Long time behavior of some Piecewise Deterministic Markov Processes

Florent Malrieu

Université de Tours

What is a PDMP?

$\rightsquigarrow$ a continuous time Markov process with

- deterministic evolution
- random jumps
- no diffusion component
- irreversible dynamics

See Davis (1984)

Where do they come from?

- communication network (TCP-IP)
- movement of bacteria (chemotaxis)
- size of small organisms
- chemical reactions
- approximation of Markov chains...

What is the topic today?

Long time behavior of some PDMPs

- qualitative properties (ergodicity, recurrence...)
- quantitative bounds for the rate of convergence

Motivations

- efficiency of a simulation algorithm
- parameter estimation...

Several distances

$$W_p(\mu, \tilde{\mu}) = \inf_{X \sim \mu, \tilde{X} \sim \tilde{\mu}} \left(\mathbb{E}\left(\left| X - \tilde{X} \right|^p \right) \right)^{1/p}$$

$$\|\mu - \tilde{\mu}\|_{\mathrm{TV}} = \inf_{X \sim \mu, \tilde{X} \sim \tilde{\mu}} \mathbb{P}(X \neq \tilde{X})$$

Parameters of a PDMP

Vector field F and flow Φ

The function $t \mapsto \Phi_t(x)$ is solution of the ODE

$$\dot{x}_t = F(x_t), \quad x_0 = x$$

Jump rate λ

Starting at x, the first jump time is given by

$$\mathbb{P}_x(T > t) = \exp\left(-\int_0^t \lambda(\Phi_s(x))\,ds\right)$$

Jump measure Q

$$X_{T_-} = \Phi_T(x) \quad \text{and} \quad X_T \sim Q(X_{T_-}, dy)$$

Infinitesimal generator

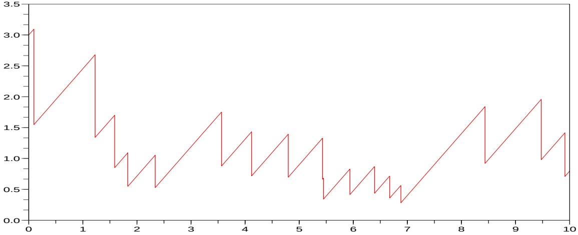
$$Lf(x) = F(x) \cdot \nabla f(x) + \lambda(x) \int (f(y) - f(x))\, Q(x, dy)$$

Example 1: the TCP window size process on $[0, +\infty)$

$$Lf(x) = f'(x) + \lambda(x)(f(x/2) - f(x))$$

Jump mechanism

$$\lambda(x) = \begin{cases} \lambda & \text{(simple)} \\ x & \text{(more realistic)} \end{cases} \quad \text{and} \quad Q(x, dy) = \delta_{x/2}(dy)$$



Example 2: a telegraph process on the circle

See Miclo-Monmarché (2013)

- **State space:** $E = S^1 \times \{-1, +1\}$
- **Generator**

$$Lf(x,i) = i\partial_x f(x,i) + \lambda(f(x,-i) - f(x,i))$$

- **Invariant measure** π**:** uniform measure on E
- **Explicit expression for** $\|P_t - \pi\|_{L^2(\pi)}$

Reminiscent of

- persistent random walks
- kinetic processes (Langevin dynamics)

Example 3: lifted Metropolis-Hastings algorithm

See Turitsyn-Chertkov-Vucelja (2011)

Ingredients:
T^+ and T^- two sub-Markov transition matrices on S s.t.

$$\pi(x)T^+(x,y) = \pi(y)T^-(y,x) \quad \text{for all } x \neq y$$

New object: T Markov kernel on $S \times \{-,+\}$ given by

$$T((x,-),(y,-)) = T^-(x,y), \qquad T((x,+),(y,+)) = T^+(x,y)$$

$$T((x,+),(x,-)) = T^{+-}(x), \qquad T((x,-),(x,+)) = T^{-+}(x)$$

$$T((x,-),(y,+)) = T((x,+),(y,-)) = 0 \quad \text{for } x \neq y$$

Example 3: LMH for Curie-Weiss model

Target measure on $S^n = \{-1, +1\}^n$

$$\pi^n(x) = Z_n^{-1} \exp(-\beta H^n(x))$$

where

$$H^n(x) = -\frac{1}{2n} \sum_{i,j}^{n} x_i x_j - h \sum_{i=1}^{n} x_i = -n\big(m^n(x)^2/2 + hm^n(x)\big)$$

Theorem (Bierkens-Robert (2015))

With a correct scaling, the LMH for CWM converges to the process on $\mathbb{R} \times \{-1, +1\}$ *driven by*

$$Lf(m,i) = i\partial_m f(m,i) + \max(0, im)(f(m,-i) - f(m,i))$$

Back to TCP process

$$Lf(x) = f'(x) + \lambda(x)(f(x/2) - f(x))$$

Motivations

- Idealized rate of a connection in a network
- Size of bacteria

Difficulties

- the randomness only appears via the jump times
- irreversibility
- no regularization via diffusion
- $\mathcal{L}(X_t)$ has atoms as soon as $\mathcal{L}(X_0)$ does
- if λ is increasing, X is not monotonous

Constant jump rate: invariant regime

- **Its inv. prob. is abs. cont. with smooth density** and

$$\int x^n \, \mu(dx) = \frac{n!}{\lambda^n \prod_{k=1}^{n}(1 - 2^{-n})}$$

- **Spectral decompostion of the inf. gen.**

$$\lambda_n = \lambda\left(1 - \frac{1}{2^n}\right) \quad \leftrightsquigarrow \quad P_n \text{ with } \deg(P_n) = n$$

- **But the process is not reversible...**

Constant jump rate: convergence

Theorem (Bardet-Christen-Guillin-M.-Zitt (2013))

For any $p \geqslant 1$,

$$W_p(\nu_t, \tilde{\nu}_t) \leqslant e^{-\lambda_p t} W_p(\nu_0, \tilde{\nu}_0) \quad \textit{with} \quad \lambda_p = \frac{\lambda(1 - 2^{-p})}{p}$$

and

$$\|\nu_t - \tilde{\nu}_t\|_{\mathrm{TV}} \leqslant \lambda e^{-\lambda t/2} W_1(\nu_0, \tilde{\nu}_0) + e^{-\lambda t} \|\nu_0 - \tilde{\nu}_0\|_{\mathrm{TV}}$$

See also Perthame-Ryzhik (2005)

Sketch of proof

- **trivial coupling** for the Wasserstein distances (same jump times)

$$\left|X_t - \tilde{X}_t\right|^p = \left(\frac{1}{2}\right)^{pN_t} |x - \tilde{x}|^p \quad \text{where} \quad N_t \sim \mathcal{P}(\lambda t)$$

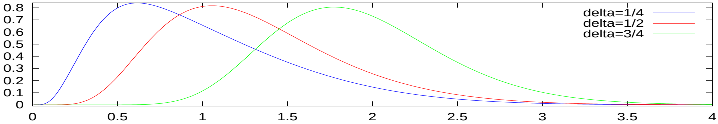
- **"one shot" coalescent coupling** for total variation distance
  - previous coupling until the penultimate jump time
  - a unique try to stick the particles

Remark: different of the classical approach by Meyn-Tweedie

Linear jump rate: invariant regime

- The eigenvalues of L are not explicit anymore
- Density of the inv. prob.

$$x \mapsto \frac{\sqrt{2/\pi}}{\prod_{n\geqslant 0}\left(1-2^{-(2n+1)}\right)} \sum_{n\geqslant 0} \frac{(-1)^n 2^{2n}}{\prod_{k=1}^{n}\left(2^{2k}-1\right)} e^{-2^{2n-1}x^2}$$



Linear jump rate: convergence

Theorem

For any initial measures and $p \geqslant 1$,

$$W_p(\nu_t, \tilde{\nu}_t) \leqslant C \exp\left(-\frac{\alpha}{p} t\right)$$

and

$$\|\nu_t - \tilde{\nu}_t\|_{\mathrm{TV}} \leqslant C \exp\left(-\frac{2\alpha}{3} t\right)$$

with $\alpha \sim 0.12$.

Main difficulty: the choice of a Wasserstein coupling

Idea: maximize the number of simultaneous jumps

The associated generator is given (for $x > \tilde{x}$) by

$$
\begin{aligned}
\mathcal{L}f(x,\tilde{x}) =& \partial_x f(x,\tilde{x}) + \partial_{\tilde{x}} f(x,\tilde{x}) && \text{linear increase} \\
& + \tilde{x}(f(x/2,\tilde{x}/2) - f(x,\tilde{x})) && \text{simultaneous jumps} \\
& + (x-\tilde{x})(f(x/2,\tilde{x}) - f(x,\tilde{x})) && \text{single jump}
\end{aligned}
$$

This provides

$$
\frac{d}{dt}\mathbb{E}\left(\left|X_t - \tilde{X}_t\right|\right) \leqslant -\frac{3}{2}\mathbb{E}\left(\left|X_t - \tilde{X}_t\right|^2\right)\ldots
$$

Several comments

- **Many other models:**
  - growth/division of a cell

$$Lf(x) = \tau(x)f'(x) + \lambda(x) \int_0^x (f(y) - f(x))Q(x, dy)$$

  - storage or pharmacokinetic models

$$Lf(x) = -\tau(x)f'(x) + \lambda(x) \int (f(x + y) - f(x))Q(x, dy)$$

- **Functional inequalities:** see Monmarché (2013)

Back to the telegraph process on $\mathbb{R} \times \{-1, +1\}$

Generator

$$Lf(x,i) = i\partial_x f(x,i) + \lambda(x,i)(f(x,-i) - f(x,i))$$

Several approaches

- coupling (Fontbona-Guérin-M. (2012))
- functional inequalities (Monmarché (2014-15))
- Meyn-Tweedie strategy (Bierkens-Roberts (2015))

BUT...

BUT, as Tony is used to say...

"These proofs are very 1-d..."

Lelièvre (2002, 2003,..., 2014, 2015, 2016,...)

Switched flows

$Z = (X, I) \in \mathbb{R}^d \times E$ with E finite

Framework

- Smooth flows $(F^i)_{i \in E}$ on $\mathbb{R}^d$
- The jump rates λ of I depend on (x, i)
- Infinitesimal generator

$$Lf(x,i) = F^i(x) \cdot \nabla_x f(x,i) + \sum_{j \in E} \lambda(x,i,j)(f(x,j) - f(x,i))$$

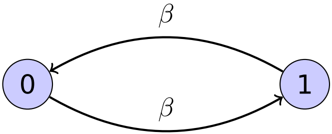
Questions

- Ergodicity
- Rate of convergence to equilibrium

The linear case may be tricky

$Z = (X, I)$ belongs to $\mathbb{R}^2 \times \{0, 1\}$

I is Markov on its own



$$\nu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$$

Dynamics of X

$$\dot{X}_t = A_{I_t} X_t \quad \text{with} \quad A_0 = \begin{pmatrix} -1 & 2b \\ -2/b & -1 \end{pmatrix} \text{ and } A_1 = A_0^T$$

A_0 and A_1 are "Hurwitz" (eigenvalues have neg. real parts)

Lyapunov exponent

Theorem

If $\mathbb{P}(X_0 \neq 0) = 1$ *then*

$$\frac{1}{t} \log \|X_t\| \xrightarrow[t \to \infty]{p.s.} \chi(b, \beta)$$

If $b - 1/b > 1$ *then* $\beta \mapsto \chi(b, \beta)$ *is increasing and*

$$\chi(b, 0) = -1 \quad \textit{and} \quad \lim_{\beta \to \infty} \chi(b, \beta) = b - \frac{1}{b} - 1 > 0$$

Sketch of proof

Introduce the polar coordinates (R_t, U_t) of X_t:

$$\dot{R}_t = R_t \langle U_t, A_{I_t} U_t \rangle$$
$$\dot{U}_t = AU_t - \langle U_t, A_{I_t} U_t \rangle U_t.$$

Notice that (U, I) is Markovian and

$$R_t = R_0 \exp\left(\int_0^t \langle U_s, A_{I_s} U_s \rangle \, ds \right)$$

(U, I) is ergodic and its probability measure μ is known explicitly and

$$\chi(b, \beta) = \lim_{t \to \infty} \frac{1}{t} \int_0^t \langle U_s, A_{I_s} U_s \rangle \, ds = \int \langle u, A_i u \rangle \, d\mu(u, i)$$

Why may $\chi(b, \beta)$ be positive for large β?

Even if A_0 and A_1 are Hurwitz matrices, $b - 1/b - b$ is an eigenvalue of

$$A_{1/2} = \frac{1}{2}A_0 + \frac{1}{2}A_1$$

For large βs, the inv. meas. concentrates near the instable direction of $A_{1/2}$.

General case. Same situation as soon as $\nu = p\delta_1 + (1-p)\delta_0$ and A_p has a positive eigenvalue.

Yet another example

$$A_0 = \begin{pmatrix} -\alpha & 1 \\ 0 & -\alpha \end{pmatrix} \quad \text{and} \quad A_1 = \begin{pmatrix} -\alpha & 0 \\ -1 & -\alpha \end{pmatrix}$$

Theorem (Lawley-Mattingly-Reed-(2013))

If $\mathbb{P}(X_0 = 0) = 0$ *then*

$$\frac{1}{t} \log \|X_t\| \xrightarrow[t\to\infty]{a.s.} \chi(\alpha, \beta)$$

- *If* β *is small or large enough then* $\chi(\alpha, \beta) < 0$ *and* $X_t \to 0$.
- *If* α *is small enough, there exists* β_0 *such that* $\chi(\alpha, \beta_0) > 0$ *and* $\|X_t\| \to \infty$.

The Lyapunov exponent $\chi(\alpha, \beta)$

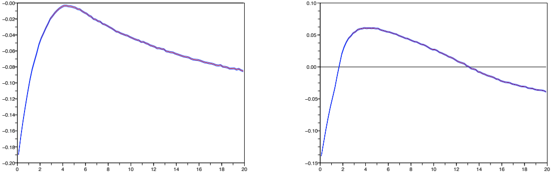


Figure: $\beta \mapsto \chi(\alpha, \beta)$ for $\alpha = 0.2$ (on the left) and $\alpha = 0.15$ (in the right).

- **Link with deterministic control**
  See Balde-Boscain-Mason (2009)

The system is unbounded iff

$$T(\alpha^2) := \frac{1 + 2\alpha^2 + \sqrt{1 + 4\alpha^2}}{2\alpha^2} e^{-2\sqrt{1+4\alpha^2}} > 1.$$

A (positive) quantitative result

Theorem (Benaïm-Le Borgne-M.-Zitt (2012))

Assume that

- *there exists* $\alpha > 0$ *such that*

$$\left\langle F^i(x) - F^i(\tilde{x}), x - \tilde{x} \right\rangle \leqslant -\alpha |x - \tilde{x}|^2, \quad x, \tilde{x} \in \mathbb{R}^d,\ i \in E$$

- *the jump rates are Lipschitz (w.r.t. x) and positive.*

Then

$$W_1(\mu_t, \tilde{\mu}_t) \leqslant c e^{-\rho t}$$

where ρ *depends explicitly on the parameters.*

Sketch of proof: once again, construction of a coupling...

Thank you
for your attention!