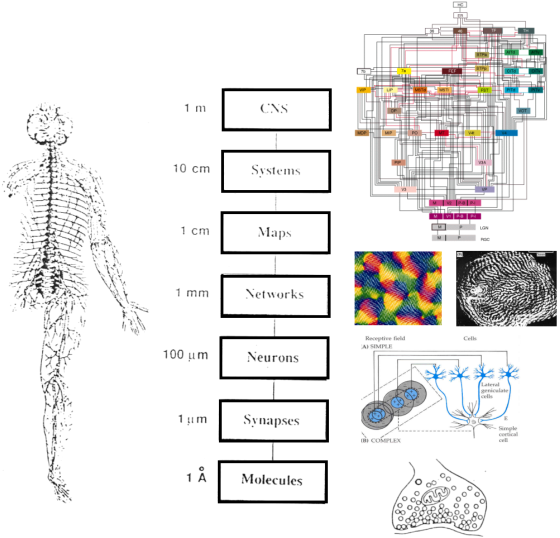
Toward a Statistical Neuroscience

Olivier Faugeras

NeuroMathComp Laboratory - INRIA Sophia/UNSA LJAD

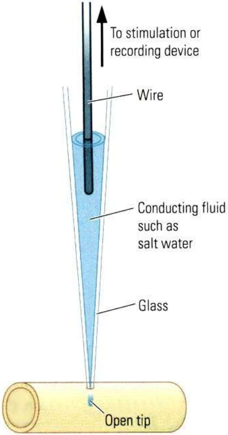
From Medical Images to Computational Medicine
Collège de France
June 24 2014

Scales in the CNS

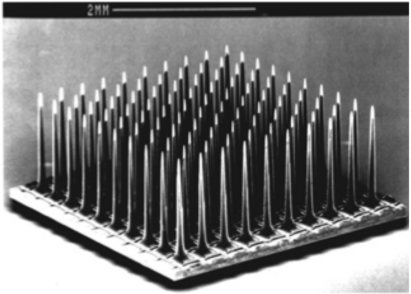


Measuring at different scales

Microelectrode

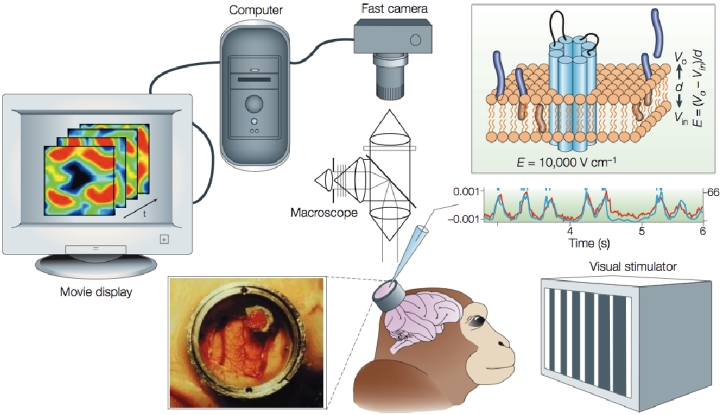


Utah Multi Electrode Array (MEA)



Mesoscopic and macroscopic models

Optical imaging



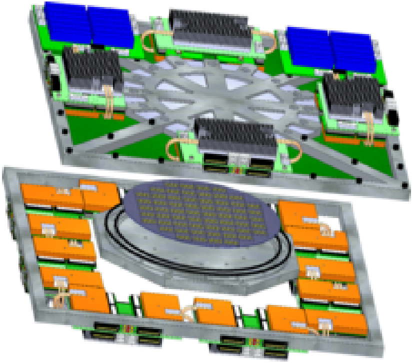
Grinvald-Hildesheim Nature Reviews Neuroscience 04

Mesoscopic and macroscopic models

Electroencephalography (EEG) and Magnetoencephalography (MEG)

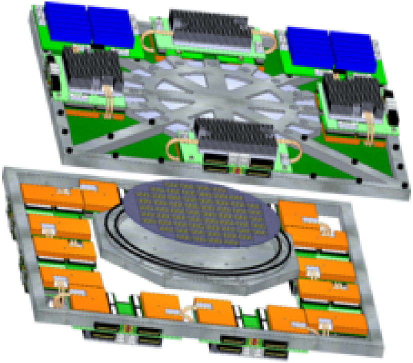
Mesoscopic and macroscopic models

The Brain-Scales project hardware

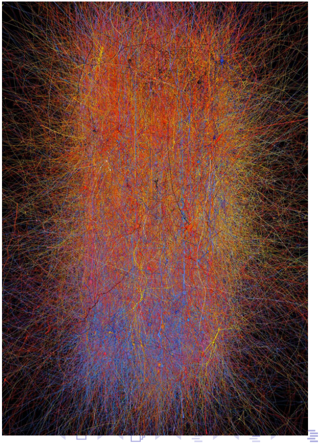


Mesoscopic and macroscopic models

The Brain-Scales project hardware



The Blue Brain project simulator (EPFL)



Mesoscopic and macroscopic models

- These two projects continue with the Human Brain Project (HBP) EC Flagship
- and the Brain project in the US

Introduction

Motivations: Sparse representations, emerging phenomena

Large Deviations

Take home messages

Motivations

- Represent the neuronal activity at different scales (**sparsity**)

Motivations

- Represent the neuronal activity at different scales (**sparsity**)
- Predict the occurence of new, **emerging**, neural phenomena

Motivations

- Represent the neuronal activity at different scales (**sparsity**)
- Predict the occurence of new, **emerging**, neural phenomena
- Understand the role of **randomness**

An example: a toy model of V1

- For each neuron write its equations (2×10^7)

An example: a toy model of V1

- For each neuron write its equations (2×10^7)
- Include the synaptic connections, e.g. chemical ($\simeq 10^{11}$)

An example: a toy model of V1

- For each neuron write its equations (2×10^7)
- Include the synaptic connections, e.g. chemical ($\simeq 10^{11}$)
- You end up with

An example: a toy model of V1

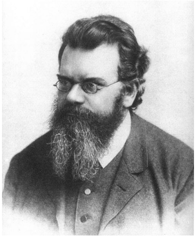
$$\begin{cases} dV_t^i &= \left(V_t^i - \frac{(V_t^i)^3}{3} - w_t^i + I(t) \right) dt + \frac{1}{N} \sum_{j=1}^N \bar{J}(V_t^i - V_{\rm rev}) y_t^j dt + \\ & \quad \frac{1}{N} \left(\sum_{j=1}^N \sigma(V_t^i - V_{\rm rev}) y_t^j \right) dB_t^i + \sigma_{\rm ext}\, dW_t^i \\ dw_t^i &= a\left(b\, V_t^i - w_t^i \right) dt \\ dy_t^i &= \left(a_r S(V_t^i)(1 - y_t^i) - a_d y_t^i \right) dt + \sigma(V_t^i, y_t^i) dW_t^{i,y} \\ J_{ij}(t) &= \frac{\bar{J}}{N} + \frac{\sigma}{N} \xi^i(t) \end{cases}$$

$\simeq$ **10 millions times!**

What can be done?

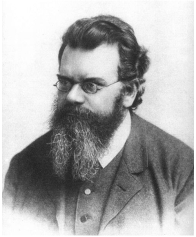
- Develop a statistical neuroscience

What can be done?



- Develop a statistical neuroscience
- Ludwig Boltzmann: Inventor of statistical mechanics

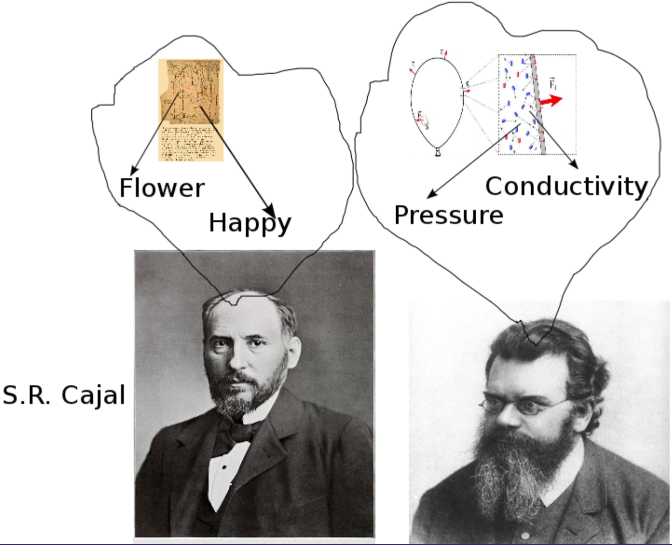
What can be done?



- Develop a statistical neuroscience

- Ludwig Boltzmann: Inventor of statistical mechanics
- Explains and predicts how the properties of atoms (mass, charge, and structure . . .) determine the visible properties of matter (viscosity, thermal conductivity, diffusion . . .)

Statistical mechanics and statistical neuroscience



Introduction

Motivations: Sparse representations, emerging phenomena

Large Deviations

Take home messages

Large Deviation Principle in a nutshell

- Deals with rare events,
- and the asymptotic computation of their probability on a log scale.
- Requires a rate function H.

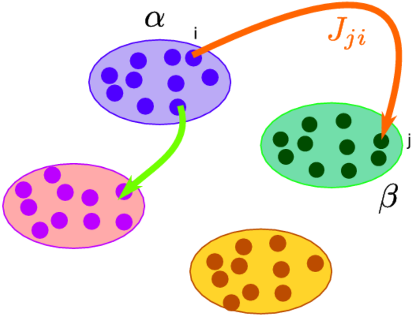
An intuitive approach

- Consider a sequence of random variables $\{X_n\}_{n \geq 0}$.
- It satisfies a Large Deviation Principle with rate function H if

$$P(X_n \simeq x) \simeq e^{-nH(x)}$$

- The random variables X_n concentrate on the points x such that $H(x) = 0$.
- If H reaches 0 at a unique point x^* then the law of X_n converges in law toward the Dirac mass δ_{x^*}: **concentration of measure** phenomenon..

The network



The model

- N neurons, discrete and finite time (T)

The model

- N neurons, discrete and finite time (T)
- Intrinsic dynamics:

$$U^i(t) = \gamma U^i(t-1) + \sigma B^i(t-1) \quad t = 1 \cdots T$$

The model

- N neurons, discrete and finite time (T)
- Intrinsic dynamics:

$$U^i(t) = \gamma U^i(t-1) + \sigma B^i(t-1) \quad t = 1 \cdots T$$

- Coupled dynamics

$$U^i(t) = \gamma U^i(t-1) + \sum_{j=1}^{N} J_{ij}^N f(U^j(t-1)) + B^i(t-1) \quad t = 1 \cdots T$$

The parameters

- B^i: i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$: intrinsic noise on the membrane potentials

The parameters

- B^i: i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$: intrinsic noise on the membrane potentials
- J_{ij}^N: stationary Gaussian field: random synaptic weights

$$\mathbb{E}[J_{ij}^N] = \frac{\overline{J}}{N}$$

$$cov(J_{ij}^N J_{kl}^N) = \frac{\Lambda(k-i, l-j)}{N}$$

The parameters

- B^i: i.i.d. Gaussian $\mathcal{N}(0, \sigma^2)$: intrinsic noise on the membrane potentials
- J_{ij}^N: stationary Gaussian field: random synaptic weights

$$\mathbb{E}[J_{ij}^N] = \frac{\overline{J}}{N}$$

$$cov(J_{ij}^N J_{kl}^N) = \frac{\Lambda(k-i, l-j)}{N}$$

- f is a sigmoid defining the firing rate

Probabilistic framework

- A trajectory is a point of $\mathbb{R}^{T+1} \equiv \mathcal{T}$

Probabilistic framework

- A trajectory is a point of $\mathbb{R}^{T+1} \equiv \mathcal{T}$
- A solution of the networks equations is a (random) probability measure $Q^N(J^N)$ on $\mathcal{T}^N$

Probabilistic framework

- A trajectory is a point of $\mathbb{R}^{T+1} \equiv \mathcal{T}$
- A solution of the networks equations is a (random) probability measure $Q^N(J^N)$ on $\mathcal{T}^N$
- The goal is to characterize the limit (if it exists) of

$$Q^N = \mathbb{E}^{J^N}[Q^N(J^N)],$$

when $N \to \infty$

The tools of the trade

- Let $\mathcal{T}^{\mathbb{Z}}$ be the set of doubly infinite sequences of trajectories

$$u \in \mathcal{T}^{\mathbb{Z}}, \quad u^i \quad \text{its} \quad i\text{th coordinate} \ \in \mathcal{T}$$

The tools of the trade

- Let $\mathcal{T}^{\mathbb{Z}}$ be the set of doubly infinite sequences of trajectories

$$u \in \mathcal{T}^{\mathbb{Z}}, \quad u^i \quad \text{its} \quad i\text{th coordinate} \ \in \mathcal{T}$$

- The shift operator:

$$S(u)^i = u^{i+1}$$

The tools of the trade

- Let $\mathcal{T}^{\mathbb{Z}}$ be the set of doubly infinite sequences of trajectories

$$u \in \mathcal{T}^{\mathbb{Z}}, \quad u^i \quad \text{its} \quad i\text{th coordinate} \ \in \mathcal{T}$$

- The shift operator:

$$S(u)^i = u^{i+1}$$

- $\mathcal{M}_1^+(\mathcal{T}^{\mathbb{Z}})$ the set of probability measures on $\mathcal{T}^{\mathbb{Z}}$

The tools of the trade

- Let $\mathcal{T}^{\mathbb{Z}}$ be the set of doubly infinite sequences of trajectories

$$u \in \mathcal{T}^{\mathbb{Z}}, \quad u^i \quad \text{its} \quad i\text{th coordinate} \ \in \mathcal{T}$$

- The shift operator:

$$S(u)^i = u^{i+1}$$

- $\mathcal{M}_1^+(\mathcal{T}^{\mathbb{Z}})$ the set of probability measures on $\mathcal{T}^{\mathbb{Z}}$
- $\mathcal{M}_{1,S}^+(\mathcal{T}^{\mathbb{Z}})$ the set of stationary probability measures on $\mathcal{T}^{\mathbb{Z}}$ (shift invariant)

Empirical measure

- Let $u^N = (u^{-n}, \cdots, u^n)$ be an element of $\mathcal{T}^N$ $(N = 2n + 1)$

Empirical measure

- Let $u^N = (u^{-n}, \cdots, u^n)$ be an element of $\mathcal{T}^N$ $(N = 2n+1)$
- Define the empirical measure

$$\hat{\mu}^N(u^N) = \frac{1}{N} \sum_{i=-n}^{n} \delta_{S^i(u^N)}$$

Empirical measure

- Let $u^N = (u^{-n}, \cdots, u^n)$ be an element of $\mathcal{T}^N$ $(N = 2n+1)$
- Define the empirical measure

$$\hat{\mu}^N(u^N) = \frac{1}{N} \sum_{i=-n}^{n} \delta_{S^i(u^N)}$$

- **What it does:**
  counts the relative frequency of u^N, and all its shifted "friends".

Empirical measure

- Let $u^N = (u^{-n}, \cdots, u^n)$ be an element of $\mathcal{T}^N$ $(N = 2n+1)$
- Define the empirical measure

$$\hat{\mu}^N(u^N) = \frac{1}{N} \sum_{i=-n}^{n} \delta_{S^i(u^N)}$$

- **What it does:**
  counts the relative frequency of u^N, and all its shifted "friends".
- **What it is:**
  a mapping from $\mathcal{T}^N$ to $\mathcal{M}^+_{1,S}(\mathcal{T}^{\mathbb{Z}})$

The search for a neuronal LDP

- **Key mathematical idea:**
  Consider the image law Π_N of Q^N through

$$\hat{\mu}^N : \mathcal{T}^N \to \mathcal{M}_{1,S}^{+}(\mathcal{T}^{\mathbb{Z}})$$

$$\Pi_N = Q^N \circ (\hat{\mu}^N)^{-1}$$

Pushed forward measure

The search for a neuronal LDP

- **Key mathematical idea:**
  Consider the image law Π_N of Q^N through

$$\hat{\mu}^N : \mathcal{T}^N \to \mathcal{M}^+_{1,S}(\mathcal{T}^{\mathbb{Z}})$$

$$\Pi_N = Q^N \circ (\hat{\mu}^N)^{-1}$$

  Pushed forward measure

- It is a probability measure on $\mathcal{M}^+_{1,S}(\mathcal{T}^{\mathbb{Z}})$ (**probability measure on a set of probability measures!**):

$$\forall B \in \mathcal{B}(\mathcal{M}^+_{1,S}(\mathcal{T}^{\mathbb{Z}})), \quad \Pi^N(B) = Q^N(\hat{\mu}^N \in B)$$

Main result

Theorem (O.F., J. MacLaurin)

Π^N *is governed by a large deviation principle with a good rate function* H.

O.F., J. MacLaurin, 2013.

Main result

Theorem (O.F., J. MacLaurin)

Π^N *is governed by a large deviation principle with a good rate function* H*.*

O.F., J. MacLaurin, 2013.

It generalizes previous work by

- H. Sompolinsky (Hebrew University)
- G. BenArous and A. Guionnet (Courant Institute and MIT)
- O. Moynot and M. Samuelides (Toulouse University)

The unique minimum of the rate function

- One proves that there is a unique $\mu_e \in \mathcal{M}_{1,s}^{+}(\mathcal{T}^{\mathbb{Z}})$ which minimizes H (remember the **concentration of measure** phenomenon):

The unique minimum of the rate function

- One proves that there is a unique $\mu_e \in \mathcal{M}_{1,s}^{+}(\mathcal{T}^{\mathbb{Z}})$ which minimizes H (remember the **concentration of measure** phenomenon):
- μ_e is essentially an infinite dimensional Gaussian measure (a Gaussian process)

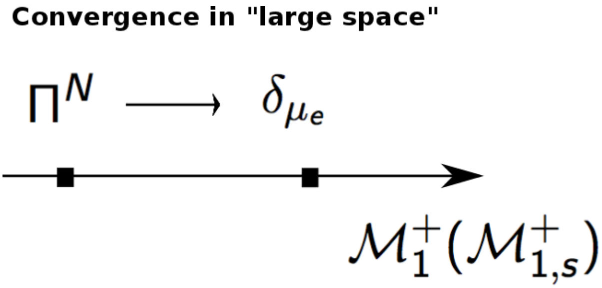
The unique minimum of the rate function

- One proves that there is a unique $\mu_e \in \mathcal{M}_{1,s}^{+}(\mathcal{T}^{\mathbb{Z}})$ which minimizes H (remember the **concentration of measure** phenomenon):
- μ_e is essentially an infinite dimensional Gaussian measure (a Gaussian process)
- μ_e can be effectively computed: numerical experiments are possible!

The unique minimum of the rate function

- One proves that there is a unique $\mu_e \in \mathcal{M}_{1,s}^{+}(\mathcal{T}^{\mathbb{Z}})$ which minimizes H (remember the **concentration of measure** phenomenon):
- μ_e is essentially an infinite dimensional Gaussian measure (a Gaussian process)
- μ_e can be effectively computed: numerical experiments are possible!
- It is non-Markov

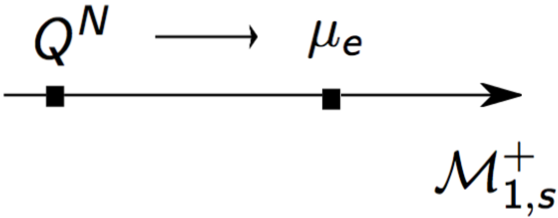
Consequences: Convergence results



This is the concentration of measure phenomenon.

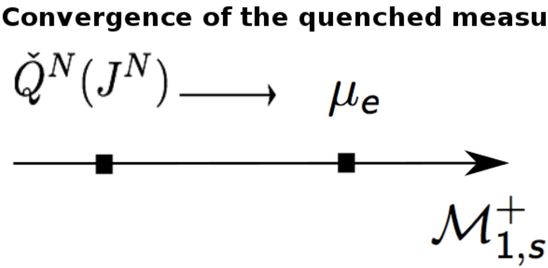
Consequences: Convergence results

Convergence of the averaged measure



The convergence holds when one averages over **all possible networks**

Consequences: Convergence results



The convergence is for **almost all synaptic weights**: it is unnecessary to average over many different networks.

Conclusion: LDP

- Large Deviations are essential in:
  1. Proving the existence of a limit law for the dynamics
  2. Characterizing the law
  3. Establishing averaged and quenched results

Conclusion: LDP

- Large Deviations are essential in:
  1. Proving the existence of a limit law for the dynamics
  2. Characterizing the law
  3. Establishing averaged and quenched results
- Situation is a natural generalization of the i.i.d. case:
  1. Maximum correlation distance of the synaptic weights is d.
  2. The mean-field neurons are uncorrelated if they are separated by more than d: the representation is **sparse**
  3. If $d = 0$, then we observe propagation of chaos.

Take home messages

- Independent synaptic weights hypothesis produces nice mathematical results **but**
- it is somewhat unrealistic.

Take home messages

- Independent synaptic weights hypothesis produces nice mathematical results **but**
- it is somewhat unrealistic.
- Taking into account correlations between synaptic weights is more realistic **but**
- the mathematical description is more complex.

Take home messages

- Independent synaptic weights hypothesis produces nice mathematical results **but**
- it is somewhat unrealistic.
- Taking into account correlations between synaptic weights is more realistic **but**
- the mathematical description is more complex.
- **Emerging phenomena** may be contained in the non-Markov description

Other potential applications of Large Deviations techniques

- Epilepsy: usually involves very large neuronal populations.

Other potential applications of Large Deviations techniques

- Epilepsy: usually involves very large neuronal populations.
- Piecewise deterministic processes:

Other potential applications of Large Deviations techniques

- Epilepsy: usually involves very large neuronal populations.
- Piecewise deterministic processes:
  1. Stochastic ion channels,

Other potential applications of Large Deviations techniques

- Epilepsy: usually involves very large neuronal populations.
- Piecewise deterministic processes:
  1. Stochastic ion channels,
  2. motor-driven intracellular transport,

Other potential applications of Large Deviations techniques

- Epilepsy: usually involves very large neuronal populations.
- Piecewise deterministic processes:
  1. Stochastic ion channels,
  2. motor-driven intracellular transport,
  3. gene networks.