On oscillating systems of interacting Hawkes processes

Susanne Ditlevsen Eva Löcherbach

Bielefeld, November 2015

Outline

We will consider large systems of randomly interacting point processes presenting intrinsic oscillations.

1. Introduction of the model : Multi class systems of interacting nonlinear Hawkes processes : several populations of particles (individuals, neurons...) which interact. Within each population, all particles behave in the same way.
2. Propagation of chaos and associated CLT.
3. Erlang kernels allow to develop the memory. Associated PDMP's.
4. Study of the oscillatory behavior of the limit system.
5. And of the finite size system.

Outline

We will consider large systems of randomly interacting point processes presenting intrinsic oscillations.

1. Introduction of the model : Multi class systems of interacting nonlinear Hawkes processes : several populations of particles (individuals, neurons...) which interact. Within each population, all particles behave in the same way.
2. Propagation of chaos and associated CLT.
3. Erlang kernels allow to develop the memory. Associated PDMP's.
4. Study of the oscillatory behavior of the limit system.
5. And of the finite size system.

The **second part** is deeply based on results of *Delattre, Fournier and Hoffmann (2015)* on high dimensional Hawkes processes (in the one-class frame).

Hawkes processes

- Point process model : for each individuum, we model the random times of appearance of an event we are interested in (spikes for neurons, transaction events in economic models, etc)
- Counting process associated to particle i :

$$Z_i(t) = \text{ number of events occuring for } i \text{ during } [0, t]$$

Hawkes processes

- Point process model : for each individuum, we model the random times of appearance of an event we are interested in (spikes for neurons, transaction events in economic models, etc)
- Counting process associated to particle i :

$$Z_i(t) = \text{ number of events occuring for } i \text{ during } [0, t]$$

with intensity process $\lambda_i(t)$ defined by

$$P(Z_i \text{ has a jump during }]t, t+dt] | \mathcal{F}_t) = \lambda_i(t) dt.$$

- $\lambda_i(t)$ is a **stochastic process, depending on the whole history before time t.**

Hawkes intensity

- Hawkes (1971), Hawkes and Oakes (1974), Brémaud and Massoulié (1996) : each past event **triggers** future events : self-exciting processes (or better : self influencing)
- Intensity of $Z_i(t)$ of form

$$\lambda_i(t) = f \left(\int_{]0,t]} h(t-s) dZ_i(s) \right).$$

↑ rate fct ↑ loss fct ↑ past event

– **rate function** $f : \mathbb{R} \to \mathbb{R}_+$ Lipschitz, increasing.
– **loss term** $h(t-s)$ describes how an event **lying back** $t-s$ **time units in the past** influences the present time t.
– if h is not of compact support, then : truly infinite memory process.

• Questions like : Existence and stability, longtime behavior etc have been answered in the literature (Brémaud and Massoulié 1996)

• We are interested here in a **large system of interacting Hawkes processes**, describing each one individual (neuron, particle, ...).

• This system is made of **several populations** $k = 1, \ldots, n$.

• Each population k consists of N_k **particles** described by their counting processes

$$Z_{k,i}(t), 1 \leq i \leq N_k.$$

• Within a population, all particles behave in the same way. This is a mean-field assumption.

• Intensity of $i-$th particle belonging to population k :

$$\lambda_{k,i}(t) = f_k \left(\sum_{l=1}^{n} \frac{1}{N_l} \sum_{1 \le j \le N_l} \int_{]0,t[} h_{kl}(t-s) dZ_{l,j}(s) \right).$$

• f_k = jump rate function of population k; *supposed to be Lipschitz continuous.*

• h_{kl} measures the influence of a typical particle of population l on a typical particle of population k; *supposed to be in* $L^2_{loc}(\mathbb{R}_+, \mathbb{R})$.

• Intensity of $i-$th particle belonging to population k :

$$\lambda_{k,i}(t) = f_k \left(\sum_{l=1}^{n} \frac{1}{N_l} \sum_{1 \le j \le N_l} \int_{]0,t[} h_{kl}(t-s) dZ_{l,j}(s) \right).$$

• f_k = jump rate function of population k; *supposed to be Lipschitz continuous.*

• h_{kl} measures the influence of a typical particle of population l on a typical particle of population k; *supposed to be in* $L^2_{loc}(\mathbb{R}_+, \mathbb{R})$.

• We are in a **mean field frame :** population l influences population k only through its **empirical measure.**

Mean field limit

- What happens in the large system size limit ?

- I.e. $N = N_1 + \ldots + N_n$ total number of particles $\to \infty$ such that for each population

$$\lim_{N \to \infty} \frac{N_k}{N} > 0.$$

- Remember the intensity

$$\lambda_{k,i}(t) = f_k \left(\sum_{l=1}^{n} \int_{]0,t[} h_{kl}(t-s) \left[\frac{1}{N_l} \sum_{1 \le j \le N_l} dZ_{l,j}(s) \right] \right)$$

$$\uparrow \ \text{LLN} \to d\mathbb{E}(\bar{Z}_l(s)),$$

where $\bar{Z}_l$ is the counting process of a typical particle belonging to population l in the $N \to \infty-$limit.

Limit system

• Limit system : family of counting processes $\bar Z_k(t), 1 \le k \le n$ (one for each population), solution of an inhomogeneous equation

$$\bar Z_k(t) = \int_0^t \int_{\mathbb{R}_+} 1_{\{z \le f_k(\sum_{l=1}^n \int_0^s h_{kl}(s-u) d\mathbb{E}(\bar Z_l(u))\}} N^k(ds, dz),$$

N^k i.i.d. PRM on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $dsdz$.

Limit system

- Limit system : family of counting processes $\bar{Z}_k(t), 1 \leq k \leq n$ (one for each population), solution of an inhomogeneous equation

$$\bar{Z}_k(t) = \int_0^t \int_{\mathbb{R}_+} 1_{\{z \leq f_k(\sum_{l=1}^n \int_0^s h_{kl}(s-u) d\mathbb{E}(\bar{Z}_l(u)))\}} N^k(ds, dz),$$

N^k i.i.d. PRM on $\mathbb{R}_+ \times \mathbb{R}_+$ with intensity $dsdz$.

- Taking expectations yields : $m_t^k = \mathbb{E}(\bar{Z}_k(t)), 1 \leq k \leq n$, solves

$$m_t^k = \int_0^t f_k \left(\sum_{l=1}^n \int_0^s h_{kl}(s-u) dm_u^l \right) ds, 1 \leq k \leq n.$$

Convergence to limit system

• Existence of a pathwise unique solution of the limit system standard ; follows ideas of Delattre, Fournier and Hoffmann (2015) in the one-population case.

• Convergence of the finite size system (of the collection of empirical measures of each population) to the limit as well : We take empirical measures within each population and obtain

Theorem (Propagation of chaos)

$$\Big(\frac{1}{N_1}\sum_{1\le i\le N_1}\delta_{(Z^N_{1,i}(t))_{t\ge 0}}, \ldots, \frac{1}{N_n}\sum_{1\le i\le N_n}\delta_{(Z^N_{n,i}(t))_{t\ge 0}}\Big) \to \mathcal{L}((\bar Z_1(t), \ldots, \bar Z_n(t))_{t\ge 0})$$

in probability, as $N \to \infty$.

• Multi-population frame : reminiscent of Graham (2008), see also Graham and Robert (2009), who has invented the notion of "multi-chaoticity".

• Note that in the limit the different populations are independent. Interactions of classes do only survive in law.

Associated CLT

What is the speed of convergence in the above limit theorem ?

Theorem

Under suitable assumptions : For any fixed $\ell_1 \leq N_1, \ldots, \ell_n \leq N_n$,

$$\left(\left(\frac{Z_{1,i}(t) - m_t^1}{\sqrt{m_t^1}} \right)_{1 \leq i \leq \ell_1}, \ldots, \left(\frac{Z_{n,i}(t) - m_t^n}{\sqrt{m_t^n}} \right)_{1 \leq i \leq \ell_n} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I_{\ell_1 + \ldots + \ell_n})$$

as $N, t \to \infty$, *where we recall that*

$$m_t^i = \mathbb{E}(\bar{Z}^i(t)) = \text{ mean number of events in population } i.$$

Have to impose conditions on the way $N, t \to \infty$: depends on spectral properties of offspring matrix.

Remark

1) Result similar to the one obtained by Delattre, Fournier and Hoffmann (2015), but extension to the non-linear case (the rate functions f_k are not supposed to be linear) : we have to use old results on **matrix renewal equations** *obtained by Crump (1970) and Athreya and Murthy (1976).*

Remark

1) Result similar to the one obtained by Delattre, Fournier and Hoffmann (2015), but extension to the non-linear case (the rate functions f_k are not supposed to be linear) : we have to use old results on **matrix renewal equations** *obtained by Crump (1970) and Athreya and Murthy (1976).*
2) Rate of convergence given by $\sqrt{m_t^k}, 1 \le k \le n$.
3) Main difficulty : We do not dispose of equivalents of m_t^k as $t \to \infty$.

Remark

1) Result similar to the one obtained by Delattre, Fournier and Hoffmann (2015), but extension to the non-linear case (the rate functions f_k are not supposed to be linear) : we have to use old results on **matrix renewal equations** *obtained by Crump (1970) and Athreya and Murthy (1976).*

2) Rate of convergence given by $\sqrt{m_t^k}, 1 \leq k \leq n$.

3) Main difficulty : We do not dispose of equivalents of m_t^k as $t \to \infty$.

4) Result only holds assuming that m_t^k is at least of linear growth, within all populations. (In other words, within each population, there is always some minimal strictly positive spiking intensity - we will come back to this point later).

Remark

1) Main assumption is on the spectral properties of the "upper" offspring matrix Λ *given by*

$$\Lambda_{ij} = L \int_0^\infty |h_{ij}|(t) dt, 1 \leq i, j \leq n.$$

Here, L *is the Lipschitz constant of the rate functions* $f_1, \ldots, f_n$.
2) In the subcritical case, nothing has to be imposed on the way $N, t \to \infty$. *Main ingredient of the proof in this case is*

$$\mathbb{E}(|Z_{k,i}(t) - \bar{Z}_k(t)|) \leq CtN^{-1/2}.$$

Remark

1) Main assumption is on the spectral properties of the "upper" offspring matrix Λ *given by*

$$\Lambda_{ij} = L \int_0^\infty |h_{ij}|(t)dt, 1 \leq i, j \leq n.$$

Here, L *is the Lipschitz constant of the rate functions* $f_1, \ldots, f_n$.
2) In the subcritical case, nothing has to be imposed on the way $N, t \to \infty$. *Main ingredient of the proof in this case is*

$$\mathbb{E}(|Z_{k,i}(t) - \bar{Z}_k(t)|) \leq CtN^{-1/2}.$$

3) Supercritical case more difficult, in this case

$$\mathbb{E}(|Z_{k,i}(t) - \bar{Z}_k(t)|) \leq Ce^{\alpha t} N^{-1/2},$$

and we have to suppose that $t, N \to \infty$ *in such a way that* $e^{\alpha t} N^{-1/2} \to 0$.

Developing the memory

• Hawkes processes are truly infinite memory processes - the intensity depends on the whole history.

• Sometimes possible to "develop" the memory : Suppose $n = 1$ (only one population) with intensity

$$\lambda(t) = f\left(\int_0^t h(t-s)d\bar{Z}_N(s)\right), \quad \bar{Z}_N(s) = \frac{1}{N}\sum_{i=1}^N Z_i(s).$$

Erlang kernel :

$$h(t) = c\frac{t^m}{m!}e^{-\nu_t}, \nu > 0, m \in \mathbb{N}_0, c \in \mathbb{R}.$$

• The delay of influence of the past is distributed. It takes its maximum at about m/ν time units back in the past.
• The higher the order of the delay m, the more the delay is concentrated around its mean value $(m+1)/\nu$.

Developping the memory - continued

- Suppose e.g. $h(t) = cte^{-\nu t}$ (short memory of length $m = 1$)
-
$$h'(t) = -\nu h(t) + ce^{-\nu t} =: -\nu h(t) + h_1(t),$$

and

$$h_1'(t) = -\nu h_1(t) : \text{ system closed !}$$

Developping the memory - continued

- Suppose e.g. $h(t) = cte^{-\nu t}$ (short memory of length $m = 1$)
-
$$h'(t) = -\nu h(t) + ce^{-\nu t} =: -\nu h(t) + h_1(t),$$

and

$$h_1'(t) = -\nu h_1(t) : \text{ system closed !}$$

- In terms of the intensity process : $\lambda(t) = f(X(t))$ where

$$X(t) = \int_0^t h(t-s) d\bar{Z}_N(s), \quad Y(t) = \int_0^t h_1(t-s) d\bar{Z}_N(s)$$

is a two dimensional system of PDMPs solving

$$dX_t = -\nu X_t + dY_t, \quad dY_t = -\nu Y_t dt + c\, d\bar{Z}_N(t).$$

Diffusion approximation

Each $Z_i(t)$ jumping at rate $f(X(t))$ gives rise to the approximation

$$\left\{ \begin{array}{lll} d\tilde{X}(t) & = & -\nu \tilde{X}(t)dt + \tilde{Y}_t dt \\ d\tilde{Y}(t) & = & -\nu \tilde{Y}(t)dt + cf(\tilde{X}(t))dt \\ & & + \frac{c}{\sqrt{N}} \sqrt{f(\tilde{X}(t)} dB_t \end{array} \right\}.$$

• Can be extended to the general case of n populations and of general Erlang memory kernels.

• We obtain a diffusion of high dimension driven by only few (actually, n) Brownian motions.

• We have the control on the weak error

$$\| P_t \varphi - \tilde{P}_t \varphi \|_\infty \le C t \frac{\|\varphi\|_{4,\infty}}{N^2}.$$

• We will show in some cases that this approximating diffusion oscillates in the long run - in the general case of n populations driven by Erlang kernels.

• Frame of a **monotone cyclic feedback system** in the sense of Mallet-Paret and Smith 1990 - which is somehow the oscillatory system "per se".
– Cyclic means : population k is only influenced by population $k+1$, for all k.
– Feedback : population n is only influenced by population 1.
– Monotone : we suppose that all rate functions f_k are non-decreasing.

Erlang kernels for the synaptic weight functions

- "Cyclic" : Intensity of the $i-$th particle belonging to population k is given by

$$\lambda_{k,i}(t) = f_k\left(\int_{]0,t[} h_k(t-s)\frac{1}{N_{k+1}}\sum_{j=1}^{N_{k+1}} dZ_{k+1,j}(s)\right).$$

- Here, the h_k are given by Erlang kernels

$$h_k(t) = c_k\frac{t^{m_k}}{m_k!}e^{-\nu_k t}, \nu_k > 0, m_k \in \mathbb{N}_0, c_k \in \mathbb{R}.$$

- If $c_k > 0$, then population $k+1$ is excitatory for population k. Else : inhibitory.

- Put $\delta := \prod_{k=1}^n c_k$. If $\delta > 0$, the system is of positive feedback, else, it is of negative feedback.

Erlang kernels for the synaptic weight functions

• "Cyclic" : Intensity of the $i-$th particle belonging to population k is given by

$$\lambda_{k,i}(t) = f_k \left(\int_{]0,t[} h_k(t-s) \frac{1}{N_{k+1}} \sum_{j=1}^{N_{k+1}} dZ_{k+1,j}(s) \right) .$$

• Here, the h_k are given by Erlang kernels

$$h_k(t) = c_k \frac{t^{m_k}}{m_k!} e^{-\nu_k t}, \nu_k > 0, m_k \in \mathbb{N}_0, c_k \in \mathbb{R}.$$

• If $c_k > 0$, then population $k+1$ is excitatory for population k. Else : inhibitory.

• Put $\delta := \prod_{k=1}^n c_k$. If $\delta > 0$, the system is of positive feedback, else, it is of negative feedback. We will consider the **negative feedback case.**

Monotone cyclic feedback system

• For the limit system and $m_t^k =$ mean number of jumps of population k before time t : $dm_t^k = f_k(x_t^k)dt$, where

$$x_t^k = \int_0^t h_k(t-s)dm_s^{k+1}, 1 \leq k \leq n.$$

• Deriving successively the Erlang kernel functions with respect to time, it is possible to develop the above system into a **high dimensional ODE** of dimension $\kappa := n$ (number of populations) $+ \sum_{k=1}^n m_k$ (memory length).

• This is a monotone cyclic negative feedback system as considered by *Mallet-Paret and Smith (1990)*, see also *Benaïm and Hirsch (1999)*.

Consequences

Under the condition that $\delta < 0$ (negative feedback) and that the f_k are bounded Lipschitz functions :

Theorem (Mallet-Paret and Smith)

1) $\exists!$ *equilibrium point* x^* *of the above system.*
2) $\exists$ *easily verifiable condition implying that* x^* *is unstable. In this case, there exists at least one non constant periodic orbit which is attracting.*

Consequences

Under the condition that $\delta < 0$ (negative feedback) and that the f_k are bounded Lipschitz functions :

Theorem (Mallet-Paret and Smith)

1) $\exists!$ equilibrium point x^* of the above system.
2) $\exists$ easily verifiable condition implying that x^* is unstable. In this case, there exists at least one non constant periodic orbit which is attracting.
3) If the dimension of the system is 3, then there exists a globally attracting invariance surface, and x^* is a repellor for the system.
The theorem of Poincaré-Bendixson implies that in this case, any solution will be attracted to a non constant periodic orbit.

• non constant periodic orbit = oscillations

• In which sense are these oscillations also felt by the finite size system ?

• non constant periodic orbit = oscillations

• In which sense are these oscillations also felt by the finite size system ?

• We will give an answer to this question for the associated diffusion approximation (noise is easier to handle !)

• Diffusion in dimension $n + \sum_{k=1}^{n} m_k$, driven by $n-$dimensional Brownian motion.

- Due to the **cascade structure** of the drift - coming from the development of the memory - it is easy to show that the diffusion satisfies the weak Hörmander condition.

- Hence it is strong Feller (*Ichihara and Kunita 1974*).

• Due to the **cascade structure** of the drift - coming from the development of the memory - it is easy to show that the diffusion satisfies the weak Hörmander condition.

• Hence it is strong Feller (*Ichihara and Kunita 1974*).

• Using a convenient Lyapunov-function and the control theorem (and ideas inspired by the work we did with Michèle Thieullen and Reinhard Höpfner on the stochastic Hodgkin-Huxley system), we obtain the following theorem.

Theorem

Let Γ be a non constant periodic orbit of the limit system which is asymptotically orbitally stable. Then for all $\varepsilon > 0$ and for all $T > 0$, for all starting configurations x, P_x−almost surely,

the approx diffusion visits $B_\varepsilon(\Gamma)$ during a time period of length T,

infinitely often.

Hence the diffusion approximation visits the oscillatory region infinitely often.

Theorem

Let Γ be a non constant periodic orbit of the limit system which is asymptotically orbitally stable. Then for all $\varepsilon > 0$ and for all $T > 0$, for all starting configurations x, P_x-almost surely,

the approx diffusion visits $B_\varepsilon(\Gamma)$ during a time period of length T,

infinitely often.

Hence the diffusion approximation visits the oscillatory region infinitely often.

The same result should hold true for the original PDMP (the intensities of the system, plus the developments of the memory) - but we did not prove this yet (we have a control on the weak approximation error when replacing the PDMP by the diffusion, but only on finite time horizon - and we did not yet look at support properties of the PDMPs).

Thank you for your attention.