Spike train statistics: from mathematical models to software to experiments

Bruno Cessac

NeuroMathComp Team, INRIA Sophia Antipolis, France.

Marseille, 27-03-14
Context
Spike coding

What is the response given a stimulus?  
Decoding  
Can we guess the stimulus given the response?  
Rate coding, spike count, population coding ...

Response variability

Bruno Cessac  
Spike train statistics
Multi Electrodes Array

Figure: Multi-Electrodes Array.

Figure: Spike train.
Stimulus $S \rightarrow$ spike response $R$.

Try to infer $P[S | R]$.

Bayesian inference $\Rightarrow$ needs to know $P[R | S]$.

Can we have a reasonable idea of what $P[R | S]$ could be?
To fit

Extract a probabilistic model of $P[R \mid S]$, $P[S \mid R]$ from data.
To fit
Extract a probabilistic model of $P[R|S]$, $P[S|R]$ from data.

⇒

To predict
Apply the probabilistic model to predict the behaviour of test samples.

(René Thom)
To fit
Extract a probabilistic model of $P[R|S]$, $P[S|R]$ from data.

To predict
Apply the probabilistic model to predict the behaviour of test samples.

To explain
How does a neural network ”encode” a stimulus.
To fit
Extract a probabilistic model of $P[R|S]$, $P[S|R]$ from data.

⇓

To predict
Apply the probabilistic model to predict the behaviour of test samples.

⇓

To explain
How does a neural network "encode" a stimulus.

To predict is not to explain.
(René Thom)
This spike train has been generated by a hidden dynamics / stochastic process. Can we infer this process from the spike train's analysis?
This spike train has been generated by an hidden dynamics / stochastic process.

Can we infer this process from the spike train’s analysis?
Some features of the hidden dynamics

Bruno Cessac
Spike train statistics
Some features of the hidden dynamics

1. **Nonlinear and multiscale dynamics;**

- [Diagram 1: Nonlinear dynamics]
- [Diagram 2: Multiscale dynamics]

---

Bruno Cessac  
Spike train statistics
Some features of the hidden dynamics

1. Nonlinear and multiscale dynamics;

2. Long memory tail (conductance, plasticity, adaptation).
Some features of the hidden dynamics

1. Nonlinear and multiscale dynamics;

2. Long memory tail (conductance, plasticity, adaptation).

3. Collective dynamics
Probability of characteristic spatio-temporal patterns

1. Empirical probability (experiments).
2. Theoretical probability (statistical model).

Bruno Cessac
Spike train statistics
Probability of characteristic spatio-temporal patterns

1. Empirical probability (experiments).
2. Theoretical probability (statistical model).

Bruno Cessac
Spike train statistics
1. Empirical probability (experiments).
2. Theoretical probability (statistical model).

---

**Probability of characteristic spatio-temporal patterns**

- **1. Empirical probability (experiments).**
- **2. Theoretical probability (statistical model).**
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
The probability of a spike pattern depends on the network history (transition probabilities).
Probability of characteristic spatio-temporal patterns

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

Probability of characteristic spatio-temporal patterns
Probability of characteristic spatio-temporal patterns

Given a set of hypotheses on transition probabilities there exists a mathematical framework to solve the problem.
Mathematics
Spike events
Spike events

Figure: 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

**Spike state**

\[ \omega_k(n) \in \{0, 1\} \]

**Spike pattern**

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

**Figure:** Spike pattern.
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
\[ \omega_m^n = \{\omega(m) \omega(m + 1) \ldots \omega(n)\} \]
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
\[ \omega_m^n = \{ \omega(m) \omega(m+1) \ldots \omega(n) \} \]

\[
\begin{pmatrix}
1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 \\
\end{pmatrix}
\]
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

\[ \omega_m^n = \{\omega(m) \omega(m+1) \ldots \omega(n)\} \]

Raster plot

\[ \omega \overset{\text{def}}{=} \omega_0^T \]
Main idea

Construct:

\[ P_n \left[ \omega(n) \, | \, \omega^{n-1}_{n-D} \right] \]

from data.
Spatial models ... let’s forget about the past
Spatial models ... let’s forget about the past

To remove time correlations, bin data at a time scale larger that the typical correlation time between cells.
Spatial models … let’s forget about the past

\[ P_n \left[ \omega(n) \mid \omega_{n-D}^{n-1} \right] = P_n[\omega(n)] \]
Spatial models ... let’s forget about the past

\[ P \left[ \omega_n^{n+3} \right] \]
Spatial models ... let’s forget about the past

\[
P \left[ \omega_{n+3}^n \right] = P \left[ \omega(n) \right] P \left[ \omega(n+1) \right] P \left[ \omega(n+2) \right] P \left[ \omega(n+3) \right]
\]
Spatial models … let’s forget about the past

Useful to estimate the *spatial* distribution of patterns.
Spatial models ... let’s forget about the past

Useful to estimate the *spatial* distribution of patterns.
Most statistical models of Maximum Entropy are of this type.
Markov chains
Markov chains

The actual state is influenced by a past with a fixed depth.
Markov chains ... fixed past depth.
Markov chains ... fixed past depth
Markov chains ... fixed past depth
Markov chains ... fixed past depth

- Conditional intensities models (GLM, LN, ...).
Markov chains ... fixed past depth

- **Mathematics**: Perron-Frobenius theorem, general theory of Markov chains.
- **Numerics**: Statistical inference of Markov chain, Monte-Carlo.
Variable Length Markov chains
Variable Length Markov chains

The actual state is influenced by a past with a variable but *bounded* depth.
Variable Length Markov chains

- Conditional intensities models.
Chains with complete connections ... unbounded memory
Chains with complete connections ... unbounded memory

The actual state is influenced by a past with an *unbounded* depth.
Chains with complete connections ... unbounded memory


Bruno Cessac

Spike train statistics
Chains with complete connections … unbounded memory


Chains with complete connections \(\sim\) Gibbs distributions.
Examples
Example 1: The Generalized-Linear Model (GLM)

Paradigms of rates and receptive fields.

\[ \lambda_k(t|H_t) = f(b_k + (K_k \ast x)(t) + \sum_j (h_{kj} \ast y_j)(t)) \]

\[ \theta = \{k, h_{kj}\} \]

**Figure:** Generalized Linear Models.
Example 2: Maximum Entropy models

Measuring the statistics of characteristic spike events
Example 2: Maximum Entropy models

Measuring the statistics of characteristic spike events

single spikes,
Example 2: Maximum Entropy models

Measuring the statistics of characteristic spike events

single spikes, pairs,
Example 2: Maximum Entropy models

Measuring the statistics of **characteristic spike events**

single spikes, pairs, triplets,
Example 2: Maximum Entropy models

Measuring the statistics of characteristic spike events

single spikes, pairs, triplets, . . . , what else?
Example 2: Maximum Entropy models

- Assume data is **stationary**.
Assume data is stationary.

Choose a set of characteristic spike events.
Assume data is **stationary**.

Choose a set of characteristic spike events.

Compute their empirical mean.

\( \pi_l \): emp. prob. of spike event \( l \).
Assume data is **stationary**.

Choose a set of characteristic spike events.

Compute their empirical mean.

\[ \pi_l : \text{emp. prob. of spike event } l. \]

Find a probability distribution \( \mu \) which *reproduces the empirical averages*:

\[ \mu [ \text{Spike event } l ] = \pi_l \quad (C) \]
Example 2: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events.
- Compute their empirical mean.  
  \( \pi_l \): emp. prob. of spike event \( l \).
- Find a probability distribution \( \mu \) which reproduces the empirical averages:
  \[
  \mu [ \text{Spike event } l ] = \pi_l \quad (C)
  \]
  without any further assumption.

\( \iff \)

*Maximize the entropy under constraints \( C \).*
Example 2: Maximum Entropy models

- Assume data is **stationary**.
- Choose a set of characteristic spike events.
- Compute their empirical mean.
  \[ \pi_l : \text{emp. prob. of spike event } l. \]
- Find a probability distribution \( \mu \) which **reproduces the empirical averages**:
  \[ \mu [\text{Spike event } l] = \pi_l \quad (C) \]
  without any further assumption.

\[ \iff \]

*Maximize the entropy under constraints \( C \).*

- In good cases this probability exists and is unique. It is called a **Gibbs distribution**.
Example 2: Maximum Entropy models

Most authors consider spatial events (no memory) because it fits with the usual definition of Gibbs distributions $\frac{e^{\beta H}}{Z}$.
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 (Marre et al, 2009; Vasquez et al, 2012; Nasser-Cessac, 2014)
Can we have a *reasonable idea* of what could be the spike statistics by studying a *neural network model*?
An Integrate and Fire neural network model with chemical and electric synapses

What could be the hidden process?

Sub-threshold dynamics:

\[ C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L) \]
What could be the hidden process?

Sub-threshold dynamics:

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

\[- \sum_j g_{kj}(t, \omega)(V_k - E_j)\]
What could be the hidden process?

Sub-threshold dynamics:

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

\[- \sum_j g_{kj}(t, \omega)(V_k - E_j) \]
Sub-threshold dynamics:

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

\[- \sum_j g_{kj}(t, \omega)(V_k - E_j)\]
What could be the hidden process?

Sub-threshold dynamics:

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

\[ - \sum_j g_{kj}(t, \omega)(V_k - E_j) \]
What could be the hidden process?

Sub-threshold dynamics:

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

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

\[- \sum_j \bar{g}_{kj}(V_k - V_j) + i_{ext}(t) + \sigma_B \xi_k(t) \]
What could be the hidden process?

Sub-threshold dynamics:

\[
C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L) \\
- \sum_j g_{kj}(t, \omega)(V_k - E_j) \\
- \sum_j \tilde{g}_{kj}(V_k - V_j) \\
+ i_k^{(ext)}(t) + \sigma_B \xi_k(t)
\]
Mathematical answers

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

**Mathematical answers**

- 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 explicitly computed. The form is similar to *GLM* ⇒ conditional independence, interpretation of parameters, sigmoidal transfer function.
Mathematical answers

- 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 explicitly computed. 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 has (not yet) been computed.
What could be the hidden process?

**Mathematical answers**

- 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 explicitly computed. 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 has (not yet) been computed.

- 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.
What could be the hidden process?

- We have now a *reasonable model* where we can have a mathematical characterization of spike statistics given a stimulus.
- *Gap junctions* might play a crucial role in shaping statistics.
- Possible *new paradigm* to fit data (beyond GLM and MaxEnt).
Relations between models (with R. Cofré).

Neuromimetic models

Conditional Intensities Models

Maximum Entropy Models
Relations between models (with R. Cofré).

Neuromimetic models

$O(N^2)$ parameters
Non stationary

Conditional Intensities Models

$O(N^2)$ parameters
Non stationary

Maximum Entropy Models

$O(N^{D+1})$ parameters
Stationary
Relations between models (with R. Cofré).

- **Neuromimetic models**
  - $O(N^2)$ parameters
  - Non stationary
  - Analytical results limited to IF models.

- **Conditional Intensities Models**
  - $O(N^2)$ parameters
  - Non stationary
  - (Conditional independence; exponential transfer function)

- **Maximum Entropy Models**
  - $O(C^{N(D+1)})$ parameters
  - Stationary
  - How to choose the constraints?
Relations between models (with R. Cofré).

Neuromimetic models

Analytical results limited to IF models.

Cessac, 2010, 2012

Conditional Intensities Models

O(N^2) parameters
Non stationary

(C Conditional independence; exponential transfer function)

Cofré, Cessac 2013

Maximum Entropy Models

O(N^{D+1}) parameters
Stationary

How to choose the constraints?

Cofré, Cessac 2014
Relations between models (with R. Cofré).

- **Neuromimetic models**
  - $O(N^2)$ parameters
  - Non stationary
  - Analytical results limited to IF models
  - Cessac, 2010, 2012

- **Conditional intensities Models**
  - $O(N^2)$ parameters
  - Non stationary
  - (Conditional independence; exponential transfer function)
  - Cofré, Cessac 2013

- **Maximum Entropy Models**
  - Max Ent models span too huge a space
  - They could be useful if some hidden laws make most parameters vanish (as in physics)
  - How to catch the shape of a MaxEnt potential?
  - Cofré, Cessac 2014
Real versus "functional" connectivity (with R. Cofré).
Numerics
The ENAS software

Freely downloadable at
http://enas.gforge.inria.fr/v3/
The ENAS software

Freely downloadable at
http://enas.gforge.inria.fr/v3/

**Goal:** Provide a tool to neuroscientists community to analyze their experiments and tests their hypotheses on spike statistics.
The ENAS software

Freely downloadable at
http://enas.gforge.inria.fr/v3/

Goal: Provide a tool to neuroscientists community to analyze their experiments and tests their hypotheses on spike statistics.

- Stimulus and raster loading (retina).
The ENAS software

Freely downloadable at
http://enas.gforge.inria.fr/v3/

**Goal:** Provide a tool to neuroscientists community to analyze their experiments and tests their hypotheses on spike statistics.

- Stimulus and raster loading (retina).
- Spike train statistics analysis with spatio-temporal Max Ent and conditional intensities model (in progress).
The ENAS software

Freely downloadable at
http://enas.gforge.inria.fr/v3/

Goal: Provide a tool to neuroscientists community to analyze their experiments and tests their hypotheses on spike statistics.

- Stimulus and raster loading (retina).
- Spike train statistics analysis with spatio-temporal Max Ent and conditional intensities model (in progress).
- Interface with the Virtual Retina simulator (INRIA Neuromathcomp).
Experiments
Gibbs distributions for retina ganglion cells
(with U. Princeton, U. Valaparaiso, Institut de la Vision)

- Comparing Spatial Max Ent models to Spatio-temporal in the retina.
- Small number of neurons (∼5).
- The new version of Enas allows (∼50) neurons.
The KEOPS project (ANR)
(with U. Valparaiso, U. Santiago and INRIA Mnemosyne)

- Analyse of non-standard retinal cell behaviour
- Natural image scenarios
- Designing innovative visual operators.
The Renvision project (FP7)

**Action 1:** novel large-scale & high-resolution experimental platform and analysis tools

**Action 2:** pan-retinal experiments & computational modeling

**Action 3:** retina-inspired computation to solve artificial visual problems
Acknowledgment

Neuromathcomp team
- Rodrigo Cofré
- Pierre Kornprobst
- Sélim Kraria
- Gaia Lombardi
- Hassan Nasser (Startup)
- Geoffrey Portelli
- Vivien Robinet (Kourou)
- Horacio Rostro (Mexico)
- Wahiba Taouali (INT)
- Juan-Carlos Vasquez (Bogota)

CN Valparaiso
- Adrian Palacios
- Maria-José Escobar

Institut de la Vision
- Olivier Marre
- Serge Picaud

Mnemosyne team
- Thierry Viéville

Princeton University
- Michael J. Berry II