Infinite systems of interacting chains with memory of variable length - a stochastic model for biological neural nets

Antonio Galves   Eva Löcherbach

Utrecht
September, 2013

- This is a joint work with Antonio Galves.

- This is a joint work with Antonio Galves.
- It is part of Fapesp Project NeuroMat

- This is a joint work with Antonio Galves.
- It is part of Fapesp Project NeuroMat
- and of the project "Stochastic modeling of cortical plasticity" (Brazil-The Netherlands cooperation Capes-Nuffic, coordinated by R. Fernández).

- This is a joint work with Antonio Galves.
- It is part of Fapesp Project NeuroMat
- and of the project "Stochastic modeling of cortical plasticity" (Brazil-The Netherlands cooperation Capes-Nuffic, coordinated by R. Fernández).
- Journal of Statistical Physics, June 2013

- This is a joint work with Antonio Galves.
- It is part of Fapesp Project NeuroMat
- and of the project "Stochastic modeling of cortical plasticity" (Brazil-The Netherlands cooperation Capes-Nuffic, coordinated by R. Fernández).
- Journal of Statistical Physics, June 2013
- *This work is dedicated to Errico Presutti.*

Spike trains

- Neurons "talk to each other" by firing sequences of action potentials.

Spike trains

- Neurons "talk to each other" by firing sequences of action potentials.
- One emission of such an action potential is called **spike**.

Spike trains

- Neurons "talk to each other" by firing sequences of action potentials.
- One emission of such an action potential is called **spike**.
- Duration of spikes is very short (about 1 ms) - so : report if in a given time interval (of about 3 ms) there is **presence or absence of spike.**

Spike trains

- Neurons "talk to each other" by firing sequences of action potentials.
- One emission of such an action potential is called **spike**.
- Duration of spikes is very short (about 1 ms) - so : report if in a given time interval (of about 3 ms) there is **presence or absence of spike.**
- If we report for any neuron the discrete times of appearance of a spike $\rightarrow$ spike trains.
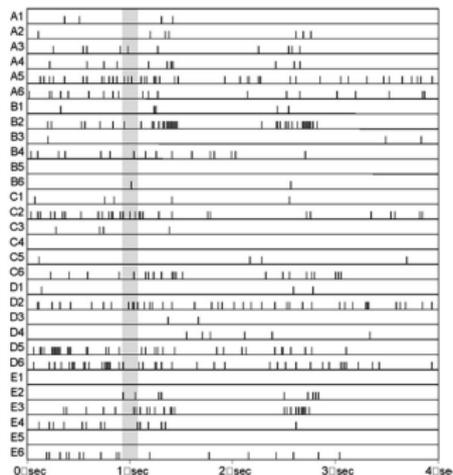
Spike trains

FIGURE: Spike trains of several neurons - Picture by W. Maass

Important - and open - questions

- How is information encoded in such patterns?
- How can we see an external stimulus?
- How to explain the appearance of synchronized spiking patterns ($\rightarrow$ evoked potential)?
- Are successive interspike intervalls independent?

Important - and open - questions

- How is information encoded in such patterns ?
- How can we see an external stimulus ?
- How to explain the appearance of synchronized spiking patterns ($\rightarrow$ evoked potential) ?
- Are successive interspike intervalls independent ?

Try to find a model in which we will be able to answer such kinds of questions. The model we present is partly inspired by Bruno Cessac (2011).

The model

- Huge system with $N \approx 10^{11}$ neurons that interact.

- Spike train : for each neuron i we indicate if there is a spike or not at time t, $t \in \mathbb{Z}$.

  $X_t(i) \in \{0, 1\}, X_t(i) = 1 \Leftrightarrow$ neuron i has a spike at time t .

- t is an index of the time window in which we observe the neuron. In the data we considered, the width of this window is typically 3 ms.

Background

• The membrane potential process of one neuron accumulates the stimulus coming from the other neurons. It spikes depending on the height of the accumulated potential.

• Then : reset to a resting potential. Restart accumulating potentials coming from other neurons.

Background

• The membrane potential process of one neuron accumulates the stimulus coming from the other neurons. It spikes depending on the height of the accumulated potential.

• Then : reset to a resting potential. Restart accumulating potentials coming from other neurons.

• Hence : **Variable length memory :** the memory of the neuron goes back up to its last spike – at least at a first glance.

Background

• The membrane potential process of one neuron accumulates the stimulus coming from the other neurons. It spikes depending on the height of the accumulated potential.

• Then : reset to a resting potential. Restart accumulating potentials coming from other neurons.

• Hence : **Variable length memory :** the memory of the neuron goes back up to its last spike – at least at a first glance.

• This is the framework considered by Cessac (2011) - but only for a **finite** number of neurons.

The model

Chain $X_t \in \{0,1\}^{\mathcal{I}}$,

$$X_t = (X_t(i), i \in \mathcal{I}), X_t(i) \in \{0,1\}, t \in \mathbb{Z},$$

$\mathcal{I}$ countable is the set of neurons. **We will work in the case where $\mathcal{I}$ is infinite**.

Time evolution : At each time step, neurons update independently from each other :

The model

Chain $X_t \in \{0,1\}^{\mathcal{I}}$,

$$X_t = (X_t(i), i \in \mathcal{I}), X_t(i) \in \{0,1\}, t \in \mathbb{Z},$$

$\mathcal{I}$ countable is the set of neurons. **We will work in the case where $\mathcal{I}$ is infinite**.

Time evolution : At each time step, neurons update independently from each other : For any finite subset J of neurons,

$$P(X_t(i) = a_i, i \in J | \mathcal{F}_{t-1}) = \prod_{i \in J} P(X_t(i) = a_i | \mathcal{F}_{t-1}),$$

where

$\mathcal{F}_{t-1}$ is the past history up to time $t-1$.

The model II

$$P(X_t(i) = 1 | \mathcal{F}_{t-1}) = \Phi \left(\sum_j W_{j \to i} \sum_{s=L_t^i}^{t-1} g(t-s) X_s(j), t - L_t^i \right).$$

Here :

- $W_{j \to i} \in \mathbb{R}$: **synaptic weight** of neuron j on i.
- $L_t^i = \sup\{s < t : X_s(i) = 1\}$ last spike strictly before time t in neuron i.
- $g : \mathbb{N} \to \mathbb{R}_+$ describes an aging effect. If there is no aging, then $g \equiv 1$.

Excitatory versus inhibitory influence

Neurons who have a direct influence on i are those belonging to

$$\mathcal{V}_{\cdot \to i} := \{j : W_{j \to i} \neq 0\} :$$

Excitatory versus inhibitory influence

Neurons who have a direct influence on i are those belonging to

$$\mathcal{V}_{\cdot \to i} := \{ j : W_{j \to i} \neq 0 \} :$$

Either excitatory : $W_{j \to i} > 0$.
Or inhibitory : $W_{j \to i} < 0$.

This is a new class of non Markovian processes having a countable number of interacting components.

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems (which are Markovian).

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems (which are Markovian).
- It also extends Rissanen's stochastic chains with memory of variable length

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems (which are Markovian).
- It also extends Rissanen's stochastic chains with memory of variable length (it is only *locally* of variable length).

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems (which are Markovian).
- It also extends Rissanen's stochastic chains with memory of variable length (it is only *locally* of variable length).
- It is a chain of infinite order with a non countable state space.

This is a new class of non Markovian processes having a countable number of interacting components.

- It extends in a non trivial way Spitzer's interacting particle systems (which are Markovian).
- It also extends Rissanen's stochastic chains with memory of variable length (it is only *locally* of variable length).
- It is a chain of infinite order with a non countable state space.

So it is an interesting mathematical object....

- The discrete time frame is not important – a continuous time description is analogous.

- The discrete time frame is not important – a continuous time description is analogous.
- Our model is a version in discrete time of the so-called **Hawkes process** (see Brémaud& Massoulié 1991) – but : with an infinity of components and, locally, a structure of variable memory.

Basic mathematical questions

• Given $(W_{i \to j})$, Φ and g, **does a chain with the above dynamics exist and if so, is it unique ?**

Basic mathematical questions

• Given $(W_{i \to j})$, Φ and g, **does a chain with the above dynamics exist and if so, is it unique ?**

Yes - under some conditions, see in two minutes.

Basic mathematical questions

• Given $(W_{i \to j})$, Φ and g, **does a chain with the above dynamics exist and if so, is it unique ?**

Yes - under some conditions, see in two minutes.

• Are neighboring inter-spike intervals correlated ?

This is both a mathematical and a biological question,

Basic mathematical questions

• Given $(W_{i \to j})$, Φ and g, **does a chain with the above dynamics exist and if so, is it unique ?**

Yes - under some conditions, see in two minutes.

• Are neighboring inter-spike intervals correlated ?

This is both a mathematical and a biological question,

and there are experimental facts that we have to explain. . .

The proof of existence and uniqueness is based on the study of the transition probability

$$(1) \quad p_{(i,t)}(1|x) = \Phi \left(\sum_{j \neq i} W_{j \rightarrow i} \sum_{s=L_t^i(x)}^{t-1} g(t-s)x_s(j), t - L_t^i(x) \right) :$$

which depends on the space-time configuration of spike times

$x_{L_t^i}^{t-1}(\mathcal{V}_{\cdot \rightarrow i})$: locally variable length in time, infinite range in space.

Globally of **infinite range memory** !

But attention : The function $x \mapsto p_{(i,t)}(1|x)$ is not continuous !
We do not have :

$$\sup_{x,x':x\overset{k}{=}x'} |p_{(i,t)}(1|x) - p_{(i,t)}(1|x')| \to 0$$

as $k \to \infty$.

But attention : The function $x \mapsto p_{(i,t)}(1|x)$ is not continuous !
We do not have :

$$\sup_{x,x':x\overset{k}{=}x'} |p_{(i,t)}(1|x) - p_{(i,t)}(1|x')| \to 0$$

as $k \to \infty$. Since no summability is imposed on g.

But attention : The function $x \mapsto p_{(i,t)}(1|x)$ is not continuous !
We do not have :

$$\sup_{x,x':x\overset{k}{=}x'} |p_{(i,t)}(1|x) - p_{(i,t)}(1|x')| \to 0$$

as $k \to \infty$. Since no summability is imposed on g.

Continuity is usully what is required in the study of **chains having infinite order** (see work by R. Fernández, G. Maillard, ...)

Hypotheses

1) Lipschitz : There exists some $\gamma > 0$: such that for all z, z', n,

$$|\Phi(z, n) - \Phi(z', n)| \leq \gamma|z - z'|.$$

2) Uniform summability of the synaptic weights

$$\sup_i \sum_j |W_{j \to i}| < \infty.$$

3) Spontaneous spiking activity with intensity δ :

$$\Phi(\cdot, \cdot) \geq \delta > 0.$$

Theorem

Under the above hypotheses : If $\delta \geq \delta_*$ and : fast decay of synaptic weigths, then

1. *there exists a unique stationary chain $X_t(i), t \in \mathbb{Z}, i \in \mathcal{I}$, consistent with the dynamics.*

Theorem

Under the above hypotheses : If $\delta \geq \delta_*$ and : fast decay of synaptic weigths, then

1. *there exists a unique stationary chain $X_t(i), t \in \mathbb{Z}, i \in \mathcal{I}$, consistent with the dynamics.*

2. *the speed of convergence to equilibrium is bounded above :*

   $$(2) \quad |E[f(X_s^t(i))|\mathcal{F}_0] - E[f(X_s^t(i))]| \leq C(t-s+1)\|f\|_\infty \varphi(s),$$

   where $\varphi(s) \downarrow 0$ as $s \to \infty$.

Theorem

Under the above hypotheses : If $\delta \geq \delta_*$ and : fast decay of synaptic weigths, then

1. *there exists a unique stationary chain $X_t(i), t \in \mathbb{Z}, i \in \mathcal{I}$, consistent with the dynamics.*

2. *the speed of convergence to equilibrium is bounded above :*

   $$(2) \quad |E[f(X_s^t(i))|\mathcal{F}_0] - E[f(X_s^t(i))]| \leq C(t - s + 1)\|f\|_\infty \varphi(s),$$

   where $\varphi(s) \downarrow 0$ as $s \to \infty$.

3. *If moreover*

   $$g(n) < Ce^{-\beta n},$$

   then we have in (2) that $\varphi(s) \leq C\varrho^s$ for some $\varrho \in]0, 1[$, if $\beta >> 1$.

Proof : Conditional Kalikow-decomposition

- $\Phi(\cdot, \cdot) \geq \delta \Rightarrow$ **Coupling** with i.i.d. field

$$\xi_t(i), t \in \mathbb{Z}, i \in \mathcal{I}, \xi_i(t) \sim \mathcal{B}(\delta) :$$

$X_t(i) \geq \xi_t(i)$ for all t, i.

- We have to work in the configuration space conditioned on the realization of ξ :

$$\mathcal{S}^\xi = \{x \in \{0,1\}^{\mathbb{Z} \times \mathcal{I}} : x_t(i) \geq \xi_t(i) \forall t, \forall i\}.$$

- Now : We **have** continuity of $x \mapsto p_{(i,t)}(a|x)$ in restriction to $\mathcal{S}^\xi$.

Continuation of the proof

Each site (i, t) has its memory bounded by

$$R_t^i = \sup\{s < t : \xi_s(i) = 1\}.$$

Introduce : $V_i(0) := \{i\}, \ V_i(k) \uparrow \mathcal{V}_i = \{j : W_{j \to i} \neq 0\} \cup \{i\}.$

Proposition

$$p_{(i,t)}(a|x) = \lambda(-1)p^{[-1]}(a) + \sum_{k \geq 0} \lambda(k)p^{[k]}(a|x_{R_t^i}^{t-1}(V_i(k))),$$

where $\lambda(k) \in [0,1], \sum \lambda(k) = 1,$

$$\lambda(k) \leq 2\gamma \sum_{s=R_t^i}^{t-1} g(t - s) \sum_{j \notin V_i(k-1)} |W_{j \to i}|, \ k \geq 1.$$

Comments

- This is a conditional decomposition, conditional on the realization of spontaneous spikes.
- The "reproduction probabilities" $\lambda(k)$ are random variables depending on ξ.
- We get uniqueness via a "dual process", the **Clan of Ancestors :** in order to decide about the value of (i, t), we have to know the values of all sites in

$$C^1_{(i,t)} = V_i(k) \times [R^i_t, t - 1], \text{ chosen with probability } \lambda(k)....$$

Comments

- This is a conditional decomposition, conditional on the realization of spontaneous spikes.
- The "reproduction probabilities" $\lambda(k)$ are random variables depending on ξ.
- We get uniqueness via a "dual process", the **Clan of Ancestors :** in order to decide about the value of (i, t), we have to know the values of all sites in

$$C^1_{(i,t)} = V_i(k) \times [R^i_t, t-1], \text{ chosen with probability } \lambda(k)....$$

Iterate ! If this process stops in finite time a.s., then we are done.

Comments

- This is a conditional decomposition, conditional on the realization of spontaneous spikes.
- The "reproduction probabilities" $\lambda(k)$ are random variables depending on ξ.
- We get uniqueness via a "dual process", the **Clan of Ancestors :** in order to decide about the value of (i, t), we have to know the values of all sites in

  $$C^1_{(i,t)} = V_i(k) \times [R^i_t, t-1], \quad \text{chosen with probability } \lambda(k)....$$

  Iterate ! If this process stops in finite time a.s., then we are done. This is granted by a comparison with a multi-type branching process in random environment.

Back to neuroscience

Goldberg et al. (1964) in their article "Response of neurons of the superior olivary complex of the cat to acoustic stimuli of long duration" observe :

In many experimental setups the empirical correlation between successive inter-spike intervals is very small –

Back to neuroscience

Goldberg et al. (1964) in their article "Response of neurons of the superior olivary complex of the cat to acoustic stimuli of long duration" observe :

In many experimental setups the empirical correlation between successive inter-spike intervals is very small –

"indicating that a description of spiking as a **stationary renewal process** is a good approximation" (Gerstner and Kistler 2002).

In the same direction :

The statistical analysis of the activity of several (but not all !) neurons in the hippocampus selects as best model a

renewal process.

− *Data registered by Sidarta Ribeiro (Brain Institute UFRN), in 2005.*

− *Data analyzed by Karina Y. Yaginuma, using the SMC (smallest maximiser criterion).*

HOWEVER :

Nawrot et al. (2007) in their article "Serial interval statistics of spontaneous activity in cortical neurons in vivo and in vitro"

find statistical evidence that **neighboring inter-spike intervals are correlated**, having negative correlation ! ! !

HOWEVER :

Nawrot et al. (2007) in their article "Serial interval statistics of spontaneous activity in cortical neurons in vivo and in vitro"

find statistical evidence that **neighboring inter-spike intervals are correlated**, having negative correlation ! ! !

Can we account for these apparently contradictory facts with our model ?

Random synaptic weights

We must describe in a more precise way the **directed graph defined by the synaptic weights :**

Random synaptic weights

We must describe in a more precise way the **directed graph defined by the synaptic weights :**

Vertices = neurons.

Random synaptic weights

We must describe in a more precise way the **directed graph defined by the synaptic weights :**

Vertices $=$ neurons.

There is a directed edge from neuron i to neuron j iff $W_{i \to j} \neq 0$.

Random synaptic weights

We must describe in a more precise way the **directed graph defined by the synaptic weights :**

Vertices = neurons.

There is a directed edge from neuron i to neuron j iff $W_{i \to j} \neq 0$.

In what follows this graph will be a realization of a critical directed Erdös-Rényi graph. In such a graph there is a unique giant cluster, and we work in this giant cluster.

Critical directed Erdös-Rényi random graph

- Large but finite system of neurons with
  $\mathcal{I} = \{1, \ldots, N\}, N \approx 10^{11}$.

Critical directed Erdös-Rényi random graph

- Large but finite system of neurons with
  $\mathcal{I} = \{1, \ldots, N\}, N \approx 10^{11}$.
- Random synaptic weights : $W_{i \to j}, i \neq j$, are independent
  random variables taking values 0 or 1

Critical directed Erdös-Rényi random graph

- Large but finite system of neurons with
  $\mathcal{I} = \{1, \ldots, N\}, N \approx 10^{11}$.
- Random synaptic weights : $W_{i \to j}, i \neq j$, are independent
  random variables taking values 0 or 1
  with $P(W_{i \to j} = 1) = 1 - P(W_{i \to j} = 0) = p$.

Critical directed Erdös-Rényi random graph

- Large but finite system of neurons with
  $\mathcal{I} = \{1, \ldots, N\}, N \approx 10^{11}$.
- Random synaptic weights : $W_{i \to j}, i \neq j$, are independent
  random variables taking values 0 or 1
  with $P(W_{i \to j} = 1) = 1 - P(W_{i \to j} = 0) = p$.
- Here, $p = \lambda/N$ and $\lambda = 1 + \vartheta/N, \vartheta > 0$.

Critical directed Erdös-Rényi random graph

- Large but finite system of neurons with
  $\mathcal{I} = \{1, \ldots, N\}, N \approx 10^{11}$.
- Random synaptic weights : $W_{i \to j}, i \neq j$, are independent
  random variables taking values 0 or 1
  with $P(W_{i \to j} = 1) = 1 - P(W_{i \to j} = 0) = p$.
- Here, $p = \lambda/N$ and $\lambda = 1 + \vartheta/N$, $\vartheta > 0$.
- Observe that $W_{i \to j}$ and $W_{j \to i}$ are distinct and independent :
  being influenced by neuron i is different from influencing
  neuron i....

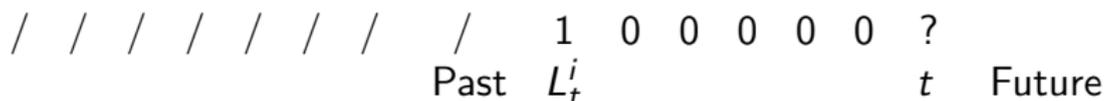# Does the past before the last spike of a neuron influence the future ?

$$/ \quad / \quad / \quad / \quad / \quad / \quad / \quad \quad / \quad \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad ?$$
$$\text{Past} \quad L_t^i \qquad\qquad\qquad\qquad t \quad \text{Future}$$

Does the past before the last spike of a neuron influence the future ?

$$/ \quad / \quad / \quad / \quad / \quad / \quad / \quad / \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad ?$$
$$\text{Past} \quad L_t^i \qquad\qquad\qquad t \quad \text{Future}$$

Does it affect the future whether the last spike before L_t^i took place immediately before L_t^i or whether it took place many steps before ?

Does the past before the last spike of a neuron influence the future ?

$$/ \ / \ / \ / \ / \ / \ / \ \quad / \quad \underset{\text{Past} \ L_t^i}{1} \ 0 \ 0 \ 0 \ 0 \ 0 \ \underset{t}{?} \quad \text{Future}$$

Does it affect the future whether the last spike before L_t^i took place immediately before L_t^i or whether it took place many steps before ?

The point is : the last spike of neuron i before time L_t^i affects many neurons – different from i, which in turn affect other neurons and so on. How long does it take until this influence returns to the starting neuron i ?

This time is a sort of *recurrence time* in the random graph :

$$C_1^i = \{j : W_{j \to i} \neq 0\}, \ldots, C_n^i = \{j : \exists k \in C_{n-1}^i : W_{j \to k} \neq 0\}.$$

Then the recurrence time is

$$T_i = \inf\{n : i \in C_n^i\}.$$

Proposition

$$P(\text{recurrence time} \leq k) \leq \frac{k}{N} e^{\vartheta k / N}.$$

$N =$ number of neurons.

$\vartheta =$ parameter appearing in the definition of the synaptic weight probabilities, $Np = 1 + \vartheta/N$.

This implies

Theorem

On a "good set" of random synaptic weights :

$$|Covariance\ of\ neighboring\ inter\text{-}spike\ intervals| \leq C\frac{1}{\delta^2}N(1-\delta)^{\sqrt{N}}.$$

Moreover,

$$P(good\ set) \geq 1 - CN^{-1/2},$$

where δ is the spontaneous spiking activity.

This implies

Theorem

On a "good set" of random synaptic weights :

$$|Covariance\ of\ neighboring\ inter\text{-}spike\ intervals| \leq C\frac{1}{\delta^2}N(1-\delta)^{\sqrt{N}}.$$

Moreover,

$$P(good\ set) \geq 1 - CN^{-1/2},$$

where δ is the spontaneous spiking activity.

This conciliates the empirical results both of Goldberg et al. (1964) and of Nawrot et al. (2007) !

Thanks for your attention !