

# Gibbs distribution: from neural network dynamics to spike train statistics estimation

Bruno Cessac

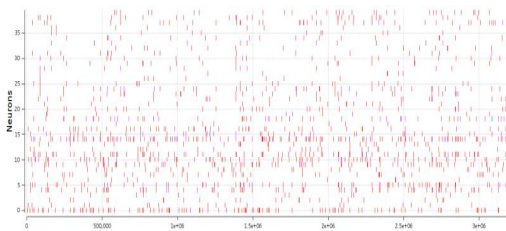
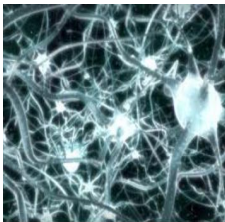
Biovision Team, INRIA Sophia Antipolis, France.

12-09-2017

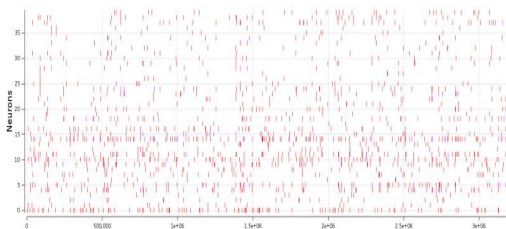
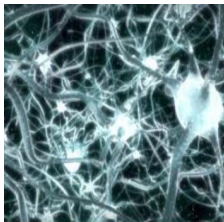
- 1 Introduction
- 2 Mathematical setting
- 3 Gibbs distributions
- 4 Examples
- 5 Response to stimuli
- 6 On the mathematical consequences of binning
- 7 Conclusion

# Introduction

# Spike trains



# Spike trains



Local dynamics

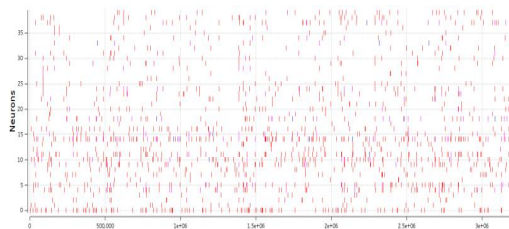
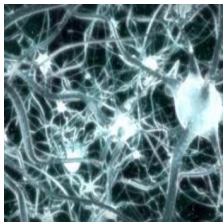
Network interactions

Stimulus

⇒

*specific spatio-temporal correlations*

# Spike trains



Local dynamics

Network interactions

Stimulus



*specific spatio-temporal correlations*

# A challenge: "Reading" "the" neural "code"

# A challenge: "Reading" "the" neural "code"

- **Fundamental challenge:** Understanding how the nervous system "encodes" / "decodes" "information"



# A challenge: "Reading" "the" neural "code"

- **Fundamental challenge:** Understanding how the nervous system "encodes" / "decodes" "information"
- **Experimental challenge:** Elaborate better techniques to record neuronal activity

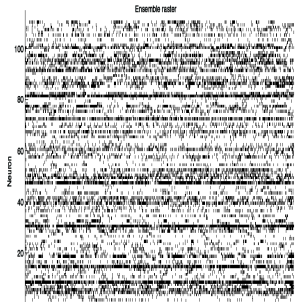
# A challenge: "Reading" "the" neural "code"

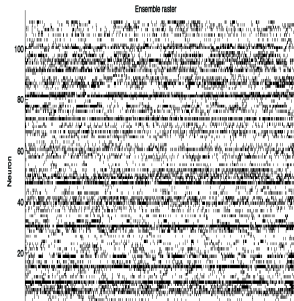
- **Fundamental challenge:** Understanding how the nervous system "encodes" / "decodes" "information"
- **Experimental challenge:** Elaborate better techniques to record neuronal activity
- **Mathematical challenge:** Improve existing mathematical tools / Elaborate new ones

# A challenge: "Reading" "the" neural "code"

- **Fundamental challenge:** Understanding how the nervous system "encodes" / "decodes" "information"
- **Experimental challenge:** Elaborate better techniques to record neuronal activity
- **Mathematical challenge:** Improve existing mathematical tools / Elaborate new ones
- **Technological challenge:**
  - Neural stimulation (prostheses...)
  - Bio-inspired technologies
  - Computer vision
  - ....

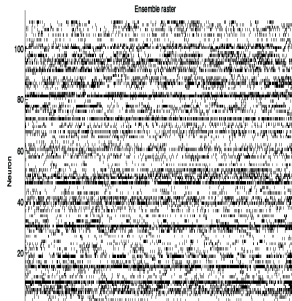
## Mathematical setting





This spike train has been generated by an hidden *dynamics / stochastic process.*

*Can we infer this process from the spike train's analysis ?*



## What do we learn about:

- the stimulus;
- the underlying network;
- the response of that network to a stimulus.

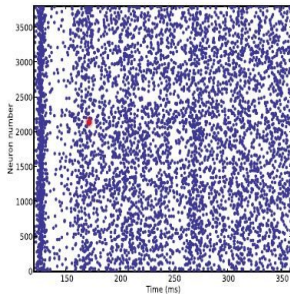


Figure: Spike state.

Spike state

$$\omega_k(n) \in \{0, 1\}$$



# Spike events

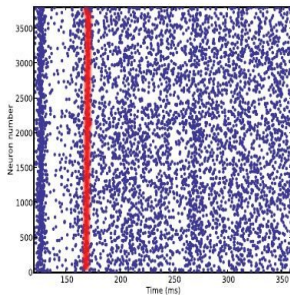


Figure: Spike pattern.

Spike state

$$\omega_k(n) \in \{0, 1\}$$

Spike pattern

$$\omega(n) = (\omega_k(n))_{k=1}^N$$

# Spike events

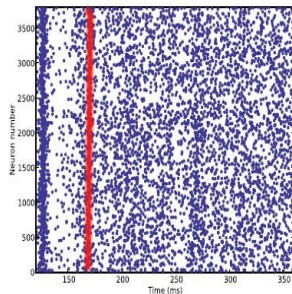


Figure: Spike pattern.

## Spike state

$$\omega_k(n) \in \{0, 1\}$$

## Spike pattern

$$\omega(n) = (\omega_k(n))_{k=1}^N \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

# Spike events

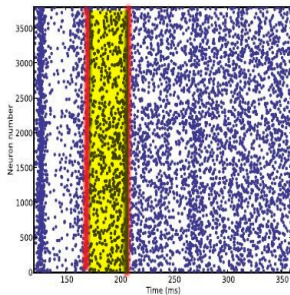


Figure: Spike block.

## Spike state

$$\omega_k(n) \in \{0, 1\}$$

## Spike pattern

$$\omega(n) = (\omega_k(n))_{k=1}^N$$

## Spike block

$$\omega_m^n = \{\omega(m)\omega(m+1)\dots\omega(n)\}$$

# Spike events

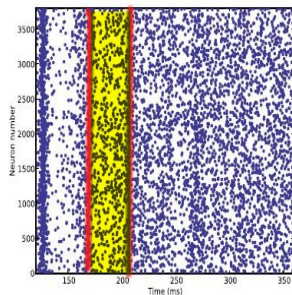


Figure: Spike block.

## Spike state

$$\omega_k(n) \in \{0, 1\}$$

## Spike pattern

$$\omega(n) = (\omega_k(n))_{k=1}^N$$

## Spike block

$$\omega_m^n = \{\omega(m)\omega(m+1)\dots\omega(n)\}$$
$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

# Spike events

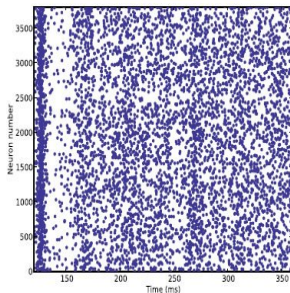


Figure: Raster plot/Spike train.

## Spike state

$$\omega_k(n) \in \{0, 1\}$$

## Spike pattern

$$\omega(n) = (\omega_k(n))_{k=1}^N$$

## Spike block

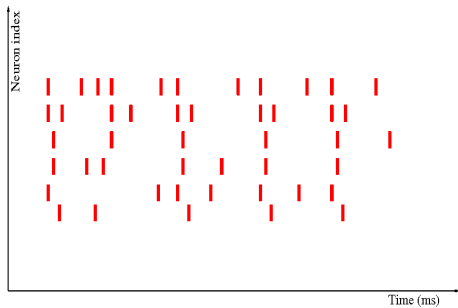
$$\omega_m^n = \{\omega(m)\omega(m+1)\dots\omega(n)\}$$

## Raster plot

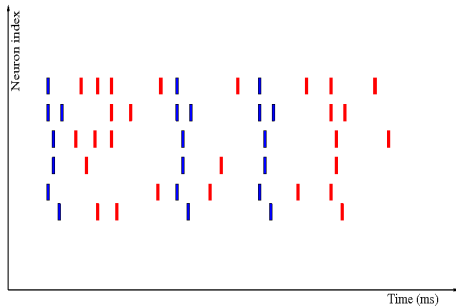
$$\omega \stackrel{\text{def}}{=} \omega_0^T$$

# Handling temporality and memory

# Handling temporality and memory

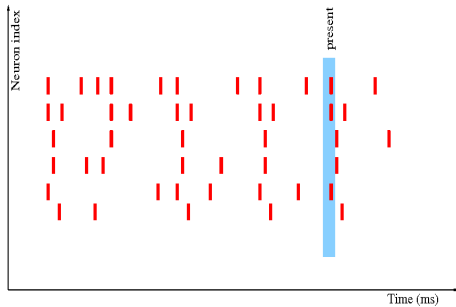


# Handling temporality and memory

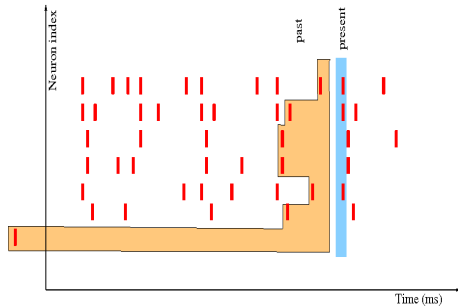




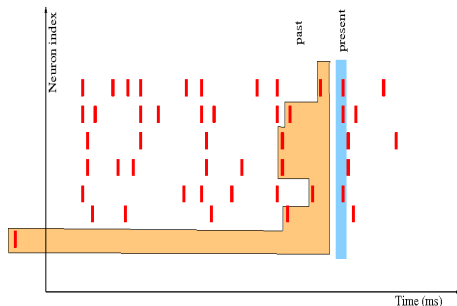
# Handling temporality and memory



# Handling temporality and memory



# Handling temporality and memory



## Probability of characteristic spatio-temporal patterns

The probability of a spike pattern ....  
depends on the **network history** (transition probabilities).

Estimate:

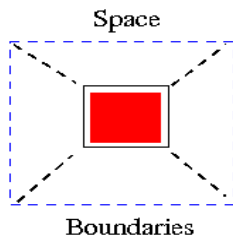
$$P_n [\omega(n) \mid \omega_{n-D}^{n-1}]$$

where  $D$  can be

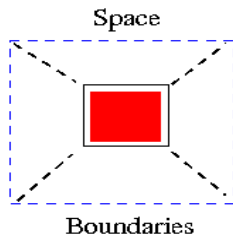
- Finite and constant (Markov chain).
- Finite and variable (Variable length Markov chain).
- Infinite (Chain with complete connection).

## Gibbs distributions

# Gibbs distribution in the space domain

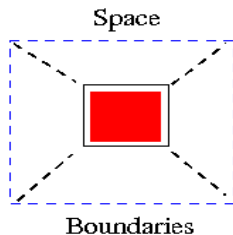


# Gibbs distribution in the space domain



$$\forall \Lambda \subset \mathbb{Z}^d, \mu(\{S\} \mid \partial\Lambda) = \frac{1}{Z_{\Lambda, \partial\Lambda}} e^{-\beta H_{\Lambda, \partial\Lambda}(\{S\})}$$

# Gibbs distribution in the space domain



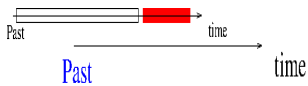
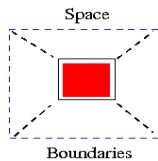
$$\forall \Lambda \subset \mathbb{Z}^d, \mu(\{S\} \mid \partial\Lambda) = \frac{1}{Z_{\Lambda, \partial\Lambda}} e^{-\beta H_{\Lambda, \partial\Lambda}(\{S\})}$$

$$f_{\partial\Lambda}(\beta) = \lim_{\Lambda \uparrow \infty} \frac{1}{|\Lambda|} \log Z_{\Lambda, \partial\Lambda}$$

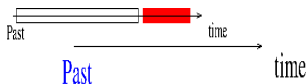
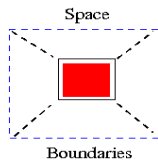
(free energy density)



# Gibbs distribution

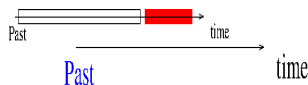
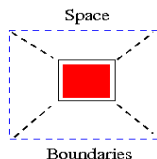


# Gibbs distribution



$$\mu[\omega_m^n] = \prod_{l=m+D}^n P_l[\omega(l) \mid \omega_{l-D}^{l-1}] \mu[\omega_m^{m+D-1}], \quad \forall m < n \in \mathbb{Z}$$

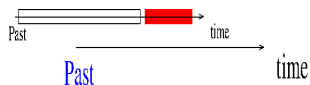
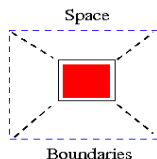
# Gibbs distribution



$$\mu[\omega_m^n] = \prod_{l=m+D}^n P_l[\omega(l) \mid \omega_{l-D}^{l-1}] \mu[\omega_m^{m+D-1}], \quad \forall m < n \in \mathbb{Z}$$

$$\phi_l(\omega_{l-D}^l) = \log P_l[\omega(l) \mid \omega_{l-D}^{l-1}] \quad (\text{normalized potential})$$

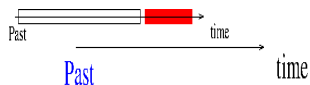
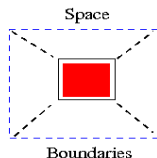
# Gibbs distribution



$$\mu[\omega_m^n] = \exp \sum_{l=m+D}^n \phi_l(\omega_{l-D}^l) \mu[\omega_m^{m+D-1}], \quad \forall m < n \in \mathbb{Z}$$

$$\phi_l(\omega_{l-D}^l) = \log P_l[\omega(l) \mid \omega_{l-D}^{l-1}] \quad (\text{normalized potential})$$

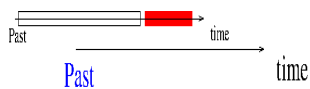
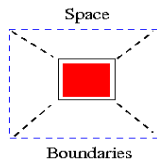
# Gibbs distribution



$$\mu \left[ \omega_m^n \mid \omega_m^{m+D-1} \right] = \exp \sum_{l=m+D}^n \phi_l \left( \omega_{l-D}^l \right), \quad \forall m < n \in \mathbb{Z}$$

$$\phi_l \left( \omega_{l-D}^l \right) = \log P_l \left[ \omega(l) \mid \omega_{l-D}^{l-1} \right] \quad (\text{normalized potential})$$

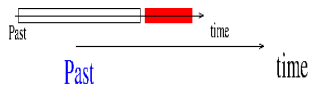
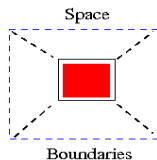
# Gibbs distribution



$$\mu \left[ \omega_m^n \mid \omega_m^{m+D-1} \right] = \exp \sum_{l=m+D}^n \phi_l \left( \omega_{l-D}^l \right), \quad \forall m < n \in \mathbb{Z}$$

$$\forall \Lambda \subset \mathbb{Z}^d, \mu(\{S\} \mid \partial \Lambda) = \frac{1}{Z_{\Lambda, \partial \Lambda}} e^{-\beta H_{\Lambda, \partial \Lambda}(\{S\})}$$

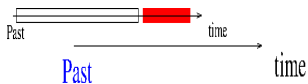
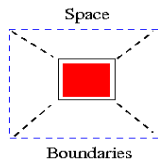
# Gibbs distribution



$$\mu \left[ \omega_m^n \mid \omega_m^{m+D-1} \right] \sim \exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right) \exp -(n-m)P(H), \quad \forall m < n$$

$$A < \frac{\mu \left[ \omega_m^n \right]}{\exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right) \exp -(n-m)P(H)} < B, \quad \forall m < n \in \mathbb{Z}$$

# Gibbs distribution

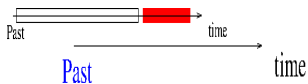
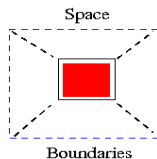


$$\mu \left[ \omega_m^n \mid \omega_m^{m+D-1} \right] \sim \exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right) \exp - (n - m) P(H), \quad \forall m < n$$

$$P(H) = \limsup_{m \rightarrow -\infty} \frac{1}{n - m} \log \sum_{\omega_m^n} \exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right)$$



# Gibbs distribution



$$\mu \left[ \omega_m^n \mid \omega_m^{m+D-1} \right] \sim \exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right) \exp -(n-m)P(H), \quad \forall m < n$$

$$P(H) = \limsup_{m \rightarrow -\infty} \frac{1}{n-m} \log \sum_{\omega_m^n} \exp \sum_{l=m+D}^n H \left( \omega_{l-D}^l \right)$$

$$f_{\partial \Lambda}(\beta) = \lim_{\Lambda \uparrow \infty} \frac{1}{|\Lambda|} \log Z_{\Lambda, \partial \Lambda}$$

A Gibbs distribution is a probability measure  $\mu$  such that:

- (i) for all  $n \in \mathbb{Z}$  and all  $\mathcal{F}_{\leq n}$ -measurable functions  $f$ :

$$\int f(\omega_{-\infty}^n) \mu(d\omega) = \int \sum_{\omega(n) \in \mathcal{A}} f(\omega_{-\infty}^{n-1} \omega(n)) P_n[\omega(n) | \omega_{-\infty}^{n-1}] \mu(d\omega)$$

- (ii)  $\forall n \in \mathbb{Z}, \forall \omega_{-\infty}^{n-1} \in \mathcal{A}_{-\infty}^{n-1}, P_n[\omega(n) | \omega_{-\infty}^{n-1}] > 0$ .
- (iii) For each  $n \in \mathbb{Z}$ ,  $P_n[\omega(n) | \omega_{-\infty}^{n-1}]$  is continuous with respect to  $\omega$ .

## Examples

# Example 1: Maximum Entropy models

- Assume data is **stationary**.

## Example 1: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events e.g. "neuron  $i_1$  is firing at time  $t_1$ , ... neuron  $i_k$  is firing at time  $t_k$ "

$$m_I(\omega) = \omega_{i_1}(t_1) \dots \omega_{i_k}(t_k) \quad (\text{"spikes interaction"}).$$

## Example 1: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events e.g. "neuron  $i_1$  is firing at time  $t_1$ , ... neuron  $i_k$  is firing at time  $t_k$ "

$$m_I(\omega) = \omega_{i_1}(t_1) \dots \omega_{i_k}(t_k) \quad (\text{"spikes interaction"}).$$

- Compute  $\pi_I$ : empirical probability of spike event  $I$ .

# Example 1: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events e.g. "neuron  $i_1$  is firing at time  $t_1$ , ... neuron  $i_k$  is firing at time  $t_k$ "

$$m_I(\omega) = \omega_{i_1}(t_1) \dots \omega_{i_k}(t_k) \quad (\text{"spikes interaction"}).$$

- Compute  $\pi_I$ : empirical probability of spike event  $I$ .
- Find a probability distribution  $\mu$  which *reproduces the empirical averages*:

$$\mu[\text{Spike event } I] = \pi_I \quad (C)$$

without any further assumption.

*Maximize the entropy under constraints  $C$ .*

## Example 1: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events e.g. "neuron  $i_1$  is firing at time  $t_1$ , ... neuron  $i_k$  is firing at time  $t_k$ "

$$m_I(\omega) = \omega_{i_1}(t_1) \dots \omega_{i_k}(t_k) \quad (\text{"spikes interaction"}).$$

- Compute  $\pi_I$ : empirical probability of spike event  $I$ .
- Find a probability distribution  $\mu$  which *reproduces the empirical averages*:

$$\mu[\text{Spike event } I] = \pi_I \quad (C)$$

without any further assumption.

*Maximize the entropy under constraints  $C$ .*

- In good cases this probability exists and is unique. *Gibbs distribution.*



## Example 1: Maximum Entropy models

$$\mathcal{H} = \sum_I h_I m_I$$

Most authors consider spatial events (no memory) because it fits with the usual definition of Gibbs distributions  $\frac{e^{\beta H}}{Z}$ .

## Example 1: Maximum Entropy models

$$\mathcal{H} = \sum_I h_I m_I$$

Most authors consider spatial events (no memory) because it fits with the usual definition of Gibbs distributions  $\frac{e^{\beta H}}{Z}$ .

Extension to spatio-temporal events (memory) can be achieved to the price of a bit more mathematical complexity (Vasquez et al, 2012; Nasser-Cessac, 2014)

## Example 1: Maximum Entropy models

$$\mathcal{H} = \sum_I h_I m_I$$

Most authors consider spatial events (no memory) because it fits with the usual definition of Gibbs distributions  $\frac{e^{\beta H}}{Z}$ .

Extension to spatio-temporal events (memory) can be achieved to the price of a bit more mathematical complexity (Vasquez et al, 2012; Nasser-Cessac, 2014)

Gibbs distribution extends to (i) non stationarity; (ii) infinite memory.

# Example 2: The Generalized-Linear Model (GLM)

Paradigms of rates and receptive fields.

$$\lambda_k(t|H_t^{hist}) = f(b_k + (K_k * x)(t) + \sum_j (h_{kj} * y_j)(t))$$

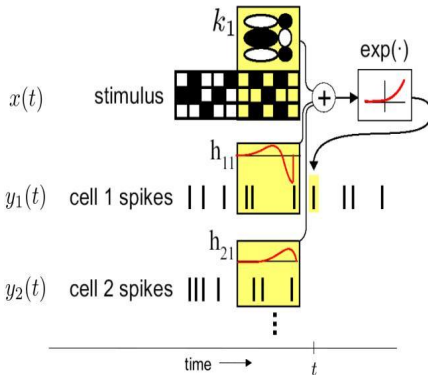
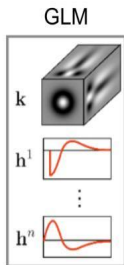
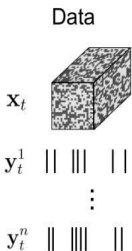
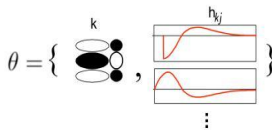
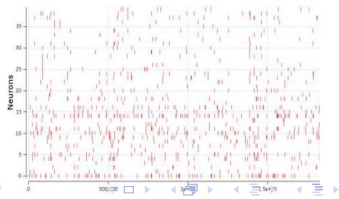
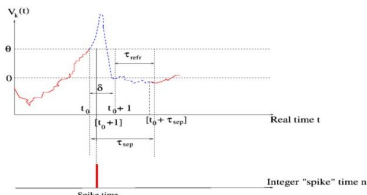
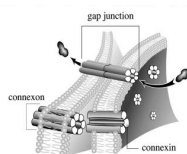
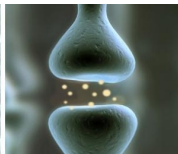
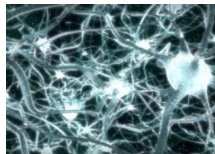


Figure: Generalized Linear Models.

# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

*R. Cofré, B. Cessac: "Dynamics and spike trains statistics in conductance-based Integrate-and-Fire neural networks with chemical and electric synapses", Chaos, Solitons and Fractals, 2013.*  
(Inspired from Rudolph, Destexhe, 2006).



# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

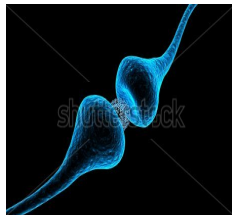
$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$

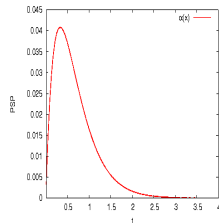


# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$



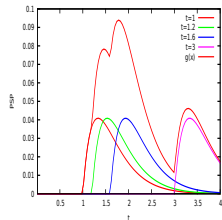


# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$

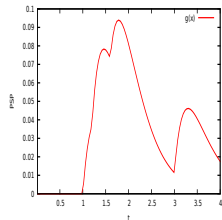


# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$



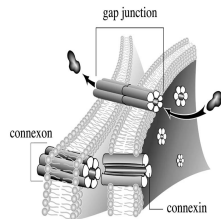
# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$

$$- \sum_j \bar{g}_{kj}(V_k - V_j)$$



# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

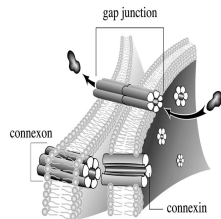
Sub-threshold dynamics:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L)$$

$$- \sum_j g_{kj}(t, \omega)(V_k - E_j)$$

$$- \sum_j \bar{g}_{kj}(V_k - V_j)$$

$$+ i_k^{(ext)}(t) + \sigma_B \xi_k(t)$$



# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

## Spike generation process

- In this example, the hidden process is *non Markovian*: it has an *infinite memory*, although it can be well approximated by a Markov process.

# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

## Spike generation process

- In this example, the hidden process is *non Markovian*: it has an *infinite memory*, although it can be well approximated by a Markov process.
- Without *gap-junctions* the transition probabilities can be approximated. The form is similar to *GLM*  $\Rightarrow$  conditional independence, interpretation of parameters, sigmoidal transfer function.

# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

## Spike generation process

- In this example, the hidden process is *non Markovian*: it has an *infinite memory*, although it can be well approximated by a Markov process.
- Without *gap-junctions* the transition probabilities can be approximated. The form is similar to *GLM*  $\Rightarrow$  conditional independence, interpretation of parameters, sigmoidal transfer function.
- With gap-junctions the *conditional independence breaks down*. The explicit form of the transition probabilities is unknown.

# Example 3: An Integrate and Fire neural network model with chemical and electric synapses

## Spike generation process

- In this example, the hidden process is *non Markovian*: it has an *infinite memory*, although it can be well approximated by a Markov process.
- Without *gap-junctions* the transition probabilities can be approximated. The form is similar to *GLM*  $\Rightarrow$  conditional independence, interpretation of parameters, sigmoidal transfer function.
- With gap-junctions the *conditional independence breaks down*. The explicit form of the transition probabilities is unknown.
- The statistics of spike is described by a *Gibbs distribution* (even in the non stationary case). In the stationary case, it obeys a Maximum Entropy Principle.



## Response to stimuli

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t)$   $\Rightarrow$  *non stationary Gibbs distribution*.

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$  , (linear response).

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$  , (linear response).

$$\kappa_{k,f}(t - t_1) = \frac{1}{C_k} \sum_{r=-\infty}^{[t-t_1]} C^{sp} \left[ f(t - t_1, \cdot), \frac{\mathcal{H}_k^{(1)}(r, \cdot)}{\sigma_k(r - 1, \cdot)} \Gamma_k(0, r - 1, \cdot) \right]$$

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$  , (linear response).

$$\kappa_{k,f}(t - t_1) = \frac{1}{C_k} \sum_{r = -\infty}^{[t-t_1]} C^{SP} \left[ f(t - t_1, \cdot), \frac{\mathcal{H}_k^{(1)}(r, \cdot)}{\sigma_k(r - 1, \cdot)} \Gamma_k(0, r - 1, \cdot) \right]$$

*History dependence.*



# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$  , (linear response).

$$\kappa_{k,f}(t - t_1) = \frac{1}{C_k} \sum_{r=-\infty}^{[t-t_1]} C^{sp} \left[ f(t - t_1, \cdot), \frac{\mathcal{H}_k^{(1)}(r, \cdot)}{\sigma_k(r - 1, \cdot)} \Gamma_k(0, r - 1, \cdot) \right]$$

*History dependence, observable*  $\square$

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$  , (linear response).

$$\kappa_{k,f}(t - t_1) = \frac{1}{C_k} \sum_{r=-\infty}^{[t-t_1]} C^{sp} \left[ f(t - t_1, \cdot), \frac{\mathcal{H}_k^{(1)}(r, \cdot)}{\sigma_k(r-1, \cdot)} \Gamma_k(0, r-1, \cdot) \right]$$

*History dependence, observable, network dynamics*

# Response to stimuli

B. Cessac, R. Cofré, Linear Response of Gibbs measures from Spiking Neuronal Network Models, submitted

GIF model without stimulus  $\Rightarrow$  *stationary Gibbs distribution*.

*time-dependent* stimulus  $S(t) \Rightarrow$  *non stationary Gibbs distribution*.

*How is the average of an observable  $f(\omega, t)$  affected by the stimulus ?*

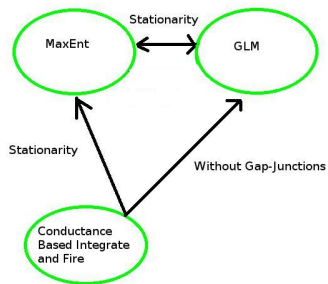
If  $S$  is weak enough:  $\delta\mu[f(t)] = [\kappa_f * S](t)$ , (linear response).

$$\kappa_{k,f}(t - t_1) = \frac{1}{C_k} \sum_{r=-\infty}^{[t-t_1]} C^{SP} \left[ f(t - t_1, \cdot), \frac{\mathcal{H}_k^{(1)(r,\cdot)}}{\sigma_k(r-1,\cdot)} \Gamma_k(0, r-1, \cdot) \right],$$

*History dependence, (spontaneous) correlation between observable and network dynamics*

# Link between these approaches

B. Cessac and R. Cofré, Spike train statistics and Gibbs distributions, *J. Physiol. Paris*, Volume 107, Issue 5, Pages 360-368 (November 2013). Special issue: Neural Coding and Natural Image Statistics.



# Maximum entropy potentials with spatio-temporal constraints are generic

Rodrigo Cofré, Bruno Cessac, "Exact computation of the maximum-entropy potential of spiking neural-network models", Phys. Rev. E 89, 052117 (2014).

Given a set of transition probabilities  $P \left[ \omega(D) \mid \omega_0^{D-1} \right] > 0$  there is a unique, up to a constant, equivalent MaxEnt potential, written as a linear combination of constraints (average of spike events) with a minimal number of terms. This potential can be explicitly (and algorithmically) computed.

# Combinatorial Explosion of constraints

# Combinatorial Explosion of constraints

- A GLM like or a neuronal network model has typically  $O(N^2)$  parameters where  $N$  is the number of neurons.

# Combinatorial Explosion of constraints

- A GLM like or a neuronal network model has typically  $O(N^2)$  parameters where  $N$  is the number of neurons.
- The equivalent MaxEnt potential has *generically*  $2^{N(D+1)} - 2^{ND}$  parameters, non linear and redundant functions of the GLM parameters.

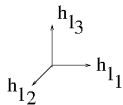


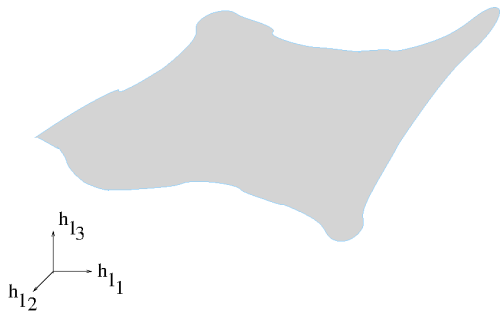
# Combinatorial Explosion of constraints

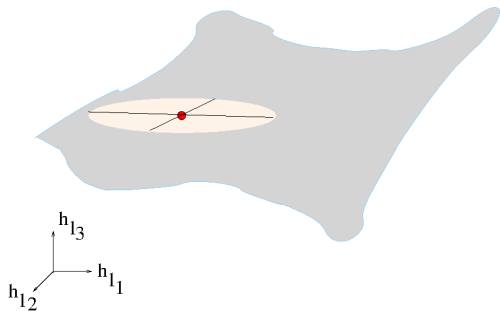
- A GLM like or a neuronal network model has typically  $O(N^2)$  parameters where  $N$  is the number of neurons.
- The equivalent MaxEnt potential has *generically*  $2^{N(D+1)} - 2^{ND}$  parameters, non linear and redundant functions of the GLM parameters.

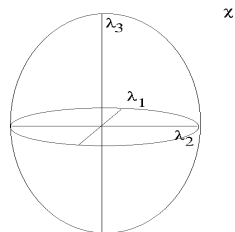
$\Rightarrow$

- Intractable determination of parameters;
- Stimulus dependent parameters;
- Overfitting.









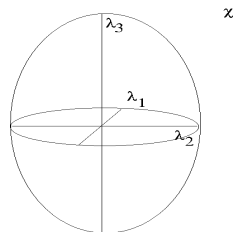
$\chi$  Fisher metric

$$\chi_{ll'} = \frac{\partial^2 d_{KL}(\pi, \mu)}{\partial h_l \partial h_{l'}}$$

$$\chi_{ll'} = \frac{\partial^2 \mathcal{P}}{\partial h_l \partial h_{l'}};$$
$$\mathcal{P} = \sup_{\mu \in M_{inv}} h(\mu) + \mu(\mathcal{H})$$

$$\chi_{ll'} = C_{ll'}(0) + \sum_{n=0}^{+\infty} C_{ll'}(n)$$

$$\delta\mu[m] = \chi\delta h \Rightarrow \delta h = \chi^{-1}\delta\mu[m]$$



$\chi$  Fisher metric

$$\chi_{ll'} = \frac{\partial^2 d_{KL}(\pi, \mu)}{\partial h_l \partial h_{l'}}$$

$$\chi_{ll'} = \frac{\partial^2 \mathcal{P}}{\partial h_l \partial h_{l'}};$$
$$\mathcal{P} = \sup_{\mu \in M_{inv}} h(\mu) + \mu(\mathcal{H})$$

$$\chi_{ll'} = C_{ll'}(0) + \sum_{n=0}^{+\infty} C_{ll'}(n)$$

$$\delta\mu[m] = \chi\delta h \Rightarrow \delta h = \chi^{-1}\delta\mu[m]$$

Dimensionality Reduction and Reliable Observations on Maximum Entropy Models on Spiking Networks, R. Herzog, B. Cessac, M.J. Escobar, A. Palacios, submitted

## On the mathematical consequences of binning

# Binning

B. Cessac, A. LeNy, E. Loecherbach, On the mathematical consequences of binning spike trains, Neural Computation, 2017

Does binning affect the statistics ?



# Binning

B. Cessac, A. LeNy, E. Loecherbach, On the mathematical consequences of binning spike trains, Neural Computation, 2017

Does binning affect the statistics ?

- Consider a (stationary) Markov process generating spike trains, with transition probabilities  $P \left[ \omega(D) \mid \omega_0^{D-1} \right]$ .

# Binning

B. Cessac, A. LeNy, E. Loecherbach, On the mathematical consequences of binning spike trains, Neural Computation, 2017

Does binning affect the statistics ?

- Consider a (stationary) Markov process generating spike trains, with transition probabilities  $P \left[ \omega(D) \mid \omega_0^{D-1} \right]$ .
- Consider the process generated by binning. Is it still Markovian ?

# Binning

B. Cessac, A. LeNy, E. Loecherbach, On the mathematical consequences of binning spike trains, Neural Computation, 2017

Does binning affect the statistics ?

- Consider a (stationary) Markov process generating spike trains, with transition probabilities  $P \left[ \omega(D) \mid \omega_0^{D-1} \right]$ .
- Consider the process generated by binning. Is it still Markovian ?

**NO !**

# Binning

**Example.** Take 1 neuron,  $D = 1$ , and bin with 2 time steps. Call  $\omega$  the initial process and  $\varpi$  the binned process. Let  $\mathbb{P}^{(b)}$  the probabilities in the binned process.

$$\varpi = 0 \Rightarrow (0, 0)$$

$$\varpi = 1 \Rightarrow (0, 1), (1, 0), (1, 1)$$

# Binning

**Example.** Take 1 neuron,  $D = 1$ , and bin with 2 time steps. Call  $\omega$  the initial process and  $\varpi$  the binned process. Let  $\mathbb{P}^{(b)}$  the probabilities in the binned process.

$$\varpi = 0 \Rightarrow (0, 0)$$

$$\varpi = 1 \Rightarrow (0, 1), (1, 0), (1, 1)$$

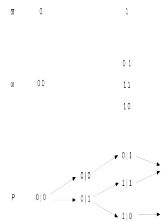
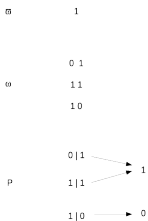
Let us first show that:

$$\mathbb{P}^{(b)}[\varpi(2) = 0 | \varpi(1) = 1, \varpi(0) = 1] \neq \mathbb{P}^{(b)}[\varpi(2) = 0 | \varpi(1) = 1].$$

# Binning

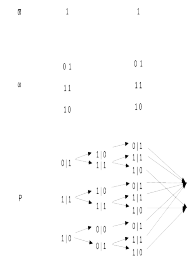
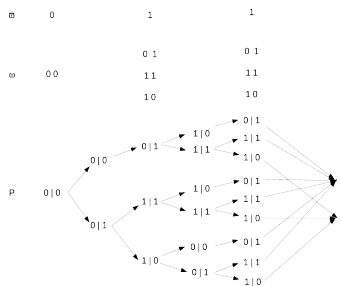
$$P^{(b)}[\varpi(2) = 0 | \varpi(1) = 1] = \frac{P^{(b)}[\varpi(2)=0, \varpi(1)=1]}{P^{(b)}[\varpi(1)=1]}$$

$$P^{(b)}[0|1] = \frac{P[0|0]P[0|0]P[0|1]P[1]+P[0|0]P[0|1]P[1]+P[0|1]P[1]+P[0|0]P[0|1]P[1|0]P[0]}{P[0|1]P[1]+P[1|1]P[1]+P[1|0]P[0]}$$



# Binning

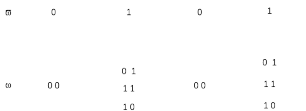
$$P^{(b)}[0|11] = \frac{P^{(b)}[011]}{P^{(b)}[11]}$$



# Binning

But:

$$\mathbb{P}^{(b)}[0|101] = \mathbb{P}^{(b)}[0|10]$$





# Binning

In general, the conditional probability

$$\mathbb{P}^{(b)}[\varpi(s+1)|\varpi_r^s] = \mathbb{P}^{(b)}[\varpi(s+1)|\varpi_l^s],$$

where  $l$  is the first occurrence of the symbol 0 when going from  $s$  to  $r$  (we set  $l = r$ , if  $\varpi$  does not contain the symbol 0).

$\Rightarrow$

The binned process is a Variable Length Markov Chain (VLMC) with unbounded memory.

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.
- 2 This can lead to wrong causal interpretation and fallacious long range space correlations.

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.
- 2 This can lead to wrong causal interpretation and fallacious long range space correlations.

### Proposition (Theorem 3.1 of chazottes-ugalde, 2011)

*The transition kernel  $\mathbb{P}^{(b)}$  of the binned chain is continuous with an exponential rate  $\alpha(N) > 0$ , depending on  $N$ , the number of neurons.*

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.
- 2 This can lead to wrong causal interpretation and fallacious long range space correlations.

Theorem (B. Cessac, A. LeNy, E. Loecherbach, 2017)

*The law  $\mathbb{P}^{(b)}$  of the binned chain is a Gibbs measure.*

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.
- 2 This can lead to wrong causal interpretation and fallacious long range space correlations.

Theorem (B. Cessac, A. LeNy, E. Loecherbach, 2017)

*The law  $\mathbb{P}^{(b)}$  of the binned chain is a Gibbs measure.*

- 3 Nice measure when conditioning upon past, but also upon future (anticipation).

# Binning

## Consequences:

- 1 Binning generate fallacious long-range time correlations.
- 2 This can lead to wrong causal interpretation and fallacious long range space correlations.

Theorem (B. Cessac, A. LeNy, E. Loecherbach, 2017)

*The law  $\mathbb{P}^{(b)}$  of the binned chain is a Gibbs measure.*

- 3 Nice measure when conditioning upon past, but also upon future (anticipation).
- 4 **Open question:** behaviour of the binned process when the number of neurons tends to  $\infty$  (fallacious phase transitions ?)

## Conclusion



# A challenge: "Reading" "the" neural "code"

- **Mathematical challenge:** Improve existing mathematical tools / Elaborate new ones

# A challenge: "Reading" "the" neural "code"

- **Fundamental challenge:** Understanding how visual system "encodes" / "decodes" "information"
- **Experimental challenge:** Elaborate better and better techniques to record neuronal activity
- **Mathematical challenge:** Improve existing mathematical tools / Elaborate new ones
- **Technological challenge:** New techniques in
  - Improving vision for visually impaired people (retinal prostheses, tools for reading, ...)
  - Computer vision
  - ....

Part of these challenges relies on:

Part of these challenges relies on:

- Reliable estimations of spike trains statistics (finite samples, induced biases, finding the right model, estimating its parameters ...),

Part of these challenges relies on:

- Reliable estimations of spike trains statistics (finite samples, induced biases, finding the right model, estimating its parameters ...),
- Generate proper synthetic data to calibrate/tests methods,

Part of these challenges relies on:

- Reliable estimations of spike trains statistics (finite samples, induced biases, finding the right model, estimating its parameters ...),
- Generate proper synthetic data to calibrate/test methods,
- Inferring/guessing/extrapolating hidden relations (laws ?) relating a stimulus to the collective response of neurons.

Part of these challenges relies on:

- Reliable estimations of spike trains statistics (finite samples, induced biases, finding the right model, estimating its parameters ...),
- Generate proper synthetic data to calibrate/test methods,
- *Inferring/guessing/extrapolating hidden relations (laws ?) relating a stimulus to the collective response of neurons.*