Kalikow decomposition for counting processes with stochastic intensity

Tien Cuong PHI
supervised by: Patricia Reynaud-Bouret, Eva Löcherbach

Université Côte d'Azur

June 2021

Introduction

- We are interested in simulation of the point process has long mermory and in stationary regime.
- We present a new Kalikow decomposition to make Perfect simulation: an algorithm that allows to simulate in finite time what happens on a node in the potentially infinite network in stationary regime.
- New Kalikow decomposition $\rightarrow$ modified Perfect algorithm

Definitions and Notations

A counting process $Z^i, i \in \mathbf{I}$, can be described by its sequence of jump times in $\mathbb{R}$, $(T_n^i)_{n\in\mathbb{Z}}$ [Bre81].

Canonical path space of a counting process

$$\mathcal{X}_\infty = \{(\{t_n^i\}_{n\in\mathbb{Z}})_{i\in\mathbf{I}} \quad \text{such that} \, \forall n \quad t_n^i < t_{n+1}^i, \, t_0^i \leq 0 < t_1^i\}.$$

where $\{t_n^i\}_{n\in\mathbb{Z}}$ denotes a possible realization of $(T_n^i)_{n\in\mathbb{Z}}$.

- Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*, Springer-Verlag.

Definitions and Notations

Denote $\mathcal{X}_t$ the canonical path space of $Z = (Z_i)_i$ before time t:

$$\mathcal{X}_t = \mathcal{X} \cap (-\infty, t)^{\mathbf{I}}.$$

A past configuration x_t is an element of $\mathcal{X}_t$ which is a realization of arrival times of Z before t.

In addition, the evolution of the point process Z^i with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}}$ is fully characterized by its stochastic intensity which depends on the past configuration, see Proposition 7.2.IV [Daley, Vere-Jones]

Stochastics intensity

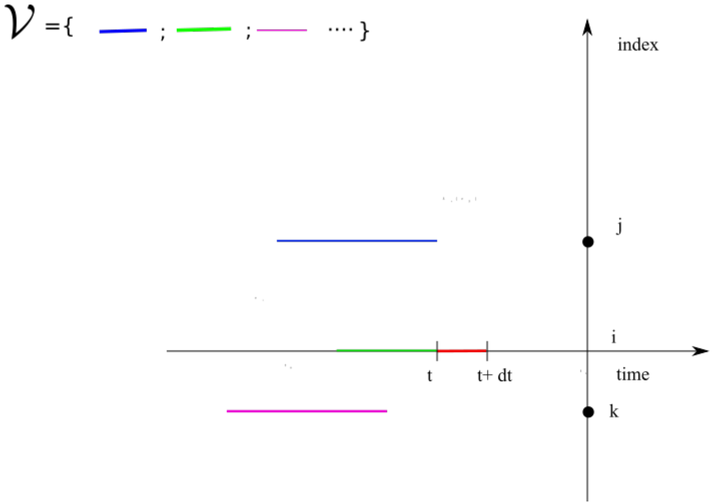
$$\mathbb{P}\left(Z^i \text{ has jump in } [t, t+dt) \mid \text{past before time } t = x_t\right) = \phi_{i,t}(x_t)dt.$$

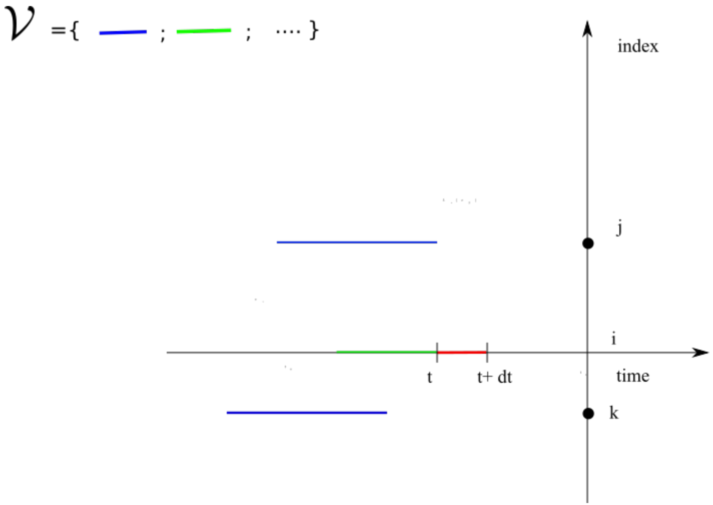
- DALEY, D.J, VERE-JONES, D. (2003). *An introduction to the theory of Point Processes*, 2nd edn. Volume **1**, Springer-Verlag, New York.

ATTENTION

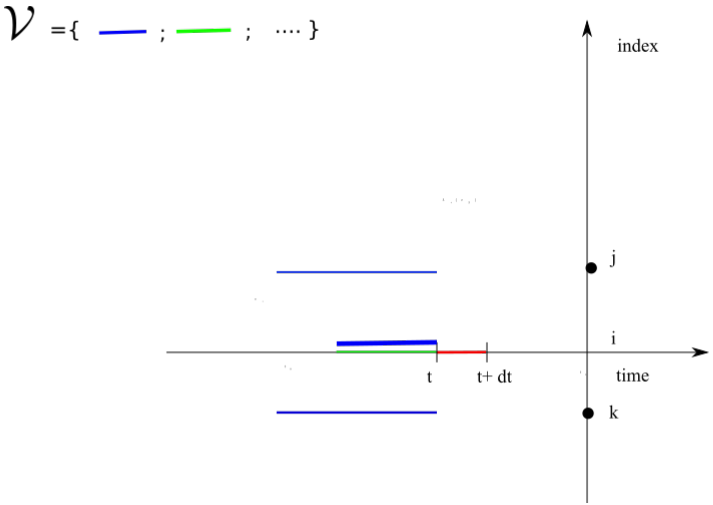
It is however the stochastic intensity is complex in most of the time. In practice, computing the stochastic intensity is highly computational task.

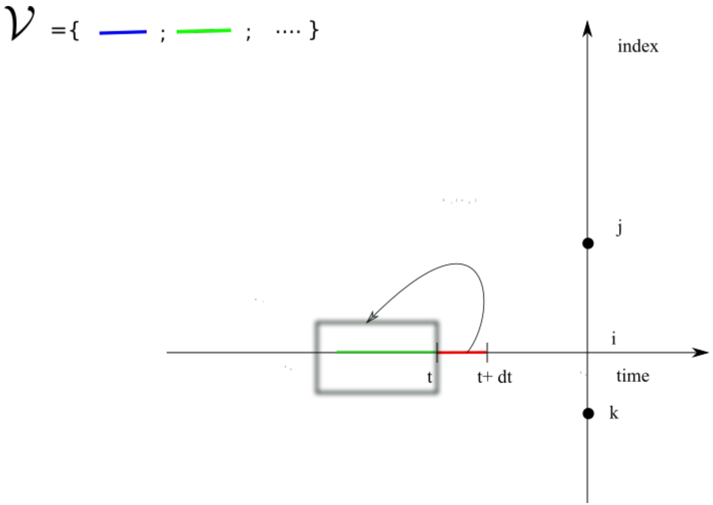
Random neighborhood



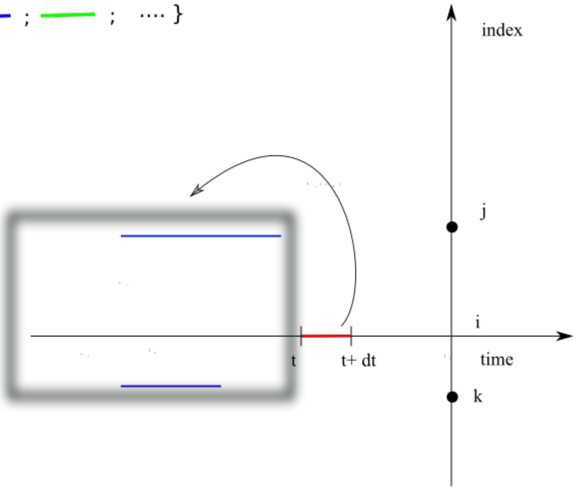


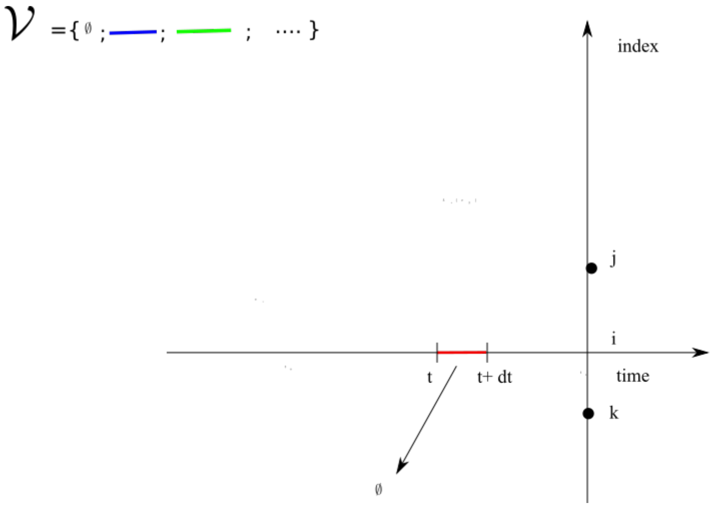
$\mathcal{V}$ ={ ; ; .... }
index
j
i
t
t+ dt
time
k


$\mathcal{V}$ ={ ; ; .... }
index
j
i
time
t
t+ dt
k


$\mathcal{V}$ ={ ; ; .... }
index
j
i
time
k
t
t+ dt

$\mathcal{V} = \{$; ; .... $\}$





$\mathcal{V} = \{ \emptyset ; \quad ; \quad ; \quad \cdots \}$
index
j
i
time
t
t+ dt
k
∅

Kalikow decomposition

For any $x_t \in \mathcal{X}_t \cap \mathcal{Y}_t$

$$\phi_{i,t}(x_t) = \lambda_{i,t}(\emptyset)\phi_{i,t}^{\emptyset} + \underbrace{\sum_{v_t \in \mathbf{V}_t, v_t \neq \emptyset}}_{\text{neighborhood family}} \lambda_{i,t}(v_t)\phi_{i,t}^{v_t}(x_t)$$

- Comets, F., Fernández, R., Ferrari, P. A. (2002). Processes with long memory: Regenerative construction and perfect simulation. *Ann. of Appl. Probab.*
- Ferrari, P. A., Maass, A., Martínez, S., Ney, P. (2000). Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory Dyn. Syst.*.

prob. to pick a neigh.
(sum up to 1)

$$\phi_{i,t}(x_t) = \lambda_{i,t}(\emptyset)\phi_{i,t}^{\emptyset} + \sum_{v_t \in \mathbf{V}_t, v_t \neq \emptyset} \lambda_{i,t}(v_t)\phi_{i,t}^{v_t}(x_t)$$

neighborhood family

prob. to pick a neigh.
(sum up to 1)

depends on
the points inside v_t

$$\phi_{i,t}(x_t) = \lambda_{i,t}(\emptyset)\phi_{i,t}^{\emptyset} + \sum_{v_t \in \mathbf{V}_t, v_t \neq \emptyset} \lambda_{i,t}(v_t)\phi_{i,t}^{v_t}(x_t)$$

neighborhood family

Method to write a Kalikow decomposition

Assumption

There exists a non-negative sequence of functions $(\Delta_k^i(x))_k$ which are cylindrical on v_k^i such that

$$\sum_{k=0}^{n} \Delta_k^i(x) \to \phi_i(x)$$

as $n \to \infty$, for every $x \in \mathcal{X} \cap \mathcal{Y}$.

The Kalikow decomposition of $\phi_i(x)$ with neighborhood family $\mathbf{V}$ and the subspace $\mathcal{Y}$:

$$\begin{cases} \lambda_i(\emptyset) = \eta_0^i \\ \phi_i^{\emptyset} = \dfrac{\Delta_0^i}{\eta_0^i} \\ \lambda_i(v_k^i) = \eta_k^i \\ \phi_i^{v_k^i}(x) = \dfrac{\Delta_k^i(x)}{\eta_k^i} \end{cases}$$

for any choice of non negative weights $(\eta_k^i)_k$ such that

$$\sum_{k \geq 0} \eta_k^i = 1.$$

Attention

To perform Perfect simulation, $(\eta_k^i)_k$ can not be chosen arbitrarily. $\rightarrow$ Stopping condition + Mean number of simulated points.

Define a neighborhood family + the sequence of cylindrical function $\Delta_k^i(x)$ + a good subspace $\mathcal{Y}$ + good weights (η_k^i).

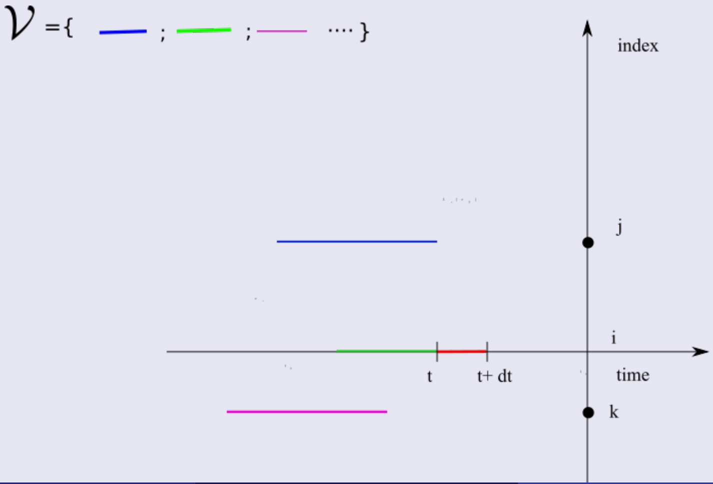
Example: Linear Hawkes processes

$$\phi_i(x) = \mu_i + \sum_{j \in \mathbf{I}} \int_{-\infty}^{0} h_{ji}(-s) dx_s^j = \mu_i + \sum_{j \in \mathbf{I}} \sum_{n \in \mathbb{N}^*} \int_{-n\epsilon}^{-n\epsilon+\epsilon} h_{ji}(-s) dx_s^j.$$

- HAWKES, A. G. (1971). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*.
- HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*.

we consider the neighborhood family $\mathbf{V}^{atom}$ as follows: for $l, n \geq 1$

$$v^i_{(l,n)} := \{j_l\} \times [-n\epsilon, -n\epsilon + \epsilon).$$



We define, for all $l, n \geq 1$:

$$\Delta^i_{(l,n)}(x) := \int_{-n\epsilon}^{-n\epsilon+\epsilon} h_{j_l i}(-s) dx^{j_l}_s \geq 0,$$

and $\Delta^i_{(0,0)} := \Delta^i_{(0,0)}(x) := \mu_i$.

Applying the Monotone Convergence theorem, we obtain

$$\Delta^i_{(0,0)} + \sum_{(l,n)=(1,1)}^{(L,N)} \Delta^i_{(l,n)}(x) \to \phi_i(x),$$

when $(L, N) \to (\infty, \infty)$.

V^{atom}

$$\begin{cases} \lambda_i(\emptyset) = \eta_{(0,0)} \\ \phi_i^{\emptyset} = \dfrac{\mu_i}{\eta_{(0,0)}} \\ \lambda_i\left(v^i_{(l,n)}\right) = \eta_{(l,n)} \\ \phi_i^{v^i_{(l,n)}}(x) = \dfrac{\int_{-n\epsilon}^{-n\epsilon+\epsilon} h_{j_l i}(-s)dx_s^{j_l}}{\eta_{(l,n)}}, \end{cases}$$

for any choice of non negative weights $\left(\eta^i_{(l,n)}\right)_{(l,n)\in\mathbb{N}^*\times\mathbb{N}^*}$ such that

$$\eta^i_{(0,0)} + \sum_{l,n\geq 1} \eta^i_{(l,n)} = 1.$$

Example: Age dependent Hawkes processes with hard refractory period

$$\phi_i(x) = \psi_i \left(\sum_{j \in \mathbf{I}} \int_{-\infty}^{0} h_{ji}(-s) dx_s^j \right) \mathbb{1}_{a_0^i(x) > \delta} \qquad (1)$$

With the convention that $a_t^i(x) = t - L_t^i(x)$, where L_t^i is the last jump before t of process Z^i.

By the definition, this implies that there does not exist any two jumps those distance is smaller than δ.

- RAAD, M.B, DITLEVSEN, S., LÖCHERBACH, E. , (2020). Stability and mean-field limits of age dependent Hawkes processes. *Ann. Inst. H. Poincar Probab. Statist.*.
- CHEVALLIER, J. (2017). Mean-field limit of generalized Hawkes processes. *Stoch. Proc. Appl.* .

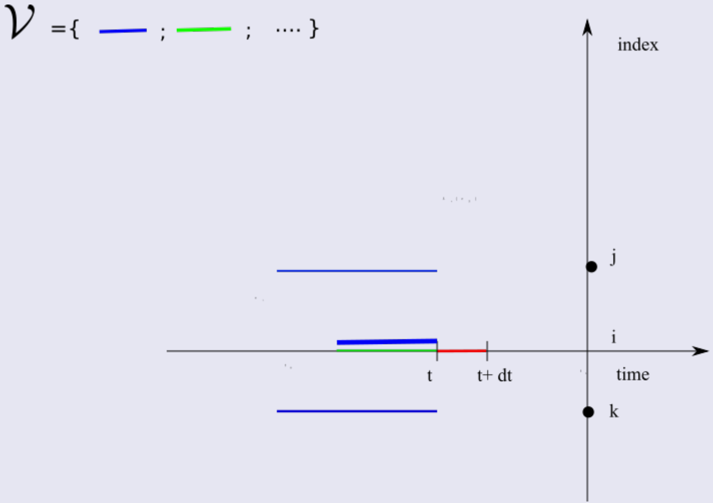
We build a neighborhood family $\mathbf{V}^{nested}$ by introducing a non decreasing sequence $(V_i(k))_{k\geq 0}$ of finite subsets of $\mathbf{I}$ such that

$$V_i(0) = \emptyset, V_i(1) = \{i\}, V_i(k-1) \subset V_i(k) \quad \text{and} \quad \cup_k V_i(k) = \mathcal{V}_{\cdot\to i} \cup \{i\}. \tag{2}$$

with

$$\mathcal{V}_{\cdot\to i} := \{j \in \mathbf{I} \quad \text{such that } i \text{ is locally dependent on } j\}$$

We set $v_k^i = V_i(k) \times [-k\delta, 0)$, then we obtain by construction an increasing, nested neighborhood sequence $(v_k^i)_{k\in\mathbb{N}}$.



$$\Delta_k^i(x) = \psi_i \left(\sum_{j \in V_i(k)} \int_{-k\delta}^{0} h_{ji}(-s) dx_s^j \right) \mathbb{1}_{a_0^i(x) > \delta} - \psi_i \left(\sum_{j \in V_i(k-1)} \int_{-k\delta+\delta}^{0} h_{ji}(-s) dx_s^j \right) \mathbb{1}_{a_0^i(x) > \delta} \quad (3)$$

$\mathbf{V}^{nested}$ and $\mathcal{Y} = \mathcal{X}^{>\delta}$

$$\begin{cases} \lambda_i(v_k^i) = \eta_k^i \\ \phi_i^{v_k^i}(x) = \dfrac{\Delta_k^i(x)}{\eta_k^i} \end{cases}$$

Remark

Consider $D_i^k(x) = \{z \in \mathcal{X}^{>\delta} : z \overset{v_k^i}{=} x\}$.

$$\Delta_k^i(x) = \inf_{z \in D_i^k(x)} \psi_i \left(\sum_{j \in \mathbf{I}} \int_{-\infty}^{0} h_{ji}(-s) dz_s^j \right) \mathbb{1}_{a_0^i(x) > \delta} - \inf_{z \in D_i^{k-1}(x)} \psi_i \left(\sum_{j \in \mathbf{I}} \int_{-\infty}^{0} h_{ji}(-s) dz_s^j \right) \mathbb{1}_{a_0^i(x) > \delta}.$$

The above prescription corresponds to the classical method of obtaining a Kalikow decomposition in discrete time.

- Hodara, P. and Löcherbach, E., (2017). Hawkes Processes with variable length memory and an infinite number of components, *Adv. Appl. Prob.*.
- Galves, A., Löcherbach, E. (2013). Infinite systems of interacting chains with memory of variable length–a stochastic model for biological neural nets. *Journal of Statistical Physics.*

Other method for nonlinear Hawkes process

Consider ψ_i is an analytical function.

$$\phi_i(x) = \psi_i\left(\sum_{j\in \mathbf{I}}\sum_{n\in\mathbb{N}}\int_{-(n+1)\epsilon}^{-n\epsilon} h_{ji}(-s)dx_s^j\right)$$
$$= \psi_i\left(\sum_{j\in \mathbf{I}}\sum_{n\in\mathbb{N}} a_{jn}(x)\right) = \psi_i\left(\sum_{\alpha} a_{\alpha}(x)\right).$$

where

$$a_\alpha(x) := \int_{-(n+1)\epsilon}^{-n\epsilon} h_{ji}(-s)dx_s^j.$$

In this section, to develop our series using Taylor expansion, we choose

$$\mathcal{Y}^K = \left\{ x \,|\, \sup_i \left(\sum_j \int_{-\infty}^{0} h_{ji}(-s) dx_s^j \right) < K \right\}.$$

We construct the neighborhood family $\mathbf{V}^{Taylor}$ by defining for $\alpha_{1:k} = (\alpha_1, \ldots, \alpha_k)$,

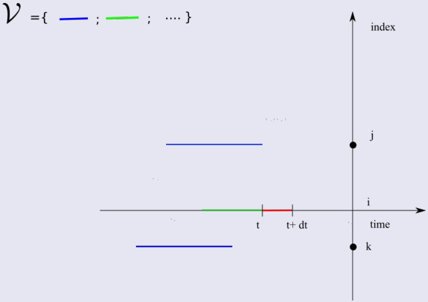
$$v^i_{\alpha_{1:k}} = v^i_{(\alpha_1,\ldots,\alpha_k)} := \cup_{l=1}^k w_{\alpha_l}, \tag{4}$$

where $w_{\alpha_l} := \{j\} \times [-(n+1)\epsilon, n\epsilon)$ if $\alpha_l = (j, n)$, and $v^i_0 := \emptyset$.

We construct the neighborhood family $\mathbf{V}^{Taylor}$ by defining for $\alpha_{1:k} = (\alpha_1, \ldots, \alpha_k)$,

$$v^i_{\alpha_{1:k}} = v^i_{(\alpha_1,\ldots,\alpha_k)} := \cup_{l=1}^k w_{\alpha_l}, \tag{5}$$

where $w_{\alpha_l} := \{j\} \times [-(n+1)\epsilon, n\epsilon)$ if $\alpha_l = (j, n)$, and $v^i_0 := \emptyset$.



Proposition

$\mathbf{V}^{Taylor}$ *and* $\mathcal{Y}^K$

$$\begin{cases} \lambda_i(\emptyset) & = 1 - \sum_{k\geq 1} \sum_{\alpha_1,\dots,\alpha_k} \eta^i_{\alpha_1} \dots \eta^i_{\alpha_k} \\ \phi_i^{\emptyset} & = \dfrac{\psi_i(0)}{\lambda_i(\emptyset)} \\ \lambda_i(v^i_{\alpha_{1:k}}) & = \eta^i_{\alpha_1} \dots \eta^i_{\alpha_k} \\ \phi_i^{v^i_{\alpha_{1:k}}}(x) & = \dfrac{\psi_i^{(k)}(0)}{k!} \dfrac{a_{\alpha_1}(x)}{\eta^i_{\alpha_1}} \dots \dfrac{a_{\alpha_k}(x)}{\eta^i_{\alpha_k}}, \end{cases}$$

with

$$a_\alpha(x) := \int_{-(n+1)\epsilon}^{-n\epsilon} h_{ji}(-s) dx^j_s.$$

Example: Exponential Hawkes processes

Consider $\psi_i() = exp(.)$
$\mathbf{V}^{Taylor}$ and $\mathcal{Y}^{\infty}$

$$\begin{cases} \lambda_i(\emptyset) & = 1 - \sum_{k \geq 1} \sum_{\alpha_1, \ldots, \alpha_k} \eta^i_{\alpha_1} \ldots \eta^i_{\alpha_k} \\ \phi_i(\emptyset) & = \dfrac{1}{\lambda_i(\emptyset)} \\ \lambda_i(v^i_{\alpha_{1:k}}) & = \eta^i_{\alpha_1} \ldots \eta^i_{\alpha_k} \\ \phi_i^{v^i_{\alpha_{1:k}}}(x) & = \dfrac{1}{k!} \times \dfrac{a_{\alpha_1}(x)}{\eta^i_{\alpha_1}} \ldots \dfrac{a_{\alpha_k}(x)}{\eta^i_{\alpha_k}}. \end{cases}$$

Perfect simulation

For this algorithm to work, we introduce a subspace $\mathcal{Y}$ as follows:

$$\mathcal{Y} = \{x \in \mathcal{X} \quad \text{such that} \quad \forall v, i : \quad \phi_i^v(x) \leq \Gamma_i\},$$

where Γ_i is a positive constant. We assume that the process generated by the algorithm stays in $\mathcal{Y}$ almost surely.

For example, in the case of age dependent Hawkes processes, we can show that $\Delta_k^i(x)$ are bounded. Thus, with an adequate choice of weights $(\lambda_i(.))_{i\in\mathbf{I}}$ and $(\Gamma_i)_{i\in\mathbf{I}}$, the algorithm remains in $\mathcal{Y}$ almost surely.

$$\phi_i(x) = \lambda_i(\emptyset)\phi_i^{\emptyset} + \sum_{\emptyset \neq v \in \mathbf{V}^i} \lambda_i(v)\phi_i^v(x) \leq \Gamma_i,$$

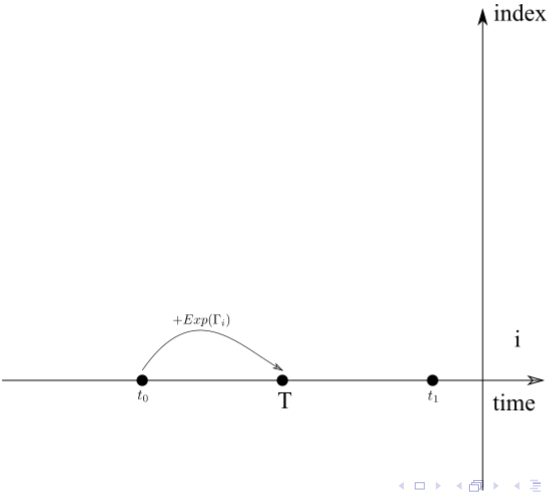
which means that the intensity is bounded.
Furthermore, for any t and $x_t \in \mathcal{Y}_t$,

$$\phi_{i,t}(x_t) = \phi_i(x_t^{\leftarrow t}) \leq \Gamma_i.$$

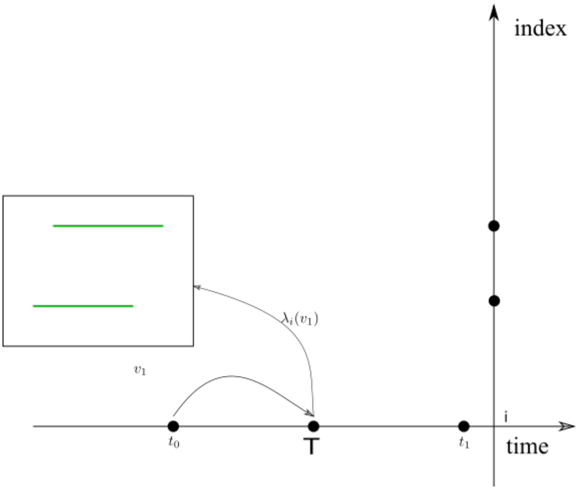
Kalikow theoretical algorithm

▷ To simulate the behavior of index i in $[t_0, t_1]$.

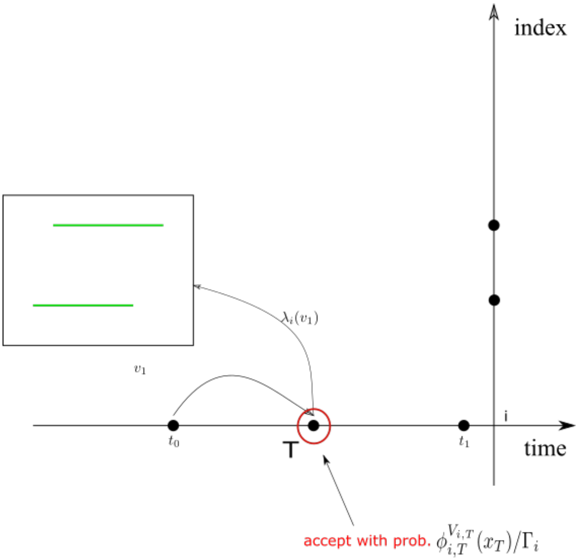
Step 1:



Step 2: pick V_T a random neighborhood according to $\lambda_i(.)$ in **Kalikow decomposition** shifted at time T.



Step 3:


index
$\lambda_i(v_1)$
v_1
i
t_0
T
t_1
time
accept with prob. $\phi_{i,T}^{V_{i,T}}(x_T)/\Gamma_i$

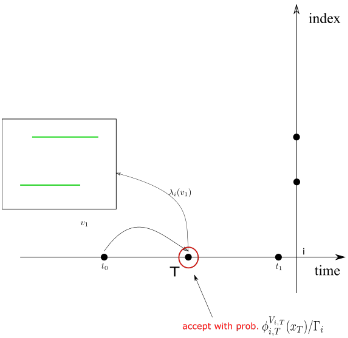
Do we construct the right intensity?

Proposition(P., Muzy, Reynaud - Bouret)

If we suppose that the obtainned process stays in $(\mathcal{Y}_t)_{t\in\mathbb{R}}$ almost surely then it admits $\phi_{i,t}(x_t)$ as predictable intensity.

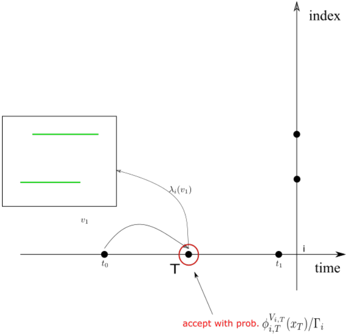
Mathematical version can be found in our paper.

- Phi, T.C, Muzy, A., Reynaud-Bouret, P. (2020). Event-scheduling algorithms for simulating potentially infinite neuronal networks, *SN Comput. Sci.*.
- Phi, T.C (2021+). Kalikow decomposition for counting processes with stochastic intensity, *submitted*.



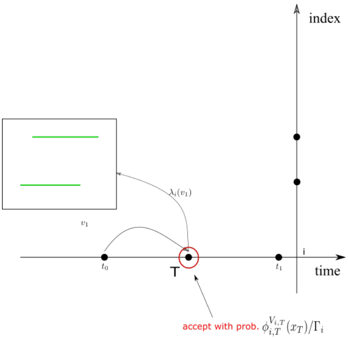
Attention

Infeasible in practice:



Attention

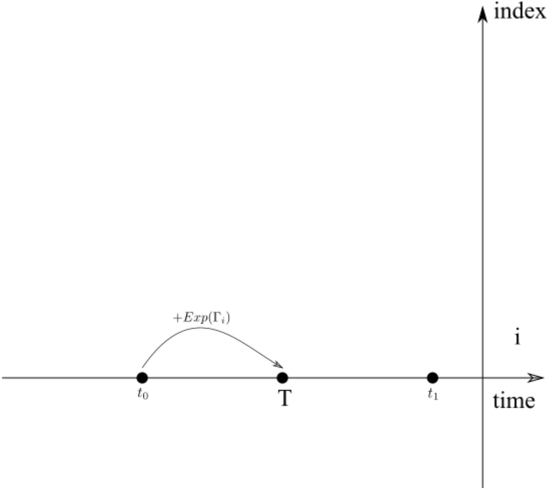
Infeasible in practice: since $\phi_{i,T}^{V_{i,T}}(x_T)$ depends on the points in $V_{i,T}$, that are not known at this stage.

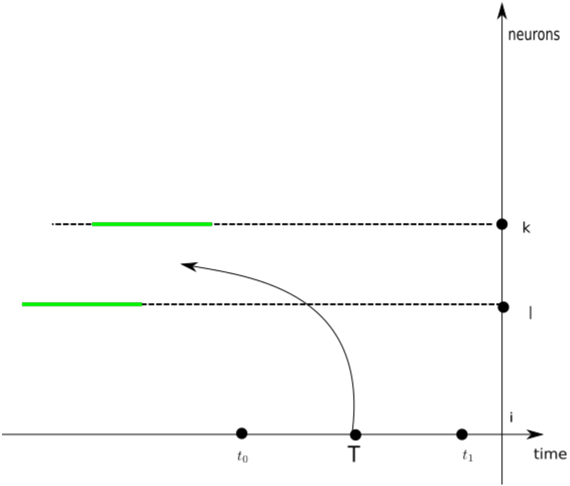


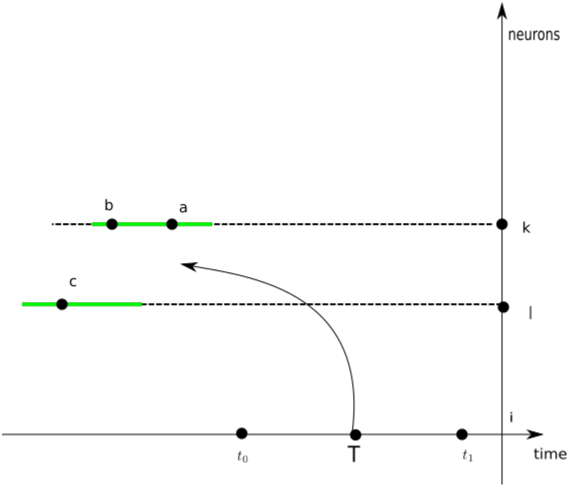
Attention

Infeasible in practice: since $\phi_{i,T}^{V_{i,T}}(x_T)$ depends on the points in $V_{i,T}$, that are not known at this stage.
$\rightarrow$ go backward in time before moving forward.

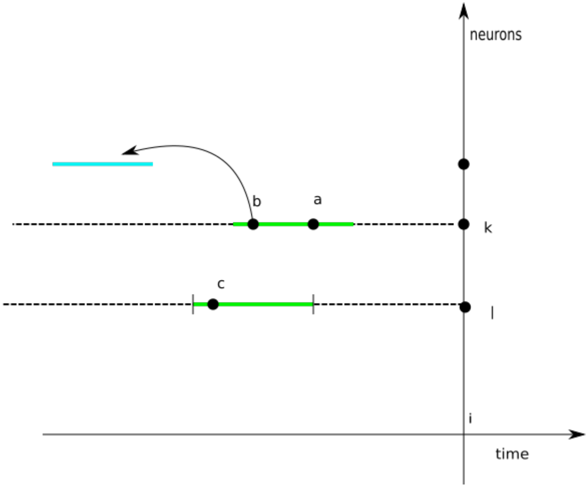
Backward steps

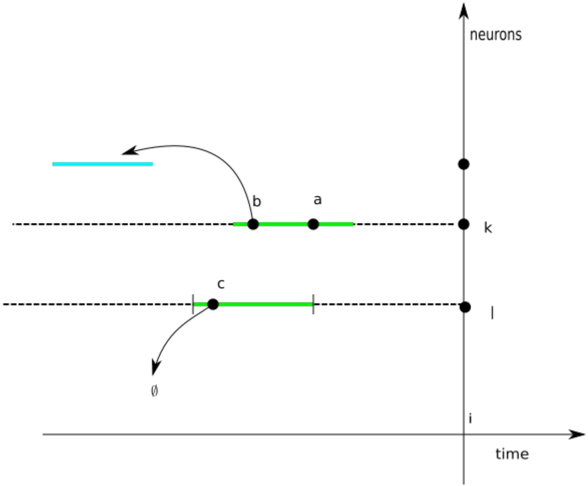


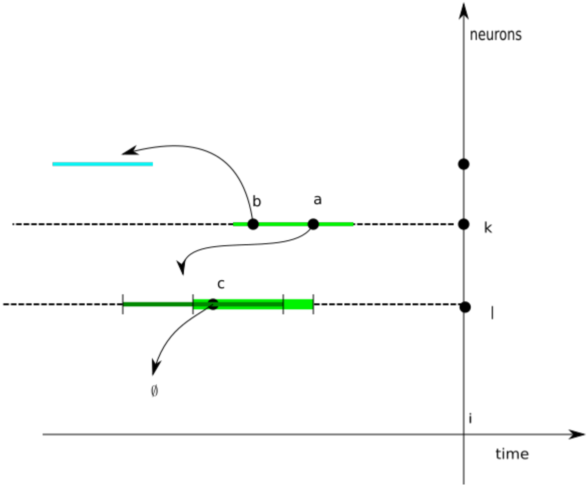

neurons
k
l
i
t_0
T
t_1
time

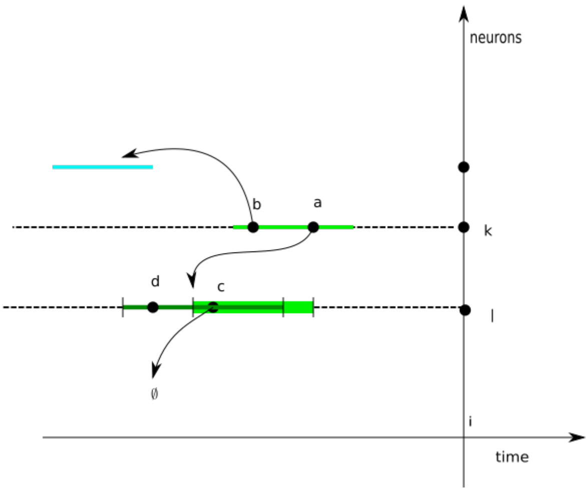


neurons
b
a
k
c
l
i
t_0
T
t_1
time

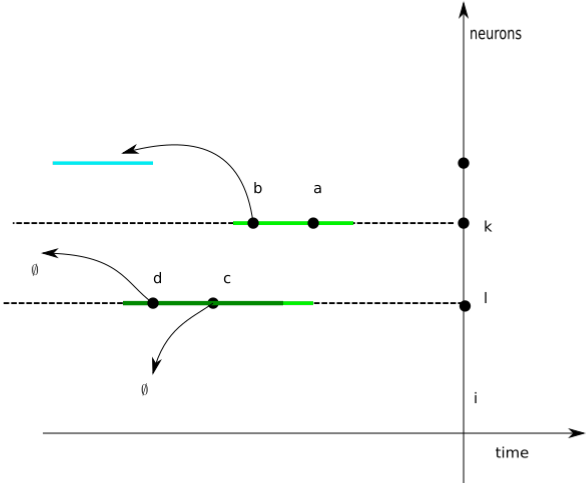

neurons
b
a
k
c
l
i
time


neurons
b
a
k
c
l
∅
i
time


neurons
b
a
k
c
l
∅
i
time

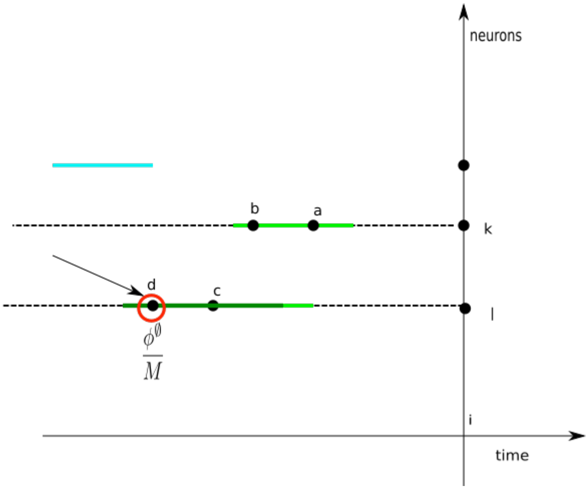

neurons
b
a
k
d
c
l
∅
i
time

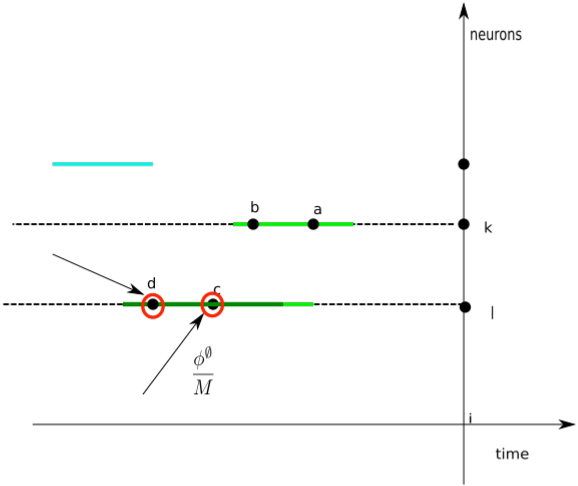
s



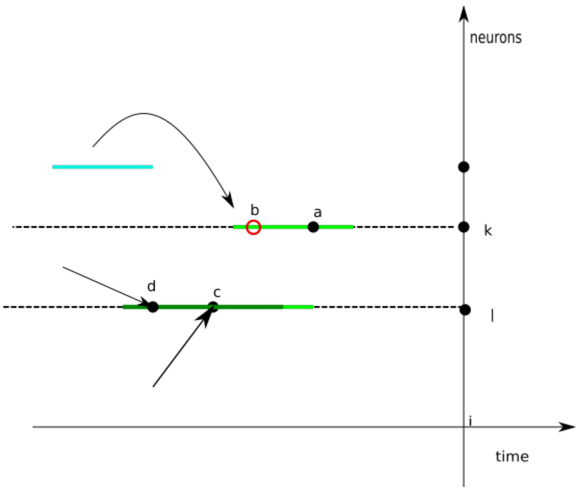
Note that $\lambda_i(\emptyset)$ does not appear in the above Kalikow decomposition. Amazingly, this does not cause any problems for the Perfect Simulation algorithm. This is one of the main differences with [P.Hodara, E. Locherbach].

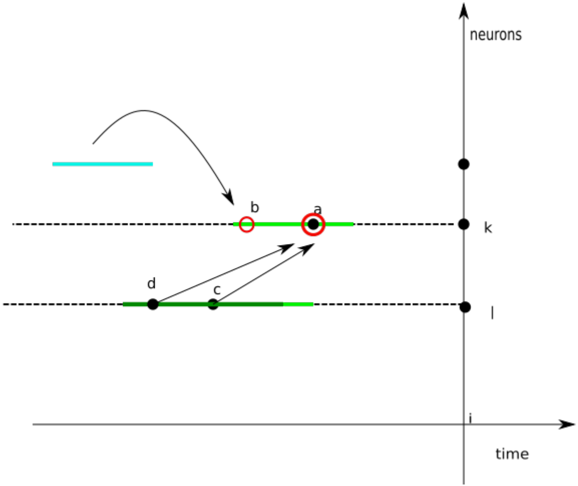
Forward steps

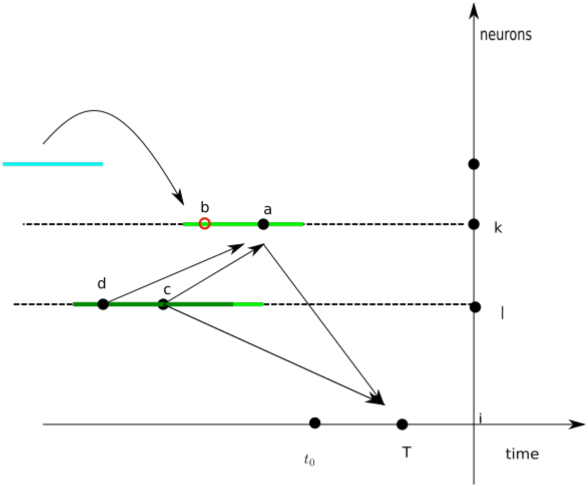




neurons
b
a
k
d
c
l
$\frac{\phi^{\emptyset}}{M}$
i
time


neurons
b
a
k
d
c
l
i
time


neurons
b
a
k
d
c
l
i
time



neurons
b
a
k
d
c
l
i
t_0
T
time

Why does the Backward steps end?

Proposition(P., Muzy, Reynaud - Bouret)

If

$$\sup_{i \in \mathbf{I}} \sum_{k \geq 1} P(v_k^i) \lambda_i(v_k^i) < 1 \tag{6}$$

then the backward part of Backward Forward Algorithm ends almost surely in finite time.

Remark

$\sum_{k \geq 1} P(v_k^i) \lambda_i(v_k^i)$ is the mean number of children issued from one point of type i.

Example: Age dependent Hawkes processes with hard refractory period

We will consider that $\mathbf{I} = \mathbb{Z}$, and for all i, we set

1. $h_{ji}(t) = \beta_{ji} \exp(-\alpha t)$ where β_{ji}, α are positive constants for all j, i. In addition, we take $\alpha = \dfrac{1}{\delta}$ and $\beta_{ji} = \dfrac{1}{2|j-i|^{\gamma}}$ for $j \neq i$ with a positive number γ and $\beta_{ii} = 1$.
2. $V_i(0) = \emptyset, V_i(1) = \{i\}, \ldots, V_i(k) = \{i-k+1, \ldots, i, i+1, \ldots, i+k-1\}$ $\forall k \geq 2$.
3. $\eta_k^i = \eta_k = c_\eta \dfrac{1}{k^p}, \forall k \geq 1$, where p is a positive constant and c_η is a normalization constant.

We are looking for values of p such that:

(a) $(\eta_k)_k$ defines a probability,

(b) the algorithm stays in $\mathcal{Y}$,

(c) the Kalikow decomposition exists ,

(d) the branching process goes extinct in finite time almost surely.

$$(a) : \sum_k \eta_k = 1 \Rightarrow p > 1$$

$$(b) : \Gamma < \infty \Rightarrow p < \gamma$$

$$(c) : \sum \int h_{ji} < \infty \Rightarrow \gamma > 1$$

$$(d) : \zeta < 1 \Rightarrow p - 2 > 1$$

Finally, we conclude that $3 < p < \gamma$.


- PHI, T.C, MUZY, A., REYNAUD-BOURET, P. (2020). Event-scheduling algorithms for simulating potentially infinite neuronal networks, *SN Comput. Sci.*.
- PHI, T.C (2021+). Kalikow decomposition for counting processes with stochastic intensity, *submitted*.
- HANSEN, N. R., REYNAUD-BOURET, P., RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli*.
- HODARA, P. AND LÖCHERBACH, E., (2017). Hawkes Processes with variable length memory and an infinite number of components, *Adv. Appl. Prob.*.
- GALVES, A., LÖCHERBACH, E. (2013). Infinite systems of interacting chains with memory of variable length–a stochastic model for biological neural nets. *Journal of Statistical Physics*.
- RAAD, M.B, DITLEVSEN, S., LÖCHERBACH, E. , (2020). Stability and mean-field limits of age dependent Hawkes processes.*Ann. Inst. H. Poincar*