Space-time correlations in spike trains and the neural code

Bruno Cessac, Rodrigo Cofre

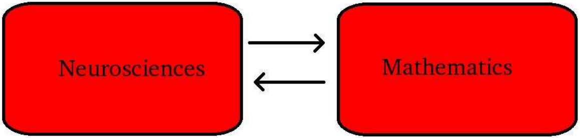
Space-time correlations in spike trains and the neural code

Bruno Cessac, Rodrigo Cofré

NeuroMathComp Team,INRIA Sophia Antipolis,France.

03-09-13

Context and motivation


Neurosciences
Mathematics

Multi Electrodes Array

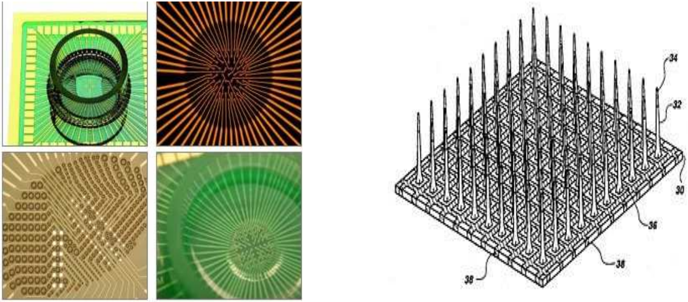


Figure: Multi-Electrodes Array.

Raster plot

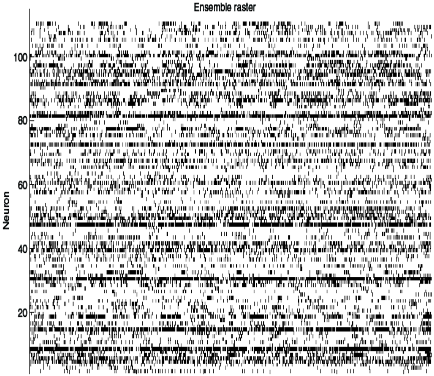


Figure: Raster plot/spike train.

Statistical decoding

Stimulus S $\longrightarrow$ spike response R.

Try to compute $P[R \mid S]$ then $P[S \mid R]$.

Ex: Moving bar

O. Marre, D. Amodei, N. Deshmukh, K. Sadeghi, F. Soo, T. E. Holy, M. J. Berry, "Mapping a Complete Neural Population in the Retina", J Neurosci. 32(43), 2012

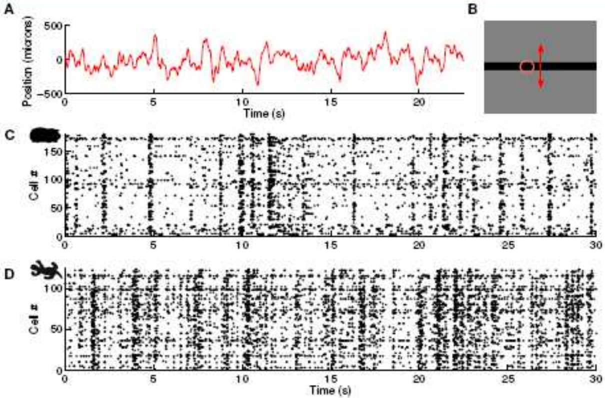


Figure 1: **Ganglion cell spike trains during random bar motion.** **A:** Position of the bar as a function of time. **B:** Example of one stimulus frame; motion is perpendicular to the bar (red arrow). Ellipse fitted to the spatial receptive field profile of one representative ganglion cell (pink). **C:** Spiking activity of 180 cells in the guinea-pig retina in response to a bar moving randomly with the trajectory shown in A. Each line corresponds to one cell, and the points represent spikes. The order of the cells along the y-axis is arbitrary. **D:** Spiking activity of 123 cells in the salamander retina responding to a bar moving randomly. Same convention as C.

Ex: Moving bar

O. Marre, D. Amodei, N. Deshmukh, K. Sadeghi, F. Soo, T. E. Holy, M. J. Berry, "Mapping a Complete Neural Population in the Retina", J Neurosci. 32(43), 2012

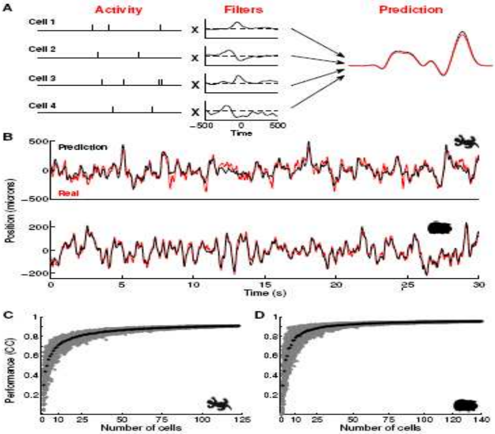


Figure 2: **High-accuracy reconstruction of the bar's trajectory.** **A:** Schematic of the linear decoding method, here for 4 cells. A temporal filter is associated with each cell. Each time the cell spikes, its filter is added to the ongoing reconstruction at the time of the spike. The filters are optimized on part of the data to have the lowest reconstruction error and then tested on the rest of the data. **B:** Prediction of the bar's position (black) from the activity of 123 cells in the salamander retina versus the real trajectory (red). **C:** Prediction of the bar's position (black) from the activity of 178 cells in the guinea pig retina versus the real trajectory (red). **D, E:** Decoding performance plotted against the number of cells in the salamander (D) and guinea pig (E). Gray

To fit

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

To fit

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

$$\Downarrow$$

To predict

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

To fit

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

$$\Downarrow$$

To predict

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

$$\Downarrow$$

To explain

How does a neural network "encode" a stimulus.

To fit

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

$\Downarrow$

To predict

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

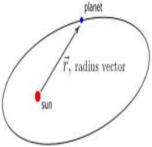
$\Downarrow$

To explain

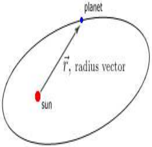
How does a neural network "encode" a stimulus.

To predict is not to explain.

(René Thom)

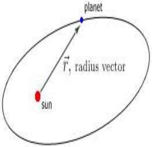


Johannes Kepler
(1571 - 1630)



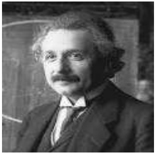

Johannes Kepler
(1571 - 1630)


Isaac Newton
(1642-1727)

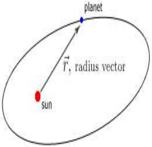


Johannes Kepler
(1571 - 1630)

Isaac Newton
(1642-1727)

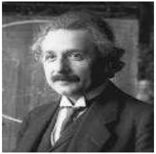


Albert Einstein
(1879-1955)

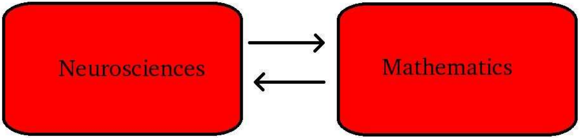



Johannes Kepler
(1571 - 1630)


Isaac Newton
(1642-1727)


Albert Einstein
(1879-1955)

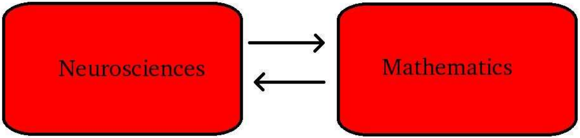

Bernhard Riemann
(1826-1866)



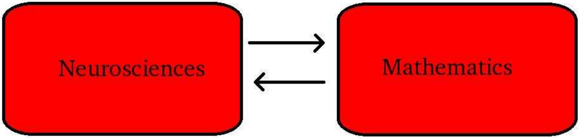
Mathematics can bring in neuroscience not only techniques to solving problems but also new concepts and questions.

Mathematics can also help to propose laws (in the same sense as in Physics) predictive and explanatory, governing the behaviour of the brain.

Proposing new paradigms.

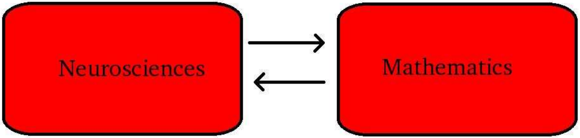


The map is not the territory. (A. Korzibsky).
A model is a representation of reality.



The map is not the territory. (A. Korzibsky).
A model is a representation of reality.

Mathematics must be fed and controlled by experiments.



The map is not the territory. (A. Korzibsky).
A model is a representation of reality.

Mathematics must be fed and controlled by experiments.
A theorem is not a sufficient justification.

From spike trains to mathematical results and questions

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

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

Spike events

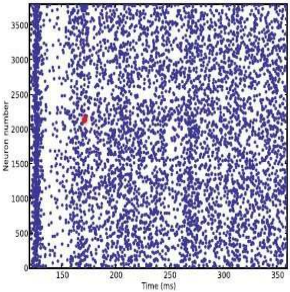


Figure: Spike state.

Spike state

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

Spike events

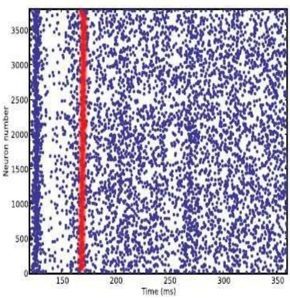


Figure: Spike pattern.

Spike state

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

Spike pattern

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

Spike events

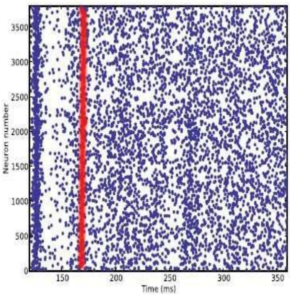


Figure: Spike pattern.

Spike state

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

Spike pattern

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

Spike events

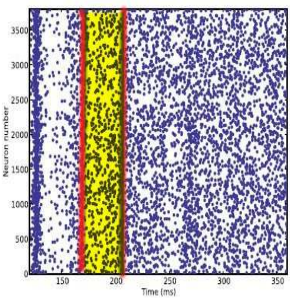


Figure: Spike block.

Spike state

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

Spike pattern

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

Spike block

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

Spike events

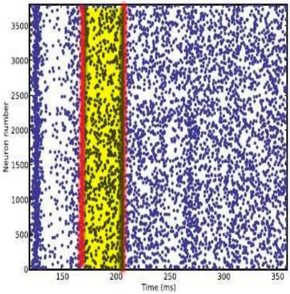


Figure: Spike block.

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

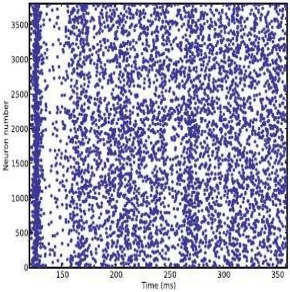


Figure: Raster plot/Spike train.

Spike state

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

Spike pattern

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

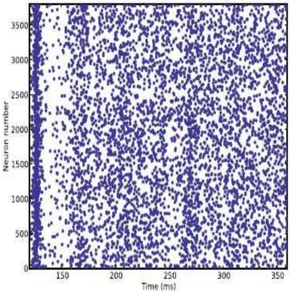
Spike block

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

Raster plot

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

Main idea



Construct:

$$P_n\left[\,\omega(n)\;\middle|\;\omega_{n-D}^{n-1}\,\right]$$

from data.

Main idea

1. Assume that spike statistics is generated by a (Markov) process.
2. Assume a parametric form for the transition probabilities of this process (Ex: Linear-Non Linear, Generalized Linear Model, . . .).
3. Fit the parameters (Maximum likelihood, Kullback-Leibler divergence minimization, learning methods, . . .).
4. Generate sample probabilities and compare to data: does the model **fit** and **predict** correctly (confidence plots, Kullback-Leibler divergence, correlations, . . .) ?
5. Handle correctly the **finite size sampling** of data (standard statistical tests, Central Limit theorem, convergence rate, . . .).

Example 1: The Generalized-Linear Model (GLM)

Paradigms of rates and receptive fields.

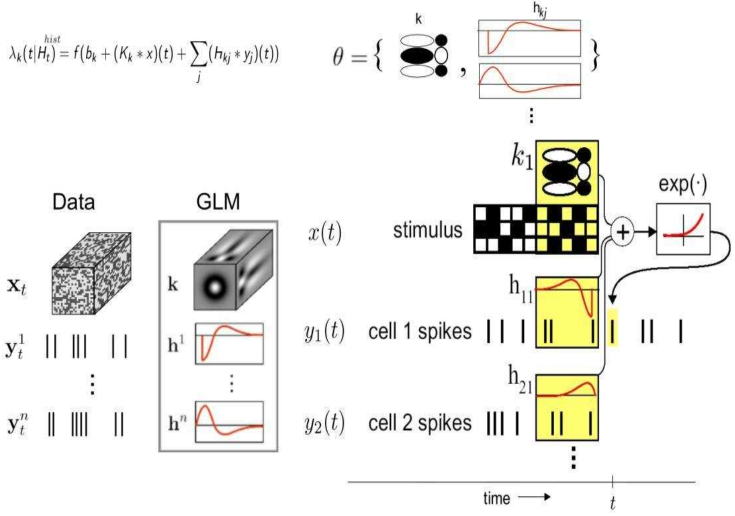


Figure: Generalized Linear Models.

The Generalized-Linear Model (GLM)

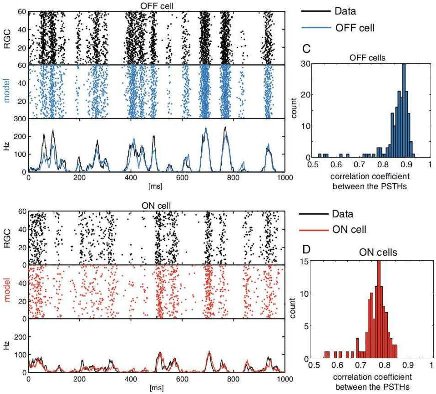
$$\lambda_k(t|H_t) \rightarrow P_n\left[\,\omega_k(n) = 1 \mid H_{n-1}\,\right] \approx \lambda_k(n|H_{n-1})\Delta t = p_k(n)$$

Considering Conditional independence:

$$P_n\left[\,\omega(n) \,\middle|\, \omega_{n-D}^{n-1}\,\right] = \prod_{k=1}^{N} p_k(n)^{\omega_k(n)}(1 - p_k(n))^{1-\omega_k(n)}$$

GLM Experimental Validation

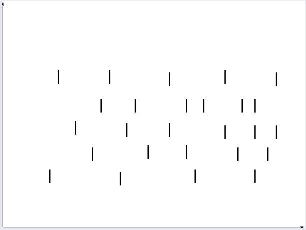
"Modeling the impact of common noise inputs on the network activity of retinal ganglion cells". M.Vidne, Y. Ahmadian, J. Shlens, J. Pillow, J.Kulkarn, A. Litke , E. J. Chichilnisky E.Simoncelli, L. Paninski. Journal of Computational Neuroscience(2011)



Example 2: Handling correlations with the Maximum Entropy Principle

Example 2: Handling correlations with the Maximum Entropy Principle

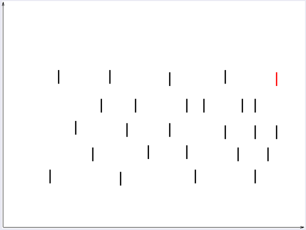
Measuring the statistics of characteristic spike events



Example 2: Handling correlations with the Maximum Entropy Principle

Measuring the statistics of characteristic spike events

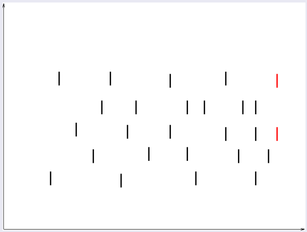
single spikes,



Example 2: Handling correlations with the Maximum Entropy Principle

Measuring the statistics of characteristic spike events

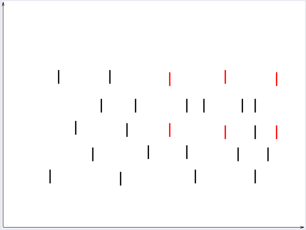
single spikes, pairs,



Example 2: Handling correlations with the Maximum Entropy Principle

Measuring the statistics of characteristic spike events

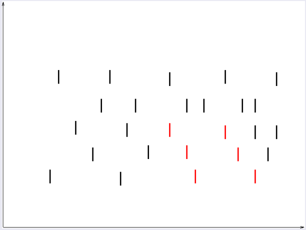
single spikes, pairs,



Example 2: Handling correlations with the Maximum Entropy Principle

Measuring the statistics of characteristic spike events

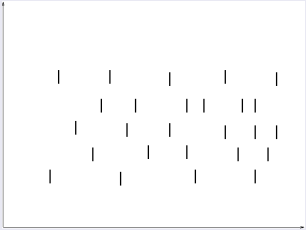
single spikes, pairs, triplets,



Example 2: Handling correlations with the Maximum Entropy Principle

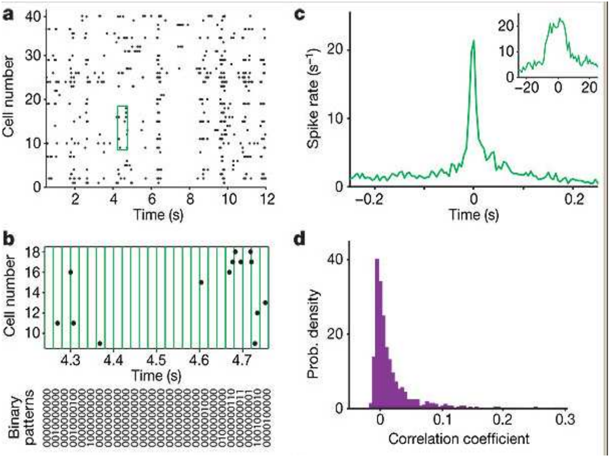
Measuring the statistics of characteristic spike events

single spikes, pairs, triplets, . . . , **what else ?**.



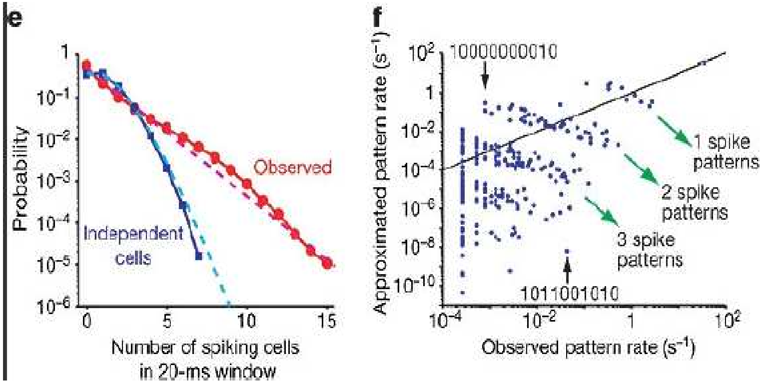
Example 2: Handling correlations with the Maximum Entropy Principle

E. Schneidman, M.J. Berry, R. Segev, and W. Bialek. "Weak pairwise correlations imply strongly correlated network states in a neural population". Nature, 440(7087):1007-1012, 2006.



Example 2: Handling correlations with the Maximum Entropy Principle

E. Schneidman, M.J. Berry, R. Segev, and W. Bialek. "Weak pairwise correlations imply strongly correlated network states in a neural population". Nature, 440(7087):1007-1012, 2006.



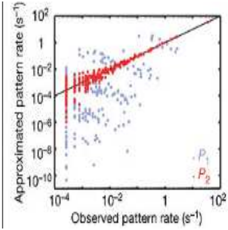
Example 2: Handling correlations with the Maximum Entropy Principle

Are pairwise correlations significant, although weak ?

1. Compute the pairwise correlations.
2. Find the probability distribution which maximizes the statistical entropy and reproduces the observed pairwise correlations $\Rightarrow$ **Gibbs distribution**.
3. Fit and predict.

Example 2: Handling correlations with the Maximum Entropy Principle

E. Schneidman, M.J. Berry, R. Segev, and W. Bialek. "Weak pairwise correlations imply strongly correlated network states in a neural population". Nature, 440(7087):1007-1012, 2006.



Example 2: Handling correlations with the Maximum Entropy Principle

Extensions:

- Ganmor-Schneidman-Segev, 2012: taking into account instantaneous triplets, quadruplets;
- Marre et al, 2009: One step memory pairwise Markov process;
- Vasquez et al, 2012: General form of events can be taken into account from general theory of Gibbs distributions and Perron-Frobenius theorem;
- Nasser et al, 2013: Monte Carlo approach to spatio-temporal Gibbs sampling.

A statistical model makes assumptions

- GLM:
  1. Assumption of conditional independence;
  2. Questionable interpretation of parameters.

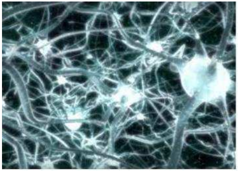
- MaxEnt:
  1. Assumption of stationarity;
  2. Questionable interpretation of parameters;
  3. Which events to choose ?
  4. Exponential complexity;
  5. Overfitting ?

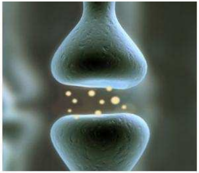
Some mathematical remarks, answers and new questions

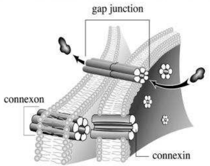
What could be the hidden process ?

What could be the hidden process ?

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.

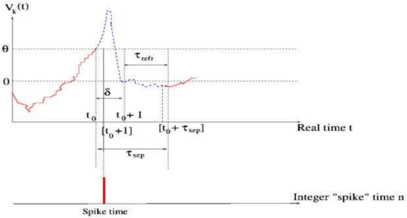


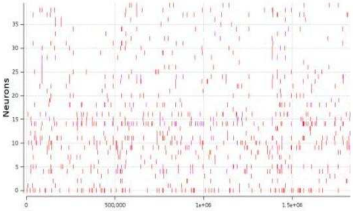




$$C_k \frac{dV_k}{dt} + g_k(t,\omega) V_k = i_k(t,\omega)$$

$$P\left[\omega(n) | \omega_{-\infty}^{n-1}\right]$$





What could be the hidden process ?

The sub-threshold variation of the membrane potential of neuron k at time t is given by:

$$C_k \frac{dV_k}{dt} = -g_{L,k}(V_k - E_L) - \sum_j g_{kj}(t,\omega)(V_k - E_j)$$

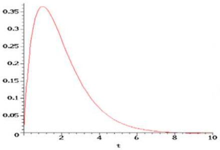
$$- \sum_j \bar{g_{kj}} (V_k - V_j) + I_k(t).$$

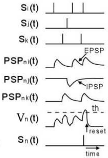
C_k is the membrane capacity of neuron k. $I_k(t) = i_k^{(ext)}(t) + \sigma_B \xi_k(t)$, where $i_k^{(ext)}(t)$ is a deterministic external current ("stimulus").

What could be the hidden process ?

$$g_{kj}(t) = g_{kj}(t_j^r(\omega)) + G_{kj}\alpha_{kj}(t - t_j^r(\omega)), \quad t > t_j^r(\omega),$$

$$\alpha_{kj}(t) = \frac{t}{\tau_{kj}} e^{-\frac{t}{\tau_{kj}}} H(t),$$





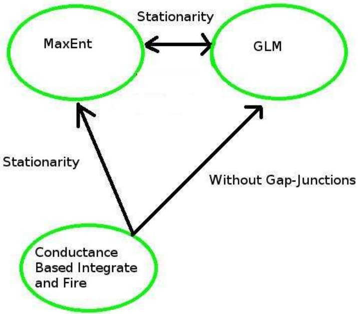
What could be the hidden process ?

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 and interpretation of parameters).
- 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.

Is there any relation between GLM like models and MaxEnt ?

Is there any relation between GLM like models and MaxEnt ?

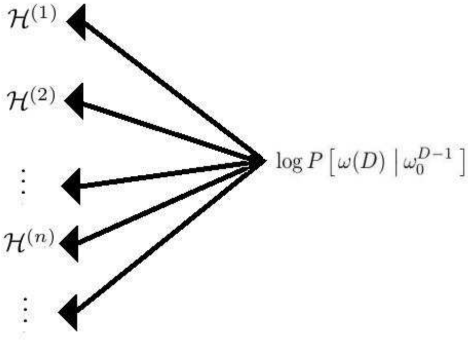


Can we hear the shape of a Maximum Entropy potential ?

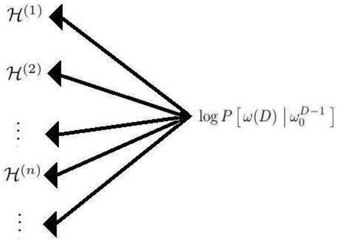
Can we hear the shape of a Maximum Entropy potential ?

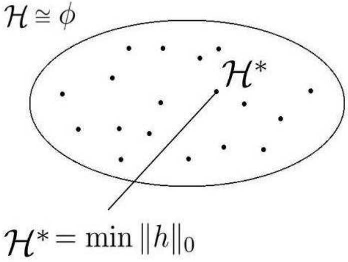
$$\mathcal{H}\left(\omega_0^D\right) \longrightarrow \log P\left[\omega(D) \,\middle|\, \omega_0^{D-1}\right]$$

Can we hear the shape of a Maximum Entropy potential ?



Can we hear the shape of a Maximum Entropy potential ?





Can we hear the shape of a Maximum entropy potential

Two distinct potentials $\mathcal{H}^{(1)}, \mathcal{H}^{(2)}$ of range $R = D + 1$ correspond to the same Gibbs distribution (are "equivalent"), if and only if there exists a range D function f such that (Chazottes-Keller (2009)):

$$\mathcal{H}^{(2)}\left(\omega_0^D\right) = \mathcal{H}^{(1)}\left(\omega_0^D\right) - f\left(\omega_0^{D-1}\right) + f\left(\omega_1^D\right) + \Delta, \qquad (1)$$

where $\Delta = \mathcal{P}(\mathcal{H}_{\boldsymbol{\beta}}{}^{(2)}) - \mathcal{P}(\mathcal{H}_{\boldsymbol{\beta}}{}^{(1)})$.

Can we hear the shape of a Maximum entropy potential

Summing over periodic orbits we get rid of the function f

$$\sum_{n=1}^{R} \phi(\omega\sigma^n l_1) = \sum_{n=1}^{R} \mathcal{H}^*(\omega\sigma^n l_1) - R\mathcal{P}(\mathcal{H}^*), \tag{2}$$

We eliminate equivalent constraints.

Can we hear the shape of a Maximum entropy potential

Conclusion

Given a set of transition probabilities $P\left[\omega(D)\,\middle|\,\omega_0^{D-1}\right] > 0$ there is a unique, up to a constant, 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 model has typically $O(N^2)$ parameters where N is the number of neurons.

Combinatorial Explosion of constraints

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

Combinatorial Explosion of constraints

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

$$\Rightarrow$$

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

Combinatorial Explosion of constraints

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

$$\Rightarrow$$

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

BUT

Combinatorial Explosion of constraints

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

$$\Rightarrow$$

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

BUT *Real neural networks are not generic*

Remark

MaxEnt approach might be useful if there is some hidden law of nature/ symmetry which cancels most of the terms of its expansion.

Finite size effects

Finite size effects

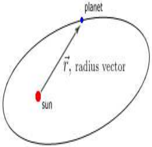
Having a nice mathematical model for spike statistics will be really efficient if one can control/ predict finite size effects:

- Fluctuations (Central Limit theorem; infinitely divisible distributions)
- Errors on parameters estimations.
- Convergence rate (Large deviations; concentrations inequalities)
- Statistical tests (Neymann-Pearson, . . .);

Paradigm changes ?

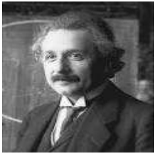
Paradigm changes ?

Thomas S. Kuhn, (1922-1996)



Johannes Kepler
(1571 - 1630)

Isaac Newton
(1642-1727)



Albert Einstein
(1879-1955)

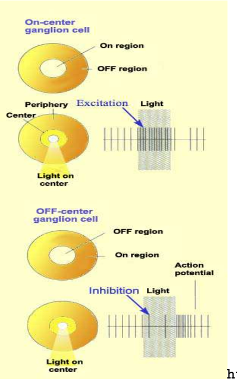
Bernhard Riemann
(1826-1866)

Receptive field

The receptive field of a sensory neuron is a region of space in which the presence of a stimulus will alter the firing of that neuron. (wikipedia)

Receptive field

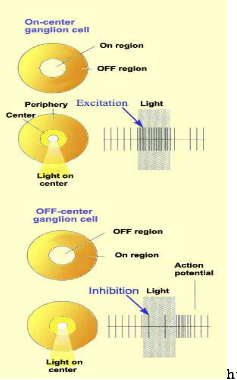
The receptive field of a sensory neuron is a region of space in which the presence of a stimulus will alter the firing of that neuron. (wikipedia)



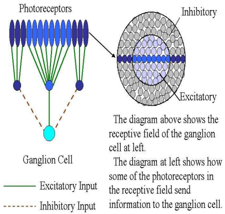
http://thebrain.mcgill.ca/flash/d/d_02/d_02_cl/d_02_cl_vis/d_02_cl_vis.html

Receptive field

The receptive field of a sensory neuron is a region of space in which the presence of a stimulus will alter the firing of that neuron. (wikipedia)



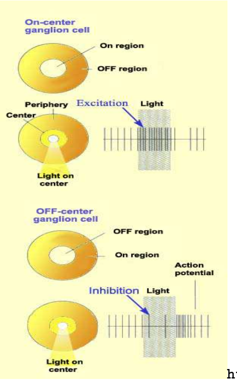
http://thebrain.mcgill.ca/flash/d/d_02/d_02_cl/d_02_cl_vis/d_02_cl_vis.html



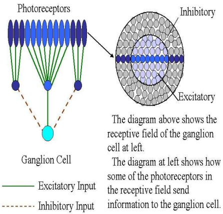
http://www.luc.edu/faculty/asutter/RecField.html

Receptive field

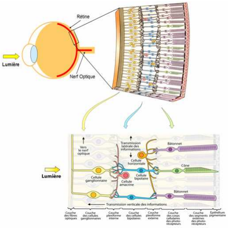
The receptive field of a sensory neuron is a region of space in which the presence of a stimulus will alter the firing of that neuron. (wikipedia)



http://thebrain.mcgill.ca/flash/d/d_02/d_02_cl/d_02_cl_vis/d_02_cl_vis.html



http://www.luc.edu/faculty/asutter/RecField.html



http://theses.ulaval.ca/archimede/fichiers/24200/ch01.html

Characterizing the collective response to stimuli.