Spatial birth-death-move processes: basic properties and estimation of their intensity functions

Frédéric Lavancier

Laboratoire de Mathématiques Jean Leray, University of Nantes

Joint work with Ronan Le Guével (Rennes 2 University)





From spatial data...

Proteins inside a living cell, acquired by fluorescence microscopy. These proteins are involved in the exocytosis process.



From spatial data...

Proteins inside a living cell, acquired by fluorescence microscopy. These proteins are involved in the exocytosis process.

... to spatio-temporal data:

7 frames per second, 1200 frames





- 2 Some probabilistic properties
- 3 Statistical problem
- Application to real data from cellular biology



Birth-death-move point processes

The state space of the birth-death-move process $(X_t)_{t\geq 0}$ is

$$E=\bigcup_{n=0}^{+\infty}E_n$$

for disjoint spaces E_n and $E_0 = \{ \varnothing \}$.

Main example: $E_n = \{ \text{point configurations in } S \subset \mathbb{R}^d \text{ with cardinality } n \}$

Starting from $X_0 \in E_n$ for some n:

- Move step: $(X_t)_{t\geq 0}$ moves according to a continuous Markov process in E_n up to a random jump time T_1
- Jump (birth or death): at $t = T_1$, there is a birth (then $X_{T_1} \in E_{n+1}$) or a death (then $X_{T_1} \in E_{n-1}$)
- Move step: $(X_t)_{t \ge T_1}$ moves according to a continuous Markov process in the new space $(E_{n+1} \text{ or } E_{n-1})$ up to a random jump time T_2
- Jump (birth or death): at $t = T_2$, there is a birth or a death
- and so on.

 ${\it E}_n$: space of point configurations in $[0,1]^2$ with cardinality n No move between jumps



 E_n : space of point configurations in $[0, 1]^2$ with cardinality *n* Independent Brownian motions between jumps



 E_n : space of point configurations in $[0,1]^2$ with cardinality n Langevin diffusion (Lennard-Jones potential) between jumps



 E_n : space of marked point configurations in $[0,1]^2 \times [0,\infty[$ with cardinality nThe move is a growing process that only applies to the marks



$$E = \bigcup_{n=0}^{+\infty} E_n$$
 for disjoint spaces E_n and $E_0 = \{\emptyset\}$.

The birth-death-move process $(X_t)_{t\geq 0}$ on (E, \mathcal{E}) depends on 3 ingredients:

 $E = \bigcup_{n=0}^{+\infty} E_n$ for disjoint spaces E_n and $E_0 = \{\emptyset\}$.

The birth-death-move process $(X_t)_{t\geq 0}$ on (E, \mathcal{E}) depends on 3 ingredients:

1. $\beta: E \to \mathbb{R}_+$ and $\delta: E \to \mathbb{R}_+$: the birth and death intensity functions

 $\longrightarrow \beta$ and δ are assumed to be continuous on E and $\delta(\emptyset) = 0$.

We let $\alpha = \beta + \delta$ be the total jump intensity function. We denote by T_1, T_2, \ldots the sequence of jump times and $T_0 = 0$.

 $E = \bigcup_{n=0}^{+\infty} E_n$ for disjoint spaces E_n and $E_0 = \{\emptyset\}$.

The birth-death-move process $(X_t)_{t\geq 0}$ on (E, \mathcal{E}) depends on 3 ingredients:

1. $\beta: E \to \mathbb{R}_+$ and $\delta: E \to \mathbb{R}_+$: the birth and death intensity functions

 $\longrightarrow \beta$ and δ are assumed to be continuous on E and $\delta(\emptyset) = 0$.

We let $\alpha = \beta + \delta$ be the total jump intensity function. We denote by T_1, T_2, \ldots the sequence of jump times and $T_0 = 0$.

2. $K_{\beta}: E \times \mathcal{E} \to [0,1]$ and $K_{\delta}: E \times \mathcal{E} \to [0,1]$: the transition kernel for a birth and for a death

 \longrightarrow For all $x \in E$, $K_{\beta}(x, E_{n+1}) = \mathbf{1}_{x \in E_n}$ and $K_{\delta}(x, E_{n-1}) = \mathbf{1}_{x \in E_n}$.

 $E = \bigcup_{n=0}^{+\infty} E_n$ for disjoint spaces E_n and $E_0 = \{\emptyset\}$.

The birth-death-move process $(X_t)_{t\geq 0}$ on (E, \mathcal{E}) depends on 3 ingredients:

1. $\beta: E \to \mathbb{R}_+$ and $\delta: E \to \mathbb{R}_+$: the birth and death intensity functions

 $\longrightarrow \beta$ and δ are assumed to be continuous on E and $\delta(\emptyset) = 0$.

We let $\alpha = \beta + \delta$ be the total jump intensity function. We denote by T_1, T_2, \ldots the sequence of jump times and $T_0 = 0$.

2. $K_{\beta}: E \times \mathcal{E} \to [0,1]$ and $K_{\delta}: E \times \mathcal{E} \to [0,1]$: the transition kernel for a birth and for a death

 $\longrightarrow \text{ For all } x \in E, \ K_{\beta}(x, E_{n+1}) = \mathbf{1}_{x \in E_n} \text{ and } \ K_{\delta}(x, E_{n-1}) = \mathbf{1}_{x \in E_n}.$

A continuous Markov process (Y^(n,j)_{t≥0} in E_n that drives the motion of X_t ∈ E_n between two jumps T_j and T_{j+1}:

$$X_t = Y_{t-T_j}^{(n,j)}$$
 for $T_j \leq t < T_{j+1}$.

 \longrightarrow We assume that the law of $Y_t^{(n,j)}$ only depends on $n: \forall j, (Y_t^{(n,j)}) \stackrel{\mathcal{L}}{=} (Y_t^{(n)})$

Starting from an initial configuration $X_0 \in E_n$,

• Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.

Starting from an initial configuration $X_0 \in E_n$,

• Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.

• Given $(Y_t^{(n)})_{t\geq 0}$, generate the first waiting time $\mathcal{T}_1-\mathcal{T}_0$ according to

$$\mathbb{P}(T_1 - T_0 < t) = 1 - \exp\left(-\int_0^t \alpha(Y_s^{(n)})ds\right)$$

Starting from an initial configuration $X_0 \in E_n$,

• Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.

• Given $(Y_t^{(n)})_{t\geq 0}$, generate the first waiting time $\mathcal{T}_1-\mathcal{T}_0$ according to

$$\mathbb{P}(T_1 - T_0 < t) = 1 - \exp\left(-\int_0^t \alpha(Y_s^{(n)}) ds\right)$$

• For $t \in [0, T_1)$, $X_t = Y_t^{(n)}$

- Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.
- Given $(Y_t^{(n)})_{t\geq 0}$, generate the first waiting time $\mathcal{T}_1-\mathcal{T}_0$ according to

$$\mathbb{P}(T_1 - T_0 < t) = 1 - \exp\left(-\int_0^t \alpha(Y_s^{(n)}) ds\right)$$

- For $t \in [0, T_1)$, $X_t = Y_t^{(n)}$
- The jump at $t = T_1$ is a birth with probability $\beta(Y_{T_1}^{(n)})/\alpha(Y_{T_1}^{(n)})$ and it is a death otherwise.

- Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.
- Given $(Y_t^{(n)})_{t\geq 0}$, generate the first waiting time $T_1 T_0$ according to

$$\mathbb{P}(T_1 - T_0 < t) = 1 - \exp\left(-\int_0^t \alpha(Y_s^{(n)}) ds\right)$$

- For $t \in [0, T_1)$, $X_t = Y_t^{(n)}$
- The jump at $t = T_1$ is a birth with probability $\beta(Y_{T_1}^{(n)})/\alpha(Y_{T_1}^{(n)})$ and it is a death otherwise.
- The post-jump configuration X_{T1} is generated according to K_β(Y⁽ⁿ⁾_{T1},.) if the jump is a birth, and according to K_δ(Y⁽ⁿ⁾_{T1},.) if it is a death.

- Generate $Y_t^{(n)}$ for $t \ge 0$ conditionally on $Y_0^{(n)} = X_0$.
- Given $(Y_t^{(n)})_{t\geq 0}$, generate the first waiting time $\mathcal{T}_1-\mathcal{T}_0$ according to

$$\mathbb{P}(T_1 - T_0 < t) = 1 - \exp\left(-\int_0^t \alpha(Y_s^{(n)}) ds\right)$$

- For $t \in [0, T_1)$, $X_t = Y_t^{(n)}$
- The jump at $t = T_1$ is a birth with probability $\beta(Y_{T_1}^{(n)})/\alpha(Y_{T_1}^{(n)})$ and it is a death otherwise.
- The post-jump configuration X_{T1} is generated according to K_β(Y⁽ⁿ⁾_{T1},.) if the jump is a birth, and according to K_δ(Y⁽ⁿ⁾_{T1},.) if it is a death.
- We start a new motion from $t = T_1$, and so on.



2 Some probabilistic properties

Theorem

A spatial birth-death-move process is a time-homogeneous Markov process.

This is because:

- There is a "renewal" of the process after each jump T_j
- $(Y_t^{(n)})_{t\geq 0}$ is a continuous Markov process
- The (conditional) waiting time before the next jump has an exponential form

$$\mathbb{P}(T_{j+1} - T_j > t) = \exp\left(-\int_0^t \alpha(Y_s^{(n)})ds\right)$$

leading to a kind of memoryless property.

Let
$$\beta_n = \sup_{x \in E_n} \beta(x)$$
 and $\delta_n = \inf_{x \in E_n} \delta(x)$.

We assume :

(H1) For all $n \ge 1$, $\delta_n > 0$ and there exists n^* such that $\beta_n = 0$ for all $n \ge n^*$.

In light of this assumption, we set $E = \bigcup_{n=0}^{n^*} E_n$.

Proposition

Under (H₁), (X_t)_{t≥0} admits a unique invariant probability measure μ_{∞} and there exist a > 0 and c > 0 such that

$$\sup_{y\in E}\sup_{F\in\mathcal{E}} \sup |\mathbb{P}(X_t\in F|X_0=y)-\mu_{\infty}(F)| \leq ae^{-ct}.$$

For birth-death process the existence of μ_{∞} is due to Preston (1975) and the rate of convergence to Møller (1989), both using coupling arguments.

The proof for general birth-death-move processes exploits the same kind of arguments.

Let N_t be the number of jumps before t, i.e. $N_t = Card\{j \ge 1 : T_j \le t\}$. Let $\mathcal{F}_t = \sigma(X_s, s \le t)$ be the natural filtration of the process $(X_t)_{t\ge 0}$.

Proposition

The intensity of N_t with respect to \mathcal{F}_t is $\alpha(X_{t^-})$.

Corollary

For any measurable bounded function g, the process $(M_t)_{t\geq 0}$ defined by

$$M_t = \int_0^t g(X_{s-})[dN_s - \alpha(X_s)ds]$$

is a martingale with respect to \mathcal{F}_t and for all $t\geq 0$

$$\mathbb{E}(M_t^2) = \mathbb{E}\left(\int_0^t g^2(X_s)\alpha(X_s)ds\right).$$

Same results for the number of jumps (with β instead of α) and for the number of deaths (with δ instead of α).

- Same kind of probabilistic results under more general assumptions
- Some connections with branching processes
- The stationary distribution μ_{∞} can be characterised in some particular cases (specific moves, specific jumps) as a Gibbs measure.



For $x \in E$, we consider the estimation of the intensities $\alpha(x)$, $\beta(x)$ and $\delta(x)$

Two possible statistical frameworks:

Ontinous time:

We observe all the process up to time T, that is $(X_s)_{0 \le s \le T}$

Oiscrete time:

We observe the process at discrete times $t_0 < \cdots < t_m$.

In the following, we focus on $\alpha(x)$ in continuous time. The estimation of $\beta(x)$ and $\delta(x)$ can be treated similarly.

Importantly: No need to specify the diffusion process between the jumps, nor the birth and death transition kernels.

Birth-death-move point processes





Statistical problem • Kernel estimator

Simulations

Application to real data from cellular biology

Assume we observe $(X_s)_{0 \le s \le T}$ (continuous-time observations).

For a given $x \in E$, the idea to estimate $\alpha(x)$ is

 $\hat{\alpha}(x) \approx \frac{\text{number of jumps when the process is in configuration } x}{\text{time spent by the process in configuration } x}$

Formally, let $(k_T)_{T\geq 0}$ be a family of bounded non-negative functions on $E \times E$. For $x \in E$ and $y \in E$, $k_T(x, y)$ quantifies the proximity of x and y.

$$\hat{\alpha}(x) = \frac{\sum_{j=1}^{N_T} k_T(x, X_{T_j^-})}{\hat{T}(x)},$$

where

$$\hat{T}(x) = \int_0^T k_T(x, X_s) ds.$$

Typically

$$k_T(x,y) = k\left(\frac{d(x,y)}{h_T}\right), \quad x,y \in E,$$

where k is a symmetric density function, d(.,.) is a (pseudo)-distance on E and $h_T > 0$ is a bandwidth.

Typically

$$k_T(x,y) = k\left(rac{d(x,y)}{h_T}
ight), \quad x,y \in E,$$

where k is a symmetric density function, d(.,.) is a (pseudo)-distance on E and $h_T > 0$ is a bandwidth.

In a pure non-parametric setting :

• d(.,.) can be the Hausdorff distance:

$$d(x, y) = \max\{\max_{u \in x} \min_{v \in y} ||u - v||, \max_{v \in y} \min_{u \in x} ||v - u||\}$$

• or an optimal matching distance (for some $\kappa > 0$): for $x = \{x_1, \dots, x_{n(x)}\}, y = \{y_1, \dots, y_{n(y)}\}$ with $n(x) \le n(y)$,

$$d_{\kappa}(x,y) = \frac{1}{n(y)} \left(\min_{\pi} \sum_{i=1}^{n(x)} (\|x_i - y_{\pi(i)}\| \wedge \kappa) + \kappa(n(y) - n(x)) \right).$$

Typically

$$k_T(x,y) = k\left(\frac{d(x,y)}{h_T}\right), \quad x,y \in E,$$

where k is a symmetric density function, d(.,.) is a (pseudo)-distance on E and $h_T > 0$ is a bandwidth.

Or assuming structural hypothesis on α :

• If we assume that $n(x) = n(y) \Rightarrow \alpha(x) = \alpha(y)$, we can choose:

•
$$d(x, y) = |n(x) - n(y)|$$

• or even $k_T(x, y) = \mathbf{1}_{n(x)=n(y)}$ in which case we recover the "standard" estimator

 $\hat{\alpha}(x) = \hat{\alpha}(n(x)) = \frac{\#\{\text{configurations with } n(x) \text{ points}\}}{\text{"Time spent in these configurations"}}$

 More generally, if f(x) = f(y) ⇒ α(x) = α(y) for some feature f ∈ ℝ^p. We can choose: d(x, y) = ||f(x) - f(y)||

Consistency

We assume

- (H1) For all $n \ge 1$, $\delta_n > 0$ and there exists n^* such that $\beta_{n^*} = 0$.
- (H2) $\alpha(.)$ and $k_T(.,.)$ bounded

(H3) Setting
$$v_T(x) = \int_E k_T(x,z) \mu_\infty(dz)$$
,

$$Tv_T(x) o \infty$$
 and $w_T(x) := rac{1}{v_T(x)} \int_E (lpha(z) - lpha(x)) k_T(x,z) \mu_\infty(dz) o 0$

Theorem

Under these assumptions,

$$\hat{\alpha}(x) - \alpha(x) = O_{\rho}\left(\frac{1}{Tv_{T}(x)} + w_{T}(x)^{2}\right) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

Remark:

 $1/(Tv_T(x))$ can be seen as a variance term and $w_T(x)$ as a bias term.

[The proof exploits Markov property + Ergodicity + Martingales properties.]

(H3) is an assumption on the rate of convergence of the bandwidth.

<u>Illustration</u>: If $k_T(y, z) = \mathbf{1}_{d(y, z) < h_T}$, then

$$v_T(x) = \int_E k_T(x,z)\mu_\infty(dz) = \int_E \mathbf{1}_{d(x,z) < h_T}\mu_\infty(dz) = \mu_\infty(B(x,h_T))$$

and if moreover $\boldsymbol{\alpha}$ is Lipshitz then

$$|w_{T}(x)| = \frac{1}{v_{T}(x)} \left| \int_{E} (\alpha(z) - \alpha(x)) k_{T}(x, z) \mu_{\infty}(dz) \right|$$

$$\leq \frac{Lip(\alpha)}{v_{T}(x)} \int_{E} d(x, z) \mathbf{1}_{d(x, z) < h_{T}} \mu_{\infty}(dz)$$

$$\leq \frac{Lip(\alpha)}{v_{T}(x)} h_{T} \int_{E} \mathbf{1}_{d(x, z) < h_{T}} \mu_{\infty}(dz)$$

$$= Lip(\alpha) h_{T}.$$

So (H3) reduces in this case to

$$h_T
ightarrow 0$$
 and $T \mu_\infty(B(x,h_T))
ightarrow \infty.$

(H3) is an assumption on the rate of convergence of the bandwidth.

<u>Illustration</u>: If $k_T(y, z) = \mathbf{1}_{d(y,z) < h_T}$, then

$$v_T(x) = \int_E k_T(x,z)\mu_\infty(dz) = \int_E \mathbf{1}_{d(x,z) < h_T}\mu_\infty(dz) = \mu_\infty(B(x,h_T))$$

and if moreover $\boldsymbol{\alpha}$ is Lipshitz then

$$|w_{\mathcal{T}}(x)| = \frac{1}{v_{\mathcal{T}}(x)} \left| \int_{E} (\alpha(z) - \alpha(x)) k_{\mathcal{T}}(x, z) \mu_{\infty}(dz) \right|$$

$$\leq \frac{Lip(\alpha)}{v_{\mathcal{T}}(x)} \int_{E} d(x, z) \mathbf{1}_{d(x, z) < h_{\mathcal{T}}} \mu_{\infty}(dz)$$

$$\leq \frac{Lip(\alpha)}{v_{\mathcal{T}}(x)} h_{\mathcal{T}} \int_{E} \mathbf{1}_{d(x, z) < h_{\mathcal{T}}} \mu_{\infty}(dz)$$

$$= Lip(\alpha) h_{\mathcal{T}}.$$

So (H3) reduces in this case to

$$h_T
ightarrow 0$$
 and $T\mu_{\infty}(B(x,h_T))
ightarrow \infty$.

<u>Remark</u>: If $f(y) = f(z) \Rightarrow \alpha(y) = \alpha(z)$ for some feature $f \in \mathbb{R}^p$, and we choose d(y,z) = ||f(y) - f(z)|| in k_T , then $v_T(x) = O(h_T^p)$.

Cross-validation to choose the bandwidth

Cross-validation to choose the bandwidth

Standard CV methods (plug-in, least-squares) rely on second order moments. \longrightarrow They are unkown for $\hat{\alpha}(x)$. Standard CV methods (plug-in, least-squares) rely on second order moments. \rightarrow They are unkown for $\hat{\alpha}(x)$.

<u>Partial likelihood cross-validation</u> Remember: the intensity of N_t with respect to \mathcal{F}_t is $\alpha(X_{t-})$.

So the log-likelihood of $(N_t)_{0 \le t \le T}$ with respect to the unit rate Poisson counting process on [0, T] is

$$T - \int_0^T lpha(X_s) ds + \sum_{j=1}^{N_T} \log lpha(X_{T_j^-}).$$

Bandwidth selection by partial likelihood amounts to choose h as

$$\hat{h} = \operatorname*{argmax}_{h} \sum_{j=1}^{N_T} \log \hat{\alpha}_h^{(-)}(X_{T_j^-}) - \int_0^T \hat{\alpha}_h^{(-)}(X_s) ds$$

where $\hat{\alpha}_h^{(-)}(X_s)$ is the estimator without using the observation X_s . To carry out this removal, we discard all observations in $[T_{N_s}, T_{N_s+1}]$. Birth-death-move point processes

2 Some probabilistic properties



3 Statistical problem

- Kernel estimator
- Simulations

Application to real data from cellular biology

We start at t = 0 with ≈ 100 points. The jump rate is

$$\alpha(x) = \exp\left(5\left(\frac{n(x)}{100} - 1\right)\right)$$



A jump is a birth with probability 1/2, otherwise it is a death. Births and deaths occur uniformly in space. We observe the following realisation. It contains 1530 jumps.





Left: $n(X_{T_i})$ for i = 1, ..., 1530. This is observed

Right: $\alpha(X_{T_i}) = \alpha(n(X_{T_i}))$ for i = 1, ..., 1530. We would like to estimate this.

Estimation in the continuous case

First estimator:

Pure non-parametric setting using the Hausdorff distance:

$$\hat{\alpha}_1(x) = \frac{1}{\hat{T}(x)} \sum_{j=1}^{N_T} k\left(\frac{d(x, X_{T_j^-})}{h_T}\right)$$

where d(.,.) is the Hausdorff distance



Estimation in the continuous case

Second estimator:

Pure non-parametric setting using the optimal matching distance:

$$\hat{lpha}_2(x) = rac{1}{\hat{T}(x)} \sum_{j=1}^{N_T} k\left(rac{d_\kappa(x, X_{\mathcal{T}_j^-})}{h_\mathcal{T}}
ight)$$

where $d_{\kappa}(.,.)$ is the optimal matching distance



Third estimator: Assuming (rightly) that $\alpha(x)$ only depends on n(x):

 $\hat{\alpha}_3(x) = \hat{\alpha}_3(n(x)) = \frac{\#\{\text{configurations with } n(x) \text{ points}\}}{\text{"Time spent in these configurations"}}$



Estimation in the continuous case

Fourth estimator: Assuming (rightly) that $\alpha(x)$ only depends on n(x):

$$\hat{\alpha}_4(x) = \hat{\alpha}_4(n(x)) = \frac{1}{\hat{T}(x)} \sum_{j=1}^{N_T} k\left(\frac{|n(x) - n(X_{T_j^-})|}{h_T}\right)$$



Estimation in the continuous case



In practice, we do not know whether $\alpha(x)$ only depends on n(x) or not. Starting from the pure non-parametric estimator, we may question this property.

Scatterplot of $(n(X_{T_i}), \hat{\alpha}_1(X_{T_i}))$ and $(n(X_{T_i}), \hat{\alpha}_2(X_{T_i}))$ for $i = 1, \dots, 1530$.



Birth-death-move point processes

2 Some probabilistic properties

Statistical problem

- Kernel estimator
- Simulations



1200 frames observed at t_0, \ldots, t_m .



Langerin proteins 1.26 jumps in average



Rab11 proteins 0.85 jumps in average

Application to data from cellular biology

The birth intensity $\beta(.)$ seems constant in both channels

For the death intensity $\delta(.)$: estimation of $\delta(X_{t_i})$ for i = 0, ..., m

- using d_{κ}
- using d(x, y) = |n(x) n(y)|



Scatterplots of $(n(X_{t_i}), \hat{\delta}(X_{t_i}))$



Application to data from cellular biology

Cross-correlation between the two types of proteins



- Birth-death-move processes: particles appear, move and disappear.
- Simple assumptions for their ergodic properties.
- We established the consistency of the kernel estimator of their intensities
- In practice, the choice of d(.,.) allows to introduce structural hypothesis.
- Partial-likelihood cross-validation to choose the bandwidth.
- Perspectives: estimation of the transition kernels; generalisation to multitype processes; testing procedures;...

- Birth-death-move processes: particles appear, move and disappear.
- Simple assumptions for their ergodic properties.
- We established the consistency of the kernel estimator of their intensities
- In practice, the choice of d(.,.) allows to introduce structural hypothesis.
- Partial-likelihood cross-validation to choose the bandwidth.
- Perspectives: estimation of the transition kernels; generalisation to multitype processes; testing procedures;...

THANK YOU!

Remember that

$$\hat{\alpha}(x) = \frac{1}{\hat{T}(x)} \int_0^T k_T(x, X_{s^-}) dN_s \quad \text{with} \quad \hat{T}(x) = \int_0^T k_T(x, X_s) ds$$

For the proof we use the decomposition

$$\hat{\alpha}(x) - \alpha(x) = \frac{1}{\hat{T}(x)} \int_0^T k_T(x, X_{s^-}) (dN_s - \alpha(X_s)ds) + \frac{1}{\hat{T}(x)} \int_0^T (\alpha(X_s) - \alpha(x)) k_T(x, X_{s^-}) ds$$

Remember that

$$\hat{\alpha}(x) = \frac{1}{\hat{T}(x)} \int_0^T k_T(x, X_{s^-}) dN_s \quad \text{with} \quad \hat{T}(x) = \int_0^T k_T(x, X_s) ds$$

For the proof we use the decomposition

$$\hat{\alpha}(x) - \alpha(x) = \frac{1}{\hat{\tau}(x)} \int_0^T k_{\tau}(x, X_{s^-}) (dN_s - \alpha(X_s)ds) + \frac{1}{\hat{\tau}(x)} \int_0^T (\alpha(X_s) - \alpha(x)) k_{\tau}(x, X_{s^-}) ds$$

We need to control

1- $M_T = \int_0^T k_T(x, X_{s^-}) dA_s$ where $A_t = N_t - \int_0^t \alpha(X_s) ds$ 2- $\int_0^T (\alpha(X_s) - \alpha(x)) k_T(x, X_s) ds$ 3- $\hat{T}(x) = \int_0^T k_T(x, X_s) ds$, the time spent by $(X_s)_{0 \le s \le T}$ "at" x. Remember that

$$\hat{\alpha}(x) = \frac{1}{\hat{T}(x)} \int_0^T k_T(x, X_{s^-}) dN_s \quad \text{with} \quad \hat{T}(x) = \int_0^T k_T(x, X_s) ds$$

For the proof we use the decomposition

$$\hat{\alpha}(x) - \alpha(x) = \frac{1}{\hat{T}(x)} \int_0^T k_T(x, X_{s-}) (dN_s - \alpha(X_s) ds) + \frac{1}{\hat{T}(x)} \int_0^T (\alpha(X_s) - \alpha(x)) k_T(x, X_{s-}) ds$$

We need to control

- 1- $M_T = \int_0^T k_T(x, X_{s^-}) dA_s$ where $A_t = N_t \int_0^t \alpha(X_s) ds$ $\longrightarrow A_t$ and M_T are martingales. We deduce $\mathbb{E}(M_t^2)$.
- 2- $\int_0^T (\alpha(X_s) \alpha(x)) k_T(x, X_s) ds$ \longrightarrow Ergodicity + (H3)
- 3- $\hat{T}(x) = \int_0^T k_T(x, X_s) ds$, the time spent by $(X_s)_{0 \le s \le T}$ "at" x. \longrightarrow Ergodicity: $\mathbb{E}(\hat{T}(x)) \sim Tv_T(x)$ and $\mathbb{V}(\hat{T}(x)) \le cTv_T(x)$

Discrete-time observations

We observe X_{t_0}, \ldots, X_{t_m} at t_0, \ldots, t_m with $t_0 = 0$ and $t_m = T$.

Remember that in the continuous case:

$$\hat{\alpha}(x) = \frac{\int_0^T k_T(x, X_{s^-}) dN_s}{\int_0^T k_T(x, X_s) ds}$$

A Riemann approximation at t_0, \ldots, t_m would give :

$$\frac{\sum_{j=0}^{m-1} (N_{t_{j+1}} - N_{t_j}) k_T(x, X_{t_j})}{\sum_{j=0}^{m-1} (t_{j+1} - t_j) k_T(x, X_{t_j})}.$$

Problem : we do not observe the exact number of jumps $N_{t_{j+1}} - N_{t_j}$.

We observe X_{t_0}, \ldots, X_{t_m} at t_0, \ldots, t_m with $t_0 = 0$ and $t_m = T$.

Remember that in the continuous case:

$$\hat{\alpha}(x) = \frac{\int_0^T k_T(x, X_{s^-}) dN_s}{\int_0^T k_T(x, X_s) ds}$$

A Riemann approximation at t_0, \ldots, t_m would give :

$$\frac{\sum_{j=0}^{m-1} (N_{t_{j+1}} - N_{t_j}) k_T(x, X_{t_j})}{\sum_{j=0}^{m-1} (t_{j+1} - t_j) k_T(x, X_{t_j})}.$$

Problem : we do not observe the exact number of jumps $N_{t_{j+1}} - N_{t_j}$. Solution : approximation by an "observed" number of jumps D_j

$$\hat{\alpha}_{(d)}(x) = \frac{\sum_{j=0}^{m-1} \mathbf{D}_j k_{\mathcal{T}}(x, X_{t_j})}{\sum_{j=0}^{m-1} (t_{j+1} - t_j) k_{\mathcal{T}}(x, X_{t_j})}$$

(H4): $D_j = N_{t_{j+1}} - N_{t_j}$ when $N_{t_{j+1}} - N_{t_j} \leq 1$, but $D_j \leq N_{t_{j+1}} - N_{t_j}$ otherwise.

Let
$$\Delta_m = \max_{j=1...m} (t_{j+1} - t_j)$$
.

(H5) $\Delta_m/v_T^2(x) \rightarrow 0$

(H6) There exist $\ell_T(x) \ge 0$ and a > 0 such that for any s, t satisfying $|s - t| < \Delta_m$ and for all $y \in E$,

$$\mathbb{E}\left[\left|k_{T}(x, Y_{s}) - k_{T}(x, Y_{t})\right| \middle| Y_{0} = y\right] \leq \ell_{T}(x)|s - t|^{a}$$

with $\lim_{T\to\infty} \frac{\Delta_m^a \ell_T(x)}{v_T^2(x)} \to 0.$

Theorem

Assume (H1)-(H6), then

$$\hat{\alpha}_{(d)}(x) - \alpha(x) = O_{\rho}\left(\frac{1}{Tv_{T}(x)} + w_{T}^{2}(x) + \frac{\Delta_{m}}{v_{T}^{2}(x)} + \frac{\Delta_{m}^{*}\ell_{T}(x)}{v_{T}^{2}(x)}\right) \stackrel{\mathbb{P}}{\longrightarrow} 0.$$

(H5)-(H6) demand that the inter-jump diffusion $(Y_t)_{t\geq 0}$ be regular enough and $\Delta_m \to 0$ fast enough.

In the discrete case, we start from the same simulation as in the continuous case but we only observe the process at m instants t_0, \ldots, t_m equally spaced between $t_0 = 0$ and $t_m = T$.



Simulation in the discrete-time case

m = 5000 observations, i.e. 0.3 jumps in average between two observations.



m = 1000 observations, i.e. 1.5 jumps in average between two observations.



m = 100 observations, i.e. 15 jumps in average between two observations.



m = 30 observations, i.e. 51 jumps in average between two observations.

