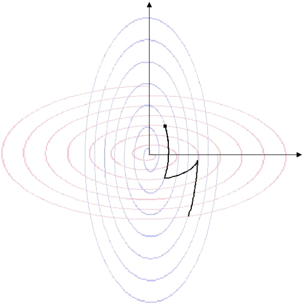

FIGURE: A typical trajectory

Explosive switched vector fields

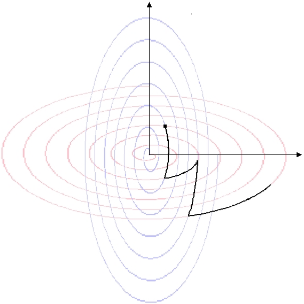

FIGURE: A typical trajectory

The most elementary example

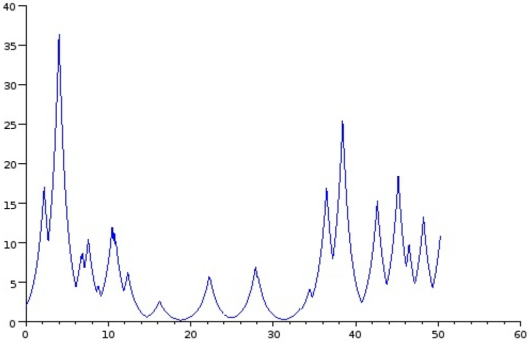

FIGURE: A trajectory of the second example

The most elementary example

Let us consider that I is a Markov Chain on $\{-1, 1\}$, the continuous component belongs to $\mathbb{R}$ and satisfies

$$\partial_t X_t = -I_t X_t.$$

The most elementary example

Let us consider that I is a Markov Chain on $\{-1,1\}$, the continuous component belongs to $\mathbb{R}$ and satisfies

$$\partial_t X_t = -I_t X_t.$$

We have

$$X_t = e^{-\int_0^t I_s ds}$$

The most elementary example

Let us consider that I is a Markov Chain on $\{-1, 1\}$, the continuous component belongs to $\mathbb{R}$ and satisfies

$$\partial_t X_t = -I_t X_t.$$

We have

$$X_t = e^{-t \times \frac{1}{t} \int_0^t I_s ds}$$

Birkhoff's ergodic theorem gives that

- $X_t \to 0$ if $\sum_i i\nu(i) = \nu(1) - \nu(-1) > 0$,
- $X_t \to +\infty$ if $\sum_i i\nu(i) = \nu(1) - \nu(-1) < 0$.

The most elementary example

Let us consider that I is a Markov Chain on $\{-1, 1\}$, the continuous component belongs to $\mathbb{R}$ and satisfies

$$\partial_t X_t = -I_t X_t.$$

We have

$$X_t = e^{-t \times \frac{1}{t} \int_0^t I_s ds}$$

Birkhoff's ergodic theorem gives that

- $X_t \to 0$ if $\sum_i i\nu(i) = \nu(1) - \nu(-1) > 0$,
- $X_t \to +\infty$ if $\sum_i i\nu(i) = \nu(1) - \nu(-1) < 0$.

$\to$ Rates of convergence ?

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0, \ \forall t \geq 0$.

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0$, $\forall t \geq 0$. In particular,

$$\forall t \geq 0, \ \|\mathcal{L}(X_t) - \delta_0\|_{TV} = \mathbb{P}(T_0 > t) = 1.$$

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0,\ \forall t \geq 0$. In particular,

$$\forall t \geq 0,\ \|\mathcal{L}(X_t) - \delta_0\|_{TV} = \mathbb{P}(T_0 > t) = 1.$$

In general

$$\lim_{t \to +\infty} E[X_t] = +\infty,$$

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0,\ \forall t \geq 0$. In particular,

$$\forall t \geq 0,\ \|\mathcal{L}(X_t) - \delta_0\|_{TV} = \mathbb{P}(T_0 > t) = 1.$$

In general

$$\lim_{t \to +\infty} E[X_t] = +\infty,$$

and then there is no convergence in L^1-norm and

$$\lim_{t \to +\infty} \mathcal{W}(\mathcal{L}(X_t), \delta_0) = +\infty.$$

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0,\ \forall t \geq 0$. In particular,

$$\forall t \geq 0,\ \|\mathcal{L}(X_t) - \delta_0\|_{TV} = \mathbb{P}(T_0 > t) = 1.$$

In general

$$\lim_{t\to+\infty} E[X_t] = +\infty,$$

and then there is no convergence in L^1-norm and

$$\lim_{t\to+\infty} \mathcal{W}(\mathcal{L}(X_t), \delta_0) = +\infty.$$

$\rightarrow$ We have to modify the distance !

Non convergence with the usual distance

If $X_0 \neq 0$ then $X_t \neq 0, \ \forall t \geq 0$. In particular,

$$\forall t \geq 0, \ \|\mathcal{L}(X_t) - \delta_0\|_{TV} = \mathbb{P}(T_0 > t) = 1.$$

In general

$$\lim_{t \to +\infty} E[X_t] = +\infty,$$

and then there is no convergence in L^1-norm and

$$\lim_{t \to +\infty} \mathcal{W}(\mathcal{L}(X_t), \delta_0) = +\infty.$$

$\to$ We have to modify the distance !
Convergence of the moment ?

$$\mathbb{E}\left[X_t^p\right] = \mathbb{E}\left[e^{-\int_0^t pI_s ds}\right], \quad p \in (0,1).$$

Moments properties

Feynman-Kac formula :

$$\mathbb{E}\left[e^{-\int_0^t pI_s ds}\right] = \mu_0 e^{t(A-pId)}\mathbf{1}$$

Moments properties

Feynman-Kac formula :

$$\mathbb{E}\left[e^{-\int_0^t pI_s ds}\right] = \mu_0 e^{t(A-pId)}\mathbf{1} \approx e^{-\lambda_p t}.$$

Moments properties

Feynman-Kac formula :

$$\mathbb{E}\left[e^{-\int_0^t p I_s ds}\right] = \mu_0 e^{t(A - pId)}\mathbf{1} \approx e^{-\lambda_p t}.$$

We have $\lambda_0 = 0$ and

$$\partial_p \lambda_p|_{p=0} = \sum_{i \in F} i\nu(i).$$

Moments properties

Feynman-Kac formula :

$$\mathbb{E}\left[e^{-\int_0^t pI_s ds}\right] = \mu_0 e^{t(A-pId)}\mathbf{1} \approx e^{-\lambda_p t}.$$

We have $\lambda_0 = 0$ and

$$\partial_p \lambda_p|_{p=0} = \sum_{i \in F} i\nu(i).$$

Hence

$$\sum_{i \in F} i\nu(i) > 0 \Rightarrow \exists p > 0, \ \lambda_p > 0.$$

Moments properties

Feynman-Kac formula :

$$\mathbb{E}\left[e^{-\int_0^t pI_s ds}\right] = \mu_0 e^{t(A-pId)}\mathbf{1} \approx e^{-\lambda_p t}.$$

We have $\lambda_0 = 0$ and

$$\partial_p \lambda_p|_{p=0} = \sum_{i \in F} i\nu(i).$$

Hence

$$\sum_{i \in F} i\nu(i) > 0 \Rightarrow \; \exists p > 0, \; \lambda_p > 0.$$

$\Rightarrow$ Convergence in "L^p-norm" and in a weaker Wasserstein distance.

More generally

Lemma

Let α be a function on F. If

$$\sum_i \alpha(i)\nu(i) > 0$$

then there exist $C, c, \lambda, p > 0$ such that

$$ce^{-\lambda t} \leq \mathbb{E}\left[e^{-\int_0^t p\alpha(I_s)ds}\right] \leq Ce^{-\lambda t}$$

More generally

Lemma

Let α be a function on F. If

$$\sum_i \alpha(i)\nu(i) > 0$$

then there exist $C, c, \lambda, p > 0$ such that

$$ce^{-\lambda t} \leq \mathbb{E}\left[e^{-\int_0^t p\alpha(I_s)ds}\right] \leq Ce^{-\lambda t}$$

$\rightarrow$ See (Bardet, Guerin, Malrieu, 2010).

Wasserstein distance

- For any probability measures μ_1, μ_2 on (E, d) :

$$\begin{aligned}\mathcal{W}_d(\mu_1, \mu_2) &= \inf_{\Pi} \int_{E \times E} d(x, y)\Pi(dx, dy) \\ &= \inf_{X_1 \sim \mu_1, X_2 \sim \mu_2} \mathbb{E}\left[d(X_1, X_2)\right].\end{aligned}$$

Wasserstein distance

- For any probability measures μ_1, μ_2 on (E, d) :

$$\begin{aligned}\mathcal{W}_d(\mu_1, \mu_2) &= \inf_{\Pi} \int_{E\times E} d(x, y)\Pi(dx, dy) \\ &= \inf_{X_1\sim\mu_1, X_2\sim\mu_2} \mathbb{E}\left[d(X_1, X_2)\right].\end{aligned}$$

- Convergence with $\mathcal{W}_d \;\Leftrightarrow\;$ Convergence in law + first moment.

Wasserstein distance

- For any probability measures μ_1, μ_2 on (E, d) :

$$\mathcal{W}_d(\mu_1, \mu_2) = \inf_{\Pi} \int_{E \times E} d(x, y) \Pi(dx, dy)$$
$$= \inf_{X_1 \sim \mu_1, X_2 \sim \mu_2} \mathbb{E}\left[d(X_1, X_2)\right].$$

- Convergence with $\mathcal{W}_d \;\Leftrightarrow\;$ Convergence in law + first moment.
- Also called Kantorovich, Mallows, Monge, Fréchet, optimal transport, coupling, minimum-L^1...

Wasserstein distance

Duality of Kantorovich-Rubinstein :

$$\mathcal{W}_d(\mu_1, \mu_2) = \sup_{\mathrm{Lip}(f) \leq 1} \int f d\mu_1 - \int f d\mu_2$$

Wasserstein distance

Duality of Kantorovich-Rubinstein :

$$\mathcal{W}_d(\mu_1, \mu_2) = \sup_{\mathrm{Lip}(f) \leq 1} \int f d\mu_1 - \int f d\mu_2$$

If $d(x, y) = \mathbf{1}_{x \neq y}$ then $\mathcal{W}_d = \| \cdot \|_{\mathrm{VT}}$ and

$$\mathcal{W}_d(\mu_1, \mu_2) = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \int f d\mu_1 - \int f d\mu_2 .$$

Wasserstein distance

Duality of Kantorovich-Rubinstein :

$$\mathcal{W}_d(\mu_1, \mu_2) = \sup_{\mathrm{Lip}(f) \leq 1} \int f d\mu_1 - \int f d\mu_2$$

If $d(x, y) = \mathbf{1}_{x \neq y}$ then $\mathcal{W}_d = \| \cdot \|_{\mathrm{VT}}$ and

$$\mathcal{W}_d(\mu_1, \mu_2) = \frac{1}{2} \sup_{\|f\|_\infty \leq 1} \int f d\mu_1 - \int f d\mu_2 .$$

If $E = \mathbb{R}$ and $d(x, y) = |x - y| \wedge 1$ then $\mathcal{W}_d = d_{\mathrm{FM}}$ and

$$\mathcal{W}_d(\mu_1, \mu_2) = \sup_{\|f\|_\infty + \|f'\|_\infty \leq 1} \int f d\mu_1 - \int f d\mu_2 .$$

Contraction Assumption

Assume that $\forall i \in F$, $\exists \rho(i) \in \mathbb{R}$ such that

$$\langle x - y, F^{(i)}(x) - F^{(i)}(y) \rangle \leq -\rho(i) \|x - y\|^2, \quad x, y \in R^d,$$

Contraction Assumption

Assume that $\forall i \in F$, $\exists \rho(i) \in \mathbb{R}$ such that

$$\langle x - y, F^{(i)}(x) - F^{(i)}(y) \rangle \leq -\rho(i) \|x - y\|^2, \quad x, y \in R^d,$$

Note that if $\rho(i) > 0$ then $\exists! x_i$ such that all the solutions $(y_t)_{t \geq 0}$ to

$$\partial_t y_t = F^{(i)}(y_t), \quad t \geq 0,$$

verify

$$\forall t \geq 0, \ \|y_t - x_i\| \leq e^{-\rho(i)t} \|y_0 - x_i\|.$$

Wasserstein exponential ergodicity

Theorem

If

$$\sum_{i\in F} \nu(i)\rho(i) > 0,$$

then $(\mathbf{X}_t)_{t\geq 0} = (X_t, I_t)_{t\geq 0}$ *admits a unique invariant probability measure* π *and there exist* $C, \lambda, t_0 > 0$ *and* $p \in (0,1)$ *such that*

$$\forall t \geq t_0,\ \mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \pi) \leq Ce^{-\lambda t}(1 + \mathcal{W}_{\|\cdot\|^p}(\mathcal{L}(\mathbf{X}_0), \pi)),$$

where

$$\mathbf{d}((x,i),(y,j)) = \mathbf{1}_{i\neq j} + \mathbf{1}_{i=j}(1 \wedge \|x-y\|^p).$$

Wasserstein exponential ergodicity

Theorem

If

$$\sum_{i\in F} \nu(i)\rho(i) > 0,$$

then $(\mathbf{X}_t)_{t\geq 0} = (X_t, I_t)_{t\geq 0}$ *admits a unique invariant probability measure* π *and there exist* $C, \lambda, t_0 > 0$ *and* $p \in (0,1)$ *such that*

$$\forall t \geq t_0,\ \mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \pi) \leq Ce^{-\lambda t}(1 + \mathcal{W}_{\|\cdot\|^p}(\mathcal{L}(\mathbf{X}_0), \pi)),$$

where

$$\mathbf{d}((x,i),(y,j)) = \mathbf{1}_{i\neq j} + \mathbf{1}_{i=j}(1 \wedge \|x-y\|^p).$$

$\rightarrow$ Already proved in (Benaïm, Le Borgne, Malrieu, Zitt 12).

A general form of Harris theorem

$\rightarrow$ Proof based on a theorem of (Hairer, Mattingly, Scheutzow, 09).

A general form of Harris theorem

$\rightarrow$ Proof based on a theorem of (Hairer, Mattingly, Scheutzow, 09). It is enough to prove :

i) There exist V and $C, K, \lambda > 0$ such that

$$\mathbb{E}[V(X_t)] \leq Ce^{-\lambda t}\mathbb{E}[V(X_0)] + K, \quad t \geq 0.$$

A general form of Harris theorem

$\rightarrow$ Proof based on a theorem of (Hairer, Mattingly, Scheutzow, 09). It is enough to prove :

i) There exist V and $C, K, \lambda > 0$ such that

$$\mathbb{E}[V(X_t)] \leq Ce^{-\lambda t}\mathbb{E}[V(X_0)] + K, \quad t \geq 0.$$

ii) For all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distributions $X_0, Y_0 \in \{V \leq A\}$.

A general form of Harris theorem

$\rightarrow$ Proof based on a theorem of (Hairer, Mattingly, Scheutzow, 09). It is enough to prove :

i) There exist V and $C, K, \lambda > 0$ such that

$$\mathbb{E}[V(X_t)] \leq Ce^{-\lambda t}\mathbb{E}[V(X_0)] + K, \quad t \geq 0.$$

ii) For all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distributions $X_0, Y_0 \in \{V \leq A\}$.

iii) There exist $\alpha \in (0, 1)$ and a time t such that

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq \alpha \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0),$$

for any starting distributions $X_0, Y_0 \in \mathbf{E} = E \times F$ verifying $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1$.

Point i)

We set $V(x,i) = V(x) = \|x\|^p$.

Point i)

We set $V(x,i) = V(x) = \|x\|^p$.Using the generator and Gronwall Lemma, we find

$$\mathbb{E}[V(X_t)] \leq K \int_0^t \mathbb{E}\left[e^{-\int_s^t p\rho(I_u)du}\right] ds + \mathbb{E}[V(X_0)]\mathbb{E}\left[e^{-\int_0^t p\rho(I_u)du}\right].$$

Point i)

We set $V(x,i) = V(x) = \|x\|^p$.Using the generator and Gronwall Lemma, we find

$$\mathbb{E}[V(X_t)] \leq K \int_0^t \mathbb{E}\left[e^{-\int_s^t p\rho(I_u)du}\right] ds + \mathbb{E}[V(X_0)]\mathbb{E}\left[e^{-\int_0^t p\rho(I_u)du}\right].$$

But we can find $p \in (0,1)$ in such a way to obtain

$$\left[e^{-\int_0^t p\rho(I_u)du}\right] \leq Ce^{-\lambda t}.$$

Point i)

We set $V(x,i) = V(x) = \|x\|^p$.Using the generator and Gronwall Lemma, we find

$$\mathbb{E}[V(X_t)] \leq K \int_0^t \mathbb{E}\left[e^{-\int_s^t p\rho(I_u)du}\right] ds + \mathbb{E}[V(X_0)]\mathbb{E}\left[e^{-\int_0^t p\rho(I_u)du}\right].$$

But we can find $p \in (0,1)$ in such a way to obtain

$$\left[e^{-\int_0^t p\rho(I_u)du}\right] \leq Ce^{-\lambda t}.$$

It gives that V is a Lyapunov function ; that is,

$$\mathbb{E}[V(X_t)] \leq Ce^{-\lambda t}\mathbb{E}[V(X_0)] + K, \quad t \geq 0.$$

Point ii)

We want to prove that for all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distribution $X_0, Y_0 \in \{V \leq A\}$.

Point ii)

We want to prove that for all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distribution $X_0, Y_0 \in \{V \leq A\}$.
Let us fix two starting point $X_0, Y_0 \in \{V \leq A\}$.

Point ii)

We want to prove that for all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distribution $X_0, Y_0 \in \{V \leq A\}$.
Let us fix two starting point $X_0, Y_0 \in \{V \leq A\}$. As there exists i_0 such that $\rho(i_0) > 0$, we easily explicit a coupling verifying

$$\mathbf{d}(\mathbf{X}_t, \mathbf{Y}_t) \leq \mathbf{1}_U C e^{-\rho(i_0)t} d(X_0, Y_0) + \mathbf{1}_{U^c}$$

Point ii)

We want to prove that for all $A > 0$ there exist $\epsilon_A > 0$ and $t_A > 0$ such that for all $t \geq t_A$,

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq 1 - \epsilon_A$$

for any starting distribution $X_0, Y_0 \in \{V \leq A\}$.
Let us fix two starting point $X_0, Y_0 \in \{V \leq A\}$. As there exists i_0 such that $\rho(i_0) > 0$, we easily explicit a coupling verifying

$$\mathbf{d}(\mathbf{X}_t, \mathbf{Y}_t) \leq \mathbf{1}_U C e^{-\rho(i_0)t} d(X_0, Y_0) + \mathbf{1}_{U^c}$$

Finally as $\{V \leq A\}$ is bounded, it ends the proof of ii).

Point iii)

We want to prove the existence of $\alpha \in (0,1)$ and a time t such that

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq \alpha \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0),$$

for any starting distribution verifying $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1$.

Point iii)

We want to prove the existence of $\alpha \in (0,1)$ and a time t such that

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq \alpha \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0),$$

for any starting distribution verifying $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1$.

But $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1 \Rightarrow I_0 = J_0$

Point iii)

We want to prove the existence of $\alpha \in (0,1)$ and a time t such that

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq \alpha \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0),$$

for any starting distribution verifying $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1$.

But $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1 \Rightarrow I_0 = J_0$ and then

$$\begin{aligned}\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) &\leq \mathbb{E}\left[e^{-\int_0^t p\rho(I_u)du}\right] \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) \\ &\leq Ce^{-\lambda t}\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0).\end{aligned}$$

Point iii)

We want to prove the existence of $\alpha \in (0,1)$ and a time t such that

$$\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) \leq \alpha \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0),$$

for any starting distribution verifying $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1$.

But $\mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) < 1 \Rightarrow I_0 = J_0$ and then

$$\begin{aligned}\mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \mathcal{L}(\mathbf{Y}_t)) &\leq \mathbb{E}\left[e^{-\int_0^t p\rho(I_u)du}\right] \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0) \\ &\leq Ce^{-\lambda t} \mathbf{d}(\mathbf{X}_0, \mathbf{Y}_0).\end{aligned}$$

$\rightarrow$ Finally i),ii),iii) holds and the theorem is proved.

Hypoellipticity assumption

By (Benaïm, Le Borgne, Malrieu, Zitt 12) and (Bakhtin, Hurth, 12), if the family $(F^{(i)})_i$ verifies an Hörmander-type condition, then the process X verifies a regularising assumption.

Hypoellipticity assumption

By (Benaïm, Le Borgne, Malrieu, Zitt 12) and (Bakhtin, Hurth, 12), if the family $(F^{(i)})_i$ verifies an Hörmander-type condition, then the process X verifies a regularising assumption.

$\rightarrow$ Hörmander-type condition + Lyapunov funtion $\Rightarrow$ Exponential convergence in $\|\cdot\|_{TV}$.

Hypoellipticity assumption

By (Benaïm, Le Borgne, Malrieu, Zitt 12) and (Bakhtin, Hurth, 12), if the family $(F^{(i)})_i$ verifies an Hörmander-type condition, then the process X verifies a regularising assumption.

$\rightarrow$ Hörmander-type condition + Lyapunov funtion $\Rightarrow$ Exponential convergence in $\|\cdot\|_{TV}$.

Lemma

If there exists V s.t.

$$F^{(i)}(x) \cdot \nabla V(x) \leq -\lambda_i V(x) + K_i,$$

where

$$\sum_{i \in F} \lambda_i \nu(i) > 0,$$

then (X, I) admits a Lyapunov function.

Wasserstein exponential ergodicity in the non-constant case

If $F = \{-1, 1\}$ and

$$\mathbf{L}f(x,i) = F^{(i)}(x) \cdot \nabla_x f(x,i) + a(x,i)(f(x,-i) - f(x,i)),$$

Wasserstein exponential ergodicity in the non-constant case

If $F = \{-1, 1\}$ and

$$\mathbf{L}f(x,i) = F^{(i)}(x) \cdot \nabla_x f(x,i) + a(x,i)(f(x,-i) - f(x,i)),$$

where we consider $\rho(1) > 0, \rho(-1) < 0$ and

$$\underline{a}(1) = \inf_x a(x,1) \;\text{ and }\; \bar{a}(-1) = \sup_x a(x,-1)$$

Theorem

If a is Lipschitz and

$$\bar{a}(-1)\rho(1) + \underline{a}(1)\rho(-1) > 0$$

then X admits an invariant probability measure and converges exponentially fast to it in a Wasserstein distance.

Wasserstein curvature

Definition

The Wasserstein curvature of a Markov semigroup $(P_t)_{t\geq 0}$ is the largest constant ρ such that

$$\mathcal{W}(\mu P_t, \nu P_t) \leq e^{-\rho t}\mathcal{W}(\mu, \nu),$$

for any probability measure μ, ν and any $t \geq 0$.

Wasserstein curvature

Definition

The Wasserstein curvature of a Markov semigroup $(P_t)_{t\geq 0}$ is the largest constant ρ such that

$$\mathcal{W}(\mu P_t, \nu P_t) \leq e^{-\rho t}\mathcal{W}(\mu, \nu),$$

for any probability measure μ, ν and any $t \geq 0$.

- introduced independently by Joulin (2007), Ollivier (2007) and Sammer (2005).
- Motivated by generalizing Bakry-Emery curvature of diffusion processes or Ricci curvature of Riemannian Manifold.
- By the Kantorovich-Rubinstein duality, we have

$$\rho = \sup_{t>0} -\frac{1}{t} \ln \|P_t\|_{\mathrm{Lip}(d)\to\mathrm{Lip}(d)}.$$

Some examples of curvature

See for instance

- (Sturm, Von Renesse, 2005) for Brownian motion on Riemannian manifold
- (Chafaï, Joulin, 2012) for birth and death processes
- (Cloez, 2012) for stochastically monotonous processes
- (Eberle, 2011) and (Cattiaux, Guillin, 2013) for inhomogeneous diffusion

Some examples of curvature

See for instance

- (Sturm, Von Renesse, 2005) for Brownian motion on Riemannian manifold
- (Chafaï, Joulin, 2012) for birth and death processes
- (Cloez, 2012) for stochastically monotonous processes
- (Eberle, 2011) and (Cattiaux, Guillin, 2013) for inhomogeneous diffusion

Now, we assume that $(\mathbf{X}_t)_{t\geq 0} = (X_t, I_t)_{t\geq 0}$ is generated by

$$\mathbf{L}f(x,i) = \mathcal{L}^{(i)}f(x,i) + \int_F (f(x,j) - f(x,i))Q(i,dj),$$

Wasserstein exponential ergodicity

Theorem

If

$$\sum_{i \in F} \nu(i)\rho(i) > 0,$$

then $(\mathbf{X}_t)_{t\geq 0} = (X_t, I_t)_{t\geq 0}$ *admits a unique invariant probability measure* π *and there exist* $C, \lambda, t_0 > 0$ *and* $p \in (0,1)$ *such that*

$$\forall t \geq t_0,\ \mathcal{W}_{\mathbf{d}}(\mathcal{L}(\mathbf{X}_t), \pi) \leq Ce^{-\lambda t}(1 + \mathcal{W}_{\|\cdot\|^p}(\mathcal{L}(\mathbf{X}_0), \pi)),$$

where

$$\mathbf{d}((x,i),(y,j)) = \mathbf{1}_{i \neq j} + \mathbf{1}_{i=j}(1 \wedge \|x-y\|^p).$$

Thank you for your attention !