# Optimal choice among a class of nonparametric estimates of the jump rate for a piecewise-deterministic Markov process

**Romain Azaïs (Inria Nancy)**

Joint work with Aurélie Muller-Gueudin

## Outline

## Piecewise-deterministic Markov process

Let $(X_t)$ be a PDMP on an open subset $E$ of $\mathbb{R}^d$, defined from:

- $\Phi(x,t)$ the deterministic motion,
- $\lambda(x)$ the jump rate,
- $\mathcal{Q}(x, \mathrm{d}y)$ the transition kernel.

Deterministic exit time from $E$:

$$t^+(x) = \inf\{t > 0 \,:\, \Phi(x,t) \in \partial E\}.$$

## Piecewise-deterministic Markov process

From the initial condition $X_0 = x$, the survival function of the jump time $T_1$ is

$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda\left(\Phi(x,s)\right) \mathrm{d}s\right) \mathbb{1}_{\{0 \leq t < t^+(x)\}}.$$

Between times $0$ and $T_1$ the process evolves according to $\Phi(x,\cdot)$,

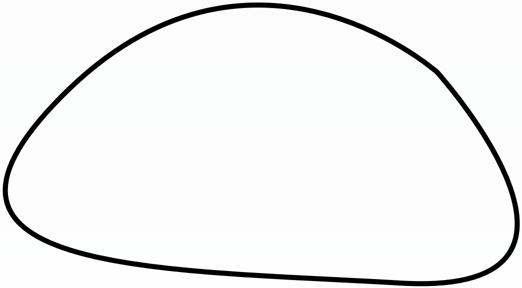
$$\forall\, 0 \leq t < T_1,\ X_t = \Phi(x,t).$$

At time $T_1$ the process "jumps" according to $\mathcal{Q}$,

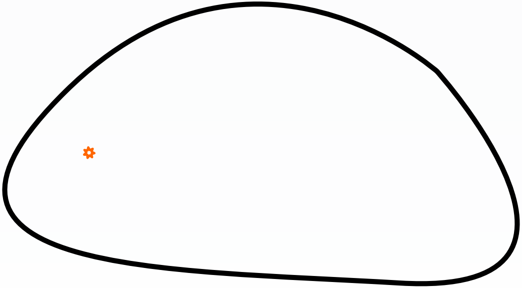
$$\mathbb{E}\left[\varphi(X_{T_1}) \,|\, \Phi(x,T_1)\right] = \int \varphi(u) \mathcal{Q}(\Phi(x,T_1), \mathrm{d}u).$$

And so on...

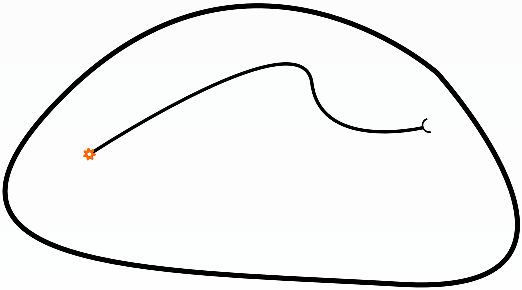
## Piecewise-deterministic Markov process



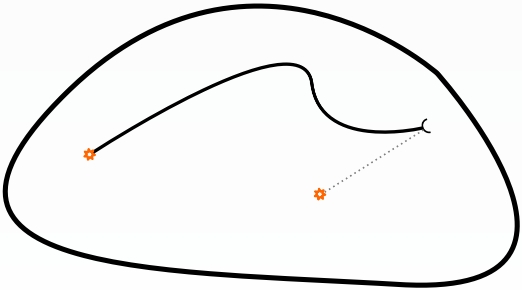
## Piecewise-deterministic Markov process



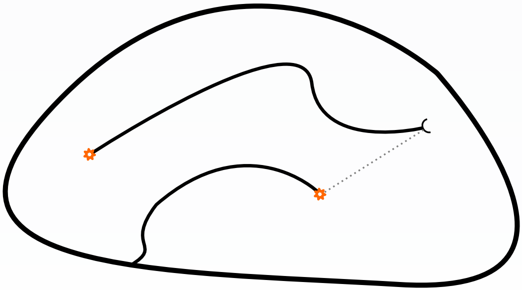
## Piecewise-deterministic Markov process



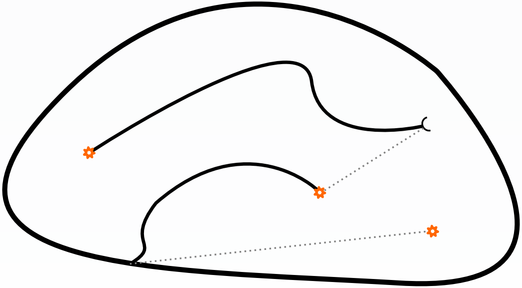
## Piecewise-deterministic Markov process



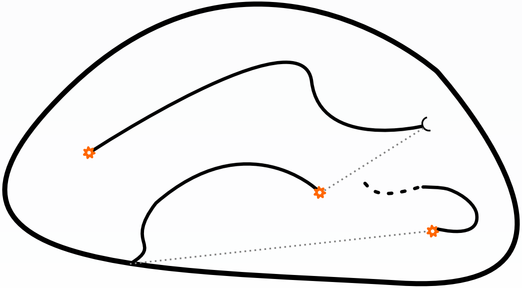
## Piecewise-deterministic Markov process



## Piecewise-deterministic Markov process



## Piecewise-deterministic Markov process



## Statistical inference

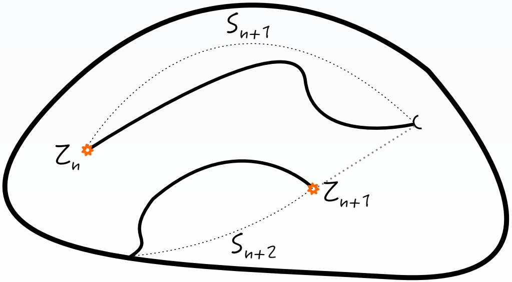
### General framework

- Process defined on $E \subset \mathbb{R}^d$
- Observation at the jump times of only one trajectory
- Deterministic equation assumed to be known (often given by the application)

### Main questions

- How to estimate the jump rate $\lambda$? The conditional density or the survival function associated with $\lambda$?
- How to estimate the transition kernel $\mathcal{Q}$?

## Statistical inference




→ Observation of the Markov chain $(Z_n, S_{n+1})$ within a long time interval

→ Some quantities of interest: $f(x,t)$, $G(x,t)$ and $\lambda \circ \Phi(x,t)$, linked by

$$\lambda \circ \Phi(x,t) = \frac{f(x,t)}{G(x,t)} \quad \text{whenever } t < t^+(x)$$

## Previous strategy

Characterization of the jump rate $\lambda$:

$$\mathbb{P}(S_{n+1} > t \,|\, Z_n = x) = \exp\left( - \int_0^t \lambda\left(\Phi(x,s)\right) \mathrm{d}s \right) \mathbb{1}_{\{0 \leq t < t^\star(x)\}}.$$

Conditionnaly to $Z_n = x$, we observe the (right-censored) time $S_{n+1}$ distributed according to the non homogeneous rate $\lambda \circ \Phi(x,t)$.

## Previous strategy

Characterization of the jump rate $\lambda$:

$$\mathbb{P}(S_{n+1} > t \,|\, Z_n = x) = \exp\left( - \int_0^t \lambda\left(\Phi(x,s)\right) \mathrm{d}s \right) \mathbb{1}_{\{0 \leq t < t^\star(x)\}}.$$

Conditionnaly to $Z_n = x$, we observe the (right-censored) time $S_{n+1}$ distributed according to the non homogeneous rate $\lambda \circ \Phi(x,t)$.

**Nelson-Aalen strategy**

$$M(t) = N(t) - \int_0^t \alpha(s) Y(s) \mathrm{d}s$$

↑ jump rate of interest

## Previous strategy

Conditionally to $Z_n$,

$$t \mapsto \mathbb{1}_{\{S_{n+1} \leq t\}} - \int_0^t \lambda(\Phi(Z_n, s))\mathbb{1}_{\{S_{n+1} \geq s\}} \mathrm{d}s$$

is a continuous-time martingale.

But the sum over $n$ is generally not a martingale.

## Previous strategy

Conditionally to $Z_n$,

$$t \mapsto \mathbb{1}_{\{S_{n+1} \leq t\}} - \int_0^t \lambda(\Phi(Z_n, s))\mathbb{1}_{\{S_{n+1} \geq s\}} \mathrm{d}s$$

is a continuous-time martingale.

But the sum over $n$ is generally not a martingale.

Solution:

- $\rightarrow$ Estimation for the double-marked renewal process $(Z_n, Z_{n+1}, S_{n+1})$.
- $\rightarrow$ Relation with the conditional density $f(x,t)$.

## Outline

## Assumptions

### Absolute continuity of the transition kernel

$$\mathcal{Q}(x, A) = \int_A \mathcal{Q}(x, y)\, \lambda_d(\mathrm{d}y)$$

### Condition of ergodicity

There exists a measure $\nu_\infty$ on $E$ such that

$$\lim_{n \to \infty} \|\nu_n - \nu_\infty\|_{TV} = 0,$$

where $\nu_n$ is the distribution of $Z_n$

## Assumptions

And some regularity conditions on $f$, $G$, $\mathcal{Q}$ (uniform Lipschitz) and $t^+$ (continuity)

Additional assumptions for the central limit theorem:

- $\mathcal{P}$ (transition kernel of $(Z_n)$) is a contraction mapping in $\mathbb{L}^a$-norm for some $a \geq 1$
- $\mathcal{Q}(\Phi(\cdot,\cdot),\cdot) \in \mathrm{Li}(r,s)$ with $2(r+s) \leq a$

## Assumptions

And some regularity conditions on $f$, $G$, $\mathcal{Q}$ (uniform Lipschitz) and $t^+$ (continuity)

Additional assumptions for the central limit theorem:

- $\mathcal{P}$ (transition kernel of $(Z_n)$) is a contraction mapping in $\mathbb{L}^a$-norm for some $a \geq 1$
- $\mathcal{Q}(\Phi(\cdot,\cdot),\cdot) \in \mathrm{Li}(r,s)$ with $2(r+s) \leq a$

But:

- $t^+$ is not assumed to be a bounded function
- The forms of the flow and of the transition kernel are not specified

## Recursive kernel estimates

$$\widehat{\mathcal{F}}^n(x,t) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d w_i}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)\mathbb{K}_1\left(\frac{S_{i+1}-t}{w_i}\right)$$

$$\widehat{\mathcal{G}}^n(x,t) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)\mathbb{1}_{\{S_{i+1}>t\}}$$

$$\widehat{\nu}_\infty^n(x) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)$$

Bandwidths: $v_k = v_0(k+1)^{-\alpha}$ and $w_k = w_0(k+1)^{-\beta}$ for some $\alpha$, $\beta > 0$

$\mathbb{K}_p$ is a kernel function on $\mathbb{R}^p$, $p \in \{1, d\}$

Inria informatics mathematics

## Recursive kernel estimates

$$\widehat{\mathcal{F}}^n(x,t) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d w_i}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)\mathbb{K}_1\left(\frac{S_{i+1}-t}{w_i}\right)$$

$$\widehat{\mathcal{G}}^n(x,t) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)\mathbb{1}_{\{S_{i+1}>t\}}$$

$$\widehat{\nu}_\infty^n(x) = \frac{1}{n}\sum_{i=0}^{n-1}\frac{1}{v_i^d}\mathbb{K}_d\left(\frac{Z_i - x}{v_i}\right)$$

Bandwidths: $v_k = v_0(k+1)^{-\alpha}$ and $w_k = w_0(k+1)^{-\beta}$ for some $\alpha$, $\beta > 0$

$\mathbb{K}_p$ is a kernel function on $\mathbb{R}^p$, $p \in \{1, d\}$

Admissible set for the bandwidth parameters:

$$\{(\alpha,\beta)\ :\ \alpha>0,\ \beta>0,\ \alpha d+\beta<1,\ \alpha d+\beta+2\min(\alpha,\beta)>1\}$$

## Consistency and asymptotic normality

**Strong consistency** $x \in E$ and $0 < t < t^+(x)$

$$\begin{bmatrix} \widehat{\mathcal{F}}^n(x,t) \\ \widehat{\mathcal{G}}^n(x,t) \\ \widehat{\nu}^n_\infty(x) \end{bmatrix} \xrightarrow{a.s} \begin{bmatrix} \nu_\infty(x) f(x,t) \\ \nu_\infty(x) G(x,t) \\ \nu_\infty(x) \end{bmatrix}$$

## Consistency and asymptotic normality

**Strong consistency** $x \in E$ and $0 < t < t^+(x)$

$$\begin{bmatrix} \widehat{\mathcal{F}}^n(x,t) \\ \widehat{\mathcal{G}}^n(x,t) \\ \widehat{\nu}^n_\infty(x) \end{bmatrix} \xrightarrow{a.s} \begin{bmatrix} \nu_\infty(x) f(x,t) \\ \nu_\infty(x) G(x,t) \\ \nu_\infty(x) \end{bmatrix}$$

**Asymptotic normality** $x \in E$ and $0 < t < t^+(x)$

$$n^{\frac{1-\alpha d-\beta}{2}} \left( \begin{bmatrix} \widehat{\mathcal{F}}^n(x,t) \\ \widehat{\mathcal{G}}^n(x,t) \\ \widehat{\nu}^n_\infty(x) \end{bmatrix} - \begin{bmatrix} \nu_\infty(x) f(x,t) \\ \nu_\infty(x) G(x,t) \\ \nu_\infty(x) \end{bmatrix} \right) \xrightarrow{d} \mathcal{N}(0_3, \Sigma(x,t,\alpha,\beta)),$$

where the variance-covariance matrix $\Sigma(x,t,\alpha,\beta)$ is degenerate with only one positive term at position $(1,1)$

**Sketch of the proof**

$$\begin{bmatrix} \widehat{\mathcal{F}}^n(x,t) \\ \widehat{\mathcal{G}}^n(x,t) \\ \widehat{\nu}_\infty^n(x) \end{bmatrix} - \begin{bmatrix} \nu_\infty(x) f(x,t) \\ \nu_\infty(x) G(x,t) \\ \nu_\infty(x) \end{bmatrix} = \frac{\mathcal{M}_n}{n} + \mathcal{R}_n$$

- Rate of the remainder term: Lipschitz mixing property
- Asymptotic behavior of the predictable variation process of $(\mathcal{M}_n)$:

$$\frac{\langle \mathcal{M} \rangle_n}{n^{1+\alpha d}} \sim \begin{bmatrix} n^\beta \triangle_1 & 0 & 0 \\ 0 & \triangle_2 & 0 \\ 0 & 0 & \triangle_3 \end{bmatrix}$$

## Two first corollaries

### Estimation of the conditional density $f$

$$\widehat{f}^n(x,t) = \frac{\widehat{\mathcal{F}}^n(x,t)}{\widehat{\nu}^n_\infty(x,t)}$$

**Pointwise convergence of $\widehat{f}^n(x,t)$** $x \in E$ and $0 < t < t^+(x)$

$$\widehat{f}^n(x,t) \xrightarrow{a.s.} f(x,t)$$

$$n^{\frac{1-\alpha d-\beta}{2}}\left(\widehat{f}^n(x,t) - f(x,t)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\tau_1^2\,\tau_d^2\,f(x,t)}{(1+\alpha d+\beta)\nu_\infty(x)}\right)$$

## Two first corollaries

### Estimation of the conditional survival function $G$

$$\widehat{G}^n(x,t) = \frac{\widehat{\mathcal{G}}^n(x,t)}{\widehat{\nu}^n_\infty(x)}$$

**Pointwise convergence of $\widehat{G}^n(x,t)$** $\quad x \in E$ and $0 < t < t^+(x)$

$$\widehat{G}^n(x,t) \xrightarrow{a.s.} G(x,t)$$

$$n^{\frac{1-\alpha d-\beta}{2}} \left(\widehat{G}^n(x,t) - G(x,t)\right) \xrightarrow{\mathbb{P}} 0$$

## Two first corollaries

**Estimation of the conditional survival function $G$**

$$\widehat{G}^n(x,t) = \frac{\widehat{\mathcal{G}}^n(x,t)}{\widehat{\nu}^n_\infty(x)}$$

**Pointwise convergence of $\widehat{G}^n(x,t)$** $x \in E$ and $0 < t < t^+(x)$

$$\widehat{G}^n(x,t) \xrightarrow{a.s.} G(x,t)$$

$$n^{\frac{1-\alpha d-\beta}{2}} \left( \widehat{G}^n(x,t) - G(x,t) \right) \xrightarrow{\mathbb{P}} 0$$

Actually, one also has

$$n^{\frac{1-\alpha d}{2}} \left( \widehat{G}^n(x,t) - G(x,t) \right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\tau_d^2\, G(x,t)}{(1+\alpha d)\nu_\infty(x)}\right)$$

## Outline

## Class of estimators indexed by the flow

**Estimation of the composed function $\lambda \circ \Phi$**

$$\widehat{\lambda \circ \Phi}^{n}(x,t) = \frac{\widehat{\mathcal{F}}^{n}(x,t)}{\widehat{\mathcal{G}}^{n}(x,t)}$$

**Pointwise convergence of $\widehat{\lambda \circ \Phi}^{n}(x,t)$** $\quad x \in E$ and $0 < t < t^{+}(x)$

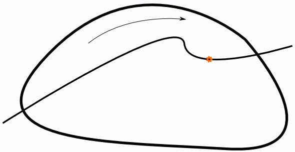
$$\widehat{\lambda \circ \Phi}^{n}(x,t) \xrightarrow{a.s.} \lambda \circ \Phi(x,t)$$

$$n^{\frac{1-\alpha d-\beta}{2}} \left( \widehat{\lambda \circ \Phi}^{n}(x,t) - \lambda \circ \Phi(x,t) \right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\tau_1^2 \tau_d^2 \lambda \circ \Phi(x,t)}{(1+\alpha d+\beta)\nu_\infty(x) G(x,t)}\right)$$

## Class of estimators indexed by the flow

Estimation of $\lambda(x)$ for some $x \in E$

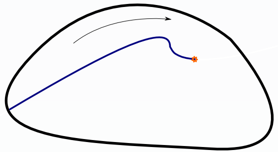
$$\mathcal{C}_x = \{\Phi(x,-t) \ : \ t \geq 0\} \cap E$$



## Class of estimators indexed by the flow

Estimation of $\lambda(x)$ for some $x \in E$

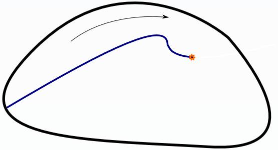
$$\mathcal{C}_x = \{\Phi(x, -t) \ : \ t \geq 0\} \cap E$$



## Class of estimators indexed by the flow

Estimation of $\lambda(x)$ for some $x \in E$

$$\mathcal{C}_x = \{\Phi(x,-t) \ : \ t \geq 0\} \cap E$$



For any $\xi \in \mathcal{C}_x$, there exists a unique time $t = \tau_x(\xi)$ such that $\Phi(\xi, \tau_x(\xi)) = x$

In particular, $\lambda \circ \Phi(\xi, \tau_x(\xi)) = \lambda(x)$.

Thus, $\widehat{\lambda \circ \Phi}^n(\xi, \tau_x(\xi))$ estimates $\lambda(x)$, for any $\xi \in \mathcal{C}_x$

## How to choose among this class?

Any element in $\Lambda^n(x) = \left\{ \widehat{\lambda}^n_\xi(x) = \widehat{\lambda \circ \Phi}^n(\xi, \tau_x(\xi)) \ : \ \xi \in \mathcal{C}_x \right\}$ is a good estimate of $\lambda(x)$ (almost sure convergence and asymptotic normality)

## How to choose among this class?

Any element in $\Lambda^n(x) = \left\{ \widehat{\lambda}^n_\xi(x) = \widehat{\lambda \circ \Phi}^n(\xi, \tau_x(\xi)) \;:\; \xi \in \mathcal{C}_x \right\}$ is a good estimate of $\lambda(x)$ (almost sure convergence and asymptotic normality)

We propose to choose among this class the estimate with the minimal asymptotic variance:

$$\frac{\tau_1^2\,\tau_d^2\,\lambda \circ \Phi(\xi, \tau_x(\xi))}{(1 + \alpha d + \beta)\nu_\infty(\xi) G(\xi, t)} \propto (\nu_\infty(\xi) G(\xi, \tau_x(\xi)))^{-1} \;=\; \kappa_x(\xi)^{-1}$$

$$\widehat{\lambda}^n(x) = \widehat{\lambda}^n_{\xi^\star}(x) \qquad \text{for} \quad \xi^\star = \arg\max_{\xi \in \mathcal{C}_x} \kappa_x(\xi)$$

## How to choose among this class?

Any element in $\Lambda^n(x) = \left\{ \widehat{\lambda}_\xi^n(x) = \widehat{\lambda \circ \Phi}^n(\xi, \tau_x(\xi)) \ : \ \xi \in \mathcal{C}_x \right\}$ is a good estimate of $\lambda(x)$ (almost sure convergence and asymptotic normality)

We propose to choose among this class the estimate with the minimal asymptotic variance:

$$\frac{\tau_1^2 \, \tau_d^2 \, \lambda \circ \Phi(\xi, \tau_x(\xi))}{(1 + \alpha d + \beta)\nu_\infty(\xi) G(\xi, t)} \propto (\nu_\infty(\xi) G(\xi, \tau_x(\xi)))^{-1} \ = \ \kappa_x(\xi)^{-1}$$

$$\widehat{\lambda}^n(x) = \widehat{\lambda}_{\xi^\star}^n(x) \qquad \text{for} \ \ \xi^\star = \arg\max_{\xi \in \mathcal{C}_x} \kappa_x(\xi)$$

$$\widehat{\widehat{\lambda}}^n(x) = \widehat{\lambda}_{\xi^\star}^n(x) \qquad \text{for} \ \ \xi^\star = \arg\max_{\xi \in \mathcal{C}_x} \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$$

## Outline

## ISE criterion

Crucial step: choice of the bandwidth parameter $\alpha$ for $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ which estimates $\kappa_x(\xi)$

ISE criterion on the line $\mathcal{C}_x$:

$$\begin{aligned} \mathsf{ISE}^n_\kappa(\alpha) &= \int_{\mathcal{C}_x} \left( \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi)) - \kappa_x(\xi) \right)^2 \mathrm{d}\xi \qquad \text{(line integral)} \\ &= \int_{\mathcal{C}_x} \kappa_x(\xi)^2 \mathrm{d}\xi + \varepsilon^n_\kappa(\alpha) \end{aligned}$$

where

$$\varepsilon^n_\kappa(\alpha) = \int_{\mathcal{C}_x} \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))^2 \mathrm{d}\xi - 2 \int_{\mathcal{C}_x} \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))\, \kappa_x(\xi) \mathrm{d}\xi$$

But $\kappa_x(\xi)$ is unknown $\rightarrow$ cross-validation method

## Cross-validation methods

Assume that $d = 1$ ($E$ is a subset of $\mathbb{R}$)

$$\frac{1}{\widetilde{n}} \sum_{k=0}^{\widetilde{n}-1} \widehat{\mathcal{G}}^n(\widetilde{Z}_k, \tau_x(\widetilde{Z}_k)) \xrightarrow{a.s.} \int \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))\, \nu_\infty(\xi) \, \mathrm{d}\xi,$$

where $(\widetilde{Z}_k, \widetilde{S}_{k+1})$ is the embedded chain of another independent PDMP generated from the same characteristics

## Cross-validation methods

Assume that $d = 1$ ($E$ is a subset of $\mathbb{R}$)

$$\frac{1}{\widetilde{n}} \sum_{k=0}^{\widetilde{n}-1} \widehat{\mathcal{G}}^n(\widetilde{Z}_k, \tau_x(\widetilde{Z}_k)) \, \mathbb{1}_{\mathcal{C}_x}(\widetilde{Z}_k) \xrightarrow{a.s.} \int_{\mathcal{C}_x} \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi)) \, \nu_\infty(\xi) \, \mathrm{d}\xi,$$

where $(\widetilde{Z}_k, \widetilde{S}_{k+1})$ is the embedded chain of another independent PDMP generated from the same characteristics

## Cross-validation methods

Assume that $d = 1$ ($E$ is a subset of $\mathbb{R}$)

$$\frac{1}{\widetilde{n}} \sum_{k=0}^{\widetilde{n}-1} \widehat{\mathcal{G}}^n(\widetilde{Z}_k, \tau_x(\widetilde{Z}_k)) \, \mathbb{1}_{\mathcal{C}_x}(\widetilde{Z}_k) \mathbb{1}_{(\tau_x(\widetilde{Z}_k),+\infty)}(\widetilde{S}_{k+1}) \xrightarrow{a.s.} \int_{\mathcal{C}_x} \widehat{\mathcal{G}}^n(\xi, \tau_x(\xi)) \, \nu_\infty(\xi) G(\xi, \tau_x(\xi)) \, \mathrm{d}\xi,$$

where $(\widetilde{Z}_k, \widetilde{S}_{k+1})$ is the embedded chain of another independent PDMP generated from the same characteristics

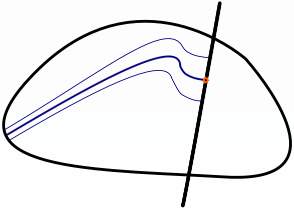
## Cross-validation methods

If $d > 1$, we use that

$$\frac{1}{\lambda_{d-1}(\mathbb{D}_{x,\rho})} \int_{\mathbb{T}_{x,\rho}} \varphi \longrightarrow \int_{\mathcal{C}_x} \varphi,$$

where

- $\mathbb{H}_x = \left\{ y \in \mathbb{R}^d \ : \ y - x \perp \frac{\partial \Phi}{\partial t}(x, 0) \right\}$
- $\mathbb{D}_{x,\rho} = B_d(x, \rho) \cap \mathbb{H}_x$
- $\mathbb{T}_{x,\rho} = \bigcup_{y \in \mathbb{D}_{x,\rho}} \mathcal{C}_y$



## Cross-validation methods

Approximation of the ISE criterion (up to an additive constant):

$$\int_{\mathcal{C}_x} \widehat{\kappa}_x^n(\xi)^2 \mathrm{d}\xi - \frac{2\,\Gamma\left(\frac{d-1}{2}+1\right)}{\widetilde{n}\,\pi^{\frac{d-1}{2}}\,\rho^{d-1}} \sum_{k=0}^{\widetilde{n}-1} \widehat{\mathcal{G}}^n\Big(\widetilde{Z}_k, \theta_x(\widetilde{Z}_k)\Big)\, \mathbb{1}_{\mathbb{T}_{x,\rho}}(\widetilde{Z}_k)\, \mathbb{1}_{(\theta_x(\widetilde{Z}_k),+\infty)}(\widetilde{S}_{k+1})$$

$$\xrightarrow{a.s.} \mathsf{ISE}_\kappa^n(\alpha) \qquad \text{when } \widetilde{n} \to +\infty,\ \rho \to 0$$

## Cross-validation methods

Approximation of the ISE criterion (up to an additive constant):

$$\int_{\mathcal{C}_x} \widehat{\kappa}_x^n(\xi)^2 \mathrm{d}\xi - \frac{2\,\Gamma\left(\frac{d-1}{2}+1\right)}{\widetilde{n}\,\pi^{\frac{d-1}{2}}\rho^{d-1}} \sum_{k=0}^{\widetilde{n}-1} \widehat{\mathcal{G}}^n\Big(\widetilde{Z}_k, \theta_x(\widetilde{Z}_k)\Big)\, \mathbb{1}_{\mathbb{T}_{x,\rho}}(\widetilde{Z}_k)\, \mathbb{1}_{(\theta_x(\widetilde{Z}_k),+\infty)}(\widetilde{S}_{k+1})$$

$$\xrightarrow{a.s.} \mathsf{ISE}_\kappa^n(\alpha) \qquad \text{when } \widetilde{n} \to +\infty,\ \rho \to 0$$

Let us recall that $\widehat{\lambda}_\xi^n(x) = \dfrac{\widehat{\mathcal{F}}^n(\xi, \tau_x(\xi))}{\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))}$

- The same choice of $\alpha$ in $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ may be applied for computing the denominator of $\widehat{\lambda}_\xi^n(x)$
- Same cross-validation procedure for choosing $\alpha$ and $\beta$ in $\widehat{\mathcal{F}}^n(\xi, \tau_x(\xi))$

## Outline

## Estimation procedure

**Estimation procedure in 4 steps**

1. Cross-validation on $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ (denominator)

Remark: steps 2 and 3 are permutable.

**Estimation procedure in 4 steps**

1. Cross-validation on $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ (denominator)
2. Compute the optimal point $\xi^\star \in \mathcal{C}_x$ which maximizes $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$

Remark: steps 2 and 3 are permutable.

## Estimation procedure

### Estimation procedure in 4 steps

1. Cross-validation on $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ (denominator)
2. Compute the optimal point $\xi^\star \in \mathcal{C}_x$ which maximizes $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$
3. Cross-validation on $\widehat{\mathcal{F}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ and $\beta$ (numerator)

Remark: steps 2 and 3 are permutable.

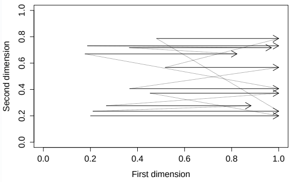
**Estimation procedure in 4 steps**

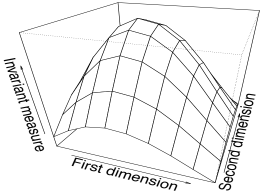
1. Cross-validation on $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ (denominator)
2. Compute the optimal point $\xi^\star \in \mathcal{C}_x$ which maximizes $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$
3. Cross-validation on $\widehat{\mathcal{F}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ and $\beta$ (numerator)
4. Compute $\widehat{\widehat{\lambda}}^n(x) = \widehat{\lambda}^n_{\xi^\star}(x)$ with the optimal parameters $\alpha$ and $\beta$

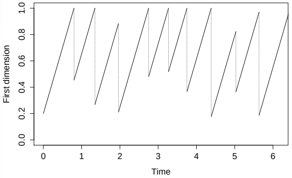
Remark: steps 2 and 3 are permutable.

## TCP-like model

- $E = (0,1)^2$
- $\Phi((x_1,x_2),t) = (x_1+t, x_2)$
- $\mathcal{Q}((x_1,x_2),\cdot) = \beta(2, 2/x_1) \otimes \beta(2,2)$
- $\lambda((x_1,x_2)) = x_1 + x_2$



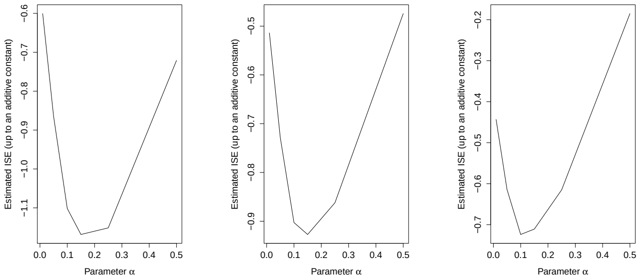
## Outline

- Aim: estimation of the jump rate $\lambda$ at $x = (0.75, 0.5)$ (not under the mode of the invariant distribution)
- $\mathcal{C}_x = (0, 0.75] \times \{0.5\}$
- The $10'000$ first jumps of the Markov chain $(Z_n, S_{n+1})$ are observed
- The $1'000$ first jumps of another Markov chain $(\widetilde{Z}_n, \widetilde{S}_{n+1})$ are observed for the cross-validation procedure
- When boxplots are presented, they have been computed over $100$ replicates

## Simulations

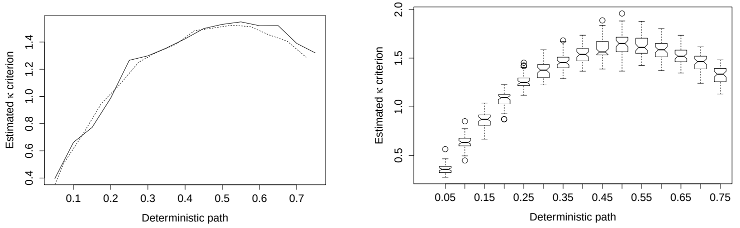
**1st step:** cross-validation on $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$
(denominator)

$\rho = 0.005$ $\rho = 0.01$ $\rho = 0.02$



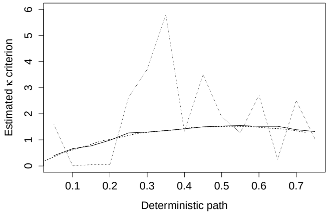
## Simulations

**2nd step:** compute the optimal point $\xi^\star \in \mathcal{C}_x$ which maximizes $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$



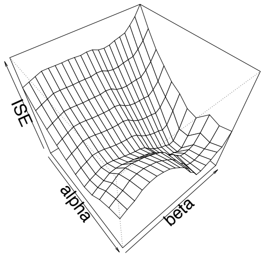
## Simulations

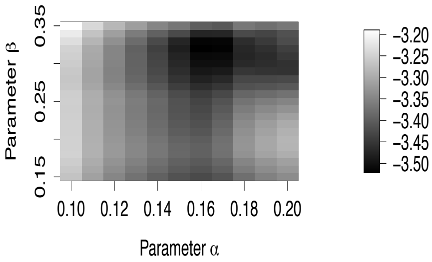
**2nd step:** compute the optimal point $\xi^\star \in \mathcal{C}_x$ which maximizes $\widehat{\mathcal{G}}^n(\xi, \tau_x(\xi))$




Poor choice of $\alpha$ ($\alpha = 0.04$)

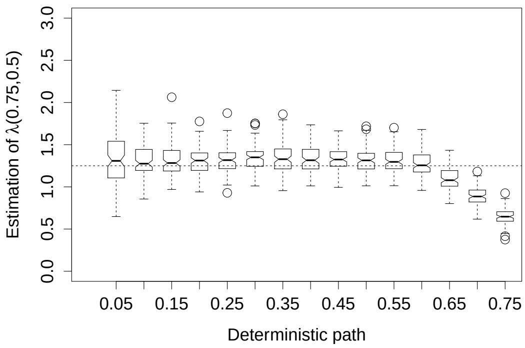
**3rd step:** cross-validation on $\widehat{\mathcal{F}}^n(\xi, \tau_x(\xi))$ for choosing $\alpha$ and $\beta$ (numerator)



## Simulations

**4th step:** compute $\widehat{\widehat{\lambda}}^{n}_{\xi^\star}(x)$ with the optimal parameters $\alpha$ and $\beta$



## Simulations

**4th step:** compute $\widehat{\widehat{\lambda}}^{n}_{\xi^\star}(x)$ with the optimal parameters $\alpha$ and $\beta$