

SPIKING NEURONS : INTERACTING HAWKES PROCESSES, MEAN FIELD LIMITS AND OSCILLATIONS*

E. LÖCHERBACH¹

Abstract. This paper gives a short survey of some aspects of the study of Hawkes processes in high dimensions, in view of modeling large systems of interacting neurons. We first present a perfect simulation result allowing for a graphical construction of the process in infinite dimension. Then we turn to the study of mean field limits and show how in some cases the delays present in the Hawkes intensities give rise to oscillatory behavior of the limit process.

Résumé. Cet article résume brièvement certains aspects de l'étude de processus de Hawkes en grande dimension, en vue d'une modélisation de systèmes de neurones en interactions. Nous présentons d'abord un résultat de simulation parfaite qui donne lieu à une construction graphique du processus en dimension infinie. Ensuite nous étudions des limites de champ moyen de processus de Hawkes et montrons comment les délais apparaissant dans le processus d'intensité engendrent des oscillations du processus limite.

INTRODUCTION

This paper provides a short survey of some aspects of the study of Hawkes processes in high dimensions in view of modeling biological neural systems and studying their long time behavior. In particular, we describe how neural oscillations might appear at a macroscopic scale even though the microscopic units of our model do not present any tendency to periodic behavior on their own. We present a microscopic model describing a large network of interacting neurons which can generate oscillations in a macroscopic frame. In our model, the activity of each neuron is represented by a point process indicating the successive times at which the neuron emits an action potential or a so-called spike. A realization of this point process is called a spike train. It is commonly admitted that the spiking intensity of a neuron, i.e., the infinitesimal probability of emitting an action potential during the next time unit, depends on the past history of the neuron and it is affected by the activity of other neurons in the network. Neurons interact mostly through chemical synapses. This means that a spike of a pre-synaptic neuron leads to an increase of the membrane potential of the post-synaptic neuron if the synapse is excitatory, or to a decrease if the synapse is inhibitory, possibly after some delay. In neurophysiological terms this feature is called synaptic integration. When the membrane potential reaches a certain upper threshold, the neuron fires a spike. Thus, excitatory inputs from the neurons in the network increase the firing intensity, and inhibitory inputs decrease it. Hawkes processes provide good models of this synaptic integration phenomenon by the structure of their intensity processes, see (1) below. We refer to Chevallier et al. (2015) [7], Chornoboy

* This research has been conducted as part of the project Labex MME-DII (ANR11-LBX-0023-01). Moreover, it is part of the activities of FAPESP Research, Dissemination and Innovation Center for Neuromathematics (grant 2013/07699-0, S. Paulo Research Foundation).

¹ CNRS UMR 8088, Département de Mathématiques, Université de Cergy-Pontoise, 2 avenue Adolphe Chauvin, 95302 Cergy-Pontoise Cedex, France. email : eva.loecherbach@u-cergy.fr

et al. (1988) [8], Hansen et al. (2015) [24] and to Reynaud-Bouret et al. (2014) [35] for the use of Hawkes processes in neuronal modeling. For an overview of point processes used as stochastic models for interacting neurons both in discrete and in continuous time and related issues, see also Galves and Löcherbach (2016) [22].

Our system consists of a large population of neurons indexed by a finite or countably infinite set I and described by a family of counting processes $Z_i(t), i \in I, t \geq 0$, recording the number of spikes of neuron i during the interval $[0, t]$. The sequence of counting processes $(Z_i)_{i \in I}$ is characterized by its intensity processes $(\lambda_i(t))_{i \in I, t \geq 0}$ which are defined through the relation

$$\mathbb{P}(Z_i \text{ has a jump in }]t, t + dt] | \mathcal{F}_t) = \lambda_i(t) dt, i \in I,$$

where $\mathcal{F}_t = \sigma(Z_j(s), s \leq t, j \in I)$ and $\lambda_i(t)$ is given as follows.

$$\lambda_i(t) = f_i \left(\sum_{k \in I} \int_{]0, t[} h_{ik}(t-s) dZ_k(s) \right). \quad (1)$$

In the above formula, $h_{ik} : \mathbb{R}_+ \rightarrow \mathbb{R}$ are *synaptic weight functions* measuring the influence of neuron k on neuron i and how this influence vanishes with the time. The function $f_i : \mathbb{R} \rightarrow \mathbb{R}_+$ is the *spiking rate function* of neuron i . Equation (1) has the typical form of the intensity of a multivariate nonlinear Hawkes process, going back to Hawkes (1971) [25] and Hawkes and Oakes (1974) [26]. We refer to Brémaud and Massoulié (1996) [5] for the stability properties of multivariate Hawkes processes, and to Delattre et al. (2016) [13] and Chevallier (2015) [6] for the study of Hawkes processes in high dimensions.

Existence and graphical construction via a perfect simulation algorithm. In the case where I is a finite set, under suitable assumptions on the parameters of the process, the existence and construction of $(Z_i)_{i \in I}$ is standard (see [5] and [13]). In the case when the number of interacting components defining the process is infinite, [13] prove path-wise existence and uniqueness of the processes, however without giving an explicit construction of the process. In Section 1 of this note, we show that a *graphical construction* of the system is possible under suitable assumptions. This graphical construction does not only imply the existence but also the possibility of a perfect simulation of a stationary version of the process. This result is achieved via a *Kalikow-type decomposition*, see Proposition 1.10 below.

Mean-field limits and oscillations. In Section 2, we study a large homogeneous population of N neurons evolving according to (1), where $I = \{1, \dots, N\}$. We assume that all neurons have identical properties which leads to a mean field description. The mean field assumption appears through the prescription that $h_{ij} = \frac{1}{N} h$ for all $1 \leq i, j \leq N$, for some integrable function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$, such that $h \in L^2_{loc}$, and that $f_i(x) = f(x)$ for some fixed spiking rate function $f : \mathbb{R} \rightarrow \mathbb{R}_+$ which is bounded and Lipschitz continuous.

We study the large population limit when $N \rightarrow \infty$ and show that in this limit self-sustained periodic behavior emerges even though each single neuron does not follow periodic dynamics. This study follows a long tradition, starting probably with Scheutzow (1985) [33, 34] in the framework of nonlinear diffusion processes, or Dai Pra, Fischer and Regoli (2015) [12], Collet, Dai Pra and Formentin (2015) [9] and Collet, Formentin and Tovazzi (2016) [10]. In all these papers, rhythmic behavior arises in the mean field limit as a consequence of the non-linear ‘‘Mc-Kean-Vlasov’’-type dynamics of the limit process. Our result extends these studies to the framework of infinite memory point processes.

The limit system is represented through a process $\bar{Z}(t)$ counting the spikes of a typical neuron, up to time t , in the limit population. Its expectation $m_t = \mathbb{E}(\bar{Z}(t))$ is solution of

$$m_t = \int_0^t f \left(\int_0^s h(s-u) dm_u \right) ds. \quad (2)$$

If the memory kernel h is an Erlang kernel, see (16) below, then our Theorem 3.2 characterizes situations in which the system (2) possesses attracting non-constant periodic orbits, that is, presents oscillatory behavior. This result goes back to deep theorems in dynamical systems, obtained by Mallet-Paret and Smith (1990) [31]

and used in a different context in Benaïm and Hirsch (1999) [1]. The order of the delay of the system is an important factor in this study; in some cases, increasing the delay will enhance oscillations, in others, higher delays tend to stabilize the system. We present in Remark 3.3 below a detailed discussion of this point.

Using Erlang kernels allows to describe the intensity process of the system in terms of an equivalent high dimensional system of *piecewise deterministic Markov processes* (PDMPs) (see also Duarte et al. (2016) [18]). In Section 4, we obtain an approximating diffusion equation in (23) which is shown to be close in a weak sense to the original PDMP defining the Hawkes process. We provide a study of the longtime behavior of this diffusion. In particular, we show to which extent the approximating diffusion presents the same oscillatory behavior as the limit system and how its invariant measure concentrates around periodic orbits of the limit system. We close this note with some final remarks in Section 5 where we discuss a few related questions that should be addressed in the future.

1. SYSTEMS OF INTERACTING HAWKES PROCESSES

Neurons generate and propagate action potentials along their axons, and they communicate by transmitting spikes. Spikes are fast transmembrane currents of K^+ / Na^+ -ions which are stimulated by ion pumps. The shape and the time duration of a spike is almost deterministic and always the same – for a given neuron, under the same experimental conditions. Moreover, the time duration of a spike is very short, of about $1ms$, and followed by a refractory period during which the neuron can not spike again. It is therefore reasonable to report for any neuron the sequence of its spike times, giving rise to what is called a neuronal *spike train*.

Therefore, in the following we adopt the following description of interacting neurons, relying on particular point processes which are the Hawkes processes. We fix a countable set I to denote the set of neurons. The activity of each neuron $i \in I$ is described by a counting process $Z_i(t), i \in I, t \geq 0$, recording the number of spikes of neuron i during the interval $[0, t]$. We suppose that $Z_i(0) = 0$ for all $i \in I$. The sequence of counting processes $(Z_i)_{i \in I}$ is characterized by its intensity processes $(\lambda_i(t))_{i \in I, t \geq 0}$ which are defined through the relation

$$\mathbb{P}(Z_i \text{ has a jump in }]t, t + dt] | \mathcal{F}_t) = \lambda_i(t) dt, i \in I,$$

where $\mathcal{F}_t = \sigma(Z_j(s), s \leq t, j \in I)$ and $\lambda_i(t)$ is given precisely in (4) below.

The intensity process $\lambda_i(t)$ should incorporate the interactions between the neurons. It should also represent the way the spiking behavior of neuron i depends on its history. Indeed, it is commonly admitted that spike trains are processes having *infinite memory*.

To specify the intensity we introduce a family of *synaptic weight functions* $h_{ij} : \mathbb{R} \rightarrow \mathbb{R}$, for $i, j \in I$, and a family of *spiking rate functions* $f_i : \mathbb{R} \rightarrow \mathbb{R}_+$, for any $i \in I$.

We assume that

Assumption 1.1. (i) *There exists a finite constant L such that for every x and x' in \mathbb{R} , for every $i \in I$,*

$$|f_i(x) - f_i(x')| \leq L|x - x'|. \quad (3)$$

(ii) *The functions $h_{kl}, k, l \in I$, belong to $L^2_{loc}(\mathbb{R}_+; \mathbb{R})$.*

The form of intensity that we shall consider is as follows.

$$\lambda_i(t) = f_i \left(\sum_{k \in I} \int_{]0, t[} h_{ik}(t-s) dZ_k(s) \right). \quad (4)$$

In the above formula, $h_{ik}(t-s)$ measures the influence of neuron k on neuron i and how this influence vanishes with the time. More precisely, $h_{ik}(t-s)$ describes how a spike of neuron k lying back $t-s$ time units in the past influences the present spiking rate at time t . If the function h_{ik} is not of compact support, then the associated process is truly of *infinite memory*.

Example 1.2. Let

$$\lambda_i(t) = f_i \left(\sum_j W_{ij} \int_{]0,t[} e^{-\alpha(t-s)} dZ_j(s) \right),$$

for some $\alpha > 0$. Here, $(W_{ij}, i, j \in I)$ is a collection of real numbers such that $W_{ii} = 0$ for all $i \in I$ and

$$\sup_i \sum_j |W_{ij}| < \infty.$$

We call W_{ij} the *synaptic weight* of neuron j on neuron i . Notice that the influence of past events decay at an exponential rate at speed α . The above structure of the intensity $\lambda_i(t)$ induces a directed graph \mathcal{E} consisting of all directed edges $e = j \rightarrow i, i, j \in I$, such that $W_{ij} \neq 0$. If $W_{ij} > 0$, we say that neuron j is excitatory for i , else, it is inhibitory. In Duarte et al. (2016) [19], a model selection procedure to estimate this interaction graph is proposed.

1.1. Discussion of the model

Writing

$$U_i(t) = \sum_{j \in I} \int_{]0,t[} h_{ij}(t-s) dZ_j(s), \quad (5)$$

we interpret $U_i(t)$ as value of the membrane potential of neuron i at time t . Indeed, the form of the intensity process in (4) implies that we adopt an *integrate-and-fire*-mechanism to describe neuronal spiking activity: the process $U_i(t)$ associated to neuron i accumulates all the past spike events of its *presynaptic neurons* up to time t , and then the neuron fires depending on the actual height of $U_i(t)$. This corresponds to the classical description of the membrane potential of a neuron in integrate-and-fire models, and therefore it is reasonable to use $U_i(t)$ as a good description of the membrane potential of neuron i a time t .

In general, $(U_i(t))_{i \in I}$ will not be a Markov process. If however $h_{ij}(t) = W_{ij}e^{-\alpha t}$ as in Example 1.2, then we formally have that

$$dU_i(t) = -\alpha U_i(t) + \sum_{j \in I} W_{ij} dZ_j(t). \quad (6)$$

We will come back to Markovian descriptions of Hawkes processes in Section 4 below.

Concerning the spiking rate functions f_i , models in computational neuroscience rely often on the assumption of a fixed *firing threshold*, i.e. a potential height K such that, once $U_i(t)$ exceeds K , neuron i spikes immediately. Loosely speaking, this would amount to taking (approximations of)

$$f_i(u) = \infty 1_{\{u \geq K\}},$$

which means that neuron i spikes immediately once its membrane potential $U_i(t)$ hits the interval $[K, \infty[$. However there seems to be statistical evidence that a fixed firing threshold does not exist, see e.g. Jahn et al. (2011) [29]. Therefore, our model does not impose a fixed firing threshold, we simply suppose that the firing rate of each neuron f_i is a Lipschitz continuous, increasing function, which is reasonable from a modeling point of view.

To summarize, our model is a huge or infinite system of interacting neurons represented by a Hawkes process in high or infinite dimension. Each neuron is represented by its membrane potential $U_i(t)$ at time t . This membrane potential integrates (sums up) all past spiking events of the other neurons directly influencing it. And each neuron i spikes randomly at rate $f_i(U_i(t))$. Finally, the resulting process is of infinite length memory.

1.2. Existence results

In the case I finite, the existence of a process $(Z_i(t))_{i \in I, t \geq 0}$, having intensity (4) on finite time intervals is not an issue, as shows the following

Proposition 1.3 (Theorem 6 of Delattre et al. (2016) [13]). *Suppose that I is finite. Then under Assumption 1.1, there exists a path-wise unique Hawkes process $(Z_i(t))_{i \in I}$, starting from $Z_i(0) = 0$ for all $i \in I$, for all $t \geq 0$.*

However in the case when I is infinite, the existence of the process is not immediately clear, since the interactions might come down from infinity. In this case we have to impose the additional assumption

Assumption 1.4. *There exist some positive weights $p_i, i \in I$, and a locally integrable function $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, such that*

$$\sum_{i \in I} f_i(0)p_i < \infty \quad (7)$$

and for all $i \in I$, for all $s > 0$,

$$\sum_{j \in I} p_j |h_{ij}(s)| \leq p_i \Phi(s). \quad (8)$$

Then the following existence theorem has been proven in [13], using simple Lipschitz arguments.

Theorem 1.5 (Theorem 6 of Delattre et al. (2016) [13]). *Grant Assumption 1.4. Then there exists a path-wise unique Hawkes process $(Z_i(t))_{i \in I, t \geq 0}$, starting from $Z_i(0) = 0$ for all $i \in I$, such that*

$$\sum_{i \in I} p_i \mathbb{E}(Z_i(t)) < \infty \quad \text{for all } t \geq 0.$$

1.3. Perfect simulation

Theorem 1.5 above gives the path-wise unique existence of a Hawkes process with intensity (4) in infinite dimension, however without giving an explicit construction of the process. In the present section we briefly show how – under suitable assumptions – a *graphical construction* of the process is possible. This graphical construction does not only imply the existence but also the possibility of a perfect simulation of a stationary version of the process. The procedure is feasible under the assumption of bounded spiking rates, that is, we suppose that $\|f_i\|_\infty < \infty$ for all $i \in I$. The construction is then achieved via a *Kalikow-type decomposition*. Kalikow-type decompositions have been largely used in the literature for perfect simulation issues and similar scopes. They have been considered first by Ferrari, Maass, Martínez and Ney (2000) [20] and in Comets, Fernández and Ferrari (2002) [11]. This type of technique was then studied in a series of papers for perfect simulation issues. See Galves and Löcherbach (2016) [22] for an application in the context of neural biological nets in discrete time.

In the following we suppose that

Assumption 1.6. *There exists an integrable function $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\int_0^\infty \Phi(s) ds < \infty$ such that*

$$\sup_{i \in I} \sum_{j \in I} |h_{ij}(s)| \leq \Phi(s),$$

for all $s \geq 0$.

We introduce $\mathcal{V}_{\rightarrow i} := \{j \in I : h_{ij} \neq 0\}$, the set of all neurons directly influencing i , and impose

Assumption 1.7. *The set I of the neurons is divided into layers $(I_n)_{n \in \mathbb{Z}}$ such that we have the partition $I = \sqcup_{n \in \mathbb{Z}} I_n$. For each $n \in \mathbb{Z}$ and for each $i \in I_n$, we suppose that*

$$\mathcal{V}_{\rightarrow i} \subset I_{n-1}. \quad (9)$$

Therefore, a neuron only receives information from neurons in the layer just above itself. This assumption does not apply to the brain's structure in general, but it can be found in simpler nervous tissues like the retina.

In order to state our perfect simulation result, we need additional notation. We introduce for any $i \in I$ a non-decreasing sequence $(V_i(k))_{k \geq 0}$ of finite subsets of I such that $V_i(0) = \emptyset$, $V_i(1) = \{i\}$, $V_i(k) \subset V_i(k+1)$, $V_i(k) \neq V_i(k+1)$ if $V_i(k) \neq \mathcal{V}_{\rightarrow i} \cup \{i\}$ and $\bigcup_k V_i(k) = \mathcal{V}_{\rightarrow i} \cup \{i\}$. We define, for all $k \geq 0$, $\partial V_i(k) := V_i(k+1) \setminus V_i(k)$, the border of the set $V_i(k)$.

Theorem 1.8. [Theorem 3 of Hodara and Löcherbach (2016) [28]] *Grant Assumptions 1.1, 1.6 and 1.7. If*

$$\sup_{i \in I} \left(\sum_{k \geq 1} \left[(k|V_i(k)| + 1) \times \left(\sum_{j \in V_i(k-1)} \int_{k-1}^k |h_{ij}(s)| ds + \sum_{j \in \partial V_i(k-1)} \int_0^k |h_{ij}(s)| ds \right) \right] \right) < \frac{1}{L}, \quad (10)$$

where L is given in Assumption 1.1, then there exists a unique stationary Hawkes process $(Z_i(t))_{i \in I, t \in \mathbb{R}}$ with intensity given by (4). Moreover, in this case, we dispose of a perfect simulation algorithm of the stationary measure.

Remark 1.9. Notice how in the above theorem, condition (10) depends on the specific choice of growing neighborhoods $(V_i(k))_{k \geq 0}$. We refer the reader to De Santis and Lissandreli (2012) [15] for a study (in the framework of stochastic chains having infinite memory) how to optimize over the choice of the sequence of growing neighborhoods in order to obtain the weakest sufficient condition (10).

Let us briefly sketch how our perfect simulation algorithm works. Recall that we work with bounded firing rate functions, i.e., $\|f_i\|_\infty < \infty$ for each $i \in I$. This allows to introduce a dominating Poisson Random Measure $N(dt, di, dz)$ on $\mathbb{R} \times I \times [0, 1]$ with intensity $dt \left(\sum_{i \in I} \|f_i\|_\infty \delta_i \right) dz$ on $\mathbb{R} \times I \times [0, 1]$ dominating the process $(Z_i(t))_{i \in I, t \in \mathbb{R}}$. In this way we will be able to select the jump times of Z_i among those of $N_i := N(dt, \{i\}, [0, 1])$, applying a thinning procedure. Note that $N_i(dt)$ is a homogeneous Poisson random measure on \mathbb{R}_+ with intensity $\|f_i\|_\infty dt$.

Since Z_i is a simple point measure, it is enough to define it through the times of its atoms. We write $T_n^i, n \in \mathbb{Z}$, for the jump times of N_i . Fix a given jump $t = T_n^i$ of N_i . Then, conditionally on N , the probability that this time is also a jump time of Z_i is given by

$$\mathbb{P}(\Delta Z_i(t) = 1 | \mathcal{F}_t) = \frac{1}{\|f_i\|_\infty} f_i \left(\sum_j \left(\int_{]-\infty, t[} h_{ij}(t-s) dZ_j(s) \right) \right) =: p_{(i,t)}(1 | \mathcal{F}_t). \quad (11)$$

In other words, given that t is a jump time of N_i , $p_{(i,t)}(1 | \mathcal{F}_t)$ is the probability that this jump is also a jump of Z_i . This probability depends on the past before time t of the process.

The main tool of our procedure is a *decomposition of $p_{(i,t)}(1 | \mathcal{F}_t)$ according to growing space-time neighborhoods* that explore the *relevant past* needed in order to determine $p_{(i,t)}(1 | \mathcal{F}_t)$. In order to do so we introduce the space-time neighborhoods

$$\mathbb{V}_t^i(k) := V_i(k) \times [t-k, t], k \geq 0,$$

and work with the state space $X = \{0, 1\}^{\mathcal{G}}$ where \mathcal{G} is the time grid $\{(i, T_n^i), (i, n) \in I \times \mathbb{Z}\}$. Thus, conditionally on the realization of N , any configuration $x \in X$ represents a possible realization of the process $(Z_i)_{i \in I}$. In the following we propose a decomposition of $p_{(i,t)}(1|x)$ where

$$p_{(i,t)}(1|x) = \frac{1}{\|f_i\|_\infty} f_i \left(\sum_j \int_{]-\infty, t[} h_{ij}(t-s) dx_j(s) \right).$$

Proposition 1.10. *[Kalikow-type decomposition; Proposition 3 of Hodara and Löcherbach (2016) [28]] Grant Assumptions 1.1, 1.6 and 1.7. Then there exists a family of positive random variables*

$$\left(p_{(i,t)}^{[k]}(a|x) \right)_{k \geq 0}, a \in \{0, 1\},$$

and

$$\mu_{(i,t)}(k), k \geq 0,$$

which depend only on the realization of the dominating Poisson Random Measure N , such that almost surely the following assertions hold.

- (1) $\mu_{(i,t)}(k), k \geq 0$, defines a probability measure on \mathbb{N} , that is, $\sum_{k=0}^{\infty} \mu_{(i,t)}(k) = 1$.
- (2) For all $a \in \{0, 1\}$, $p_{(i,t)}^{[0]}(a|x) = p_{(i,t)}^{[0]}(a)$ does not depend on x .
- (3) For all $a \in \{0, 1\}$, $k \geq 1$, $X \ni x \mapsto p_{(i,t)}^{[k]}(a|x)$ depends only on the variables $(x(j, T_n^j) : (j, T_n^j) \in \mathbb{V}_t^i(k))$.
- (4) For all $x \in X$, $k \geq 1$, $p_{(i,t)}^{[k]}(1|x) \in [0, 1]$, $p_{(i,t)}^{[k]}(1|x) + p_{(i,t)}^{[k]}(0|x) = 1$.
- (5) For all $a \in \{0, 1\}$, $x \in X$, $p_{(i,t)}^{[k]}(a|x)$ and $\mu_{(i,t)}(k)$ are \mathcal{F}_t -measurable random variables.
- (6) For all $x \in X$, we have the following convex decomposition

$$p_{(i,t)}(a|x) = \sum_{k \geq 0} \mu_{(i,t)}(k) p_{(i,t)}^{[k]}(a|x). \quad (12)$$

(12) implies that the acceptance probabilities $p_{(i,t)}(1|x)$ can be realized according to the following 2-step procedure. First, choose a space-time neighborhood $\mathbb{V}_t^i(k), k \geq 0$, with probability $\mu_{(i,t)}(k)$. Then, inspect all sites belonging to this neighborhood $\mathbb{V}_t^i(k)$ and decide to accept the jump of neuron i at time t with probability $p_{(i,t)}^{[k]}(a|x)$, depending only on the sites within this neighborhood.

This gives rise to the following perfect simulation algorithm to decide whether to accept or not a spike of neuron i at time t . We first introduce for any $(j, s) \in \mathcal{G} = \{(i, T_n^i), (i, n) \in I \times \mathbb{Z}\}$ and any $k \geq 0$,

$$C_1^{(j,s)}(k) := \mathcal{G} \cap \mathbb{V}_s^j(k),$$

where by convention $C_1^{(j,s)}(0) := \emptyset$. We then attach i.i.d. random variables $K^{(j,s)} \in \mathbb{N}$ to each site $(j, s) \in \mathcal{G}$, chosen according to

$$\mathbb{P}(K^{(j,s)} = k) = \mu_{(j,s)}(k), \text{ for all } k \geq 0.$$

These random variables allow us to define the clans of ancestors $(C_n^{(i,t)})_n \subset I \times]-\infty, t[$ for the fixed couple (i, t) as follows. We put $C_1^{(i,t)} := C_1^{(i,t)}(K^{(i,t)})$ and

$$C_n^{(i,t)} := \left(\bigcup_{(j,s) \in C_{n-1}^{(i,t)}} C_1^{(j,s)} \right) \setminus \left(C_1^{(i,t)} \cup \dots \cup C_{n-1}^{(i,t)} \right),$$

where $C_1^{(j,s)} = C_1^{(j,s)}(K^{(j,s)})$.

Notice that if $K^{(i,t)} = 0$ in the first step, then the algorithm stops immediately. Else we introduce

$$N^{Stop} := \min\{n : C_n^{(i,t)} = \emptyset\}$$

the number of steps of the algorithm, where $\min \emptyset := \infty$. The set

$$\mathcal{C}^{(i,t)} := \bigcup_{n=1}^{N^{Stop}} \mathcal{C}_n^{(i,t)}$$

contains all (sites and times of all) spikes which have to be accepted or not and whose possible presence have an influence on the acceptance/rejection decision of (i, t) itself. Under the conditions of Theorem 1.8, it can be shown that $N^{Stop} < \infty$ almost surely.

Once the clan of ancestors is determined, we can realize the acceptance/rejection procedure of the elements in this clan in a second algorithm which is a forward procedure going from the past to the present. We start with the sites for which this decision can be made independently from anything else. During the algorithm the set of all sites for which a decision has already been achieved will then progressively be updated.

- (0) At the initial stage of the algorithm, the set of sites for which the acceptance/rejection decision can be achieved is initialized by

$$D^{(i,t)} := \left\{ (j, s) \in \mathcal{C}^{(i,t)}, C_1^{(j,s)} = \emptyset \right\}.$$

The sites within this set are precisely those for which the decision can be made independently from anything else, because an interaction neighborhood of size 0 has been chosen.

- (1) For each $(j, s) \in D^{(i,t)}$, we simulate, according to the probabilities $p_{(j,s)}^{[0]}(a)$, the state of this site.
 (2) For any $(j, s) \in \mathcal{C}^{(i,t)}$ with $C_1^{(j,s)} \subset D^{(i,t)}$, we then decide, according to the probabilities $p_{(j,s)}^{[k]}(a|x)$ with $k = K^{(j,s)}$, to accept or to reject the presence of a spike of neuron j at time s . This is possible since $p_{(j,s)}^{[k]}(a|x)$ depends on the configuration x only through the space-time sites in $C_1^{(j,s)}$ whose states have already been determined since $C_1^{(j,s)} \subset D^{(i,t)}$. Then we update $D^{(i,t)}$ in the following way:

$$D^{(i,t)} \leftarrow D^{(i,t)} \bigcup \left\{ (j, s) \in \mathcal{C}^{(i,t)}, C_1^{(j,s)} \subset D^{(i,t)} \right\}.$$

- (3) The update of $D^{(i,t)}$ allows to repeat the previous step until $(i, t) \in D^{(i,t)}$.

Once we have assigned a decision to the element (i, t) itself, our *perfect simulation* algorithm stops, since we are now able to decide whether neuron i will have a spike at time t or not. We refer the reader to [28] for details.

2. MEAN FIELD LIMITS AND PROPAGATION OF CHAOS

We now turn to the second main topic of this paper which is the study of large systems of interacting Hawkes processes in a mean field frame. Suppose we observe a large homogeneous population of N neurons evolving according to (4), where $I = \{1, \dots, N\}$. Then we can assume that we are in an idealized situation where all neurons have identical properties, leading to a mean field description. The mean field assumption appears through the assumption that

$$h_{ij} = \frac{1}{N} h$$

for all $1 \leq i, j \leq N$, for some integrable function $h : \mathbb{R}_+ \rightarrow \mathbb{R}$, such that $h \in L_{loc}^2$, and that $f_i(x) = f(x)$ for some fixed spiking rate function $f : \mathbb{R} \rightarrow \mathbb{R}_+$ which is bounded and Lipschitz continuous. In order to keep track of the size of the system, we denote the process by $Z^N(t) = (Z_1^N(t), \dots, Z_N^N(t)), t \geq 0$. We identify the state of the system at time t with its empirical measure

$$\frac{1}{N} \sum_{1 \leq i \leq N} \delta_{(Z_i^N(t))_{t \geq 0}}, \quad (13)$$

taking values in the set $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}_+))$ of probability measures on the space of càdlàg functions, $D(\mathbb{R}_+, \mathbb{R}_+)$. We endow $D(\mathbb{R}_+, \mathbb{R}_+)$ with the Skorokhod topology, and $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}_+))$ with the weak convergence topology associated with the Skorokhod topology on $D(\mathbb{R}_+, \mathbb{R}_+)$.

As N , the number of neurons, tends to infinity, we have convergence in distribution

$$\frac{1}{N} \sum_{1 \leq i \leq N} \delta_{(Z_i^N(t))_{t \geq 0}} \xrightarrow{\mathcal{L}} P,$$

where the limit measure P is the distribution of a limit process $\bar{Z}(t)$, defined through

$$\bar{Z}(t) = \int_0^t \int_{\mathbb{R}_+} 1_{\{z \leq f(\int_0^s h(s-u) d\mathbb{E}(\bar{Z}(u)))\}} N(ds, dz), \quad (14)$$

and where N a Poisson random measures on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity measure $dsdz$. In particular, if we put $m_t = \mathbb{E}(\bar{Z}(t))$, then, taking expectations in (14), it follows that m_t is solution of

$$m_t = \int_0^t f \left(\int_0^s h(s-u) dm_u \right) ds. \quad (15)$$

Notice that the process $\bar{Z}(t)$ of (14), being a counting process with deterministic intensity, is necessarily a Poisson process, however, with intensity depending on its own law, as it is usually the case in mean field limits. We quote the following theorem from [13].

Theorem 2.1. [Theorem 8 of Delattre et al. (2016) [13], see also Theorem 1 of Ditlevsen and Löcherbach (2016) [17]]

Under the assumptions of this section, there exists a path-wise unique solution to (14) such that $t \mapsto \mathbb{E}(\bar{Z}(t))$ is locally bounded. Moreover, for any $i \geq 1$,

$$(Z_i^N(t))_{t \geq 0} \xrightarrow{\mathcal{L}} (\bar{Z}(t))_{t \geq 0}$$

as $N \rightarrow \infty$ (convergence in $D(\mathbb{R}_+, \mathbb{R}_+)$, endowed with the Skorokhod topology).

Remark 2.2. The equivalence between the “chaoticity” of the system and a weak law of large numbers for the empirical measures, as provided in Theorem 2.1, is well-known (see for instance Sznitman (1989) [37]). This means that in the large population limit, the neurons converge in law to independent and identically distributed copies of the same limit law. This property is usually called “propagation of chaos” in the literature.

Remark 2.3. The above theorem has been stated in [17] in a multi-class frame where neurons belong to a given collection of classes. In this multi-class frame, going back to Graham (2008) [23], interactions between neurons of the same class are of mean field type, and those of neurons belonging to different classes depend only on the label of the corresponding class that a neuron belongs to and not on its individual label. We refer to [17] for more details.

3. STUDY OF THE LIMIT PROCESS

The equation for the limit intensity (15) possesses a unique non-decreasing locally bounded solution, which furthermore is of class C^1 on $[0, \infty[$ (Lemma 24 of [13]). In the following we give conditions under which

$$\bar{\lambda}(t) = \frac{dm_t}{dt}$$

has a (unique) oscillating solution. For that sake, we study Erlang memory kernels $h(t)$. Erlang memory kernels are widely used for modeling issues, e.g. in modeling the delay in the hemodynamics in nephrons, see Ditlevsen et al (2005) [16] and Skeldon and Purvey (2005) [36]. They are of the form

$$h(t) = ce^{-\alpha t} \frac{t^n}{n!}, \quad (16)$$

where $\alpha > 0$ and $c \in \mathbb{R}$ are fixed constants and where $n \in \{0, 1, 2, \dots\}$ is an integer. In (16), the delay of influence of the past is distributed. The delay takes its maximal value at about $(n+1)/\alpha$ time steps back in the past. The higher the order n of the delay is, the more the delay is concentrated around its mean value $(n+1)/\alpha$, when keeping this ratio $(n+1)/\alpha$ fixed. We shall call the value n the *order of the memory*. Finally, the influence of the past is excitatory if $c > 0$, and inhibitory, else.

Introduce now the following variables for $0 \leq k \leq n$.

$$x_t^k = c \int_0^t e^{-\alpha(t-s)} \frac{(t-s)^{n-k}}{(n-k)!} \bar{\lambda}(s) ds. \quad (17)$$

Then

$$\bar{\lambda}(t) = f\left(\int_0^t h(t-s)\bar{\lambda}(s)ds\right) = f(x_t^0).$$

Moreover, (x_t^0, \dots, x_t^n) satisfies the following system of coupled differential equations.

$$\begin{aligned} \frac{dx_t^k}{dt} &= -\alpha x_t^k + x_t^{k+1}, \quad 0 \leq k < n, \\ \frac{dx_t^n}{dt} &= -\alpha x_t^n + cf(x_t^0), \end{aligned} \quad (18)$$

with initial conditions $x_0^k = 0$, for all $0 \leq k \leq n$. System (18) exhibits the structure of a *monotone cyclic feedback system* as considered e.g. in Mallet-Paret and Smith (1990) [31] or as (33) and (34) in Benaïm and Hirsch (1999) [1]. Cyclic means that coordinate k is only influenced by coordinate $k+1$ and itself (where we identify $n+1$ with 0). The monotonicity refers to the fact that the function f is non-decreasing. If $c > 0$, then the system (18) is of positive feedback, otherwise it is of negative feedback. We obtain the following simple first result.

Proposition 3.1. *Suppose that $c < 0$ and that f is non-decreasing. Then (18) admits a unique equilibrium x^* .*

Proof. Any equilibrium $x^* = ((x^*)^0, \dots, (x^*)^n)$ must satisfy

$$(x^*)^n = \frac{c}{\alpha} f\left(\frac{1}{\alpha^n} (x^*)^n\right).$$

Since $\frac{c}{\alpha} f\left(\frac{1}{\alpha^n}(\cdot)\right)$ is decreasing, there exists exactly one solution $(x^*)^n$ in \mathbb{R} . Once $(x^*)^n$ is fixed, we obviously have $(x^*)^{n-1} = \frac{1}{\alpha} (x^*)^n$, and the values of the other coordinates of x^* follow in a similar way. \square

We now give conditions ensuring that (18) is necessarily attracted to a non-equilibrium periodic orbit.

Theorem 3.2. *[Theorem 3 of Ditlevsen and Löcherbach (2016) [17]] Suppose that $n \geq 2$ and that f is non-decreasing, bounded and analytic. Moreover, suppose that $\rho := cf'((x^*)^0) < 0$ and that*

$$|\rho| > \frac{\alpha^{n+1}}{\left(\cos\left(\frac{\pi}{n+1}\right)\right)^{n+1}}. \quad (19)$$

- (1) Then x^* is linearly unstable, and the system (18) possesses at least one, but no more than a finite number of periodic orbits. At least one of them is orbitally asymptotically stable.
- (2) Moreover, if $n = 2$, then there exists a globally attracting invariant surface Σ such that x^* is a repeller for the flow in Σ . Every solution of (18) will be attracted to a non constant periodic orbit.

Proof. The above theorem is based on Theorem 4.3 of [31] and generalizes the result obtained by Theorem 6.3 in [1]. \square

Remark 3.3. [Influence of the order of memory on the emergence of oscillations]

In some cases, increasing the order of the memory, i.e. the value of the exponent n in (16), can lead to a bifurcation of the system (18). At the bifurcation point, a system which was stable can become unstable, and in certain cases, increasing the order even more might stabilize the system again.

If $n = 1$ (and therefore the dimension of the system (18) is 2), the fixed point is stable since eigenvalues are $\lambda_j = -\alpha \pm i\sqrt{|\rho|}$, and only damped oscillations occur. We will assume $n \geq 2$.

First note that ρ is bounded due to the Lipschitz condition on the rate function f . The right hand side of (19) goes to infinity for $\alpha \rightarrow \infty$ for all values of n , and thus, if α is large, the system will always be stable and not exhibit oscillations. For any fixed value of $\alpha > 1$, it also goes to infinity for $n \rightarrow \infty$, such that a possibly unstable system becomes stable for increasing n . Now assume that $\alpha = 1$. Then increasing n does not change the coordinate $(x^*)^0$ of the equilibrium state x^* , so ρ does not change. The right hand side of (19) decreases towards one, so if $-8 < \rho < -1$, then there exists $n_0 > 2$ minimal such that all models corresponding to $n < n_0$ have x^* as attracting equilibrium point, but for $n \geq n_0$, the equilibrium x^* becomes unstable.

Remark 3.4. In [17] we have stated the above Theorem 3.2 in multi-class frame. We were mainly thinking of two populations of interacting neurons, where the first population is inhibitory on the second and the second excitatory for the first and where oscillations occur due to the structure of the network, i.e. the fact that the network is what is usually called a *frustrated network*. However, as shows Theorem 3.2, this structure is not even necessary to explain the emergence of oscillations, it is sufficient to impose simply a sufficiently high delay of the memory.

4. DIFFUSION APPROXIMATION

Once we have well understood the limit system and the longtime behavior of the intensity process $\bar{\lambda}(t)$, a natural question to ask is to which extent the large time behavior of the limit system predicts the large time behavior of the finite size system.

We give a (preliminary) answer to this question; this answer is preliminary in the sense that we will deal with the canonical diffusion approximation of our process (see (23) below) instead of the Hawkes process (at fixed system size N) itself. In the sequel, we suppose that

$$f(x) \geq \underline{f} > 0 \quad \text{for all } x \in \mathbb{R}, \quad (20)$$

i.e., the spiking rate is strictly lower bounded. Moreover, we shall assume that the parameters of the system are such that the limit system (18) possesses a non-equilibrium periodic orbit Γ .

Recall the form of the memory kernel in (16) and consider the family of adapted càdlàg processes

$$X_k^N(t) := c \int_{]0,t]} \frac{(t-s)^{n-k}}{(n-k)!} e^{-\alpha(t-s)} d\bar{Z}^N(s), \quad (21)$$

for $0 \leq k \leq n$, where $\bar{Z}^N(s) = \frac{1}{N} \sum_{j=1}^N Z_j^N(s)$. Then each process $Z_i^N(t)$, $1 \leq i \leq N$, jumps at rate $f(X_0^N(t-))$, and we obtain the following system of stochastic differential equations which is a stochastic version of (18).

$$\begin{cases} dX_k^N(t) &= [-\alpha X_k^N(t) + X_{k+1}^N(t)]dt, \quad 0 \leq k < n, \\ dX_n^N(t) &= -\alpha X_n^N(t)dt + cdZ^N(t). \end{cases} \quad (22)$$

We call the system (22) a *cascade of memory terms*. Thus, the dynamics of the Hawkes process $(Z_i^N(t))_{1 \leq i \leq N}$ is entirely determined by the piecewise deterministic Markov process (PDMP) $(X_k^N)_{0 \leq k \leq n}$ of dimension $n + 1$. It is natural to consider the approximating diffusion process

$$\begin{cases} dY_k^N(t) &= [-\alpha Y_k^N(t) + Y_{k+1}^N(t)]dt, & 0 \leq k < n, \\ dY_n^N(t) &= -\alpha Y_n^N(t)dt + cf(Y_0^N(t))dt + c\frac{\sqrt{f(Y_0^N(t))}}{\sqrt{N}}dB(t), \end{cases} \quad (23)$$

where $B(t)$ is a standard Brownian motion, approximating the jump noise of the system. Under appropriate assumptions on f , we have the following control of the weak error,

$$\|P_t^X \varphi - P_t^Y \varphi\|_\infty \leq Ct \frac{\|\varphi\|_{4,\infty}}{N^2},$$

where P_t^X denotes the Markovian semigroup of the process (22) and P_t^Y of the diffusion approximation (23), and where for any test function φ , $\|\varphi\|_{4,\infty} := \sum_{k=0}^4 \sum_{|\alpha|=k} \|\partial^\alpha \varphi\|_\infty$ (Theorem 4 of [17]).

Some easy facts on the diffusion approximation

- (1) The diffusion approximation Y^N of (23) is a degenerate diffusion in dimension $n + 1$ (where n is the order of the memory kernel h of (16)), driven by one-dimensional Brownian motion.
- (2) The last n coordinates Y_1^N, \dots, Y_n^N are auxiliary random variables that have to be introduced in order to be able to handle the unbounded memory induced by the Hawkes intensity process. Only the last coordinate is driven by Brownian motion.
- (3) The drift of the diffusion process is of “cascade”-structure : the drift of a coordinate does only depend on the following coordinate and on the coordinate itself (here, we identify as usual $n + 1$ with 0).
- (4) Finally, the drift of the first n coordinates Y_0^N, \dots, Y_{n-1}^N is linear.

It is easy to see that Item (3) implies that the diffusion satisfies the weak Hörmander condition, since the spiking rate function f is strictly positive (Proposition 7 of [17]). Moreover, a Lyapunov type argument implies that the process is recurrent in the sense of Harris, with unique probability measure μ^N , and that it converges towards μ^N at exponential speed (Proposition 5 of [17]). For fixed N , μ^N is of full support but its mass is concentrated around the periodic orbit Γ of the limit system (18). More precisely, we can show that for any open set D with compact closure and smooth boundary,

$$\mu^N(D) \sim Ce^{-[\inf_{x \in D} W(x)]N}. \quad (24)$$

In the above formula, the cost function $W(x)$ is related to the control properties of system (23) (Theorem 3 of Löcherbach (2016) [30]), and $W(x) = 0$ for all $x \in \Gamma$, where Γ is the periodic orbit of (18). In particular, (24) shows how the invariant measure μ^N of the diffusion approximation Y^N concentrates around the periodic orbit Γ . To obtain (24) we rely on the approach of Freidlin and Wentzell [21] to sample path large deviations of diffusions, developed further in Dembo and Zeitouni [14] and extended to the case of degenerate diffusions in Rey-Bellet and Thomas [32]; we refer the interested reader to [30] for more details.

Remark 4.1. (24) shows that for all $\varepsilon > 0$ and for all starting points $x \in \mathbb{R}^{n+1}$, the diffusion Y^N , starting from $Y^N(0) = x$, visits tubes $B_\varepsilon(\Gamma) = \{y \in \mathbb{R}^{n+1} : \text{dist}(y, \Gamma) < \varepsilon\}$ infinitely often, almost surely. Moreover, for large N , the diffusion stays within such tubes around the limit cycle during long periods, before eventually leaving such a tube after a time which is of order $e^{N\bar{V}}$, where \bar{V} is related to the control problem of our diffusion and to the cost of steering the process from the limit cycle to the boundary of the tube around the limit cycle (see [30] for details).

5. CONCLUDING REMARKS

5.1. Delays and Erlang kernels

We have shown in the present paper that Erlang kernels allow to describe certain Hawkes processes via an associated high-dimensional system of Piecewise Deterministic Markov Processes (PDMP's), see (22). The stability behavior of this process can be easily analyzed, giving rise to an alternative approach to the study of the longtime behavior of Hawkes processes. We refer the interested reader to [18] for more details.

Similar ideas have been used in [10] to study rhythmic behavior in mean field Ising models where delays in the interactions are also shown to produce rhythmic behavior.

5.2. Oscillations

The oscillations in the limit system (18) appear from the non-linear “McKean-Vlasov”-type structure of the limit system, i.e. from the fact that the dynamics of the limit system depend on the law of the system. At the single particle level, the dynamics of each neuron do not include any periodic behavior. With this observation, we follow a long tradition, see e.g. Scheutzow (1985) [33, 34] in the framework of nonlinear diffusion processes, or Dai Pra, Fischer and Regoli (2015) [12] and Collet, Dai Pra and Formentin (2015) [9].

5.3. Synaptic plasticity?

How does the synaptic strength (i.e. the factor c within (16)) influence the limit dynamics (18) ? It seems promising to study systems where, in view of modeling *synaptic plasticity*, the synaptic weight c itself is a random process evolving in time.

5.4. External stimuli

The biological motivation for this paper comes from the rhythmic scratch like network activity in the turtle, induced by a mechanical stimulus, and recorded and analyzed by Berg and co-workers [2–4, 29]. Oscillations in a spinal motoneuron are initiated by the sensory input, and continue by some internal mechanisms for some time after the stimulus is terminated.

If we interpret our model as a possible microscopic description of the rhythm-generating motor circuits in the spinal cord, it would be interesting to study how sensible our system is to the fact of being exposed to an external signal. More precisely, starting with an initial configuration submitted to external stimuli and then turning these stimuli off, the question is to understand how and when the system reaches its new invariant distribution (corresponding to the absence of external stimuli).

REFERENCES

- [1] BENAÏM, M., HIRSCH, M.W. Mixed Equilibria and Dynamical Systems arising from Fictitious Play in Perturbed Games. *Games and Econom. Behaviour*, 29 (1999) 36–72.
- [2] BERG, R.W., ALABURDA, A., HOUNSGAARD, J. Balanced Inhibition and Excitation Drive Spike Activity in Spinal Half-Centers. *Science*, 315 (2007) 390–393.
- [3] BERG, R.W., DITLEVSEN, S., HOUNSGAARD, J. Intense Synaptic Activity Enhances Temporal Resolution in Spinal Motoneurons *PLoS ONE*, 3 (2008) e3218.
- [4] BERG, R.W., DITLEVSEN, S. Synaptic inhibition and excitation estimated via the time constant of membrane potential fluctuations. *J Neurophysiol*, 110 (2013) 1021–1034.
- [5] BRÉMAUD, P., MASSOULIÉ, L. Stability of nonlinear Hawkes processes. *The Annals of Probability*, 24(3) (1996) 1563–1588.
- [6] CHEVALLIER, J. Mean-field limit of generalized Hawkes processes. To appear in *Stochastic Processes and their Applications*, 2017.
- [7] CHEVALLIER, J., CACERES, M.J., DOUMIC, M., REYNAUD-BOURET, P. Microscopic approach of a time elapsed neural model. *Math. Mod. & Meth. Appl. Sci.*, 25(14) (2015) 2669–2719.
- [8] CHORNOBOY, E., SCHRAMM, L., KARR, A. Maximum likelihood identification of neural point process systems. *Biological Cybernetics* 59 (1988), 265–275.

- [9] COLLET, F., DAI PRA, P., FORMENTIN, M. Collective periodicity in mean field models of cooperative behavior. *Nonlinear Differential Equations and Applications NoDEA*, 22(5) (2015) 1461–1482.
- [10] COLLET, F., FORMENTIN, M., TOVAZZI, D. Rhythmic behavior in a two-population mean field Ising model. *Phys. Rev. E*, 94, 042139, 2016.
- [11] COMETS, F., FERNÁNDEZ, R., FERRARI, P. A. Processes with long memory: Regenerative construction and perfect simulation. *Ann. of Appl. Probab.*, 12, No 3 (2002) 921–943.
- [12] DAI PRA, P., FISCHER, M., REGOLI, D. A Curie-Weiss Model with dissipation. *J. Stat. Phys.* 152 (2015), 37–53.
- [13] DELATTRE, S., FOURNIER, N., HOFFMANN, M. Hawkes processes on large networks. *Ann. App. Probab.* 26 (2016), 216–261.
- [14] DEMBO, A., ZEITOUNI, O. Large deviation techniques and applications. *Applications of Mathematics, Vol. 38*, Springer 1989.
- [15] DE SANTIS, E., LISSANDRELLI, A. Developments in Perfect Simulation of Gibbs Measures Through a New Result for the Extinction of Galton-Watson-Like Processes. *J. Stat. Phys.* 147 (2012), 231–251.
- [16] DITLEVSEN, S., YIP, K.-P., HOLSTEIN-RATHLOU, N.-H. Parameter estimation in a stochastic model of the tubuloglomerular feedback mechanism in a rat nephron. *Math. Biosci.*, 194 (2005) 49–69.
- [17] DITLEVSEN, S., LÖCHERBACH, E. Multi-class oscillating systems of interacting neurons. *Stoch. Proc. Appl.* 127, 1840-1869, (2017).
- [18] DUARTE, A., LÖCHERBACH, E., OST, G. Stability and perfect simulation of non-linear Hawkes Processes with Erlang kernels. Submitted, arxiv.org/abs/1610.03300, 2016.
- [19] DUARTE, A., GALVES, A., LÖCHERBACH, E., OST, G. Estimating the interaction graph of stochastic neural dynamics. To appear in *Bernoulli*, available on <https://arxiv.org/abs/1604.00419>, 2016.
- [20] FERRARI, P. A., MAASS, A., MARTÍNEZ, S., NEY, P. Cesàro mean distribution of group automata starting from measures with summable decay. *Ergodic Theory Dyn. Syst.*, 20(6), (2000) 1657–1670.
- [21] FREIDLIN, M.I., WENTZELL, A.D. Random perturbations of dynamical systems. *Grundlehren der Mathematischen Wissenschaft 260*, Springer 1984.
- [22] GALVES, A., LÖCHERBACH, E. Modeling networks of spiking neurons as interacting processes with memory of variable length. *Journal de la Société Française de Statistiques* 157 (2016), 17–32.
- [23] GRAHAM, C. Chaoticity for multi-class systems and exchangeability within classes. *J. Appl. Prob.*, 45 (2008) 1196–1203.
- [24] HANSEN, N., REYNAUD-BOURET, P., RIVOIRARD, V. Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli*, 21(1) (2015) 83-143.
- [25] HAWKES, A. G. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika*, 58 (1971) 83-90.
- [26] HAWKES, A. G. AND OAKES, D. A cluster process representation of a self-exciting process. *J. Appl. Prob.*, 11 (1974) 93-503.
- [27] HIRSCH, M.W. Systems of Differential Equations which are Competitive or Cooperative, III: Competing Species. *Nonlinearity*, 1 (1988) 51–71.
- [28] HODARA, P., LÖCHERBACH, E. Hawkes Processes with variable length memory and an infinite number of components. *Adv. Appl. Probab.* 49, 84-107, 2017.
- [29] JAHN, P., BERG, R.W., HOUNSGAARD, J., DITLEVSEN, S. Motoneuron membrane potentials follow a time inhomogeneous jump diffusion process. *J Comput Neurosci*, 31 (2011) 563–579.
- [30] LÖCHERBACH, E. Large deviations for oscillating systems of interacting Hawkes processes in a mean field frame. To appear in *J. Theor. Prob.*, 2017.
- [31] MALLET-PARET, J., SMITH, H.L. The Poincaré-Bendixson Theorem for Monotone Cyclic Feedback Systems. *J. of Dynamics and Diff. Equations* 2, 4 (1990) 367–421.
- [32] REY-BELLE, L., THOMAS, L. E. Asymptotic Behavior of Thermal Nonequilibrium Steady States for a Driven Chain of Anharmonic Oscillators. *Commun. Math. Phys.*, 215 (2000), 1–24.
- [33] SCHEUTZOW, M. Some examples of nonlinear diffusion processes having a time-periodic law. *The Annals of Probability*, 13(2) (1985) 379–384.
- [34] SCHEUTZOW, M. Noise can create periodic behavior and stabilize nonlinear diffusions. *Stochastic Processes Appl.*, 20 (1985) 323–331.
- [35] REYNAUD-BOURET, P., RIVOIRARD, V., GRAMMONT, F., TULEAU-MALOT, C. Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *The Journal of Mathematical Neuroscience (JMN)* 4 (2014), 1–41.
- [36] SKELDON, A.C., PURVEY, I. The effect of different forms for the delay in a model of the nephron. *Math. Biosci. Eng.*, 2(1) (2005) 97–109.
- [37] SZNITMAN, A.-S. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Math.* Springer, Berlin, 1991, 165–251.