

NEUROENGINEERING

PART A

Author:
Gioele Migno - migno.gioele@gmail.com

02: THE NEURAL CELL – PART I

Learning objectives of the lesson

1. List the 3 main functions of the neural cell (neuron)
2. Describe the specialized structure allowing the neuron to carry out its functions and the nature of membrane potentials
3. Explain the role of the main ion families in the electrical behavior of the neuronal membrane
4. Describe the Nernst equation, the terms that appear in it and their meaning

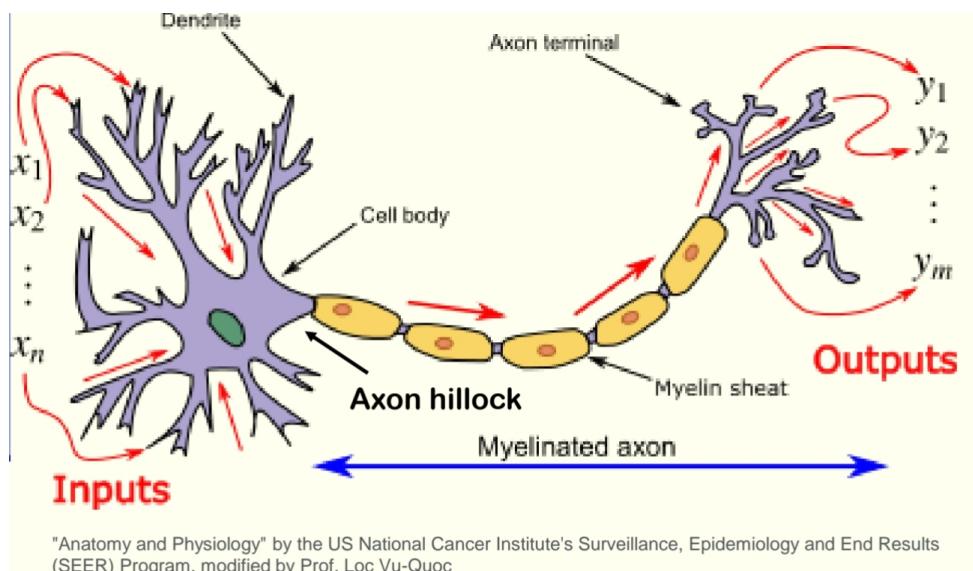
THE NEURON

The aim of a neuron is not only to transfer an electrical signal from a source to a destination, it has also to elaborate and process it, otherwise our brain would be a simple box of cables.

THE MAIN NEURON FUNCTIONS

The three main neuron functions are:

1. Collection of information from multiple sources $x_1 \dots x_n$ (other neural cells/receptors)
2. Integration (in time and space) of incoming information to provide a binary decision
3. Generation and propagation of a bit of information up to target cells $y_1 \dots y_m$ (other neural cells, muscle cells)



NOTE: The output signal is the same for all outputs y_i namely:

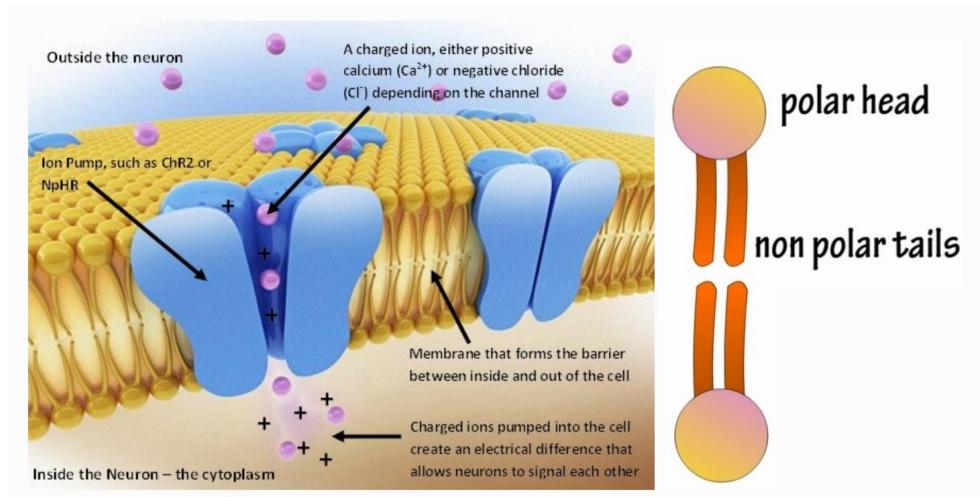
$$y_1 = y_2 = \dots = y_m$$

So, the neuron collects several inputs x_i , it adds them (integration) and finally produces an output signal that is duplicated between $y_1 \dots y_m$. This output signal is binary and it's produced in the **axon hillock** most of the time.

All of these functions are performed by **the membrane of the neuron**, so let's see now how it works and how it is able to do this.

THE NEURON MEMBRANE

The **neuronal membrane** is the main morphologically specialized structure of the neuron and as said before, it performs the three most important functions of a neuron.



The membrane is composed by a **phospholipid bilayer**, in which each layer is composed by a phosphoric head that is polar so it can connect to water (hydrophilic), and two lipidic tails that does not allow water to pass (hydrophobic). This double layer avoids the passage of charged particles (**ions**) but allows the passage of other molecules like alcohol and oxygen.

The main **ion families** we are most interested in are:

- **Sodium** (Na^+)
- **Potassium** (K^+)
- **Calcium** (Ca^{++})
- **Chloride** (Cl^-)

Membrane transport mechanisms

Inside the phospholipid bilayer there are also some proteins that allow the membrane to be **selectively permeable** to ions (electrically charged atoms or molecules).

For simplicity, we can say that there are two basic transportation mechanisms.

- **Ion channels:** Passive transportation, driven by electrochemical forces
- **Ion pumps:** Active transportation, requires an energy expenditure by the cell

These mechanisms allow ions to move into and out of the cell by opening and closing in response to voltage changes and to both internal and external signals.

OBS: For more info about membrane transport mechanisms:

<https://youtu.be/J5pWH1r3pgU>

Electrical signals

There are three different types of electrical signals that involve the neuron membrane:

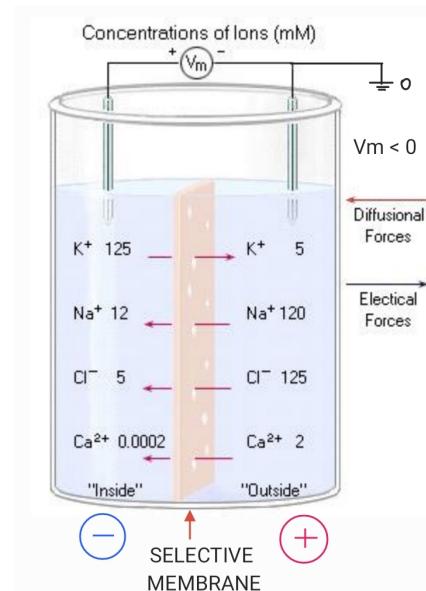
- **Membrane Potential:** Due to the different ion concentration between inside and outside the neuron
- **Membrane Current:** Due to ion movement between inside and outside the cell
- **Internal / External current:** Due to ion movement inside and outside the neuron

The three main functions introduced before are performed through the variation of these electrical signals.

Let's now see more in detail the membrane potential.

.MEMBRANE POTENTIAL

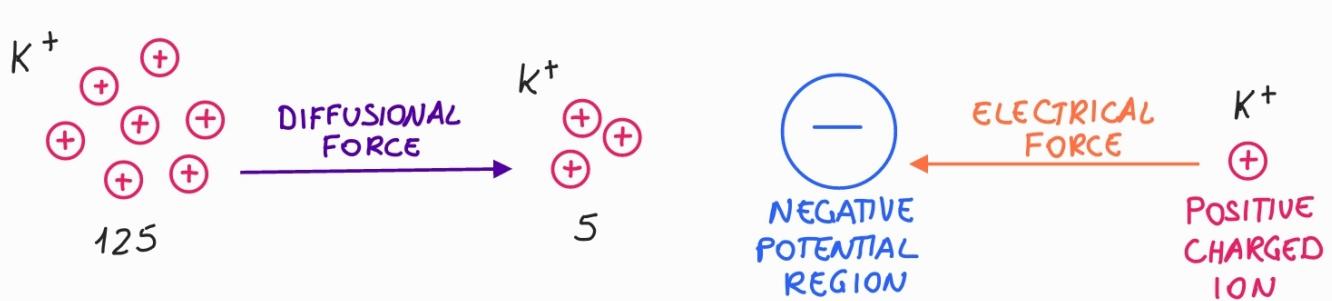
As mentioned before, the membrane potential is due to the different ion concentration between inside and outside the neuron. Since ions tend to reach an equilibrium, they are subject to two different kinds of forces, to understand that let's represent the membrane in the following way assuming a negative charge inside the neuron:



We measure the electrical potential inside the neuron by considering the outside as reference ($0V$), assuming a negative charge inside ($-$) and a positive one outside ($+$) we have a negative membrane potential $V_m < 0$. As we will see later, this is true most of the time, indeed the **membrane potential at rest** is about $-70mV$.

Each ion is subject to two different types of forces:

- **Diffusional Force:** Due to chemical gradient, namely different concentration of ions in the intra- and extra-cellular fluids
- **Electrical Force:** Due to electrical gradient, positive ions are attracted toward the region with a negative potential, and viceversa



The sum of diffusional and electrical forces leads to an electrochemical difference $\Delta\mu$.

Electrochemical difference

The **electrochemical difference** $\Delta\mu$ between two regions *A* and *B*, for a given ion family *X*, is given by:

ELECTROCHEMICAL DIFFERENCE

$$\Delta\mu = \Delta Chem + \Delta Elect = RT \ln \frac{[X]_A}{[X]_B} + zF(E_A - E_B)$$

Where

- $[X]_A$ = ion concentration on the A side of the membrane
- $[X]_B$ = ion concentration on the B side of the membrane
- R = universal gas constant
- T = temperature in Kelvin
- F = Faraday constant
- z = ion valence (e.g: +2 for Ca^{++} ; -1 for Cl^-)
- $E_A - E_B$ = membrane potential

The first term $\Delta Chem$ is related to the thermal energy of the molecules that moves ions according to the chemical gradient.

For a single ion this thermal energy is given by:

$$Therm_{ion} = K_B * T$$

Where:

- K_B is the Boltzmann constant
- T temperature in Kelvin

For one mole instead:

$$Therm_{mole} = R * T = N_A * K_B * T$$

Where:

- R is the universal gas constant
- T temperature in Kelvin
- N_A Avogadro's number
- K_B Boltzmann constant

The second term $\Delta Elect$ is instead related to the potential energy (lost or gained moving through an electric voltage equal to $(E_A - E_B)$). It moves ions according to the electrical gradient.

For a single ion:

$$Pot_{ion} = zq(E_A - E_B)$$

Where:

- q : charge of the proton
- z : valence
- $(E_A - E_B)$: Voltage

For one mole:

$$Pot_{mole} = zF(E_A - E_B)$$

Where:

- z : valence
- F : Faraday constant
- $(E_A - E_B)$: Voltage

Nernst equation

The sum and balance of diffusional and electrical forces leads to an equilibrium, when the electrochemical difference between two regions A, B is equal to zero ($\Delta\mu = 0$) we have the so called **electrochemical equilibrium** that leads to **Nernst equation**:

$$\begin{aligned}\Delta\mu = 0 &\Leftrightarrow RT \ln \frac{[X]_A}{[X]_B} + zF(E_A - E_B) = 0 \\ &\Leftrightarrow -zF(E_A - E_B) = RT \ln \frac{[X]_A}{[X]_B} \\ &\Leftrightarrow E_X \triangleq E_B - E_A = \frac{RT}{zF} \ln \frac{[X]_A}{[X]_B}\end{aligned}$$

Finally assuming *A* represents the extracellular region and *B* the intracellular region, we obtain the Nernst equation that allows to compute the **electrochemical equilibrium potential** E_j for each ion family *j*:

NERNST EQUATION

$$E_j = \frac{RT}{zF} \ln \frac{[j]_{extra}}{[j]_{intra}}$$

Where:

- $[j]_{extra}$ = extracellular concentration of ion *j*
- $[j]_{intra}$ = intracellular concentration of ion *j*
- z = valency of ion *j*

Electrochemical equilibrium potential

The electrochemical equilibrium potential E_j is the voltage at which the diffusional forces balance the electrical ones for the *j* ions, this quantity is important because it tells us in which direction *j* ions moves when the membrane potential is equal to V_m :

NOTE: *j* ions with an equilibrium voltage equal to E_j , when free to cross the membrane, will flow with net currents that move the membrane potential V_m to that value.

TYPICAL VALUES

Potassium: $E_{K^+} = -90mV$

Sodium: $E_{Na^+} = 50mV$

Calcium: $E_{Ca^{++}} = 150mV$

The goal of a j ion is to bring the membrane potential to its electrochemical equilibrium potential E_j , if the ion is positive charged then it can decrease V_m (more negative) by moving from inside to outside of the cell:

- POSITIVE ION: in \rightarrow out $\implies V_m \downarrow$ (decrease)
- POSITIVE ION: in \leftarrow out $\implies V_m \uparrow$ (increase)

If the ion is negative, then the opposite happens.

Let's see now two examples in which two positive ions have an opposite behavior

Example 1 (Potassium)(in->out)

Assuming the membrane potential equal to $V_m = -70\text{mV}$, and potassium ions K^+ free to move between inside and outside of the cell. In which direction ions will flow?

—

Since K^+ is a positive ion, usually it tends to flow towards the negative region, however in this case there is also the diffusional force that must be taken into account. Using the electrochemical equilibrium potential of potassium E_K^+ we can easily obtain the right flow direction. In this case we have:

$$E_{K^+} = -90\text{mV} < -70\text{mV} = V_m$$

Therefore ions want to reduce V_m and to do that they move **from inside to outside**

Example 2 (Sodium) (out->in)

Assuming the membrane potential equal to $V_m = -70\text{mV}$, and sodium ions Na^+ free to move between inside and outside of the cell. In which direction ions will flow?

—

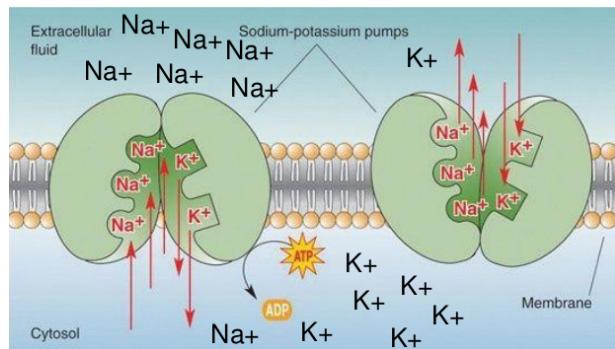
In this case we have:

$$E_{Na^+} = 50\text{mV} > -70\text{mV} = V_m$$

This time, ions want to increase V_m and to do that they move **from outside to inside**

.MEMBRANE POTENTIAL AT REST

Until now, we have seen only the passive movement of ions between inside and outside of the cell, with the term *passive* we mean the neuron has not to spend any energy for it. To work properly, the neuron membrane must have a specific potential called **membrane potential at rest** equals to about -70mV , only diffusional and electrical forces (passive mechanism) do not allow to reach this specific potential, therefore the cell uses an active mechanism based on **ion pumps**. Ion pumps are specific proteins that move ions against their electrochemical gradient by active (ATP energy consuming) transport. The most important one is the sodium-potassium pump :



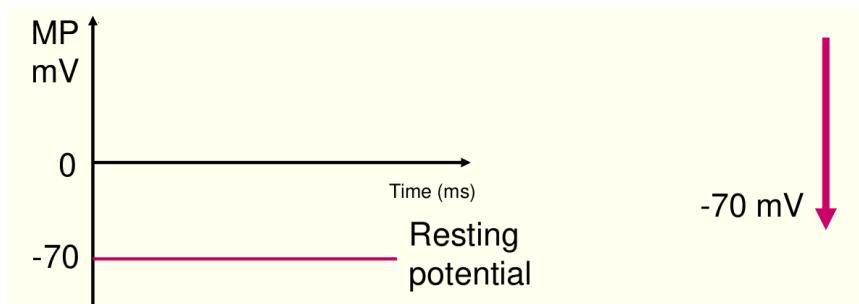
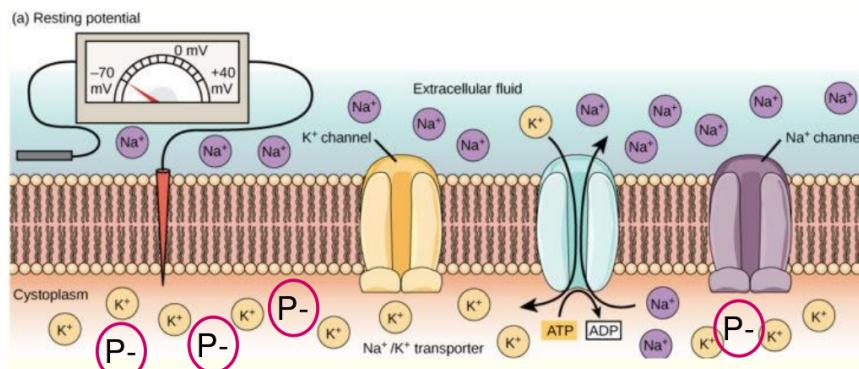
To conclude, the rest potential of the membrane is the combination of three mechanisms:

$$\text{REST POTENTIAL} = \text{DIFFUSIONAL FORCES} + \text{ELECTRICAL FORCES} + \text{ION PUMPS}$$

.VARIATION MEMBRANE POTENTIAL

As said before, the membrane potential (MP or V_m) is the difference in electrical potential between the interior of a neuron and the surrounding extracellular fluid, this last is used as reference ($0v$). The MP is due to the different ion concentrations on the two ends of the membrane, and at rest (unperturbed membrane) it's around $-70mV$, in some books or papers, this rest potential is indicated as $-65mV$ or so.

When the membrane is at rest potential (**resting potential**) then the cell membrane is said to be **polarized**:



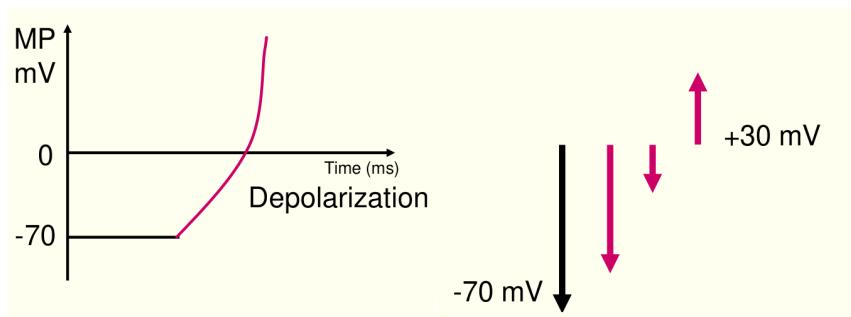
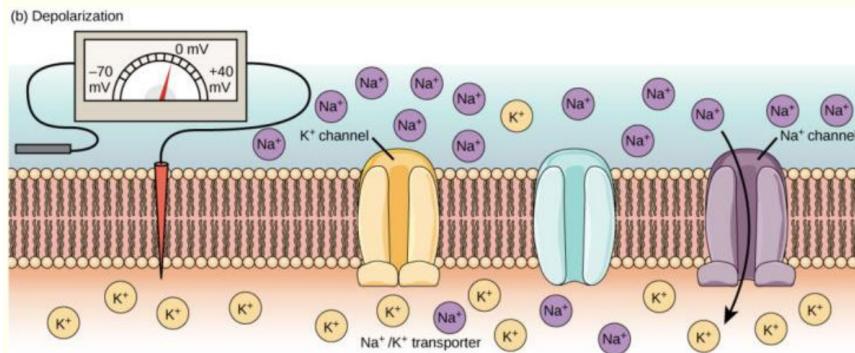
In order to perform all the functions explained before, (Collection of information - integration - Generation and propagation of a bit of information), the membrane changes its potential. The variations of the MP have specific names:

- **Membrane Depolarization:** The MP increases ($MP \uparrow$)
- **Membrane Polarization / Repolarization:** The MP decreases and **returns** to the rest potential ($MP \downarrow$)
- **Membrane Hyperpolarization:** The MP decreases **over** the rest potential ($MP \downarrow\downarrow$)

Let's see in more details how the MP increases and decreases.

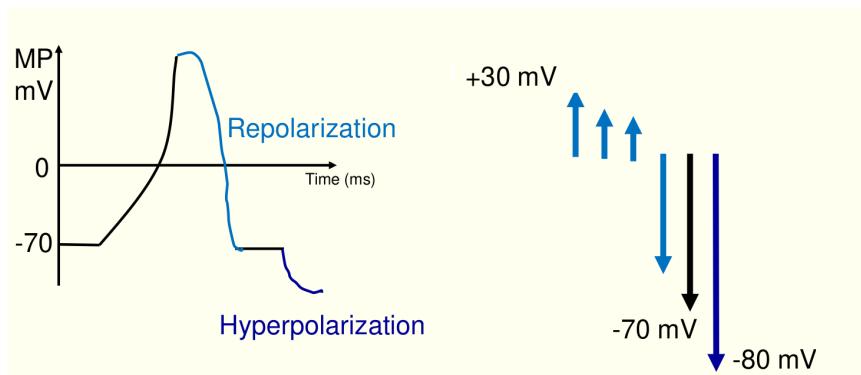
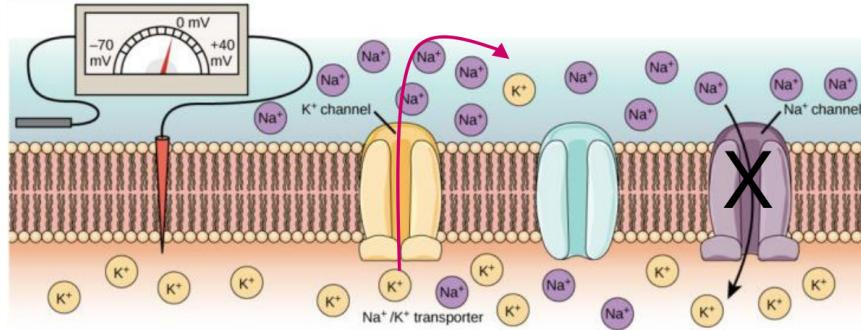
Membrane Depolarization

Membrane depolarization is due to a current in the form of positively charged ions flowing into the cell (or negatively charged ions flowing out of the cell) makes the **membrane potential less negative** or even positive.



Membrane Repolarization / Hyperpolarization

Membrane repolarization and membrane hyperpolarization are due to a current in the form of positively charged ions flowing out of the cell (or negatively charged ions flowing into the cell) makes the **membrane potential more negative**.



Self-evaluation test

1. What are the 3 main functions of the neural cell?
2. What are the four main ion families having a role in the neuron functioning?
3. How is the resting membrane potential determined? What value does it assume?

References

- Dayan & Abbott:
- Chapter 1, section 1.1
- Chapter 5, section 5.2
- Hari & Puce:
 - Section I, Chapter 2, pagg.17-18 (Sections: Communication between brain areas; Electric signaling in neurons)
- <https://youtu.be/J5pWH1r3pgU>

02: THE NEURAL CELL – PART II

Learning objectives of the lesson

1. **Understand** how the information is collected by the cell post-synaptic membrane and **tell** the difference between excitatory and inhibitory synapses (1st function)
2. **Explain** how the analog, multiple information collected by the neuron is translated into a binary decision (output) (2nd function)
3. **Illustrate** the nature of the neuronal cell output signal (action potential) (3rd function)
4. **List** the sequential events that lead to the generation of the action potential (3rd function)
5. **Describe** the two mechanisms for the propagation of the action potential along the axon (3rd function)

DENDRITIC TREE

As already said, the first important function of a neuron is to collect information from other neural cells, in this section we try to understand how this mechanism works.

The collection of information from other neural cells is performed by a part of the neuron called **dendritic tree**. This part has a tree structure in order to maximize as much as possible its surface.

The dendritic tree can be different between several kinds of neural cells, therefore the number of connections can be different.

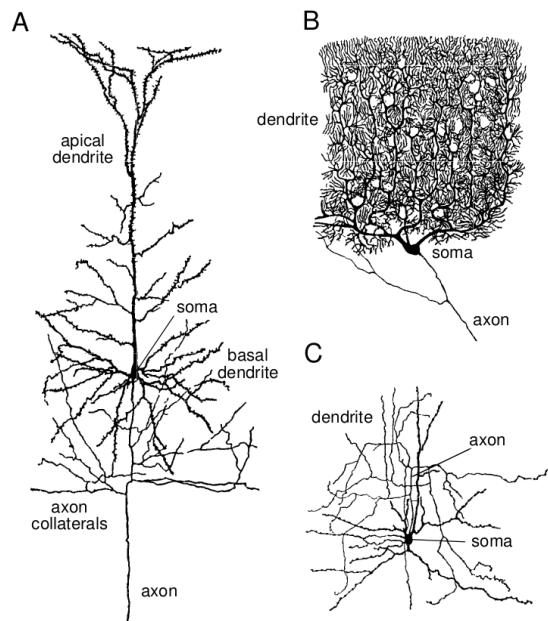


Figure 1.1 Diagrams of three neurons. (A) A cortical pyramidal cell. These are the primary excitatory neurons of the cerebral cortex. Pyramidal cell axons branch locally, sending axon collaterals to synapse with nearby neurons, and also project more distally to conduct signals to other parts of the brain and nervous system. (B) A Purkinje cell of the cerebellum. Purkinje cell axons transmit the output of the cerebellar cortex. (C) A stellate cell of the cerebral cortex. Stellate cells are one of a large class of interneurons that provide inhibitory input to the neurons of the cerebral cortex. These figures are magnified about 150-fold. (Drawings from Cajal, 1911; figure from Dowling, 1992.)

An exchange of information between two or more cells happens through the **synaptic connections**, after that, the neuron aggregates all the inputs by performing a summation in time and space. We will see this last later, let's now introduce what is a synapse, namely the structure that permits a synaptic connection.

.SYNAPSE

The dendrite tree has the same membrane of the other neuron's parts, however in addition to the structures already seen, it contains also a new kind of structure called **synapse** needed to connect two or more cells.

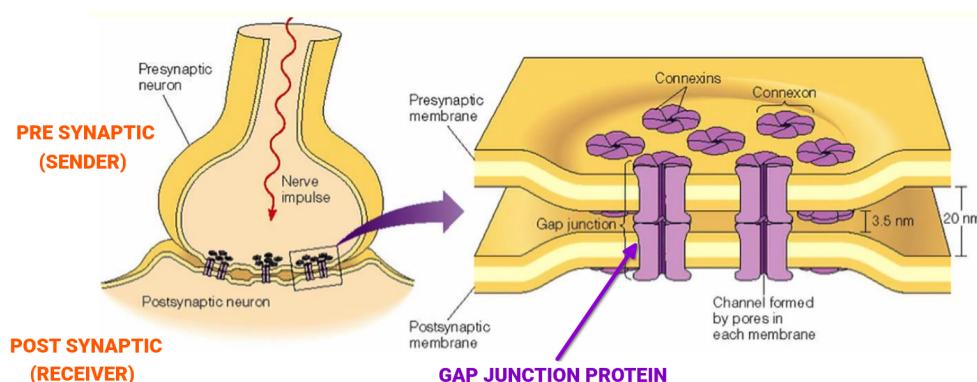
A synapse is a mechanism present not only in a neural but also in another cells, there exist two types of synapses:

- **Electrical Synapse:** Based on electrical signals (Not common in the brain)
- **Chemical Synapse:** Based on chemical signals (Common in brain)

In the brain the most common synapse present in a neuron are chemical synapses, let's see briefly also the electrical types and then we will go deeper on chemical synapse.

ELECTRICAL SYNAPSE

Electrical synapses are not so common in the brain, however they are essential in the nervous system. An electrical synapse allows a direct signal transmission from an electrical source (**PRE SYNAPTIC**) to a electrical destination (**POST SYNAPTIC**), we use this terms since in most of the cases the exchange of information is **mono-directional** (SENDER->RECEIVER). This type of synapse allows a physical contact between the two cells through a protein called **Gap Junction**.

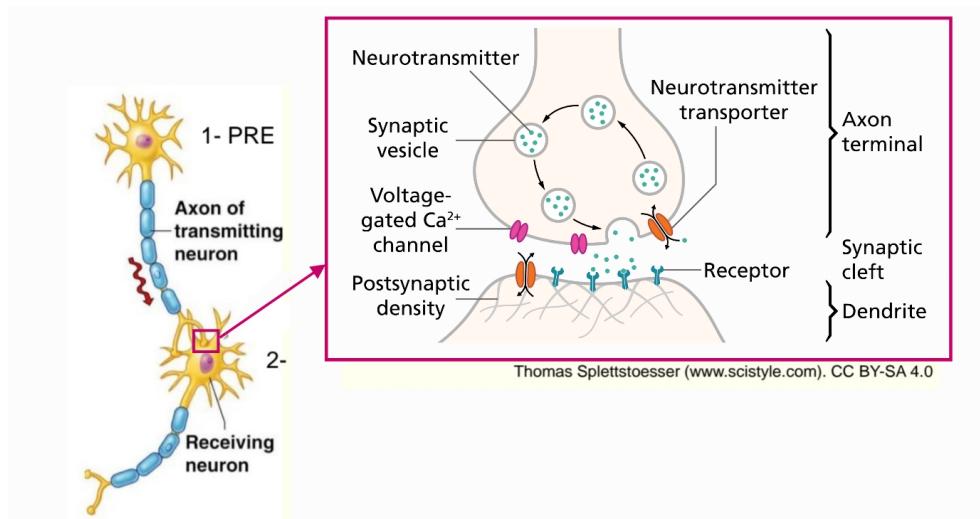


The main advantage of an electrical synapse is the speed of the transmission, this is due to the use of only pure electrical signals.

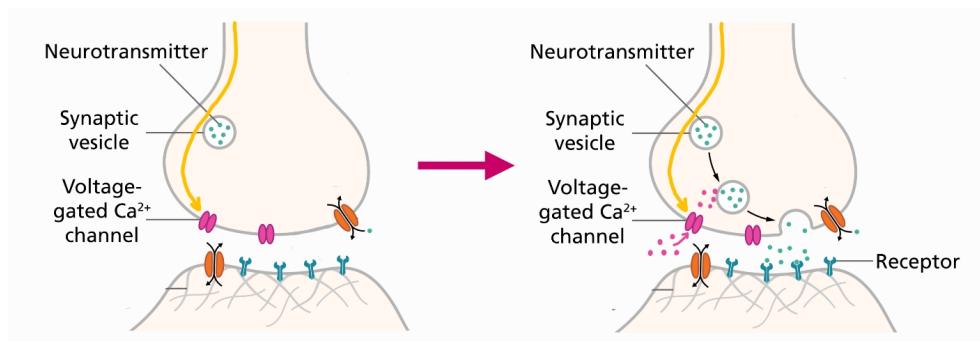
CHEMICAL SYNAPSE

The most common type of synapses in the brain is the **chemical synapse**, they are more complex and slower than the electrical ones, however they are more flexible and essential in the brain for important functions like learning and memory. A chemical synapse is slower since the signal transmission is not pure electrical but it is mediated through a chemical signal.

The two cells are not in physical contact, between them there is a gap called **synaptic cleft**, inside it, the exchange of **neurotransmitter** happens. At rest, the neurotransmitters are contained in the so called **synaptic vesicles**:



The exchange of information starts when the axon of the PRE-SYNAPTIC neuron sends an electrical signal to the **voltage gated calcium channels** (Ca^{2+}). These channels allow the passive entrance of calcium ions (Ca^{2+}) that, binding to the synaptic vesicles, permit the release of the neurotransmitters. Then the neurotransmitters bind to the **receptors** present on the dendrite of the receiver neuron, the activation of these receptors cause the opening of specific **ion gated channels** that allow the passive entrance/exit of ions from the receiver neuron.



See: https://www.youtube.com/watch?app=desktop&v=ecGEcj1tBBI&ab_channel=O2Labz

NOTE: In the following we refer to chemical synapses

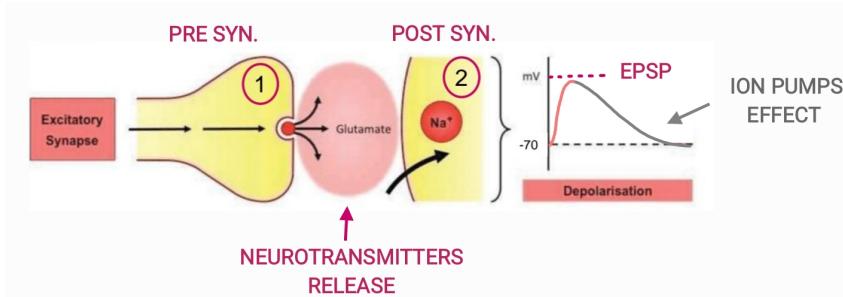
.POST SYNAPTIC POTENTIAL (CHEMICAL SYNAPSES)

Let's focus our attention on **chemical synapses**

As we already know, the movement of ions change the membrane potential of a neuron, after the activation of the receptors on the postsynaptic neuron, its MP reaches the so called **post synaptic potential (PSP)**. According to the effect (depolarization or hyperpolarization) we say that a synapse is **excitatory** or **inhibitory** respectively.

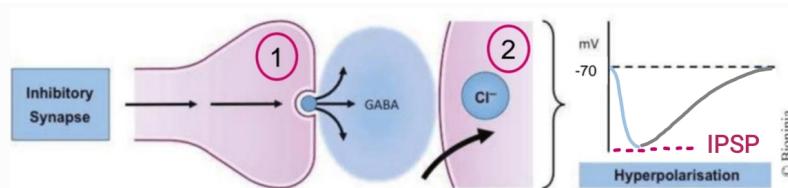
Excitatory synapse

We have an **excitatory synapse** when the PRE-SYNAPTIC neuron releases an excitatory neurotransmitter like Glutamate, after that the MP of the POST-SYNAPTIC cell reaches the **Excitatory Post Synaptic Potential (EPSP)** due to the passive entrance of potassium ions (Na^+). After a while the MP returns to the rest potential thanks to the ion pumps action.



Inhibitory synapse

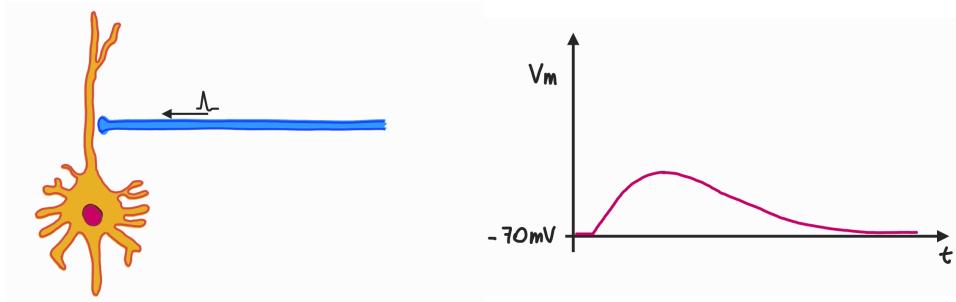
On the other hand, we have an **inhibitory synapse** when an inhibitory neurotransmitter like GABA is released, in that case the POST-SYNAPTIC reaches the **Inhibitory Post Synaptic Potential (IPSP)** due to the entrance of negative chloride ions (Cl^-)



OBS: The excitatory and the inhibitory effects refer to excitation and inhibition of the neuronal response, indeed we will later that an excitatory synapses stimulates the generation of an electrical signal that propagates through the neuron, an inhibitory synapses instead, does the opposite.

.SUMMATION OF PSPs

Assuming a single excitatory synapse we have the following scenario in which the MP (V_m) increases:



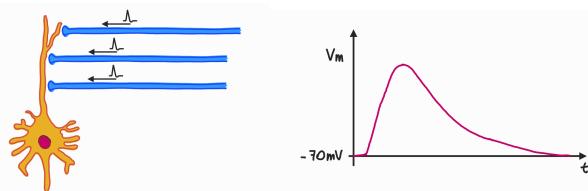
However, as previously said, a neuron is usually connected to a lot of other cells, therefore a single neuron can have multiple chemical synapses, since their effects are summed, the final PSP is given by the **summation of all PSPs**.

We distinguish two different types of PSPs summations:

- **Spatial Summation:** The effects of different synapses are summed
- **Temporal Summation:** The effects of several signals on the same synapse are summed

Spatial Summation

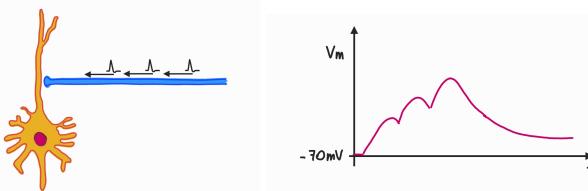
In the case of **spatial summation**, several different PRE-SYNAPTIC neurons firing at the same time but at different synapses. Their effects are summed, assuming only excitatory synapses, we have a higher PSP respect to the single synapse case:



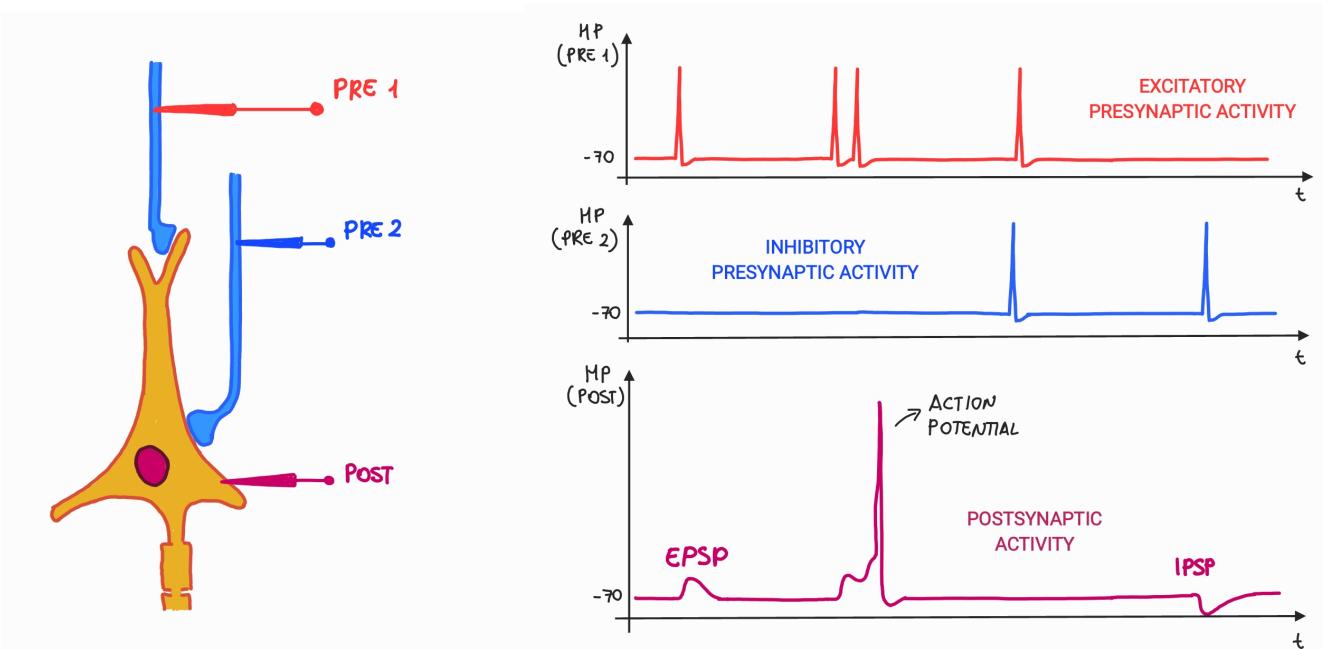
OBS: In general, we can have different synapses with different effects, therefore could be also the case in which two effects cancels each other out

Temporal Summation

The other type of summation is the **temporal summation** in which same or nearby PRE-SYNAPTIC neuron firing multiple times in close succession. Assuming excitatory synapses:



To summarize the two types of summations and the possible effects:

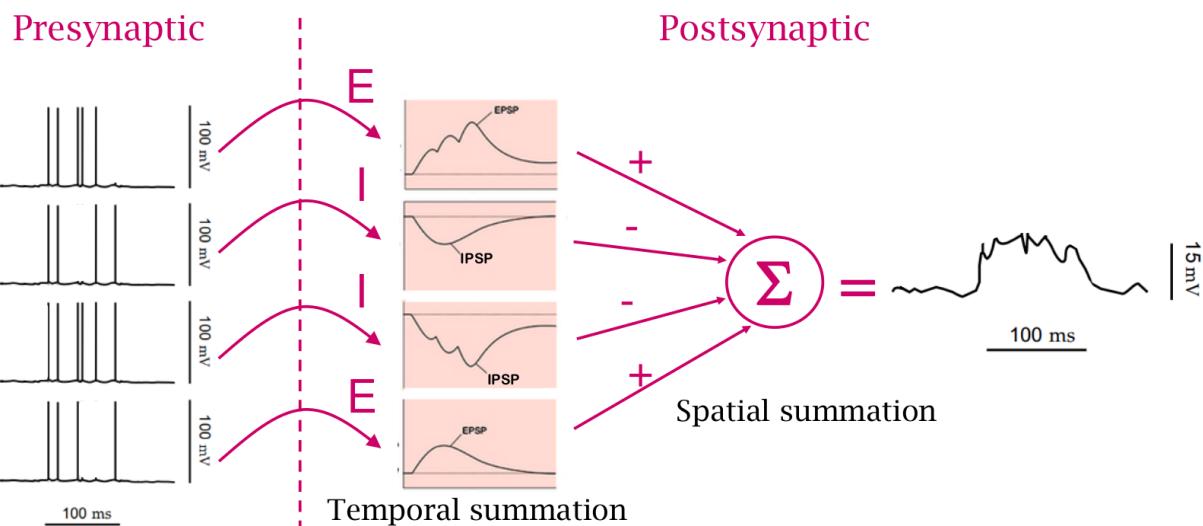


In the figure we can see a spike in the POST-SYNAAPTIC neuron MP caused by a temporal summation, this is a specific fixed signal produced by the membrane called **action potential**, we will see more about that later.

NOTE: Although it is not shown in the figure, we can have a temporal summation and a spatial summation both at the same time.

Overall summation

As just said, temporal and spatial summations can be combined, therefore from a series of binary inputs, inside the neuron, a continuous signals is generated (variation of the membrane potential), if this signal is above a specific threshold then the action potential is produced. (see later)

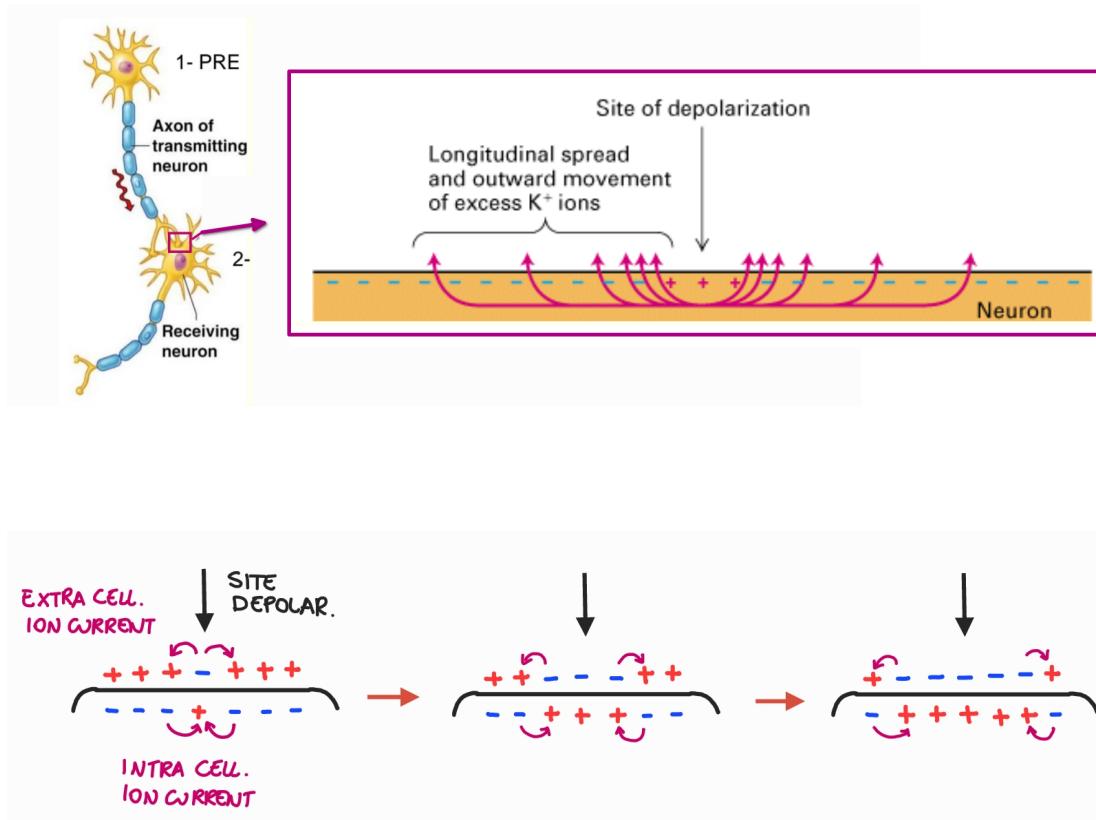


INTEGRATION OF INFORMATION

We have seen that the information processing by the neuron consists of the summation (with sign) of EPSPs and IPSPs, this summation is both spatial and temporal and it produces the membrane depolarization or hyperpolarization. This **integration of information** happens in the dendritic tree. The result (POST-SYNAPTIC potential) is then propagated along the membrane up to the **axon hillock** where the binary output decision is made. This binary decision consists of firing or not an electrical signal called **action potential**.

PROPAGATION OF THE POST-SYNPATC POTENTIAL

A local de/hyperpolarization inside the dendritic tree, causes **intracellular and extracellular ion currents**. This current is possible since the ions are free to move. Indeed, at rest the movements of ions between the membrane is not free, however it is inside each region where there is an homogeneous potential (at rest potential the inside of the neuron is negative and the outside is positive). So the POST-SYNAPTIC potential (variation of MP) is passively propagated to adjacent sections of the membrane through this ions current. This perturbation effects decrease with distance.



THE ACTION POTENTIAL

The **action potential** is a variation of the membrane potential which appears only in neural, muscular and cardiac cells. It's an all-or-none process: If the input stimulus does not reach a given threshold, it does not happen. If the threshold is reached, it has always the same shape, duration and intensity, irrespectively to the input stimulus amplitude.

So, the action potential represents the binary output of the neuron after input signals from the other cells. As already said, the acquisition and the integration of the inputs happen in the dendritic tree, the resulting POST-SYNAPTIC potential is then propagated up to the axon hillock where, according to the signal amplitude, the action potential is generated (HIGH OUTPUT) or not (LOW OUTPUT).

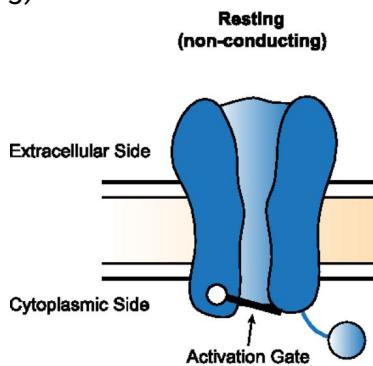
.THE VOLTAGE-GATED Na^+ AND K^+ CHANNELS

At the base of the action potential generation there are two different kinds of **voltage-gated ions channels**. Let's see now how they work.

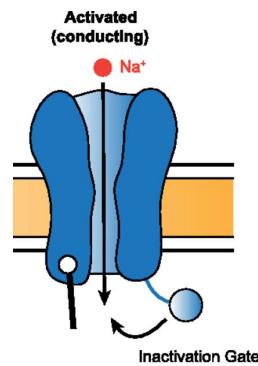
Voltage-gated Na^+ channel

The **voltage-gated Na^+ channel** is the most complicated between the two. It can be in three different states:

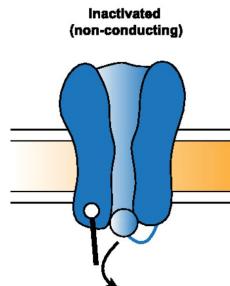
- **Resting State:** (non-conducting) In this state the **activation gate** is closed



- **Activated State:** (conducting) The activation gate is open and then it allows the ion passage

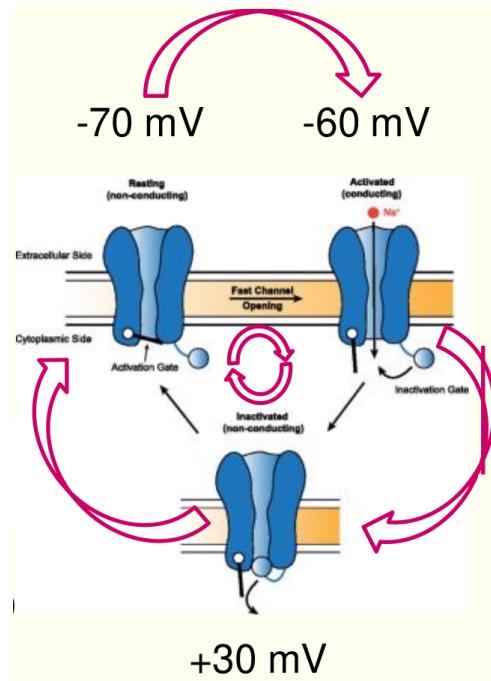


- **Inactivated State:** (non-conducting) The **inactivation gate** is closed, it does not allow the ion passage and neither the opening of the activation gate



This voltage-gated has three different voltage thresholds, let's start with the membrane at rest potential ($-70mV$), at this potential the gate is closed and then ions cannot cross it, however the gate can be open if the potential reaches a specific threshold:

- **Opening threshold potential:** (-60mV) Gate is open, ions can cross
- **Inactivation threshold potential:** ($+30\text{mV}$) (strong depolarization) The gate is inactivated (it cannot be open)
- **Closing threshold potential:** (-70mV) The gate returns at the initial closed state (ions cannot cross, but the gate can be open again)

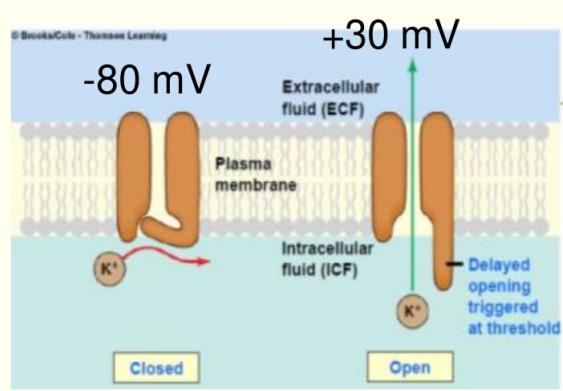


Voltage-gated K^+ channel

The other voltage-gated of interest for action potential generation is the **voltage-gated K^+ channel**. This gate is simpler and has only two states (open or close*) and two potential thresholds.

At resting potential (-70mV) The gate is closed, ions cannot cross it, but it can be open:

- **Opening threshold potential:** ($+30\text{mV}$) The gate opens, ions can cross
- **Closing threshold potential:** ($< -70\text{mV}$) The gate returns to initial close state (ions cannot cross, it can be open)



So in this case we have only two thresholds, at rest the gate is closed, then when the membrane potential depolarizes reaching $+30\text{mV}$, the gate opens and allow the ions to cross. To close the gate and come back to the initial state the membrane potential has to reach a potential **below** the rest potential (-70mV).

ACTION POTENTIAL SIGNAL

Once introduced the previous voltage-gated channels, we can see the shape of the **action potential signal** and understand each phase.

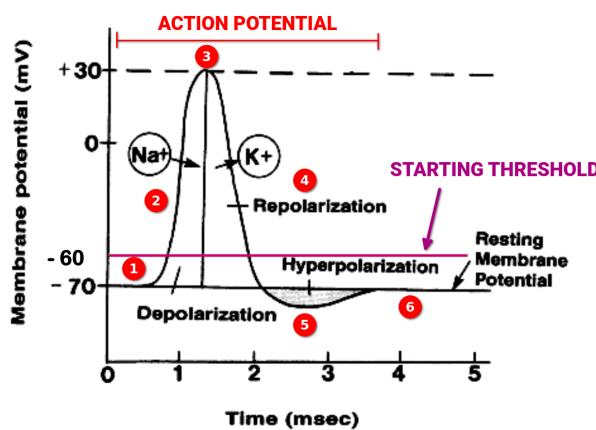
Let's recall all the events before a possible action potential firing. The neuron gets several inputs through the synapses in the dendritic tree, each signal can be an excitatory or an inhibitory one. The first cause the depolarization and the second a hyperpolarization. All of these signal is signed summed and the result is propagated up to the axon hillock, here the action potential can be fired.

At this point there are three possible scenarios according to the resulting membrane potential:

- **Depolarizing resulting potential:** (inhibitory potential) Nothing happens
- **Low polarizing resulting potential:** (low excitatory potential) The membrane polarization is not enough
- **High polarizing resulting potential:** (excitatory potential) The membrane polarization is above a threshold ($-60mV$), then the action potential starts

In the third case, the action potential starts and ends always in the same way (same timing and shape). Let's see each phase and how the voltage-gated channels intervene in the generation.

1. **Initial depolarization:** Depolarization to the Na^+ opening threshold level ($-60mV$)
2. **Fast depolarization:** Due to the Na^+ depolarizing currents (Na^+ flows into the cell)
3. **Switch channels:** Inactivation of the Na^+ channels and opening of the K^+ channels ($+30mV$)
4. **Repolarization:** Due to the K^+ repolarizing currents (K^+ flows outside of the cell)
5. **Undershoot:** (hyperpolarization) until the K^+ channels are closed
6. **Return resting potential:** Due to the sodium-potassium pump (ions pumps)



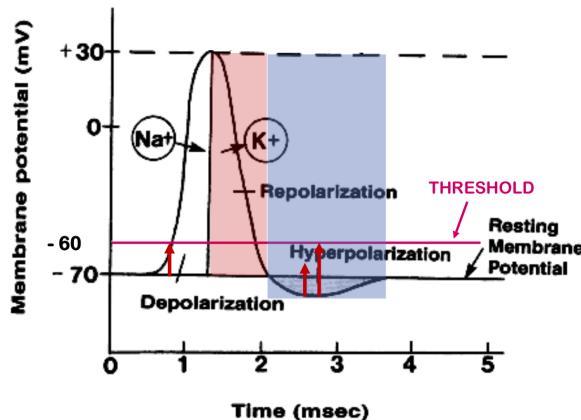
OBS: (Importance hyperpolarization)

The phase 5 (hyperpolarization) is important to inhibit other excitatory input signals and reduce the probability of another action potential firing

Absolute and Relative refractory periods

After the action potential peak, we distinguish two different regions that depend to the probability for which another action potential can occur:

- **Absolute Refractory Period:** (impossible) Due to the Na^+ voltage-gated channels inactivation, no new action potential can be produced (under any circumstances)
- **Relative Refractory Period:** (less likely to occur) Due to the K^+ voltage-gated channels. A new action potential can be produced, but it requires a stronger depolarization so it's less likely to happen.



We say that the absolute refractory period is due to the Na^+ gate since, once inactivated, it cannot be open again until reached the rest potential (-70mV), therefore it avoids an another action potential under any circumstances.

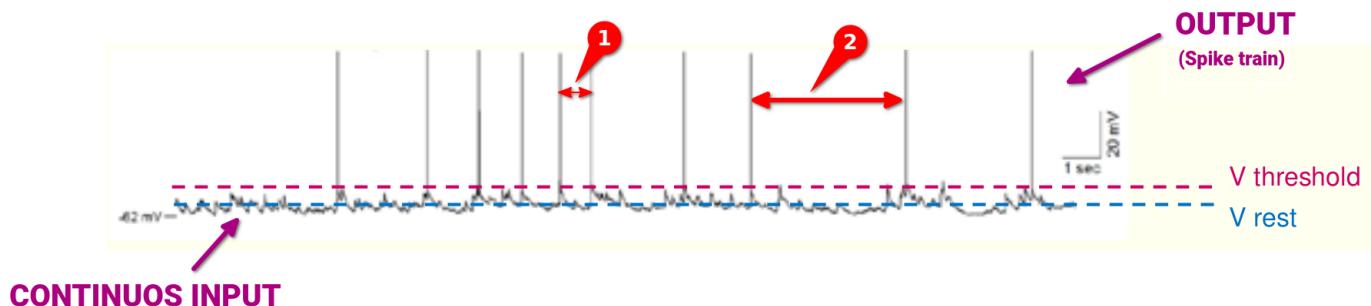
The relative refractory period is due instead to the K^+ gate since it causes the hyperpolarization that makes more difficult to reach the depolarization threshold necessary to another action potential.

NOTE: (Importance refractory periods)

We will see later that these two refractory periods are important for our applications in the brain modeling.

.INFORMATIVE CONTENT

We have seen that a neuron gets input signals from previous neurons through chemical synapses, it processes the aggregated signal (continuous signal) and according to a threshold, an action potential is produced or not (binary output). Therefore a neuron switches from a continuous input to a binary output (spike train).



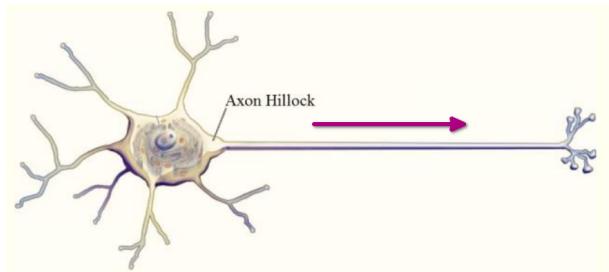
The action potential has always same shape, same duration (2ms), and same amplitude (100mV). So it is similar to a spike namely a signal very fast and very intense. Hence, the **information content** is not in its shape or amplitude, but in the temporal distance between two spikes.

The train of spikes is then propagated to the next neuron, this last, according to its resulting input, produces an action potential or not. Assuming it has only one input that corresponds to the spike train shown in the previous figure, it produces an action potential in the first case (1) and not in the second one (2) since in this last, the distance between the two spikes is too big therefore there is not a temporal summation that permits to reach the depolarization threshold.

PROPAGATION ACTION POTENTIAL

Let's come back to consider a single action potential signal.

Once generated in a specific point of the membrane (usually the axon hillock) the action potential needs to be propagated along the axon toward other cells.



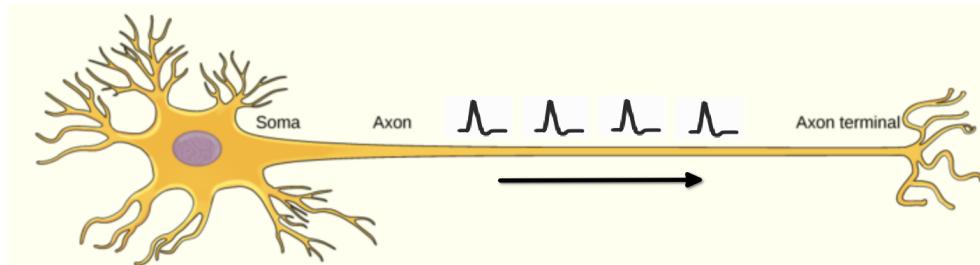
According to neuron type, there are two main mechanisms:

- **Continuous conduction** (unmyelinated fibers)
- **Saltatory conduction** (myelinated fibers)

Let's see both of them.

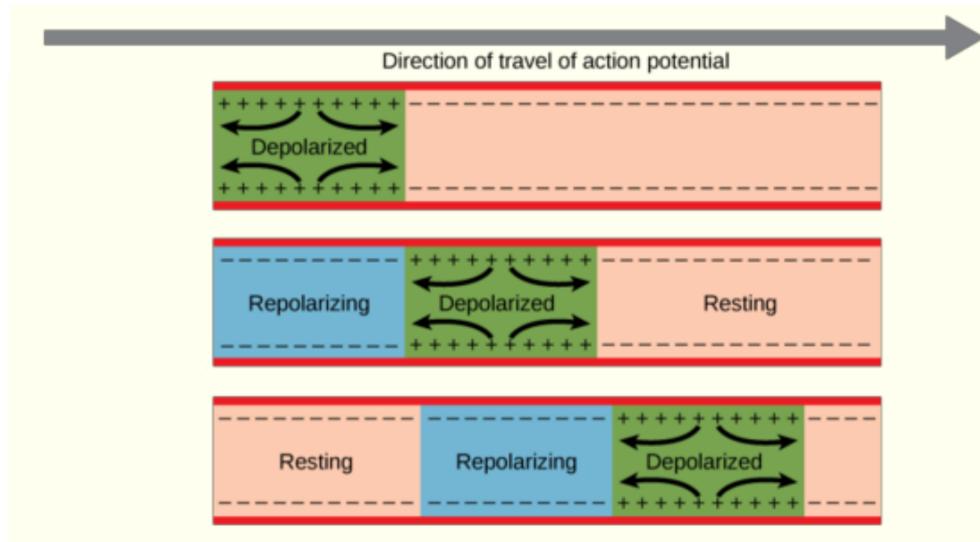
.CONTINUOUS CONDUCTION

The first mechanism is similar to the propagation of signals from dendritic tree to axon hillock, however in this case the action potential is **continuously generated** at each membrane section



The fast membrane depolarization induces intra- and extra cellular ionic currents (according to electrochemical gradients) that propagate the perturbation to close membrane regions. In the nearby regions, the membrane has the Na^+ K^+ voltage-gated channels, therefore it is able to produce an action potential. The production of a new spike depends on the current state of the region:

- **Na^+ gate ready to open:** The depolarizing perturbation produces again an action potential (active generation)
- **Na^+ gate inactive:** If the close membrane region has just produced an action potential, the Na^+ channels will be in their absolute refractory period hence a new action potential will be impossible to produce



OBS: (Importance absolute refractory period

Absolute refractory causes the unidirectionality of the AP propagation

The propagation speed of this type of conduction is limited by the activation of Na^+ and K^+ channels ($0.5 - 2m/s$). In general the speed is about $1m/s$. (slow)

OBS: (Slow and inefficient process)

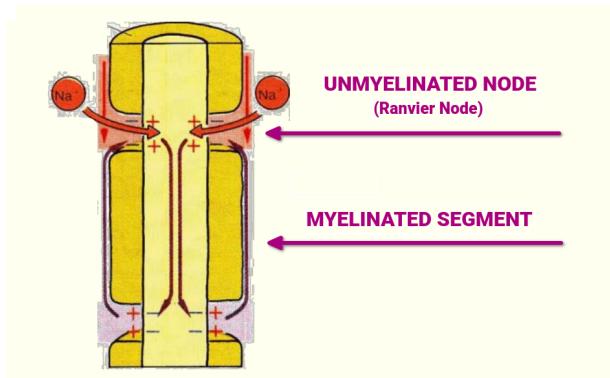
This mechanism of conduction is slow and very inefficient since it requires the generation of an action potential at each step (heavy process). For that reason, this kind of conduction is NOT so common in our neurons.

.SALTATORY CONDUCTION

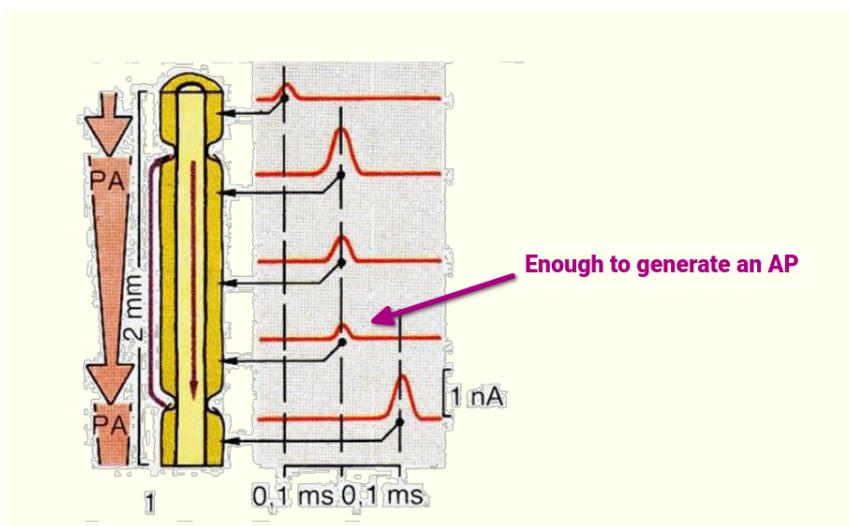
The second mechanism used to conduct a signal is called **saltatory conduction**, unlike the continuous conduction, it combines both active and passive conduction methods.

It is composed by two different parts:

- **Myelinated Segment:** Segment of axon surrounded by **myelin** that is a lipid-rich (fatty) substance able to isolate it avoiding the generation of a new action potential
- **Ranvier Nodes:** (unmyelinated nodes) Part of the axon in which there is an interruption of the myelinated sheath that allows an action potential generation



So in saltatory conduction, the action potential is generated only at the Ranvier nodes where we have interruptions of the myelinated sheath, then along the myelinated segments, the propagation is passive until the next node. Therefore along the axon we have an alternation of passive conduction, where the sheath prevents the membrane currents, and generation of the action potential at the unmyelinated nodes (active conduction).



The saltatory conduction is much faster than the continuous one and moreover allows signal propagation along longer distance, upon centimetres. However, both of them are used in our body, the continuous conduction is used to transmit delayed signals like pain, the saltatory one instead, is used to send rapid signals like commands to muscles

Self-evaluation test 1

1. How is the membrane potential modified by an excitatory synapse? And by an inhibitory one?
2. What's the difference between temporal and spatial summation? Can they occur simultaneously?
3. Why is a depolarizing post-synaptic potential called "excitatory"?
4. What is the use of an inhibitory PSP
5. In a chemical synapses, when a neurotransmitter opens the Na^+ gated channels, the resulting PSP is an inhibitory one (T/F)
6. Two PSPs can never cancel each other (T/F)
7. The effects of temporal and spatial summation are also reciprocally summed up (T/F)

Self-evaluation test 2

1. Do we need to measure the amplitude and duration of an action potential each time it occurs to understand the cell behavior?
2. Which parameter of the spike train in output to a neuronal cell is the most informative:
 - A. The amplitude of the spikes
 - B. The duration of each spike
 - C. The temporal distance between spikes
 - D. The spatial position in which the spikes are generated
3. What will the frequency of the spikes influence:
 - A. The temporal summation of the PSPs
 - B. The spatial summation of the PSPs
 - C. The amplitude of the resulting action potential in the post-synaptic cell
4. The absolute refractory period is due:
 - A. To the chemically controlled Na^+ channel
 - B. To the chemically controlled K^+ channel
 - C. To the voltage-gated Na^+ channel
 - D. To the voltage-gated K^+ channel
5. The saltatory (myelinated) conduction is faster than the continuous one (T/F)
6. A hyperpolarization of 10 mV with respect to the resting potential causes the generation of an action potential (T/F)
7. A depolarization of 20 mV causes a stronger action potential than a depolarization of 10 mV (T/F)

References

- Dayan & Abbott:
 - Chapter 1, section 1.1
 - Chapter 5, sections 5.5, 5.8
 - Chapter 6, sections 6.4
- Hari & Puce:
 - Section I, Chapter 2, pag.18-23 (Sections: Communication between brain areas; Electric signaling in neurons)

03- PRINCIPLES OF NEUROANATOMY AND BRAIN ORGANIZATION

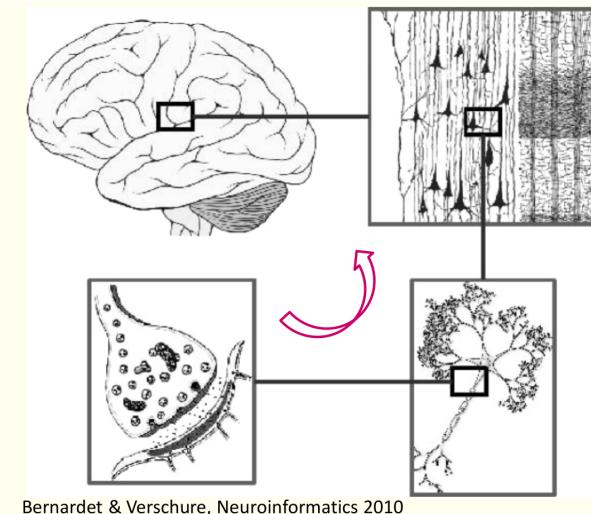
Learning objectives of the lesson

1. Illustrate the **temporal** and **spatial** scales of brain activity
2. Describe how cortical neurons organize in **layers, columns, regions** (Brodmann areas), **lobes, systems**
3. List the **main functions** of the **four cortical lobes**
4. Compare the location and structure of **cortical** and **subcortical** areas
5. Provide examples of two brain regions organized in **feature maps**
6. Illustrate the concept of **brain plasticity** and its role in the human lifespan
7. Explain the **four main neuronal mechanisms** behind brain plasticity

INTRODUCTION

.NEURAL ORGANIZATION

The brain structure is very complex and it is organized in several levels, from the entire brain upon a part of a neuron cell.



Bernardet & Verschure, Neuroinformatics 2010

.STUDYING THE HUMAN BRAIN

We started the study of the brain by exploiting subjects with brain lesion, knowing both the areas damaged and the function lost, we were able to conclude their connection. We studied the brain also by animal studies.

These primitives techniques are no more necessary thanks to neuroimaging (take picture of the brain) and neuromodulation (safely stimulate some parts of the brain). Now we are also able to link to behavioral and clinical data.



We can summarize the goal of our brain with the following statement:

The brain predicts the future on the basis of the past, helping the individual to survive and perpetuate the species

.THE BRAIN IN TIME AND SPACE

Brain functioning at the level of a single neuron, evolves in a temporal scale of milliseconds (e.g. action potential) and in a spatial scale of μm and nm (e.g. soma: μm , membrane thickness: nm). Brain evolve over time both from a collective point of view (evolution of the human species) and from an individual one. Regarding this last evolution, the brain is plastic during the entire lifespan (e.g. learning, memorization, spontaneous recovery, neurorehabilitation)

.GREY AND WHITE MATTER

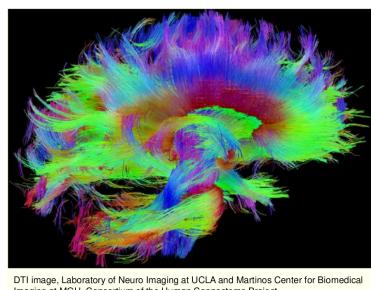
In the brain we refers to two different kinds of matters both of them are part of neuron cells:

- **Grey matter:** Cell bodies and dendrites of the neurons
- **White matter:** (fibers) myelinated axons of the neurons

These matters can be also classified according to their functionalities:

- **Grey matter:** Matter related to Information processing
- **White matter:** Matter related to information propagation

The white matter represents the information transport system of the brain, the highway of information between brain hemisphere is called *corpus callosum* and it is made of this matter. Using an in vivo procedure based on diffusion tensor imaging called DTI, it is possible to obtain the white matter tractography that represents the brain network and also direction of information flow. DTI is a non invasive procedure performed through a magnetic resonance machine.

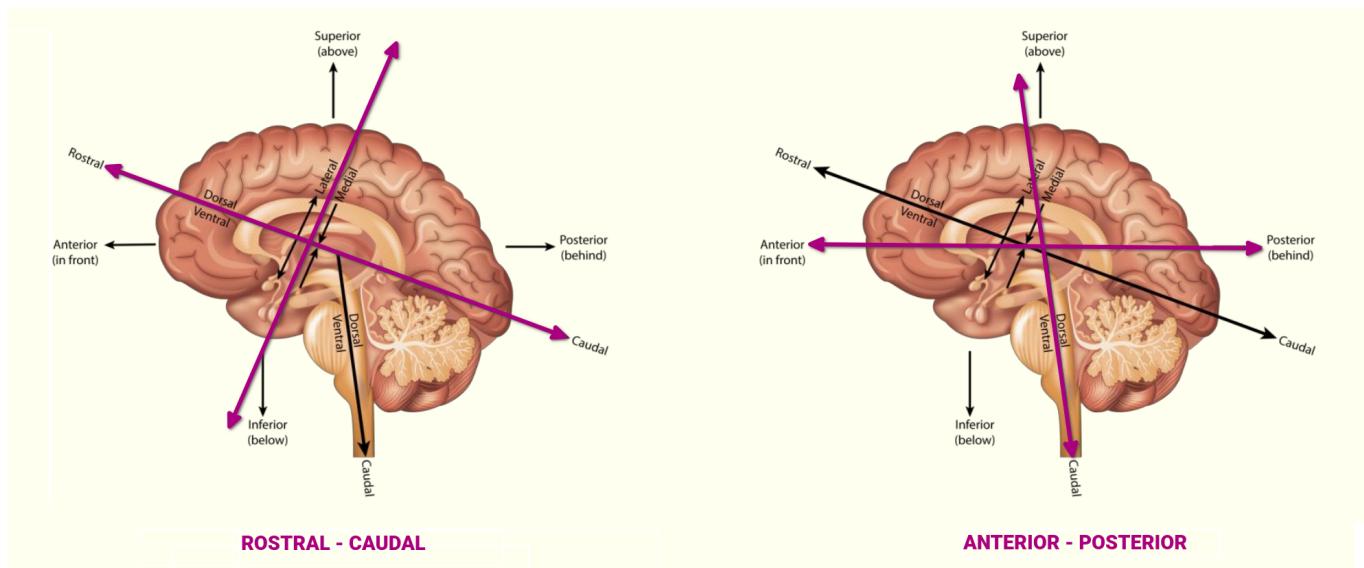


DTI Image, Laboratory of Neuro Imaging at UCLA and Martinos Center for Biomedical Imaging at MGH, Consortium of the Human Connectome Project

.HOW TO GET ORIENTED

Let's introduce some terms used in books and papers regarding brain structure.

First at all the brain can be divided in two different ways:



Then the following terms are used:

Term	Meaning
Anterior	In front of
Posterior	Behind
Superior	Above
Inferior	Below
Rostral	Towards the front of the brain
Caudal	Towards the back of the brain
Ventral	Towards the belly
Dorsal	Towards the back
Proximal	Closer to a set point
Distal	Farther from a set point
Medial	Towards midline of body
Lateral	Towards appendages
Contralateral	On the opposite side (opposite hemisphere)
Ipsilateral	On the same side (same hemisphere)

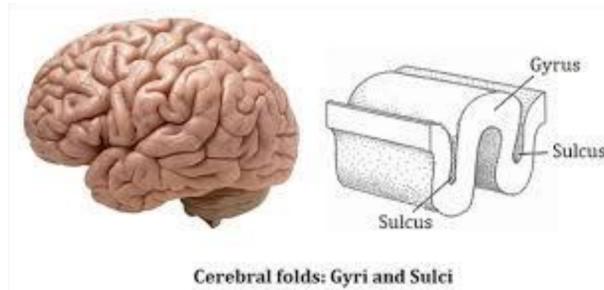
Complete glossary: <https://www.thescienceofpsychotherapy.com/glossary/>

BRAIN CORTEX

Brain cortex is the external surface of the brain, it contains most of the grey matter present in the brain. Brain cortex is one very important part of our bodies, it is about 1,5% of the body weight but it uses 15% of the total blood flow (the whole brain uses 20%).

Brain cortex is folded to increase the cortical surface in a limited volume, the two types of folds are called:

- **Gyrus:** (plural: gyri) Outwards folding
- **Sulcus:** (plural: sulci) Inward folding

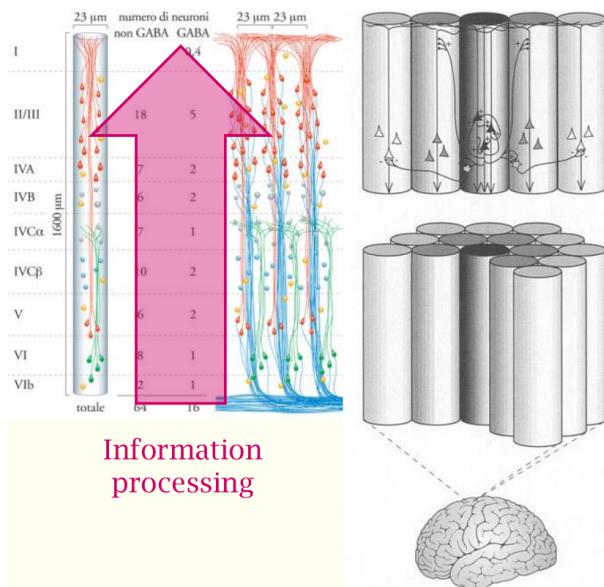


Sulci represent 2/3 of the surface, hence most of the cortical surface is hidden.

CORTICAL ORGANIZATION

Brain cortex is divided into six vertical layers, strictly interconnected. Inside each layer, neurons are organized in columns (0.5mm in diameter) perpendicular to the cortical surface. Inside each column there are same type of neurons with same function, therefore they have same behavior in response to the same stimulus.

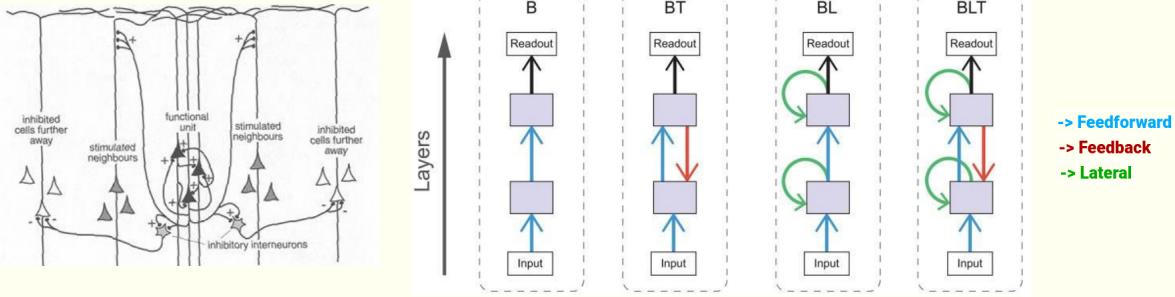
The information flow direction is shown in the following figure:



Different neural connections

We can distinguish neural connections in three main typologies:

- **Feedforward:** (or bottom-up): Directed from regions at the first processing stages to the following ones
- **Lateral:** Connections linking same stage regions
- **Feedback:** (or top-down): From advanced stages back to previous ones



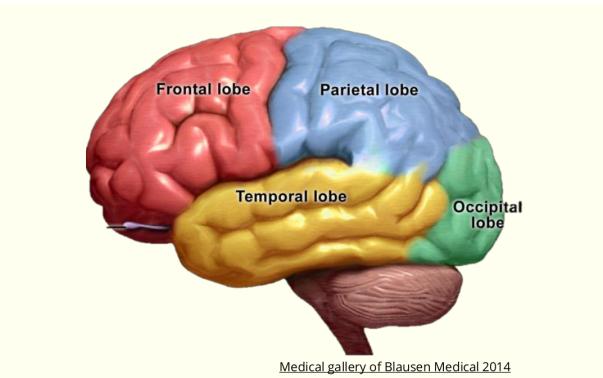
NOTE: (excitatory or inhibitory)

All these connections can be excitatory or inhibitory

The majority of connections are those that connect neurons of the same layers namely lateral ones.

BRAIN CORTEX: BRAIN LOBES

In this section we briefly introduce **brain lobes** and their functionalities.



Frontal Lobe

Control center for executive functions:

- reasoning
- decision-making
- expressive language
- higher level cognitive processes
- orientation (person, place, time, and situation integration of sensory information)
- planning and execution of movement

Parietal Lobe

- Primary and secondary somatosensory cortex
- spatial navigation
- touch, pressure, temperature, and pain

Occipital Lobe

- primary visual cortex

- processing and interpretation of visual information

Temporal Lobe

- auditory cortex
- center for receptive language
- hippocampus (memory formation and emotion)

NOTE: (*Cerebellum*)

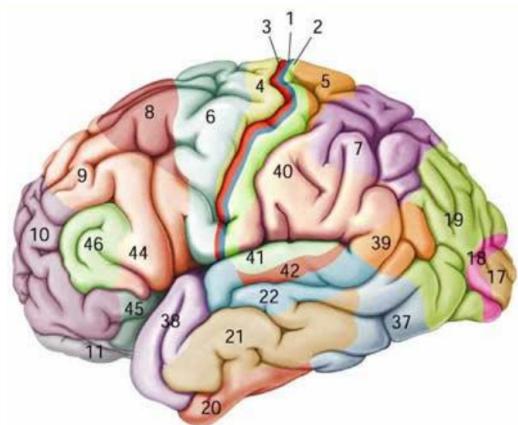
Cerebellum is not part of brain

NOTE: (*Hemispheres*)

Brain is divided into two hemispheres, they are separated but work together. Most of their parts are symmetric.

.BROADMANN AREAS

Each lobe is in turn composed of different areas called **Brodmann areas**, name come from its discover Korbinian Brodmann that described it in detail in 1909.



Each area is defined by its cytoarchitecture (neural cells type and organization), and it is associated to a function. Multiple areas/tissue types may participate in the same function: e.g., areas 3, 1, and 2 all comprise the sensory cortex.

CORTICAL FEATURE MAPS

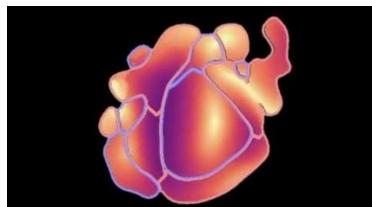
Cortical maps are collections (areas) of minicolumns in the brain cortex that have been identified as performing a specific information processing function (texture maps, color maps, contour maps, etc.). (Wikipedia)

Let's see two of them

Visual cortex: Retinotopic

Retinotopy is the mapping of visual input from the retina to neurons, particularly those neurons within the visual stream. For clarity, 'retinotopy' can be replaced with 'retinal mapping', and 'retinotopic' with 'retinally mapped'. This retinotopic mapping is a transformation of the visual image from retina to V1.

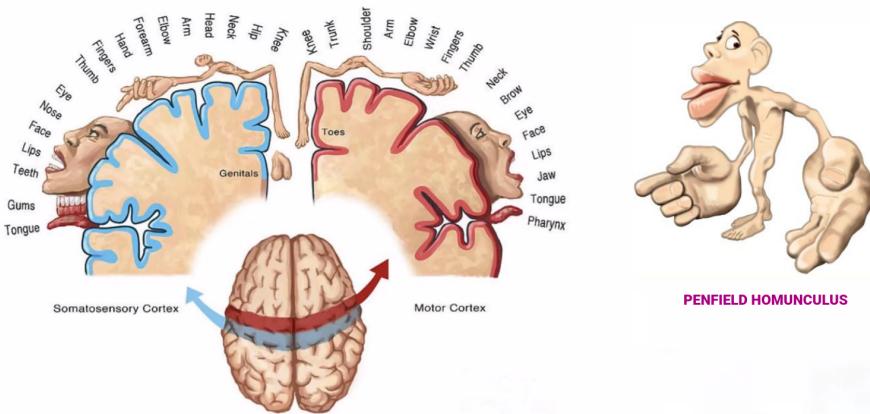
This area is highly specialized for processing information about static and moving objects and is excellent in pattern recognition, it is composed by specialized neuronal families for shape, movement, contrast and others.



Motor and somatosensory cortex

The **motor cortex** is the region of the cerebral cortex involved in the planning, control, and execution of voluntary movements, this area is located in **pre-central gyrus** of the frontal lobe (Brodmann area 4, in red in the figure below).

The **somatosensory cortex** is responsible for receiving and processing sensory information from across the body, such as touch, temperature, and pain. This area is located in **post-central gyrus** of the parietal lobe (Brodmann areas 3,1,2, in blue)



Each part of our body has a different number of neurons in motor/somatosensory cortex associated to. The difference is shown by **Penfield Homunculus** that has been derived from stimulating awake patients during epilepsy surgery. An higher number of neurons is

associated to a very sensitive area, or to body parts that require an high motion precision (e.g. hands).

NOTE: (Contralateral)

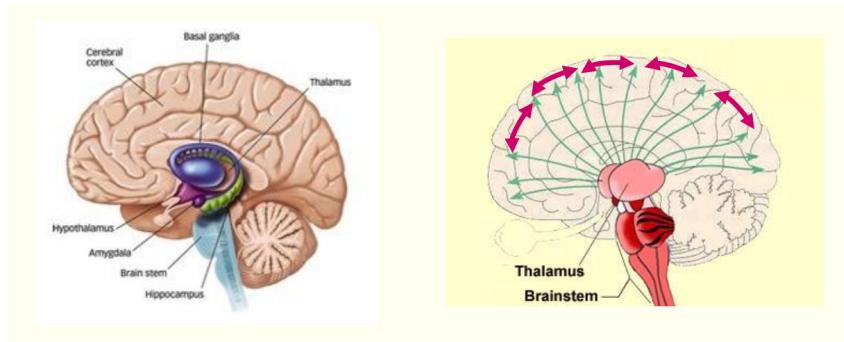
An important property of brain is its hemisphere are **Contralateral** namely left hemisphere controls right part of body and right hemisphere left one.

SUBCORTICAL AREAS

All the previous components are related to brain cortex, namely belong to the most external part of the brain. There are also several important components deeper in the so called **subcortical area**. From an neuroengineering point of view, there is an huge difference between cortical areas and subcortical ones, indeed the last is more difficult to reach especially with a non-invasive method therefore we are not able to study it easily. Both types of areas are important, cortical one contains advanced function like language, subcortical instead, more primitive functions.

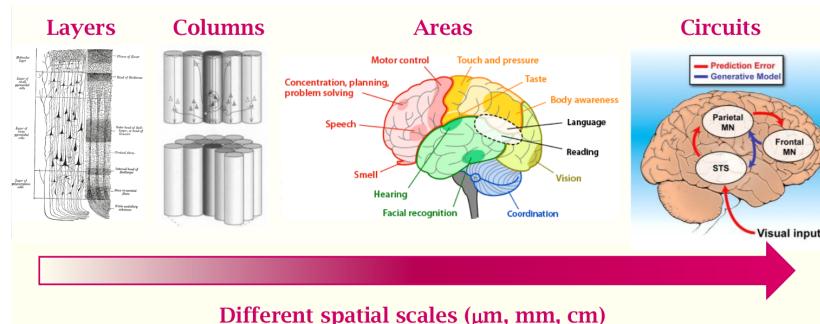
Some of the most important areas inside subcortical areas are:

- **Thalamus:** It controls traffic from most of peripheries to the cortex (sensory afferent)
- **Brainstem**
- **Basal ganglia**
- **Cerebellum**



SUMMARY BRAIN ORGANIZATION

Let's summaries the brain organization a different spatial levels with the following figure:



Each spatial level is important, indeed, when we are interested to study a specific brain function, we focus only to a specific level. For example, to understand how action potential is generated we have reasoned at the level of a single neuron, other functions like memory instead, must be studied at higher level like *areas* or *circuits*.

BRAIN EVOLUTION IN TIME: BRAIN PLASTICITY

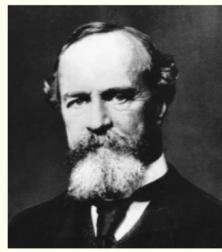
Brain plasticity is the ability of our brain to change its own structure during time. We study this concept in this course due to the strict relation to A.I. and due to the fact that it is at the base of the main applications of neuroengineering like neurorehabilitation.

William James, *Principles of Psychology*, 1890:

The phenomena of habit in living beings are due to the plasticity of the organic materials of which their bodies are composed.

Donald Hebb, *(Hebbian Learning)* 1949:

“...two cells or systems of cells that are repeatedly active at the same time will tend to become 'associated', so that activity in one facilitates activity in the other.”



William James



Donald Hebb

Hebbian learning is a form of activity-dependent synaptic plasticity where correlated activation of pre- and postsynaptic neurons leads to the strengthening of the connection between the two neurons. The learning principle was first proposed by Hebb (1949), who postulated that a presynaptic neuron A, if successful in repeatedly activating a postsynaptic neuron B when itself (neuron A) is active, will gradually become more effective in activating neuron B. The concept is widely employed and investigated in both experimental and computational neuroscience. [source](#)

Hebbian learning definition leads to an important consideration: our brain reacts to our habits and environment in which we live just like our body. The main difference is that changes in our brain are slower and not visible outside, it is not like with our muscles when do physical exercises.

Important evidences of brain plasticity are shown in neurorehabilitations when patient is able to recover brain functionalities.

.BRAIN PLASTICITY DURING THE LIFESPAN

Brain plasticity is present in our whole life although it decreases with time. We have the highest brain plasticity at age 0-2 (*critical period*) when we have maximum brain reorganization capability and

some irreversible changes.

We have several evidences of plasticity in adults and even in the elderly people:

- **Learning**
- **Memory**
- **Peripheral changes:** Consequences of sensory deprivation (e.g. amputees, acquired blindness)
- **Self-healing:** Reorganization after a brain lesion
- **Help to heal:** Effectiveness of neurorehabilitation

Let's see briefly some mechanisms at the base of brain plasticity, we will see only simplest ones, there are others more complex and others not even fully understood yet.

.SYNAPTIC PLASTICITY

The first mechanism at the base of brain plasticity is **synaptic plasticity** of chemical synapses.

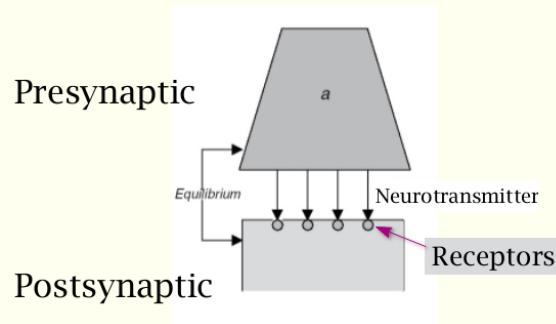
We have two different types of synaptic plasticities that depend to the *training time*:

- **Short-term plasticity:** (No structural change) Short training time required but it also have short effect duration
- **Long-term plasticity** (Structural change) Longer training time required, longer effect duration

Let's see both of them through an example.

Short-term plasticity

Suppose you would like to learn play tennis and the following chemical synapse is crucial to play tennis:



During training this synapse is used a lot, namely pre-synaptic neuron sends information to post-synaptic neuron by releasing neurotransmitters. After few hours of training there are not structural changes, however we have *changes of function*, in particular pre-synaptic neuron releases an higher number of neurotransmitter respect to before training:

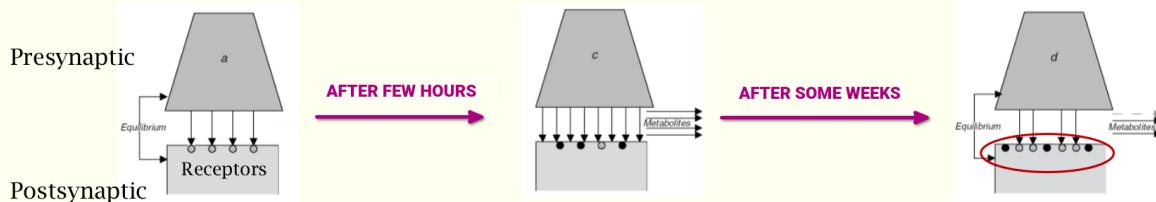


This increase of neurotransmitters released means that this synaptic becomes more important, indeed connection is stronger, this is translated into a faster opening of ions channel in post-synaptic cell, therefore we have a faster exchange of information.

However, this change of function is not stable indeed it vanishes after some days of weeks without training, namely this synaptic loses importance since pre-synaptic reduces number of neurotransmitters exchanged. For that reason this type of plasticity is called **short-term synaptic plasticity**.

Long-term plasticity

Suppose we continue to training play tennis, such as two or three times a week. The previous synaptic gains importance at each training session, after some weeks or months we have a structural change represented by an increase of receptor number on the post-synaptic cell.



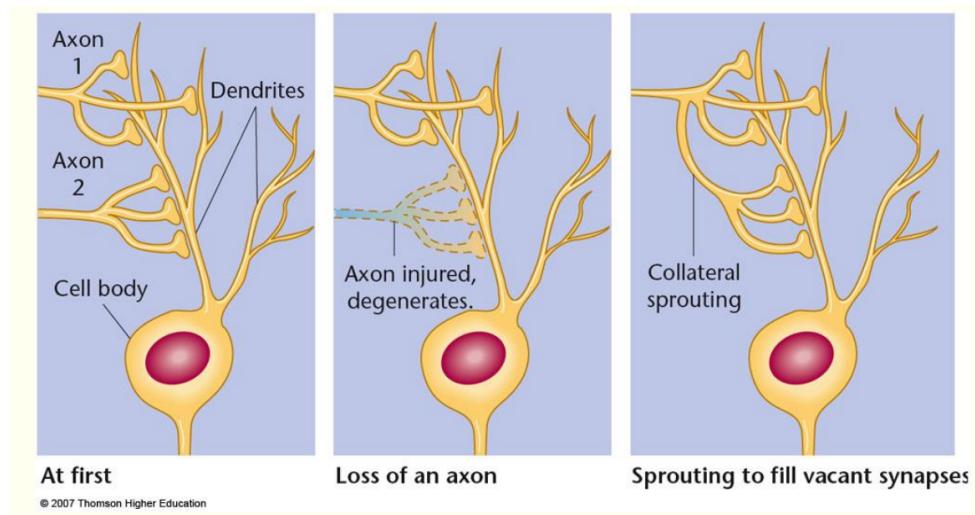
This change makes faster the connection and much more stable in time indeed it will last several months or years before vanish. For that reason is called **long-term synaptic plasticity**.

AXONAL SPROUTING AND PRUNING

Let's see another brain plasticity mechanism that is more radical respect to synaptic plasticity where a synaptic must be already present.

This new mechanism is called **axonal pruning or sprouting**:

- **Axonal Pruning:** (Loss of an axon) Pre-synaptic neuron loses an axon branch, therefore we loses their synapses
- **Axonal Sprouting:** (Sprouting of an axon) An axon of a neuron sprouts (creates new branches) to fill vacant synapses on another neuron



OBS: (Structural change)

Axonal sprouting/pruning is a stronger structural change respect to long-term synaptic plasticity

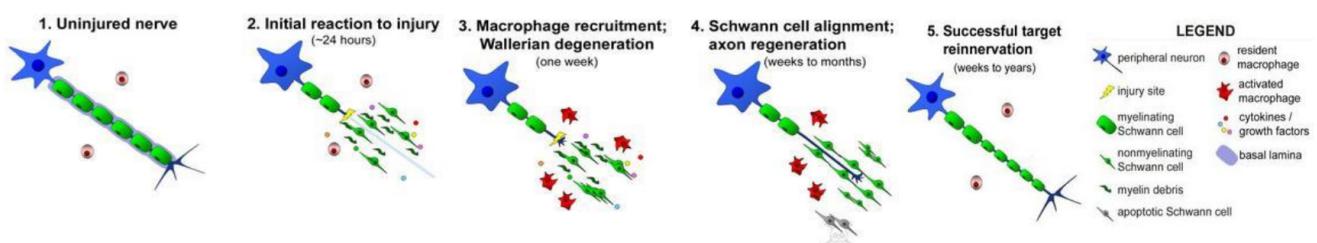
NOTE: (Axonal pruning)

Axonal pruning is not always a bad thing, indeed, at birth we have a lot of synapses necessary to discover the world but then growing up some of them are not longer necessary, therefore to save energy some of them are pruned.

Unfortunately axonal pruning may also be due to a neurodegenerative disease.

.AXONAL REGENERATION

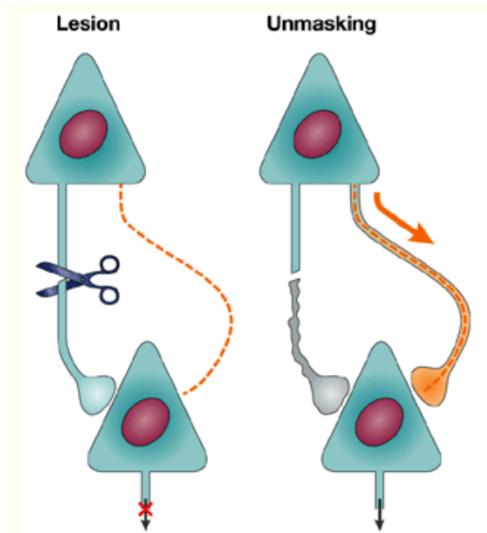
Another brain plasticity mechanism is **axonal regeneration** in which axon of a neuron regenerates after an injury. This kind of deep regeneration is more typical of early life stages and less common in adults.



The final step of axonal regeneration is re-myelination, if this process does not happen correctly, neuron cannot work properly.

.UNMASKING OF LATENT SYNAPTIC CONNECTIONS

Brain plasticity techniques include also redundancy methods, namely strategies used to increase robustness of brain functionalities after failures. One of these redundancy methods is called **unmasking latent synaptic connections** in which there are synapses that are there but are not used until they are needed because of a lesion. These unused synapses are called *latent synapses*.



All of these methods, at the base of brain plasticity, include only a variation of brain network connection, namely they do not add new neurons. For that reason in this course we will spend a lot of time in the study of brain network.

Self-evaluation test

1. At what temporal scale does the brain operate?
2. At which different spatial scales may we look at its functioning?
3. Put the following levels of cortical organization in a hierarchical order (from the smaller to the larger):
 - A. Brodmann areas
 - B. Cortical columns
 - C. Brain lobes
 - D. Cortical layers
 - E. Brain circuits
4. Indicate which of the brain lobes houses the visual function:
 - A. Frontal
 - B. Temporal
 - C. Parietal
 - D. Occipital
5. For each of the following brain areas, indicate if they are cortical or subcortical:
 - A. Thalamus
 - B. Primary motor area
 - C. Cerebellum
 - D. Broca (language) area
 - E. Brainstem
6. Does the short-term synaptic plasticity involve:
 - A. A structural change in the post-synaptic membrane
 - B. An increased number of membrane receptors
 - C. The amount of neurotransmitter released in the synaptic cleft
 - D. An irreversible change in the synaptic structure
7. List the four main neuronal mechanisms behind brain plasticity

References

- Hari & Puce:
--- Chapter 2, pagg.13-18 (Sections: Overview of the human brain; Functional structure of the human cerebral cortex; Communication between brain areas)
- Wolpaw&Wolpaw:
--- Chapter 2, pagg. 15-20
- http://learn.neurotechedu.com/neuroanatomy_and_brain_organization/

04- ELECTRICAL CORRELATES OF THE BRAIN ACTIVITY

Learning objectives of the lesson

1. Describe the **different scales** at which electrical correlates of the brain activity are produced and can be measured
2. Compare intra- and extra-cellular **single neuron recordings**
3. Illustrate the origin of the **electrical correlates** of the collective activity of groups of neurons
4. List the different **acquisition methods**
5. For each of them, **describe** their **origin**, their **spatial resolution**, and which **neuronal groups** are involved
6. Explain the **generation of the EEG signal** at the neuronal and tissutal level
7. List the **advantages** and the main **issues** of scalp EEG measures

INTRODUCTION

First at all it is important to clarify meaning of the two words in the section's title: **electrical** and **correlates**.

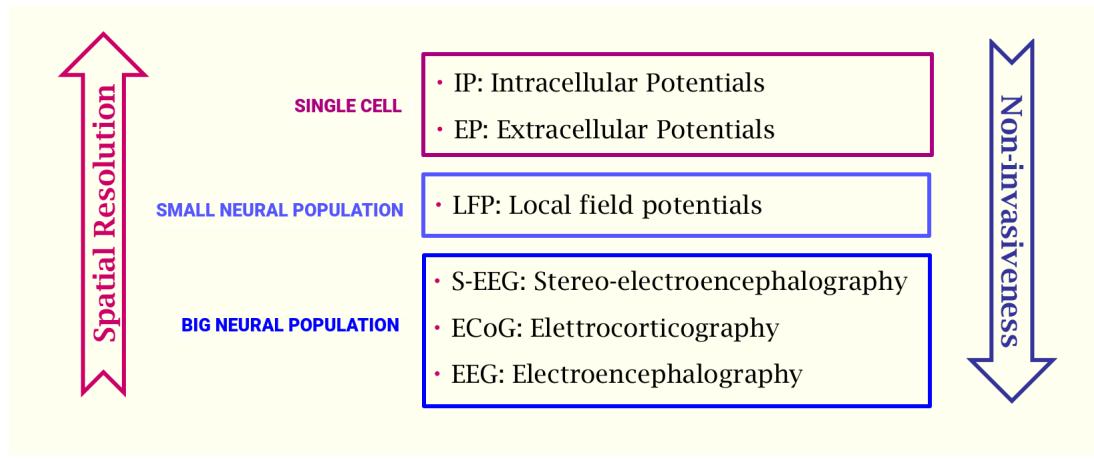
Our objective is to study brain activity, however this is not simple at all since our brain is in our scalp, therefore outside it we cannot perform a direct measurement of brain activity but only a measure of **brain activity correlates** namely signals that are *correlated* to brain activity. These correlates quantities provides to us only a partial observation to brain activity.

Then we use the term **electrical** because we measure electrical correlates of the brain activity.

.BRAIN ELECTRICAL MEASURES

There are different types of methods to measure electrical correlates of the brain activity, they essentially depend on their invasiveness. The most invasive one (Intracellular potential measurement) is even able to measure directly brain activity, it is the only one, the others can measure only an electrical correlates as said before.

The following figure summarizes different techniques for brain activity study:



As we can see, we have a tradeoff between *non-invasiveness* and *spatial resolution* of brain electrical measures

Let's see all of them.

SINGLE-CELL RECORDINGS

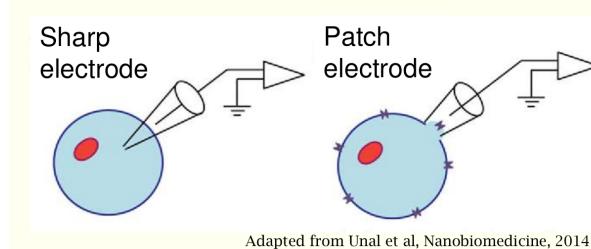
The first class of methods are very invasive, their aims are to study electrical activity of a single neuron. There are two types of methods which differ to it is necessary or not to penetrate membrane cell. (*intracellular* or *extracellular*)

.INTRACELLULAR RECORDINGS

This method is the only one able to measure directly brain activity since it performs an exactly measure of membrane potential by placing a hollow glass electrodes filled with a conducting electrolyte and a reference electrode in the extracellular medium.

Intracellular recordings can be made with two different ways that lead to the same outcome:

- **Sharp electrodes** inserted through the membrane into the cell
- **Patch electrodes** sealed to the surface of the membrane providing electrical contact with the interior of the cell

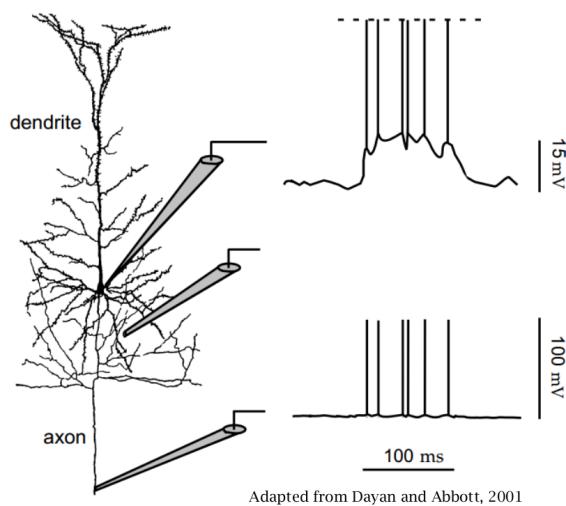


Adapted from Unal et al, Nanobiomedicine, 2014

Patch electrodes does not require an actual membrane penetration and usually is preferred since it is more easy in practice.

This technique is more complex than the extracellular recording however is the only one able to record the actual membrane potential (graduate and action potentials). It is usually performed in the

soma, sometimes in the dendrites, but very seldom along the axon due to its small thickness. This measurement is usually performed in in-vitro experiments.

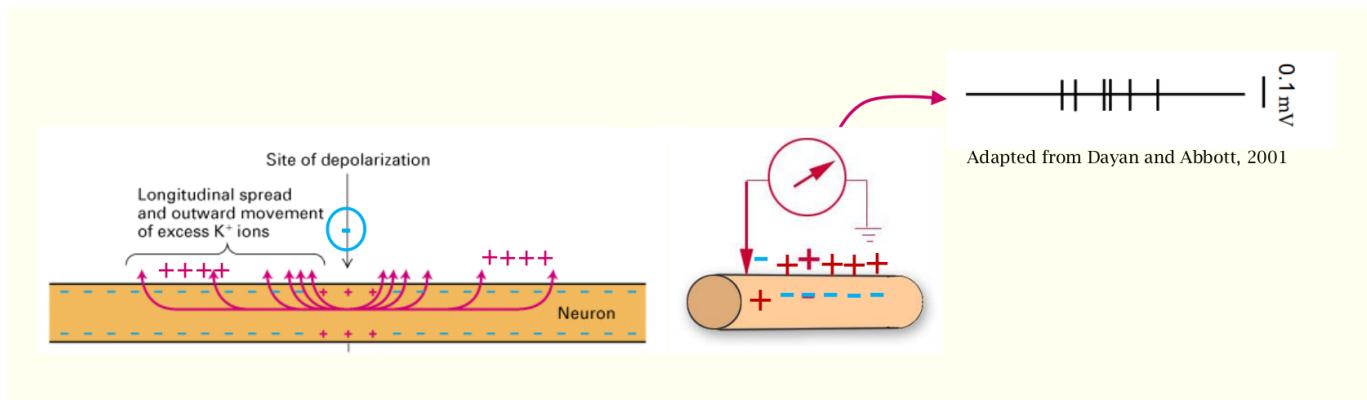


Adapted from Dayan and Abbott, 2001

.EXTRACELLULAR RECORDINGS

Intracellular recording is not simple to be performed in in-vivo, the alternative to it, feasible in in-vivo, it is **extracellular recording** in which electrodes are placed near the membrane, but do not penetrate it.

In this type of measurement we measure the potential of the extracellular fluid, this last indeed varies according to the membrane potential of the neuron as we have seen in previous sections. Extracellular recordings measure the exact moment in which an action potential occurs, but not its shape or amplitude, moreover the action potential amplitude is reduced ($0.1mV$ instead of $100mV$).



Adapted from Dayan and Abbott, 2001

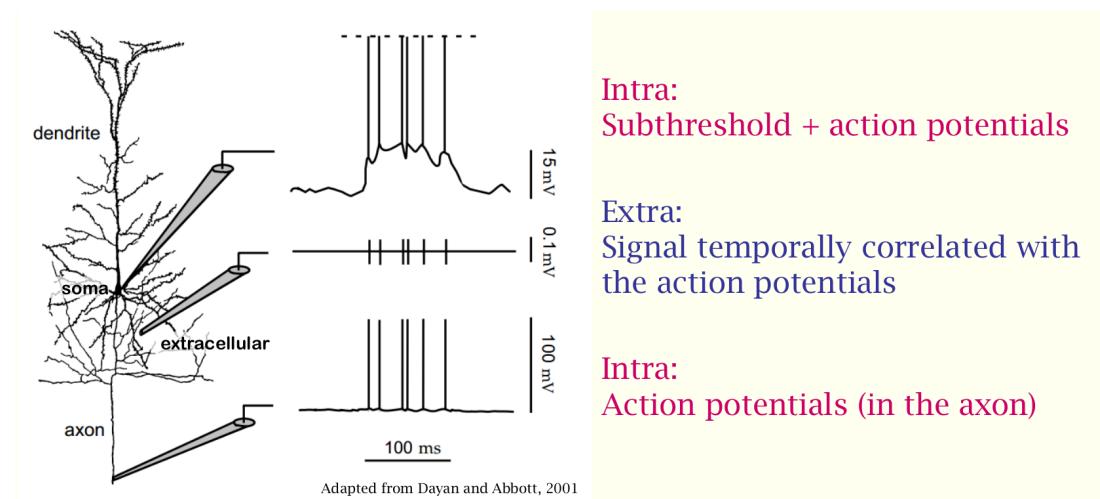
Extracellular fluid potential is much more small than membrane potential, so we are able to measure only instants of action potential firing but we do not have any information about smaller variation of membrane potential, therefore our recordings take the form of spike trains. As already said, extracellular recordings are used in in-vivo experiments.

.COMPARISON INTRA-ENTRA CELLULAR RECORDINGS

During a study, to choose between intracellular measurement and extracellular one is necessary to take in mind three different aspects:

- **Feasibility of intracellular recording:** It depends on neuron cell shape
- **Invasiveness** Level of invasiveness acceptable
- **Signals of interest** Action potential timing can be acquired by both methods

The following figure shows how each method measures the same membrane potential activity in different part of the cell:



NOTE: (No time delay

All methods provides exact action potential production timing

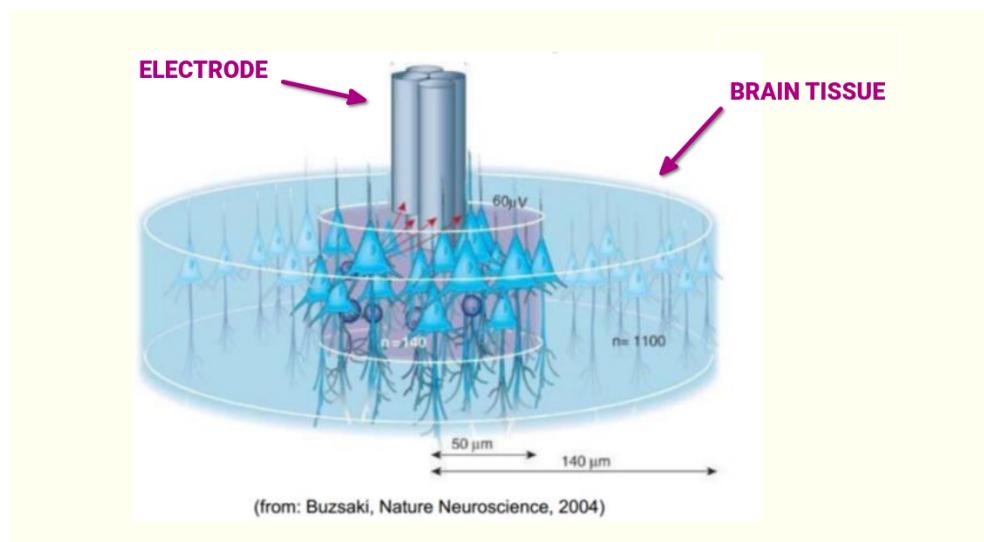
NEURAL POPULATIONS RECORDINGS

In order to understand how complex functionalities of our brain work (e.g. memory, motor functions, learning), it is necessary to study activity of group of neurons of the same layer or same column or same Brodmann area.

.LOCAL FIELD POTENTIAL (LFP)

To measure activity of a small group of neurons it is necessary to penetrate brain tissue with an electrode.

When we study a group of neurons we can only measure the overall activity given by summation of each activity. The overall activity is not given by summation of action potentials due to the fact that their durations are very small and it is practically impossible a synchronization of them. Therefore when we analyze activity of a small group of neurons we measure the extracellular current flow resulting from the **linear summation of PSP** (instead of action potential) of neuronal groups. The PSP is synchronized since it is a slower event that depend on similar inputs to each cell in the group.



This technique allows us to measure a frequency range of $0 - 100\text{Hz}$ with a spatial resolution of 10^3 to 1mm^3 .

.ELECTROENCEPHALOGRAPHY

With the term **electroencephalography** we include several types of techniques, the most common is *scalp EEG* that we usually call *EEG*.

The main differences between an **electroencephalography** and a LFP measurement are the number of neurons in the neural population observed and the electrodes distance.

The number of neuron in the neural population observed by an electroencephalography is huge, indeed recorded signal originates from synchronous **postsynaptic cortical currents** (sources) of millions/ 1 billion neurons. As said before, we do not measure action potentials summation.

Also the distance from neural population is bigger since we can measure electrical activity up to subject's scalp. The signal is propagated thanks to volume conduction provided by tissues that contain water, so ion currents can spread and the effects of the electrical activity of excitable cells can reach a large distance (cm). Electric fields produced by local currents spread instantaneously (at the light speed) and **sum up linearly**, however the amplitude is reduced indeed on the scalp we have μV that is very feeble compared to membrane potentials (mV).

The temporal resolution is very high (milliseconds) and we do not have any delay between brain activity and signal at the scalp.

OBS: (Linear sum)

The fact that all neuron activities are sum up linearly is a very positive fact.

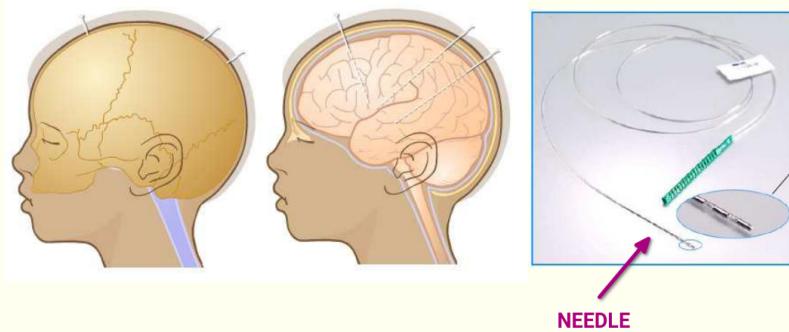
The main three type of electroencephalography are:

- **Stereo-Electroencephalography (S-EEG):** Very invasive, electrodes go inside the brain
- **Electrocorticography (ECoG):** Less invasive, electrodes on top of the brain
- **Scalp Electroencephalography (EEG):** Non invasive

Stereo-Electroencephalography (S-EEG)

Stereo-Electroencephalography is performed with electrodes (needles) implanted deeply

in the brain, these needles can measure the activity in subcortical regions. The spatial resolution of this method is very high ($1 - 4\text{mm}^3$).



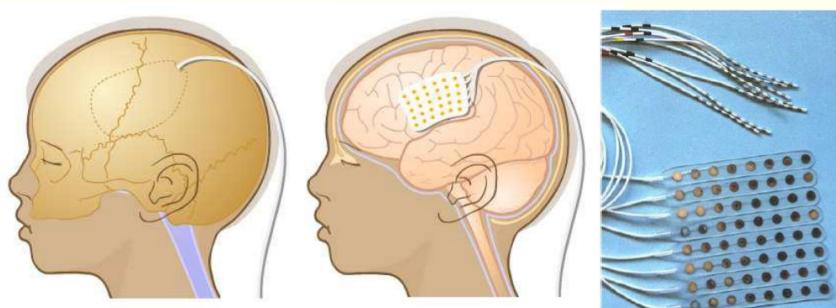
S-EEG is used only when the patient requires a neurosurgery in any cases, namely it is performed only if it is necessary to open patient skull due to a disease. For example S-EEG is used in epileptic patients when the epileptogenic zones are located in depth.

OBS: (No brain damage)

S-EEG does not cause any kind of brain damages thanks to electrodes thickness. In this kind of surgery, it is most probably an infection due to skull drilling than brain damages.

Electrocorticography (ECoG)

Electrocorticography is "less" invasive than S-EEG since the electrodes do not penetrate the brain, indeed the **electrodes arrays** are put intracranial (above *pia mater*, below *dura mater*)



This arrangement of electrodes in arrays allows to record synaptic activity in macrocolumns. Since electrodes do not penetrate the brain, this method records mainly cortical neurons activity, however due to *volume conduction*, also subcortical activity can be seen although less strong.

The spatial resolution of ECoG is $1 - 20\text{mm}^3$ and as S-EEG it is performed only on patient requiring a neurosurgery.

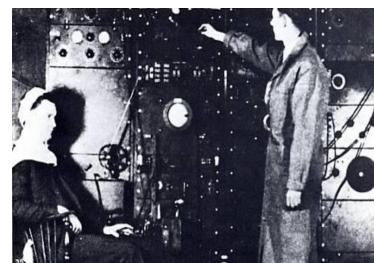
Scalp Electroencephalography (EEG)

Scalp Electroencephalography is the classical one, namely what we usually call *EEG*. This type is non-invasive.

The first scalp EEG was performed in 1929 by Hans Berger, a German psychiatrist that obtained the first measure of the electrical activity on the human scalp, he detected the alpha rhythm.

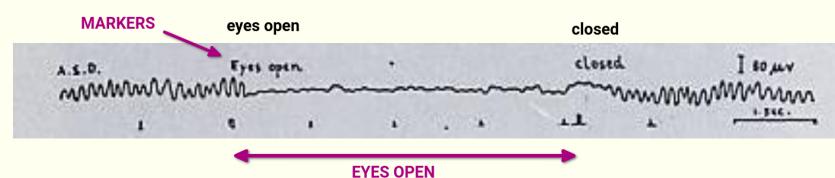


Let's see one of the first EEG recorded on paper. The following EEG was recorded in Davis Lab at Harvard medical school in 1934, it is a single EEG channel acquired thought a hypodermic needle as electrode and an head tissue band soaked with saline solution as electrical reference.

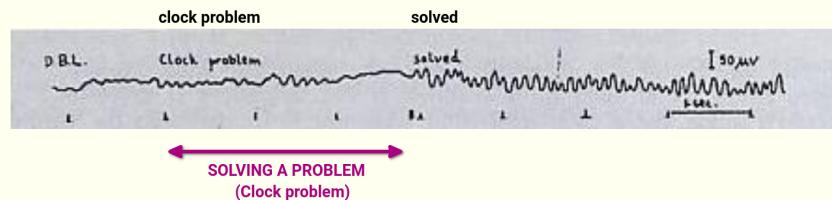


The following two EEGs have been acquired in order to study how brain activity changes when we open and close eyes, and when we put a mental effort to solve a problem (clock problem):

Open and closed eyes



Mental computation



NOTE: (Brain activity)

Looking to previous EEGs we could think that brain activity is high at rest and low when we use our eyes or solve a problem, we will see later than actually it is the opposite, namely an "high" signal respond to a low brain activity.

GENERATION OF THE EEG SIGNAL (SCALP EEG)

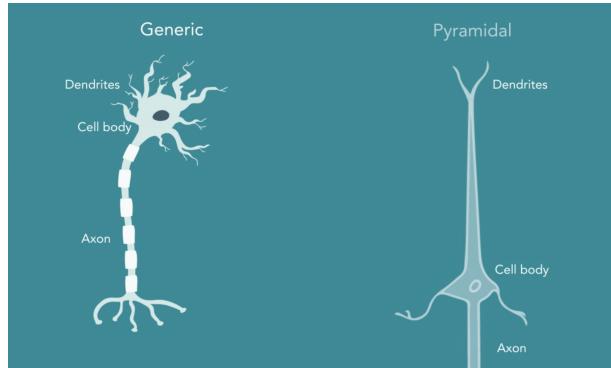
In this section we will see how signal recorded by a (scalp) EEG is generated in the brain.

OBS: (EEG)

In the follow we refer to a scalp EEG using the usual term EEG.

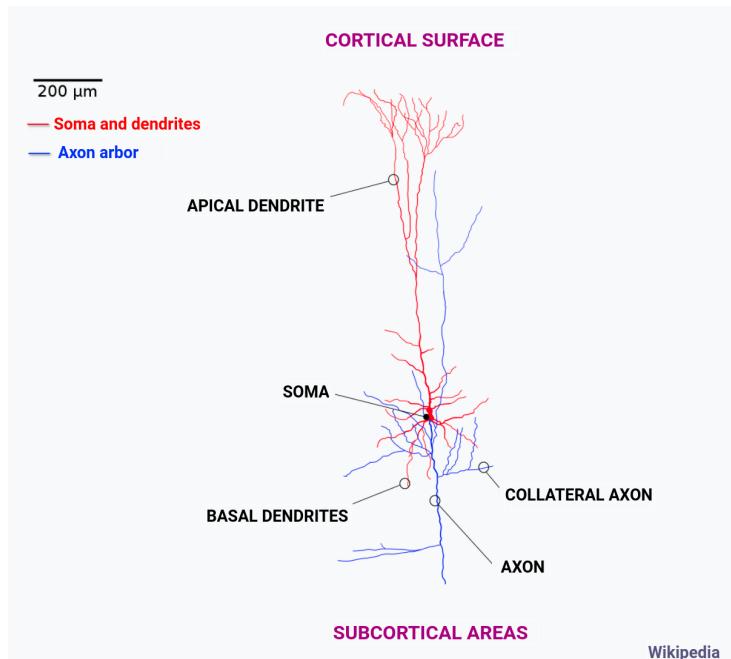
CORTICAL PYRAMIDAL NEURONS

Brain cortex is the closest region to the scalp, therefore signal acquired during an EEG comes mostly from it. So let's see the neurons inside brain cortex called **(cortical) pyramidal neurons**, this name is due to pyramidal shape of its soma.

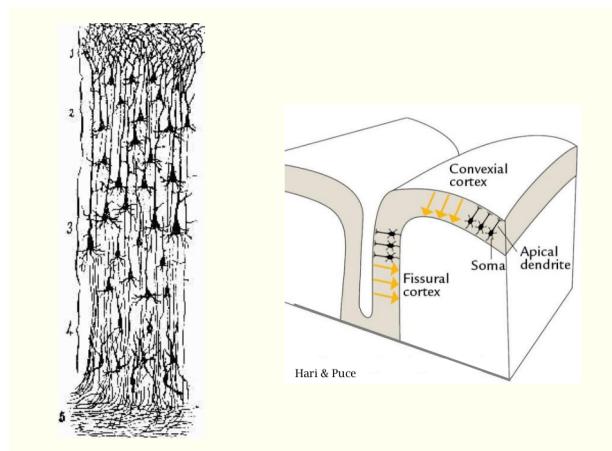


An important characteristic of a pyramidal neuron is the **longitudinal shape** of the dendritic tree that allow us to distinguish two different types of dendrites:

- **Apical dendrite:** In the upper part of the cell (above soma). It is usually projected toward the cortical surface
- **Basal dendrite:** In the lower part of the cell (below soma). It is usually projected toward the center of the brain



These kind of cells are arranged in parallel forming palisades in the brain cortex. Their orientation is normal to the cortical surface.



OBS: (Important characteristics of pyramidal neuron)

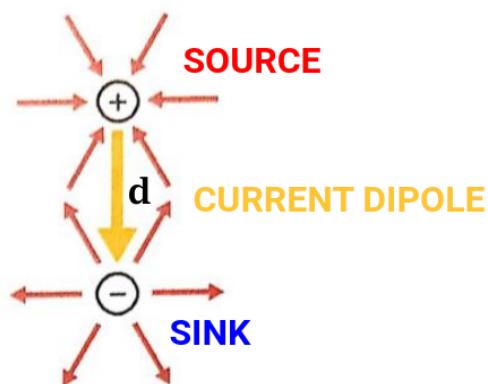
The crucial characteristics of pyramidal neuron during an EEG recording are:

- Longitudinal shape of dendritic tree
- Perpendicular orientation to cortical surface

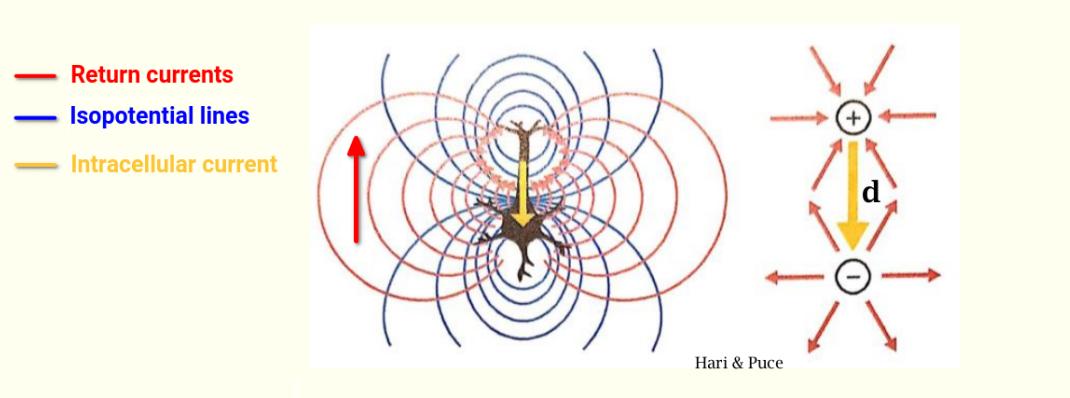
Dipolar nature of a single neuron

Current dipole

An **electrical dipole** is a physical model that describes when two point regions, with opposite charge, are separated by a distance d . The positive point region $+I$ is called **source**, the negative one $-I$ **sink** instead. Between these two points there is a current called **current dipole**.



A single pyramidal neuron can be modeled as a dipole where apical and basal dendrites act as point source and sink, the roles depend by synapses. Consider membrane potential at rest ($-70mV$) and a single excitatory synapse (+) at apical dendrite, in this case apical dendrite is the point source and the basal dendrite is the point sink since it is very close to the soma, therefore we have:



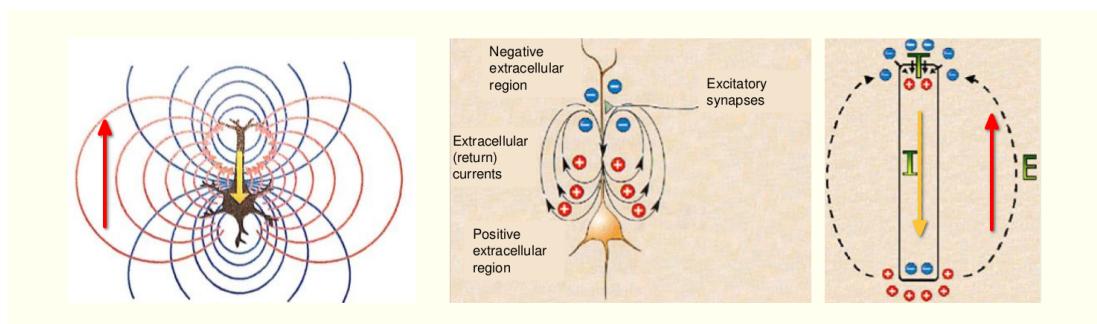
OBS: (Isopotential and return currents line)

Isopotential and return currents line are perpendicular to each other

OBS: (Dipole potential)

Any source-sink region where the total source and sink currents are equal (local current conservation) will generate a predominantly dipole potential detectable even at a large distance (scalp).

As said in previous sections, outside membrane cell there is always an opposite charge respect to the internal one, therefore we have that extracellular charges and currents are reversed with respect to intracellular ones.



NOTE: (External charges and currents)

In the following figures we will indicate **only external** charges and currents.

Dipole direction of a single neuron

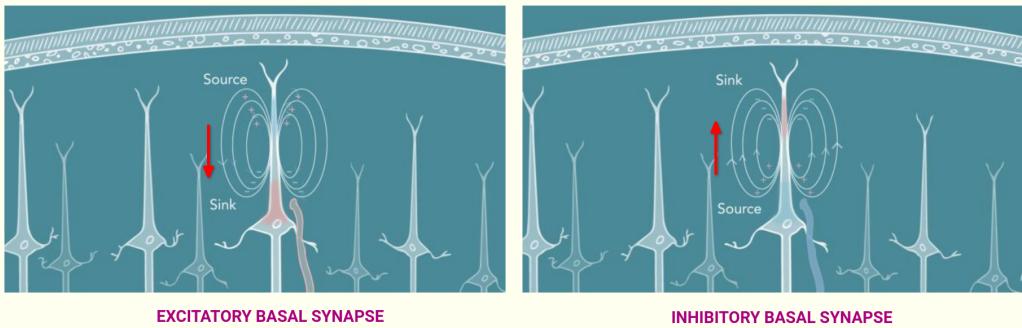
Dipole direction of a single neuron depends on two factors:

- Type of synapse (EPSP / IPSP)
- Position of synapse (Apical/Basal)

Type of synapse

As we already know, an excitatory synapse causes the increase of membrane potential (EPSP) and an inhibitory one does the opposite, therefore the synapse type determine dipole direction.

Assuming a single synapse at **basal** dendrite:



Source: <https://www.youtube.com/watch?v=rzgDOaGjjOs>

NOTE: (*External charges/current*)

In the figure are reported only **external** changes and **extracellular** current (return current), internal charges and dipole current are opposite to them

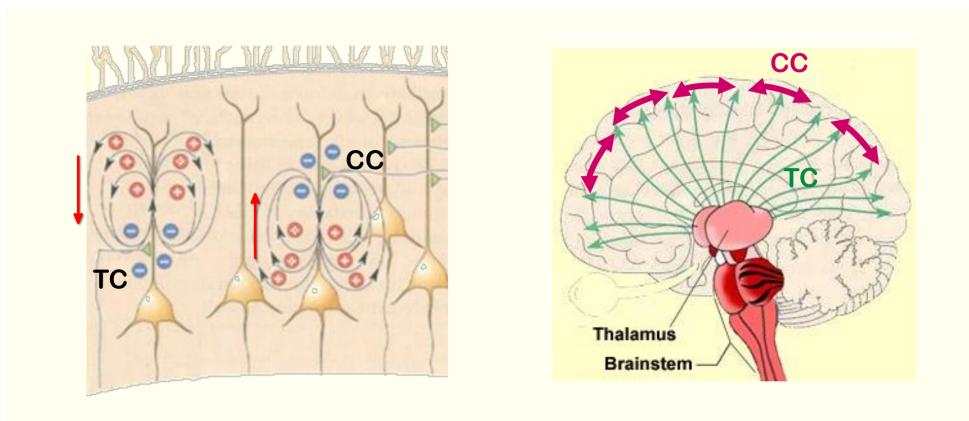
Position of synapse

Also the position of synapse changes dipole direction since it determines the location of depolarization/polarization.

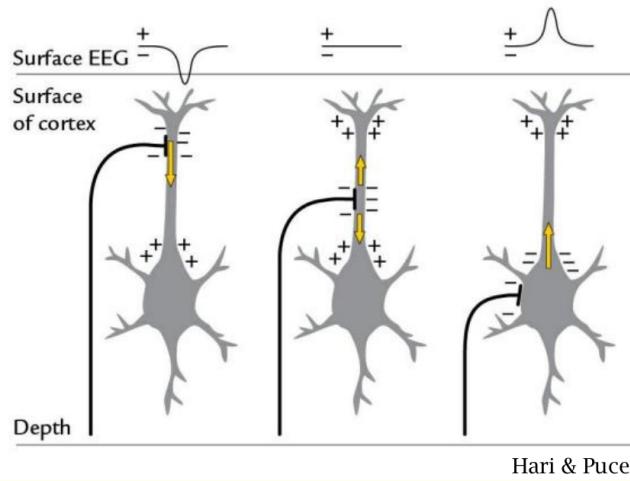
Instead to distinguish between apical and basal synapses we refer to them according to the brain parts that connect:

- **Cortico-cortical (CC)** synapses usually on the **apical** dendrites. (Connect cortical neurons)
- **Talamo-cortical (TC)** synapses usually on the **basal** dendrites. (Connect Thalamus with cortical neurons)

Assuming a single **excitatory** synapse:



In the following figure is shown how EEG signal varies according to excitatory synapse position:



Hari & Puce

To summarize, the dipole direction depends on:

- **Synapse nature:** (EPSP or IPSP)
- **Synapse position:** (CC or TC)

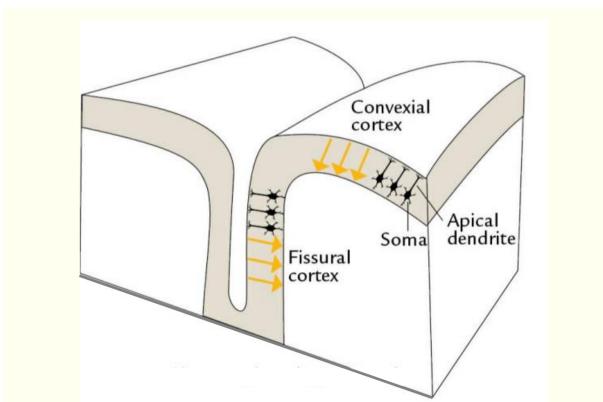
	CC	TC
E		
I		

NOTE: (First EEG limitation)

By looking at the previous table we can note the first limitation of EEG, namely since it is symmetric, we can only distinguish two cases among four.

Orientation of the pyramidal dipoles

As already said, pyramidal neurons are arranged into palisades in brain cortex and most important each palisade is perpendicular to the cortical surface making stronger the electrical signal on the scalp. Palisade dipole sources line up in parallel, creating large dipole layers.



Gyri are more efficient generators of EEG than sulci (both because of the favorable orientation of the isopotential lines with respect to the scalp and because the dipole layers in opposing sulci cortices tend to cancel each other)

.EEG SIGNAL GENERATION

Let's go deeper about how EEG signal is generated.

Effect of timing

EEG depends on **timing** since it measures the synchronous electrical activity of neural populations, namely it is able to measure the summed activity at each instant.

The amplitude of the signal is:

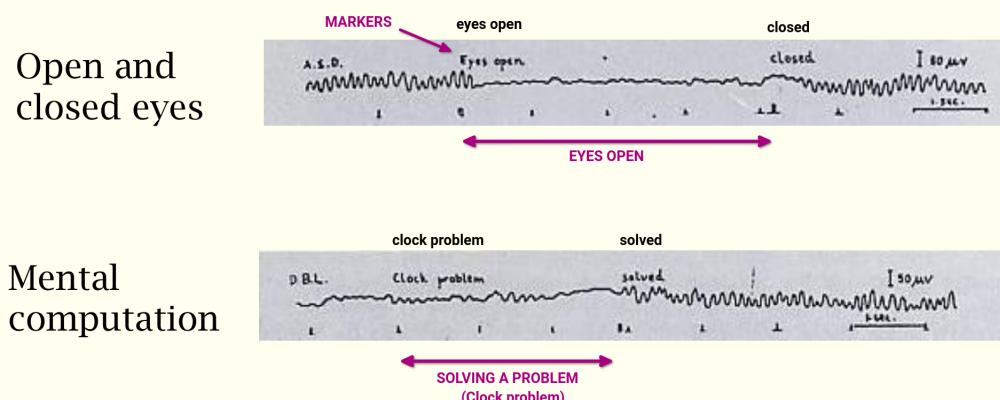
- Linearly proportional to the amount of **synchronous neurons**
- Proportional to the square root of the amount of **asynchronous neurons**, because of intracortical cancellation

So, if we have a high number of non-synchronized neurons then we obtain a very weak resulting dipole (weak EEG signal). On the other hand even if we have a small number of neurons but they are all synchronized, then we can have a strong resulting dipole.

Example: (large async vs small sync neuron population)

- $N = 10^{10}$ asynchronous dipoles \rightarrow *Signal* = 10^5
- $N = 10^8$ (1% of the previous) synchronous dipoles \rightarrow *Signal* = 10^8

Brain activity and synchronized neurons



Previously we look at this figure and we noted that seems like when our brain perform a complex task it has a low activity, actually this is not true, it is the opposite. The EEG signal recorded is weak when we perform a complex task because neurons work asynchronously and then the resulting dipole is weak. At rest instead, we have the so called **thalamus synchronization** in which neurons synchronize respect to thalamus, this cause the **idle rhythmic** (oscillation about 10Hz)

- **Complex task** \rightarrow High neuronal **asynchronism**

- Simple task -> Thalamus synchronization

Effect of orientation

Sulci and fissures produce little or no EEG signal, that is due to:

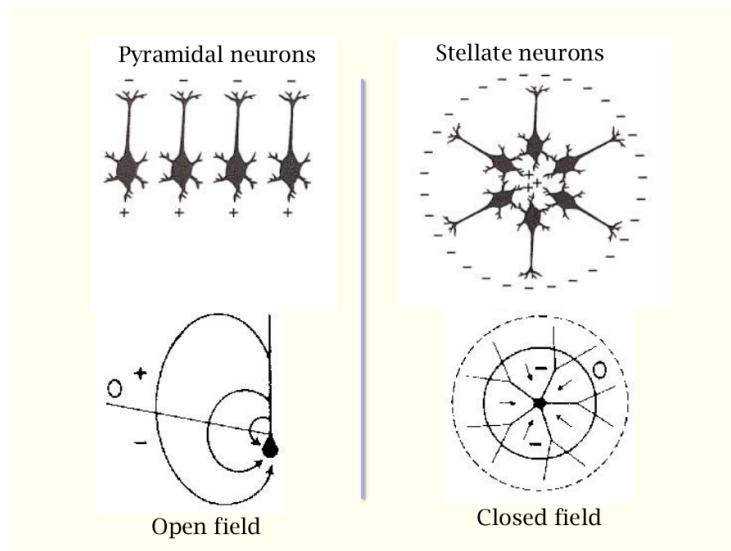
- **Orientation not favorable** to the scalp surface
- **Mutual cancellation** of opposite cortices

Gyri produce most of the EEG signal due to:

- **Favorable orientation**
- **Summation** due to the palisade disposition

Open and closed field

From the point of view of orientation of small groups of neurons, we distinguish between neurons arranged in a way that they produce an **open field** and neurons that produce a **closed field**. Only neurons that produce an open field contribute to EEG, since those that produce a closed field eliminate each other's effects.



Only neurons that produce an open field contribute to EEG

Open field (currents sum and conduct to the electrodes)
Aligned, synchronous neurons

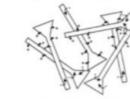


Closed field (cancellation before it reaches the electrodes):

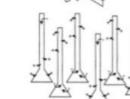
Radially symmetric neurons



Randomly oriented neurons



Asynchronously activated neurons



OBS: (EEG's blindness)

Since EEG cannot be able to acquire signal produced by closed field

neurons, we can say that it is "blind" respect to their neuron activity. Unfortunately, all of non-invasive methods has some kind of blindness.

Membrane potential studied

It is important to highlight another time which kind of membrane potential we are able to study through an EEG signal.

EEG signal is mainly constituted by **post-synaptic potentials** (PSP) because they are slower and can sum up more easily in large groups of neurons, Action potentials instead are fast and more difficult to add up in time.

EEG provides an electricsl correlates of the summation of PSPs
(NO action potential)

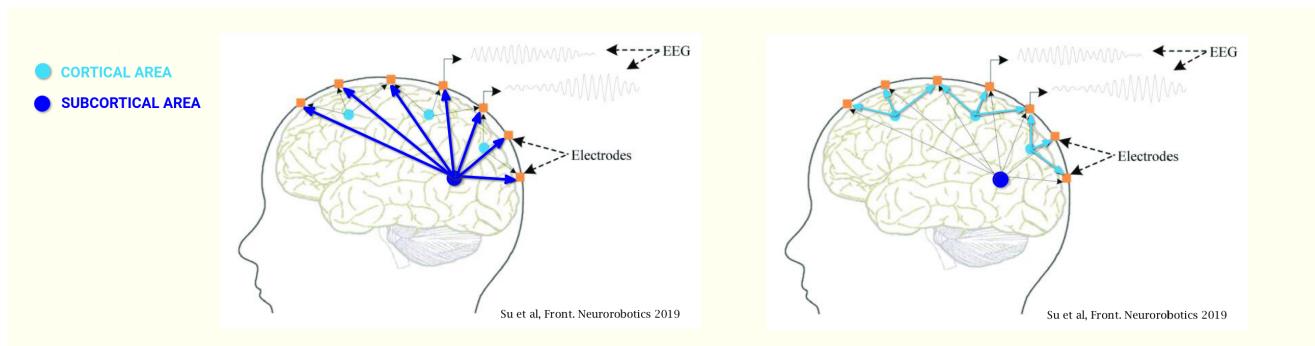
CHARACTERISTICS EEG (SCALP EEG)

Let's see in more details all limitations and advantages of a (scalp) EEG.

.LIMITATIONS

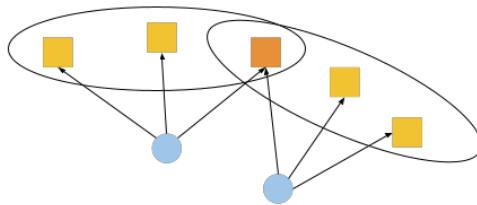
As already said we have several **limitation** in an EEG:

1. Spatial blur (attenuation and spread of the potential with distance)

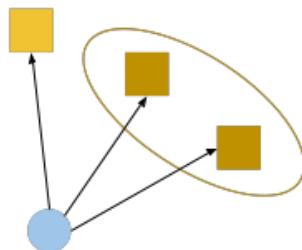


2. Low signal-to-noise ratio

3. Multiple sources contribute to the single electrode signal



4. Near electrodes record partially overlapped (correlated) signals



5. Reference choice

Then according to the source (cortical/deep):

- **Cortical sources:**
 - Open field (pyramidal neurons)
 - Closest to the scalp → stronger, more focused signals
- **Deep sources:**
 - Closed field
 - More distance
 - Attenuation
 - More spatial blur

.ADVANTAGES

Obviously EEG has also a lot of big advantages:

- Noninvasive
- Easy to use (no complex instrumentation)
- Portable
- Inexpensive
- Scalable (From 3 to n electrodes)
- Covers the entire cortical surface
- Excellent temporal resolution

Self-evaluation test

1. Put the following levels of brain electrical correlates in sequence according to their increasing spatial resolution (from the less to the more detailed):
 - A. ECoG: Electrocorticography
 - B. LFP: Local field potentials
 - C. IP: Intracellular Potentials
 - D. S-EEG: Stereo-electroencephalography
 - E. EEG: Electroencephalography
 - F. EP: Extracellular Potentials
2. To record in vitro measures of the membrane potential over the dendrites of a neural cell, you can use:
 - A. Intracellular measures
 - B. Extracellular measures
3. Describe which part of the pyramidal neuron acts as a current dipole and how
4. Indicate which of the following factors affect the amplitude of EEG signals (multiple answers):
 - A. Open/closed field
 - B. Neurons' orientation
 - C. Synchronicity of the neural activity
 - D. Distance between the neurons and the electrodes
5. Which electrical variation of the membrane potential mainly contributes to EEG?
 - A. The action potential
 - B. The spike train
 - C. The resting membrane potential
 - D. The post-synaptic potentials
6. Which regions of the brain mainly contribute to scalp EEG? Why?
7. List at least 4 limitations of scalp EEG recordings
8. List at least 5 advantages of scalp EEG recordings

References

- Dayan & Abbott:
 - Chapter 1.1 (Recording Neuronal Responses)
- Wolpaw & Wolpaw:
 - Chapter 2 (Spike Recording and Processing)
 - Chapter 3, pagg. 45-57 (Electric and magnetic fields produced in the brain)
- Hari & Puce:
 - Chapter 1, pagg. 7-11
 - Chapter 3, pagg. 25-27 (Charges and electric current) and pagg. 31-34 (Source currents)
 - Chapter 4, pagg. 38-40 (Early EEG recordings)

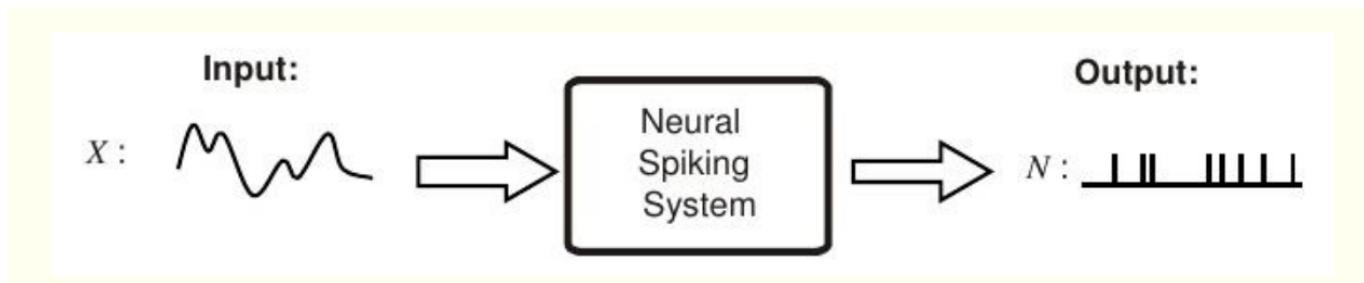
05 - NEURAL ENCODING - PART 1

Learning objectives of the lesson

1. **Understand** the definition of neural encoding and decoding
2. **Describe** the rate-coding hypothesis
3. **Define** the neural response function and the spike-count firing rate
4. **Illustrate** the experimental procedure used to build a tuning curve
5. **Interpret** a given a tuning curve, and specifically:
 - a. **Describe** the neuron behavior as a response to stimulus properties
 - b. **Compare** the neural response to specific values of the stimulus s

INTRODUCTION

We can represent a single neuron or a group of them by a black box (*neural spiking system*):



In theory, this black box can be either:

- **Deterministic model** of a neuron that contains all mechanisms behind the operation of it
- **Probabilistic model** that approximates neuron operation

However, in practice the first approach is unfeasible, therefore we build the black box as a **probabilistic model** of a neuron.

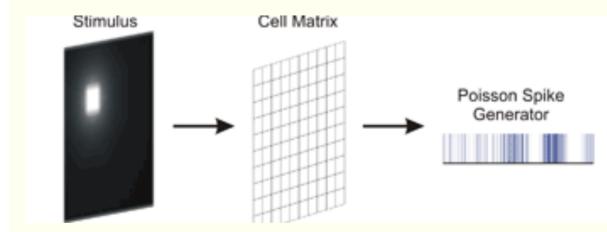
With terms **neural encoding** and **neural decoding** we refer to measure and characterization of how an **external** or **internal** stimulus (input) received by a neuron is translated into a sequence of action potentials (output). A stimulus is essentially an event that has an effect of brain activity, we can distinguish two types of stimuli:

- **External stimulus:** e.g. light or sound intensity or a picture
- **Internal stimulus:** e.g. the direction of a planned movement

.NEURAL ENCODING (From stimulus to response)

In **neural encoding** we want to find a "rule" f that links stimulus, given as input to the neuron, to its spikes train response (output).

$$\text{Neural response} = f(\text{stimulus})$$



Aim: describing how neurons react to different stimuli and trying to predict their response to a new one

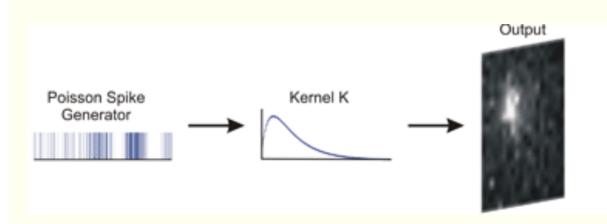
Neural encoding is important for three main reasons:

- **Understand** behavior of a neuron respect to different stimuli (e.g. visual neuron: reaction to color or orientation)
- **Replicate** neuron's behavior in a simulation
- **Predict** neuron's response to a new stimulus

.NEURAL DECODING (From response to stimulus)

In **neural decoding** we want to find a inverse "rule" f^{-1} that links neuronal response to the stimulus that induced it.

$$\text{Stimulus} = f^{-1}(\text{neural response})$$



Aim: recognize the stimulus (or its properties) that induced the spike train response

Neural decoding is important for two main reasons:

- **Understand** behavior of a neuron respect to different inputs from other neurons
- **Build** human-brain interfaces. Our prosthesis must be able to decode signal from subject's brain

OBS: (Decoding)

Our brain performs decoding at each instant, indeed each neuron must decode inputs from other cells in order to produce its response.

In this section we will study neuron **encoding**

RATE-CODING HYPOTHESIS

Everything we will see next is based on following hypothesis called **rate-coding hypothesis**:

RATE-CODING HYPOTHESIS

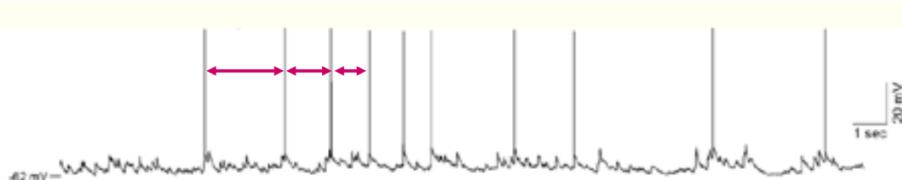
The rate-coding hypothesis suggests that spike frequency (or rate) is the fundamental mechanism of coding information

e.g.: the number of action potentials from cutaneous nerve fibers in the leg of a cat was proportional to the pressure applied to the footpad (Adrian and Zotterman, 1926)

The meaning of this hypothesis is that since action potential has always the same shape and amplitude, information content must be given by the instant of its production (firing). As we have already seen before, two close action potentials generate a higher PSP in the next cell respect to two distant ones.

There are several clues in support of this hypothesis, however it still an hypothesis.

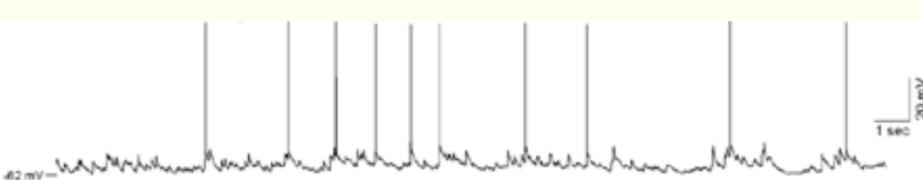
The following figure shows an intracellular recording. (We know it is an intracellular one due to sub-threshold activity)



SPIKE TRAINS REPRESENTATIONS

The first step of neuronal encoding is **spike trains representation**.

We will see two of different representations.



Neural response function

We can represent each action potential firing (spike) using Dirac δ function. Therefore, for n spikes, where i^{th} spike occurs at time t_i , the spike train can be written as:

$$rho(t) = \sum_{i=1}^n \delta(t - t_i)$$

Where:

$$i = 1, 2, \dots, n \quad 0 \leq t_i \leq T \quad T = \text{Trial duration}$$

$\rho(t)$ is called **neural response function**. It contains information about action potential timing in a limited time interval (window). Note that it does not contain sub-threshold activity.

This representation is too complex, let's see a simpler one that is used in most cases (about 90% of cases)

Spike-count firing rate

Instead of representing spikes train using Dirac functions shifted in time, we can simply count spikes occurred in a time window T , this quantity is called **spike-count firing rate**.

The spike-count firing rate r is the time average of the neural response function over the duration of the trial:

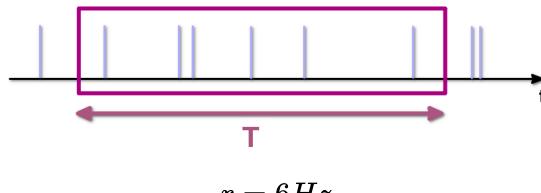
$$r = \frac{n}{T} \quad [\text{Hz}]$$

Where:

- n = Num of spikes
- T = Time window

Example

$$T = 1, n = 6$$



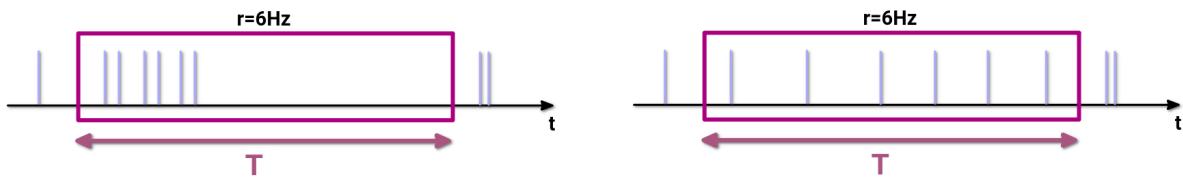
RELATION WITH NEURAL RESPONSE FUNCTION

$\rho(t)$ and r are linked:

$$\int_0^T \rho(\tau) d\tau = n \rightarrow r = \frac{1}{T} \int_0^T \rho(\tau) d\tau$$

OBS: (Temporal window T)

It is important to choose size of temporal window as small as possible, since r does not take into account distance between spikes but counts only number of occurrences in T . If T is too large, we can have a scenario similar to the following.



In this case we have same r but spike trains are very different and we know that these signals have two different effect on receiving neuron cell.

STIMULUS-RESPONSE LINK

Once defined spike-trains representations, we have to choose one of them in order to describe **stimulus-response link**. Neural response function $\rho(t)$ is a complex *deterministic* representation and for that reason in most of the cases spike-count firing rate r is used. r is useful when one wants to model neuron behavior in a probabilistic way as in our case. Another reason for use of r is the fact that neural responses can vary across repetitions (trials) even when the same stimulus is presented repeatedly, therefore we cannot describe the timing of each spike deterministically and a probabilistic approach is needed.

AVERAGE FIRING RATE

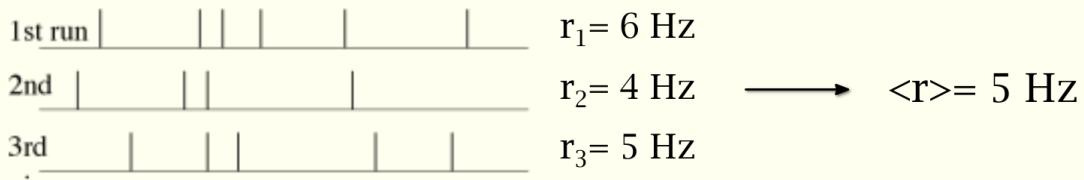
In order to define stimulus-response link, since responses change over time, we cannot compute r for a single trial but we have to perform several runs and compute an average of r collected. So assuming stationary stimuli, we define **average firing rate** $\langle r \rangle$ across trials as:

$$\langle r \rangle = \frac{\langle n \rangle}{T} = \frac{1}{T} \int_0^T r(t) dt$$

Where:

- $\langle n \rangle$: average number of spikes per trial
- $\langle r \rangle$ provides to us a more stable and robust measure of neural response to a stimulus.

Example



.TUNING CURVES

A stimulus is characterized by many properties (e.g.: shape, orientation, contrast, movement). We focus on a single property s and we assume it's stationary along each trial (repetition of the stimulation).

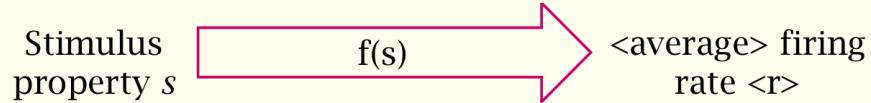
For several different stimuli $s \in S$ we compute their relative average firing rates $\langle r \rangle$ in order to define a function $f(s)$ that describes stimulus-response link. The function $f(s)$ is called **tuning curves** since it essentially describes how neural response varies according to variations of s .

$\langle r \rangle = f(s)$ (TUNING CURVE OF NEURAL RESPONSE)

WHAT HAPPENS



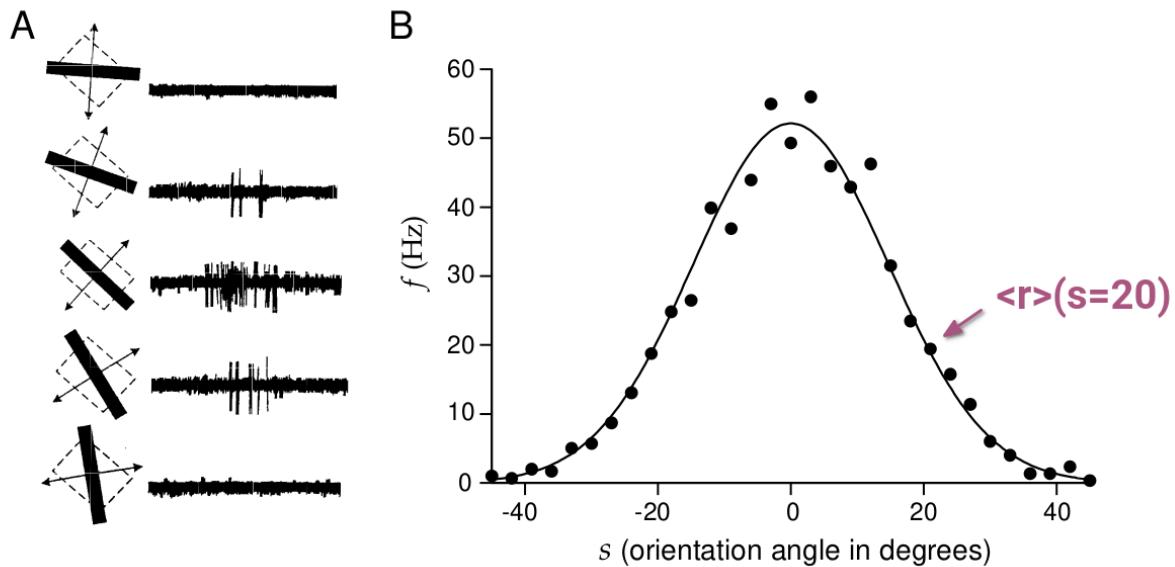
MODELLING



Let's now see some examples of tuning curves. The following are real data collected in an invasive way with an extracellular in-vivo recording

Example tuning curves - Primary visual cortex (External stimulus)

The first example comes from recording of primary visual cortex in an animal subject (monkey). Subject was stimulated by a visual stimulation, in particular a bar with different orientations, therefore s represents an orientation angle of the bar in degrees.



We can see from the figure that the neural population studied have a *preferred direction* $s = 0$.

All samples are interpolated through the following tuning curve:

$$f(s) = r_{max} \exp \left(-\frac{1}{2} \left(\frac{s - s_{max}}{\sigma_f} \right)^2 \right)$$

Where:

- r_{max} : is the highest average firing rate $\langle r \rangle$ for $s \in \mathcal{S}$
- s_{max} : angle evoking the maximum response r_{max}
- σ_f : amplitude of the tuning curve

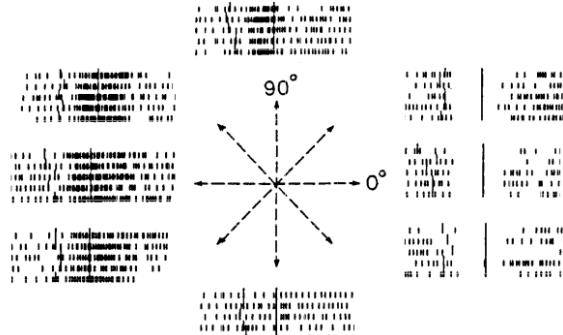
OBS: (Not unique s given response)

We can note from the figure that the encoding step is easy, given s we can obtain an unique response $\langle r \rangle$. Decoding is instead impossible for most of the values, since given a response $\langle r \rangle$, more than one stimulus are able to produce that response. The only value for which decoding is easy is for the maximum response r_{max} .

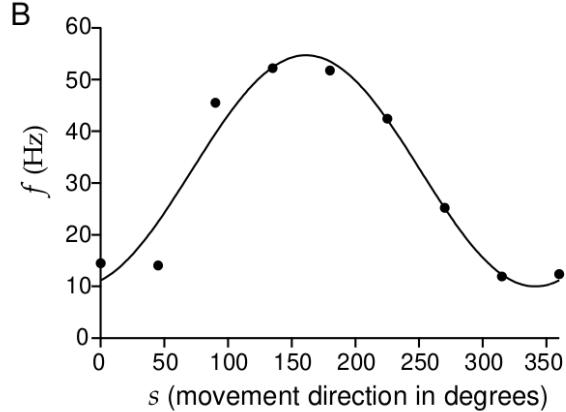
Example tuning curves - Primary motor cortex (Internal stimulus)

The second example shows responses of a neuron in the primary motor cortex of a monkey trained to reach in different directions. The firing rate of the cell is correlated with the direction of arm movement s (in degrees).

A



B



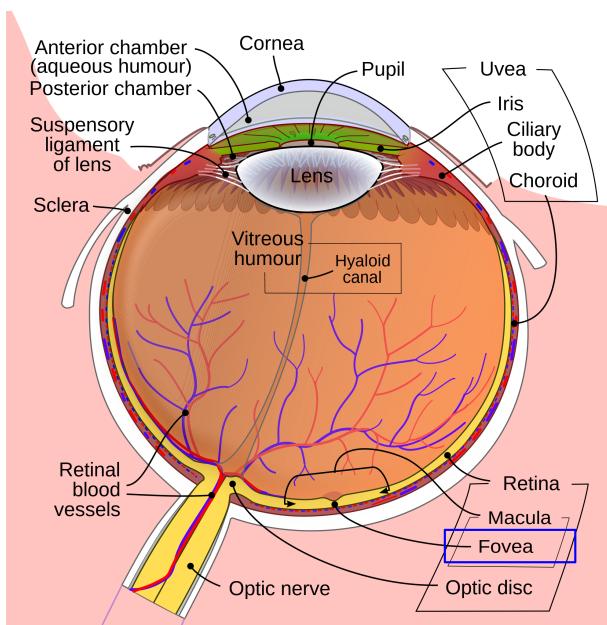
$$f(s) = r_0 + (r_{max} - r_0) \cos(s - s_{max})$$

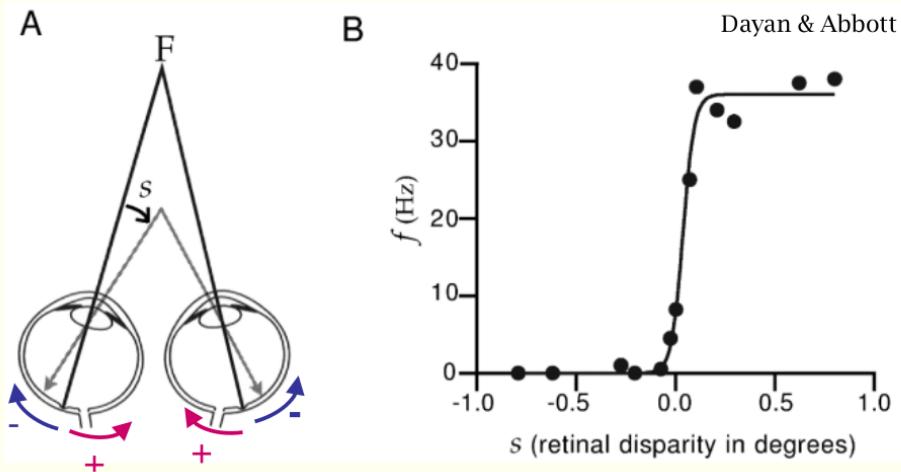
Where

- $s \in \mathcal{S}$: movement direction (degrees)
- r_{max} : is the highest average firing rate $\langle r \rangle$ for $s \in \mathcal{S}$
- s_{max} : angle evoking the maximum response r_{max}
- r_0 : offset (to avoid negative firing rate)

Example tuning curves - Retina disparity

The last example shows how primary visual cortex **fovea** neuron of a cat reacts to retinal disparity (a difference in the retinal location of an image between the two eyes)





F = Fixation point

We can see that the neuron responds only to positive s namely far objects (far-tuned neuron)

$$f(s) = \frac{r_{max}}{1 + \exp [(s_{1/2} - s) / \Delta_s]}$$

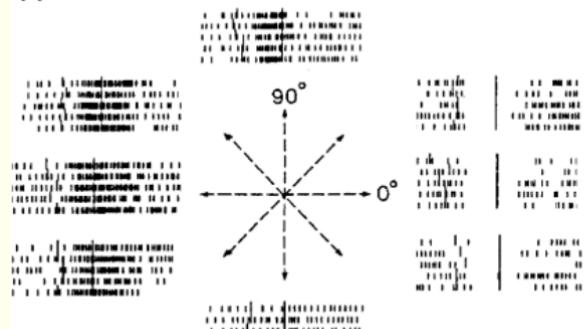
Where:

- s : binocular retinal disparity (degrees)
- $s_{1/2}$: disparity inducing a response equal to $\frac{1}{2}$ of the maximum r_{max}
- Δ_s : controls how quickly the firing rate increases as a function of s

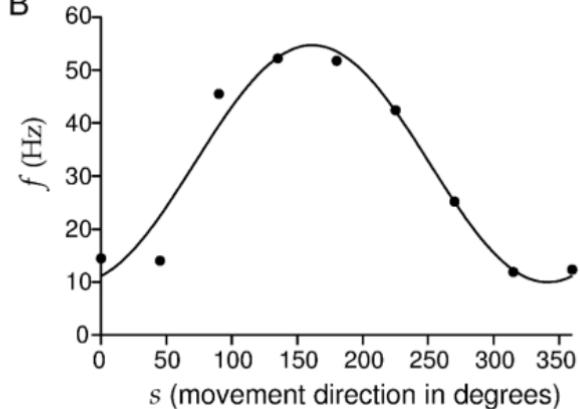
Self-evaluation test

Given the following tuning curve:

A



B



1. The neural response for a movement direction of 90 degrees is greater than for 180 degrees (T/F)
2. I will build a different tuning curve for each trial (T/F)
3. When the movement direction is 250 degrees I can expect a firing rate around 30Hz (T/F)
4. If the measured firing rate is 55Hz, I can «guess» which was the movement direction that produced that response (T/F)

References

- Dayan & Abbott:
 - Chapter 1.1 (From Stimulus to Response)
 - Chapter 1.2 (Spike Trains and Firing Rates, Tuning Curves)

05 - THE POISSON SPIKE GENERATOR - PART 2

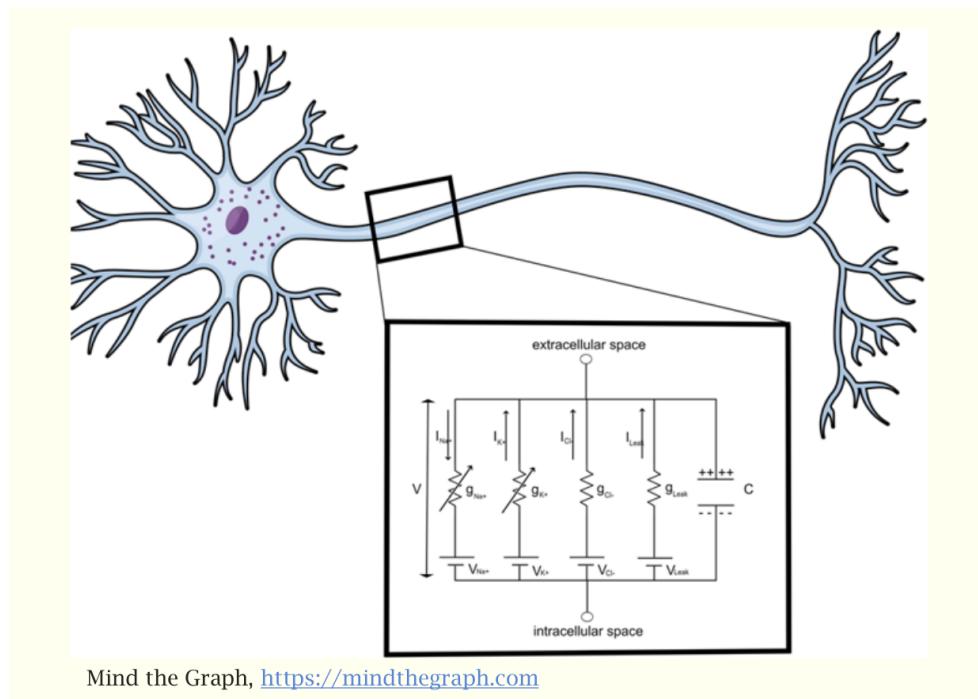
Learning objectives of the lesson

1. **Understand** the statistical properties of the spike train generated by a neuron when the firing rate is r
2. **Identify** a stochastic process able to simulate the neuronal behavior in terms of input/output relation (f)
3. **Describe** the homogeneous Poisson distribution and the inter-spike interval distribution
4. **List** the steps to build a Poisson spike generator
5. **Compare** the real data with the simulated spike trains and
 - a. **Understand** why there are such differences
 - b. **Explain** how to mitigate such differences

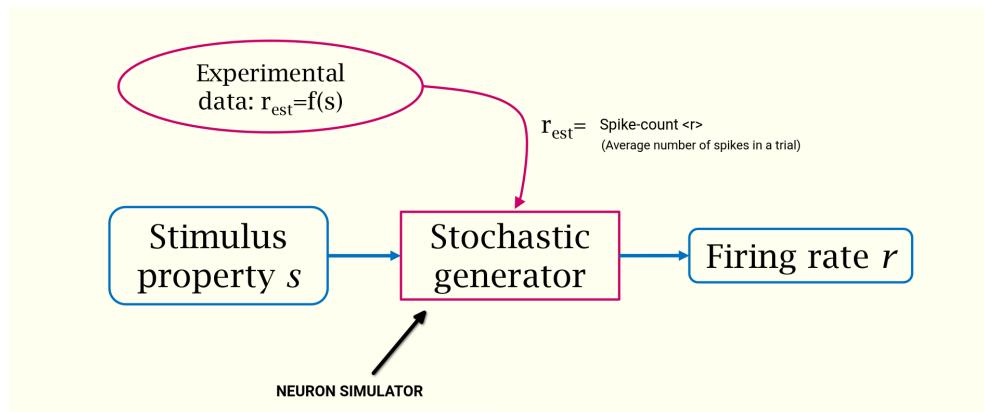
INTRODUCTION

.WHY STOCHASTIC MODEL

The relation between a given stimulus property s and the generation of a single action potential is very complex to be modeled indeed, we would need a complete electrical model of the entire membrane of each individual neuron.



For that reason, we need a statistical model that allows us to estimate the probability of an arbitrary spike sequence occurring, based on our knowledge of the responses actually recorded (empirical data).



Stochastic generator represents our neuron simulator, it must be able to simulate neuron behavior, namely given as input a stimulus s it returns as output a firing rate r . The operation of this simulator is based on collected data (pairs of $\langle s, \langle r \rangle \rangle$).

Our goal is to design a neuron simulator as simple as possible, so we rely on a stochastic model, in particular on a model based on a **Poisson process**.

THE POISSON PROCESS

Our stochastic generator is based on a **Poisson process** so let's start by introducing what it is.

Poisson Distribution

In probability theory and statistics, the Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.

A discrete random variable X is said to have a Poisson distribution, with parameter $\lambda > 0$, if it has a probability mass function given by:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where:

- k : Number of occurrences (number of events)
- λ : Expected value and variance of X (Expected number of events)

Poisson Process

In probability, statistics and related fields, a Poisson point process is a type of random mathematical object that consists of points randomly located on a mathematical space. The Poisson point process is often called simply the **Poisson process**, but it is also called a Poisson random measure, Poisson random point field or Poisson point field. This point process has convenient mathematical properties, which has led to it being frequently defined in Euclidean space and used as a mathematical model for seemingly random processes in numerous disciplines.

Poisson Process's name derives from the fact that if a collection of random points in some space forms a Poisson process, then the number of points in a region of finite size is a random variable with a Poisson distribution.

A Poisson point process is characterized via the Poisson distribution. The Poisson distribution is the probability distribution of a random variable N (called a Poisson random variable) such that the probability that N equals n is given by:

$$P[N = n] = \frac{\Lambda^n}{n!} e^{-\Lambda}$$

Where:

- Λ determines the shape of the distribution

We want to model neuronal behavior through a Poisson Process in which a single point represents a spike and the mathematical space is a time interval T . We will see more details about this in the next section.

Source: https://en.wikipedia.org/wiki/Poisson_point_process

NEURON AS POISSON PROCESS

In order to model neural behavior through a Poisson process we have to make the following very **strong approximation**:

ASSUMPTION

All action potential generations are independent from each others, namely the probability of generating an action potential is **independent** of the presence or timing of other spikes.

NOTE: (Very strong approximation/assumption)

The previous one is a huge approximation, indeed we already know that actually action potential generations are **not** independent due to **refractory periods**. We will see later on how to include this in our model.

So, assuming the probability of generating an action potential (event) is independent of the presence or timing of other spikes, the firing rate r is all we need to compute the probabilities for all possible action potential sequences.

We can use two different extremely useful approximations of stochastic neuronal firing based on a Poisson process:

- **Homogeneous Poisson process:** When the firing rate r is **constant** over time in a single trial
- **Inhomogeneous Poisson process:** For a **time-dependent** firing rate $r(t)$ that changes over a single trial

In this course we will use an **homogeneous** Poisson process, namely we assume that the firing rate r is constant over a single trial. This means that given a firing rate r , all possible sequence of spikes have the same probability:

$$r = 3$$



However, to us the previous sequences are very different since they have two different effects of the next receptor neuron. We will see later how to deal with this problem.

Inhomogeneous Poisson process

If a Poisson point process has a parameter of the form $\Lambda = \nu\lambda$, where ν is Lebesgue measure (that is, it assigns length, area, or volume to sets) and λ is a constant, then the **point process is called a homogeneous** or stationary Poisson point process.

The parameter Λ , called rate or intensity, is related to the expected (or average) number of Poisson points existing in some bounded region, where rate is usually used when the underlying space has one dimension. The parameter λ can be interpreted as the average number of points per some unit of extent such as length, area, volume, or time, depending on the underlying mathematical space, and it is also called the mean density or mean rate.

The homogeneous Poisson point process, when considered on the positive half-line, can be defined as a counting process, a type of stochastic process, which can be denoted as

$\{N(t), t \geq 0\}$ A counting process represents the total number of occurrences or events that have happened up to and including time t . A counting process is a homogeneous Poisson counting process with rate $\lambda > 0$ if it has the following three properties:

- $N(0) = 0$
- It has an independent increments
- The number of events (or points) in any interval of length t is a Poisson random variable with parameter (or mean λt)

In other words, the probability of the random variable $N(t)$ being equal to n is given by:

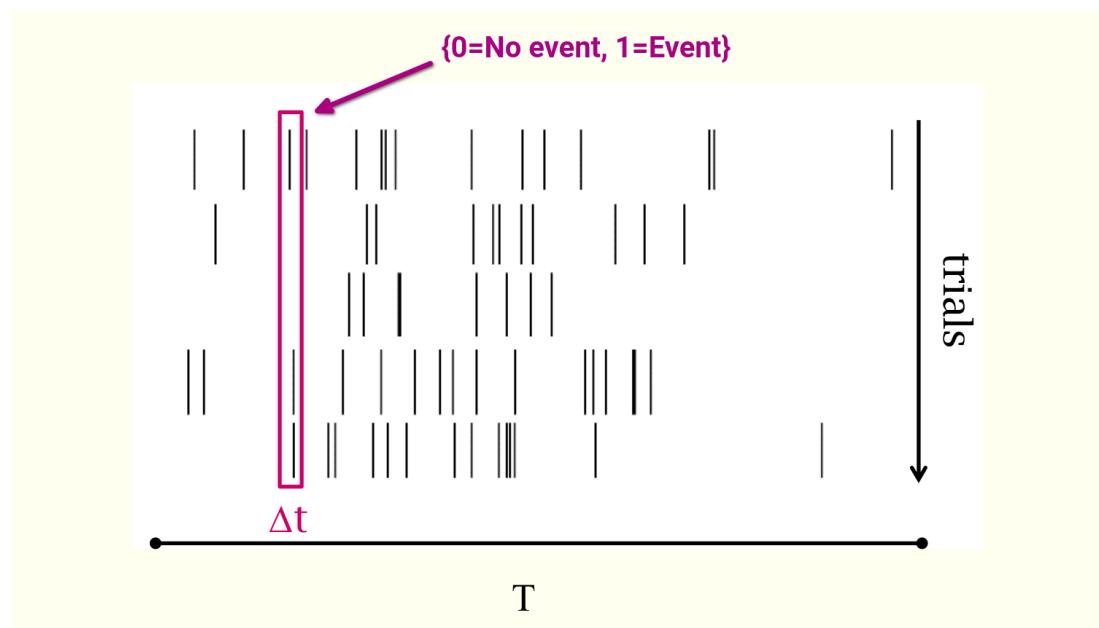
$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

NEURON AS HOMOGENEOUS POISSON PROCESS

A Poisson stochastic process produces an integer, non-negative number of statistically independent discrete events.

For the neuron:

- Each **action potential is an event**
- By **dividing the trial duration T into very small time intervals Δt ($\rightarrow 0$)** we have:
 - > **No more than a single spike** for each interval (Δt small enough)
 - > **Independent events** (spikes) at each interval (Strong assumption)



In a homogeneous Poisson process, every sequence of n spikes has an equal probability to be generated. To us, the distance between subsequent spikes is particularly interesting.

Therefore, we will focus on:

1. The probability of n spikes in a trial of duration T
2. The inter-spike interval distribution

Probability n spikes

First at all we are interested to define the **probability of n spikes** to occur during a trial when the neuron is subjected to a specific stimulus s . To do that we require the **average firing rate r** that we assume to be obtained from a tuning curve previously computed from several trials and stimuli.

So given a firing rate r computed as the average firing rate across trials (sometimes indicated as $\langle r \rangle$), and a time window T , the probability of n spikes occurring during a trial of length T is given by:

$$P_T[n] = \frac{(rT)^n}{n!} e^{-rT}$$

Where:

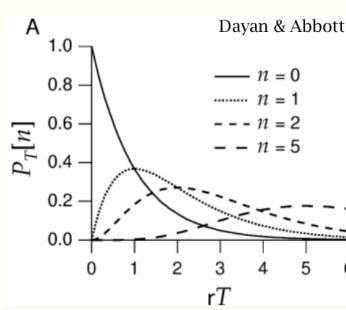
- rT is the average (across trials) number of spikes in T

NOTE: (*Probability spike in Δt*)

$r\Delta T$ can be seen as the probability of having a spike in Δt

Let's look at the shape of a Poisson distribution:

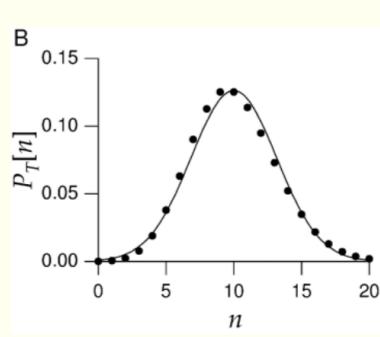
Probability of different n as a function of rT



We can make the following observations:

- For a given n , the maximum $P_T[n]$ corresponds to $rT = n$
- When T increases, higher values of n are more likely
- The same applies when r increases

Probability of different n when $rT = 10$



We can make the following observations:

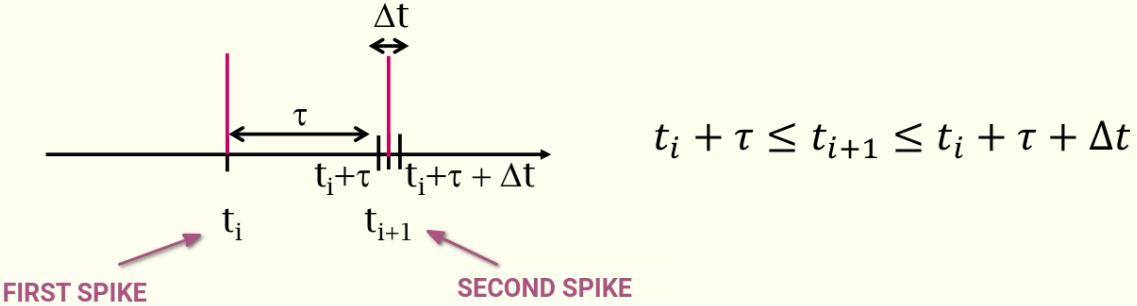
- $N = 10$ is the most likely
- When $rT \geq 10$, the distribution becomes a Gaussian with expected value and variance both equal to rT

Inter-spike interval distribution

Since to us the distance between two spikes is very interesting, let's see how to define the **inter-spike interval distribution**.

Given a spike at t_i , the probability of the following spike to be produced in the interval τ is equal to:

PROBABILITY OF NO SPIKES FOR A TIME INTERVAL τ * PROBABILITY OF A SPIKE

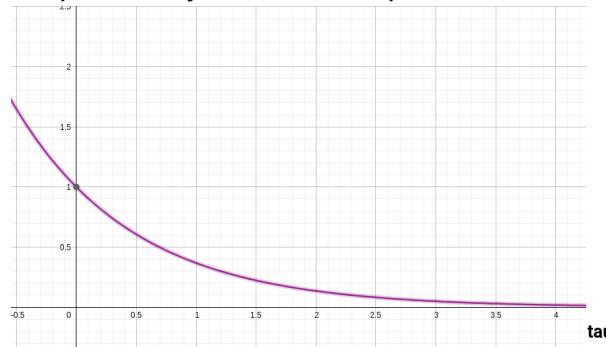


Probability of no spike in τ

Fixing $n = 0$ and $T = \tau$ in the Poisson distribution we obtain:

$$P_T[n] = \frac{(rT)^n}{n!} e^{-rT} \rightarrow P_\tau[0] = \frac{(r\tau)^0}{0!} e^{-r\tau} = e^{-r\tau}$$

So given a fixed r we have probability decreases exponential in time



Probability of a spike in Δt

As said before, the probability of a spike to occur in a small interval of time Δt is given by:

$$r\Delta t$$

Finally putting these two independent probability together we obtain the **inter-spike interval distribution**:

$$P[\tau] = P[t_i + \tau \leq t_{i+1} \leq t_i + \tau + \Delta t] = r\Delta t e^{-r\tau}$$

In this way we are able to distinguish the probability of two different sequence of spikes, indeed the most likely inter-spike intervals are short ones, and long intervals have a probability that falls exponentially as a function of their duration. So as happens in a real neuron, the sequence on the right is more likely respect to the left one:

$r = 3$



Property of inter-spike distribution

For our course we will not use the following quantities however they are very useful in practice to estimate the quality of the model.

For a homogeneous Poisson process (r constant):

- The mean of the inter-spike interval $\langle \tau \rangle = 1/r$
- The variance of the inter-spike interval $\sigma_\tau^2 = 1/r^2$

The ratio of the standard deviation to the mean $\sigma_\tau / \langle \tau \rangle$ is called the coefficient of variation (Cv) and it equals to 1 for a homogeneous Poisson process

THE POISSON SPIKE GENERATOR

Once define the inter-spike distribution we are ready to introduce **the Poisson spike generator**.

The Poisson spike generator is a stochastic model of the neural response to a stimulus property s . It produces spike trains with the same r of the neuronal response, before built it we have to collect a estimate of r , so we have the following two steps to follow:

- **Step 1:** Experimental estimation of the firing rate r_{est} (tuning curve)
- **Step 2:** Build a spike generator based on a Poisson distribution with $r = r_{est}$

We already know how to compute a tuning curve, so let's see how to build the spike generator instead.

BUILD POISSON SPIKE GENERATOR

Let's see a simple algorithm that a computer can run in order to approximate the Poisson process that models neuron's behavior.

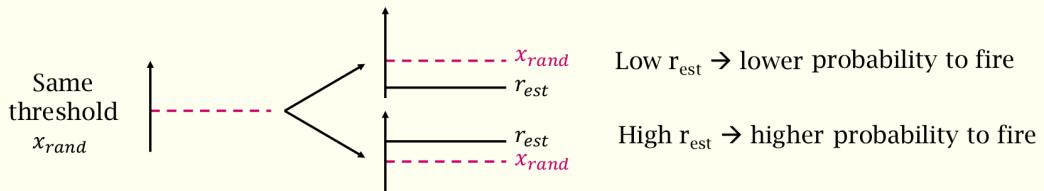
Hypothesis: The probability to generate a spike during Δt is $r_{est}\Delta t$

Algorithm:

- The program progresses through time in small steps of size Δt
- it generates, at each time step, a random number x_{rand} chosen uniformly in the range $[0, 1]$ (threshold, independent from r_{est})
- For each Δt if $r_{est}\Delta t \geq x_{rand}$ at that time step, a spike is fired, otherwise it is not

$$\text{Spike is fired} = \begin{cases} \text{True} & \text{if } r_{est}\Delta t \geq x_{rand} \\ \text{False} & \text{if } r_{est}\Delta t < x_{rand} \end{cases}$$

So given a fixed random threshold $x_{rand} \in [0, 1]$, a spike has more probability to be fired in Δt when r_{est} is higher.



OBS: (Random threshold)

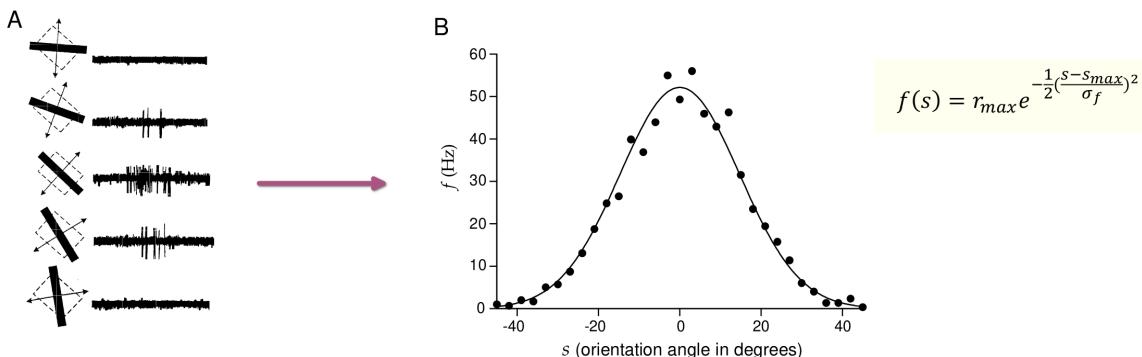
Usually when we deal with a random variable we choose a fixed deterministic threshold and then we sample the random variable and compare it with the threshold. The previous algorithm does the exact opposite.

EXAMPLE

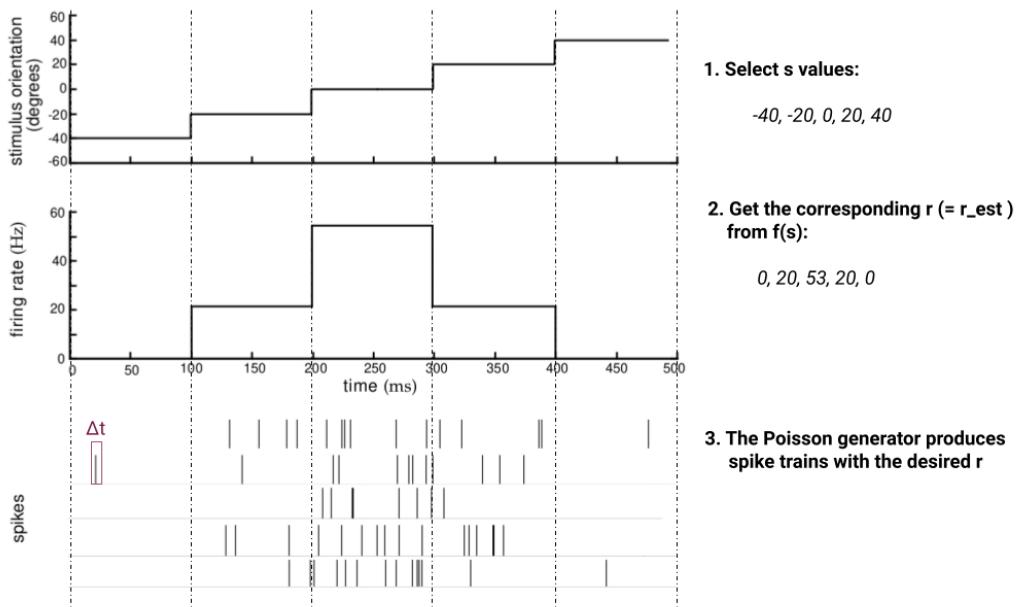
Let's see an example of a Poisson spike generator.

Step 1: (Experimental estimation of the firing rate r_{est} (tuning curve))

To imitate neuron behavior first at all we have to collect experimental data and build the tuning curves $f(s)$ that describes how the neuron responds with its firing rate r to different stimuli $s \in \mathcal{S}$. So our $r_{est}(s)$ will depend on the stimulus s in input.



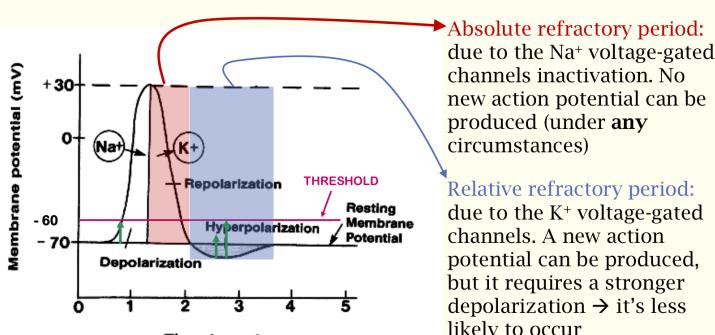
Step 2: (Build a spike generator based on a Poisson distribution with $r = r_{est}$)



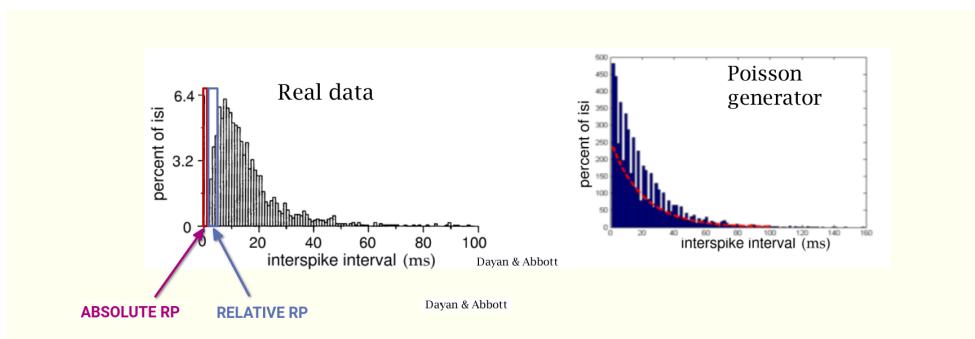
LIMITATION POISSON GENERATOR

First at all let's see a brief recall about refractory period in action potential production:

Absolute and relative refractory periods



If we make a comparison between two inter-spike interval (isi) histograms, one obtained from real data and the other generated by our model, we can note a big difference for low values:

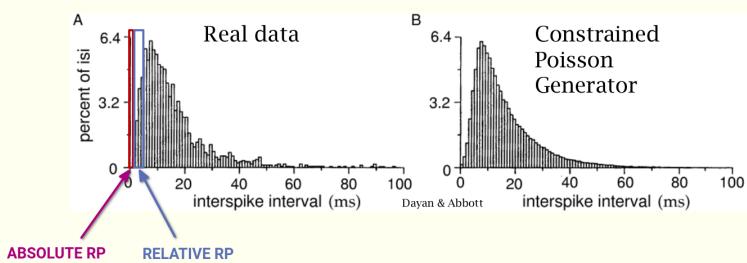


Our generator follows an exponential evolution $r\Delta t e^{-r\tau}$ that does not align with real data when isi is very low.

This difference is due to lack of refractory period in our model, in a real neuron spike generations became independent only after a time τ equals to refractory period duration. During the absolute refractory period of a spike, another spike cannot be produced and in the relative refractory period is less likely to happen, this fact is not included in our model.

Constrained Poisson Generator

In our previous neuron model we did not consider the refractory periods, in the actual neuron, the minimum inter-spike distance is constrained. We can build a more complex (and realistic) generator including such constraints based on Gamma distribution and called **constrained Poisson generator**:



Self-evaluation test

1. When the occurrence of each spike is independent from the others, the firing rate r is sufficient to compute the probabilities for all possible action potential sequences (T/F)
2. In a Poisson process, when r increases, higher values of n are more likely (T/F)
3. Long inter-spike intervals (isi) have a probability that falls with their duration according to the exponential law (T/F)
4. The differences between the distribution of isi in real data and in simulated data produced by a Poisson generator are due to the refractory periods (T/F)

References

- Dayan & Abbott:
 - Chapter 1.4 (Spike Train Statistics; The Homogeneous Poisson Process, The Poisson Spike Generator, Comparison with Data)

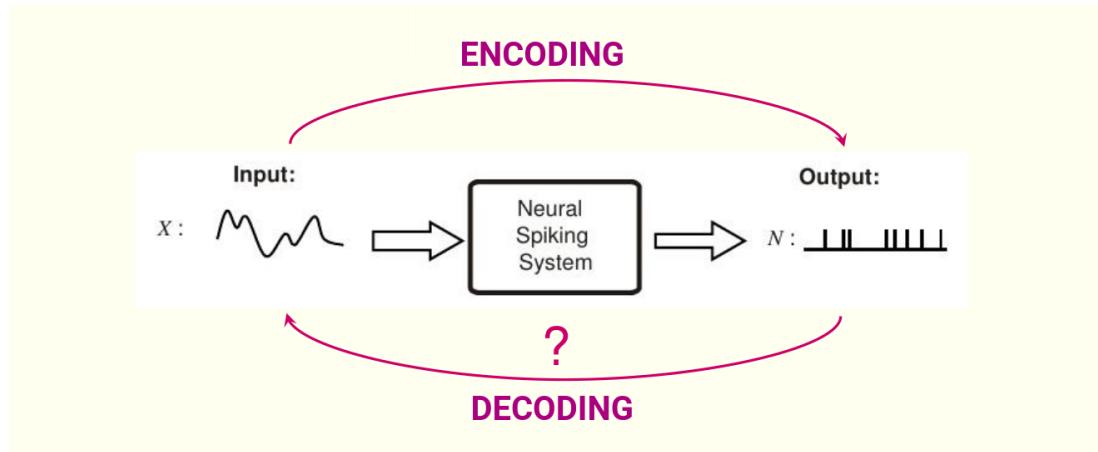
06 - NEURAL DECODING

Learning objectives of the lesson

1. **Understand** the meaning and need for the neural decoding
2. **Interpret** a distribution of firing rates as a function of s
3. **Define** the discriminability d' and its use
4. **Illustrate** how to perform a classification-based decoding based on a threshold z
5. **Introduce** FP, FN, TP and TN
6. Given a Receiving Operator Characteristic curve, **interpret** its meaning in terms of accuracy, depending on:
 - a. The experimental conditions
 - b. The classification choices
7. **Describe** what the Area Under the Curve means and the values it can assume

INTRODUCTION

Neural (en/de)coding: measuring and characterizing the link between a stimulus feature s and the train of action potentials generated as the neuronal output in response to the stimulus.



As already said in previous sections, the black box can represent a single neuron or a group of neurons with the same response to a stimulus. It is important to highlight the fact that the input (stimulus) can come from other neurons, therefore neural encoding and decoding are also very useful to understand how brain works.

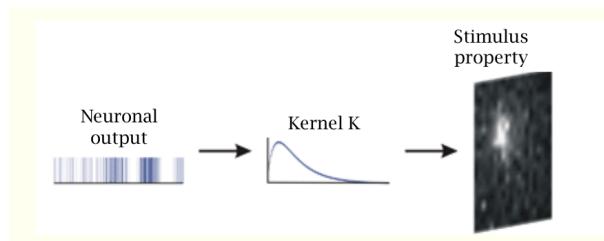
In this section we will focus on neural encoding, so let's start with a brief recall and an example of neural decoding

.AIM NEURAL DECODING (Recall)

The **aim of neural decoding** is to compute an **estimation** of the stimulus that produced a given neuronal output. We use the term **estimation** to highlight the higher variability that we have with respect to neural encoding, indeed as we have already seen in a tuning curve, several stimuli can correspond to the same output (response) therefore usually it is impossible to determine the exact stimulus that produced the spike train (output).

$$\text{Stimulus} = f^{-1}(\text{neural response})$$

Aim: to trace back the stimulus properties from the spike train that they induced



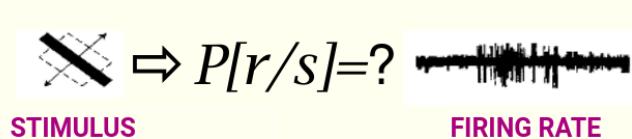
OBS: (Single trial)

In neural decoding we cannot take several trials and compute their average since we do not know the meaning of each trial at prior. Therefore we can work using only a trial.

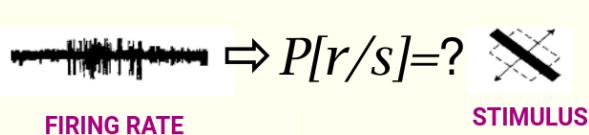
.ENCODING AND DECODING AS CONDITIONAL PROBABILITIES

As said in previous sections, our goal is to build a probabilistic model of a neuron, both neural encoding and decoding can be seen as computations of conditional probabilities:

- **Encoding** consists of computing $P[r|s]$, the probability of a response with firing rate r given a stimulus with property s



- **Decoding** consists of computing $P[s|r]$, the probability of a stimulus with property s , given that the neural response has firing rate r



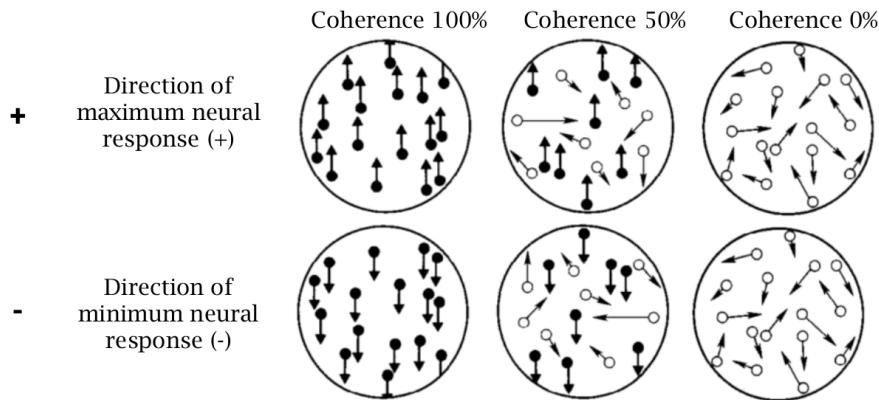
Determining what is going on in the real world from neuronal spiking patterns (neural decoding) is the natural aim of the nervous system. We need to understand how this works in order to interface an external device with the brain.

To introduce neural decoding we will use a practice example, we will start by studying how subject's brain response to certain stimuli (firing rates from other neurons) then we will try to define a model able to have similar decoding performance respect to the subject.

EXAMPLE OF DECODING (VISUAL DISCRIMINATION)

In order to introduce neural decoding we will use the following **visual discrimination task** in which the subject or our model has to make a binary classification.

The task is simple, given an image similar to the following ones, subject has to determine the overall direction of the image (+) or (-).



So the task consists to recognize the motion direction of dots on the screen, we have two possible directions named in the following way:

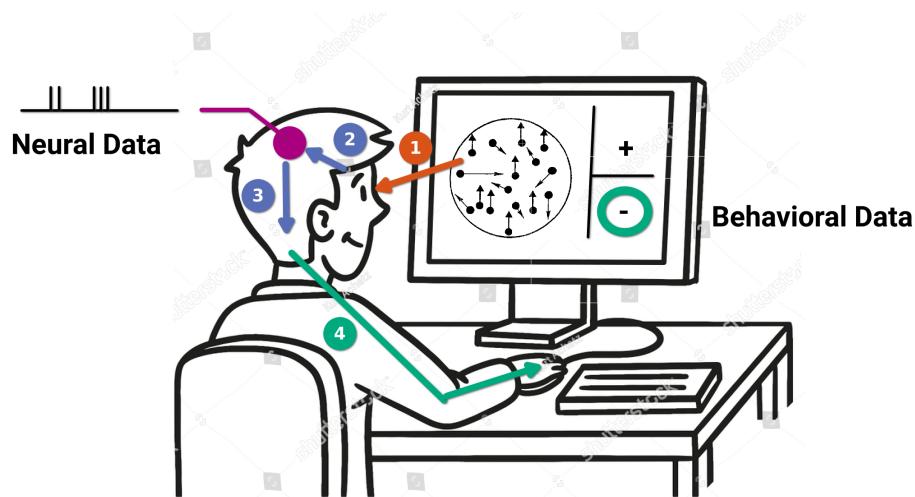
- +: that corresponds to the maximum neural response
- -: that corresponds to a minimum neural response

We use the terms maximum and minimum responses since we will see that brain neurons studied react only to one direction (preferred direction)

As we can see in the previous image, we can have a very different quantity of "noise". We control this noise thought the **coherence level** defined as the percentage of dots moving in the same direction. So changing the coherence level means controlling noise. When we have zero coherence it means that dots move randomly and subject decides direction randomly.

.EXPERIMENTAL SETUP

The **experimental setup** used for this example is the following one, we show images with different coherence levels to the subject and we ask him to choose the correct direction (+) or (-):



In this setup we can distinguish four steps:

- **Step 1:** Image reaches subject's eyes
- **Step 2:** Image signal passes from eyes to a neuron on visual system, this neuron is shown in magenta
- **Step 3:** Subject's brain processes output of the "magenta" neuron
- **Step 4:** Subject's brain sends signal to muscle in order to make the classification choice

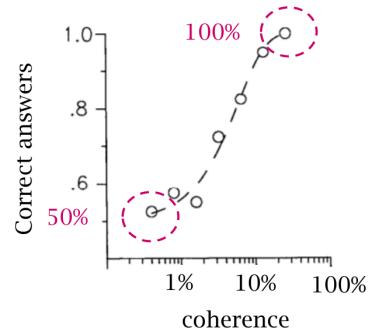
In neuronal decoding we are interested into understand the effect of the output "magenta" neuron to the classification choice (Step 3). Our goal will be to build a model that imitates the decoding processing of subject's brain starting by output of the "magenta" neuron.

During the experiment we collect two kinds of data:

- **Behavioral data:** Classification choices made by the subject
- **Neuronal data:** Firing rate collected in an invasive way from "magenta" neuron output

Behavioral data

We represent **behavioral data** as percentage of correct recognition of the motion direction (between + and -) as a function of coherence:

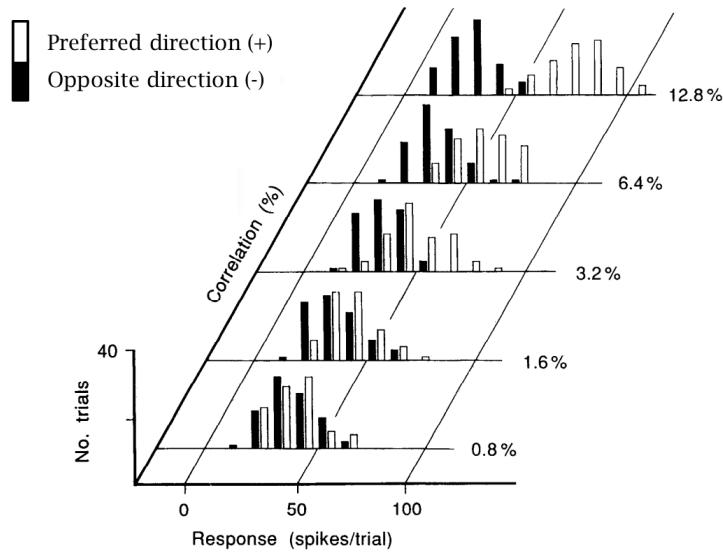


Note that coherence axis uses a logarithmic scale. We are mainly interested in results in middle range since when the coherence is high we have a perfect classification and when it is low we have the random baseline.

Neural data

During the experiment we also collect **neural data** namely the output of "magenta" neuron. The following figure shows neuron output (firing rate) respect to combination of the two factors:

- Movement direction [+,-]
- Coherence level [0.8%, 1.6%, ..., 12.8%]



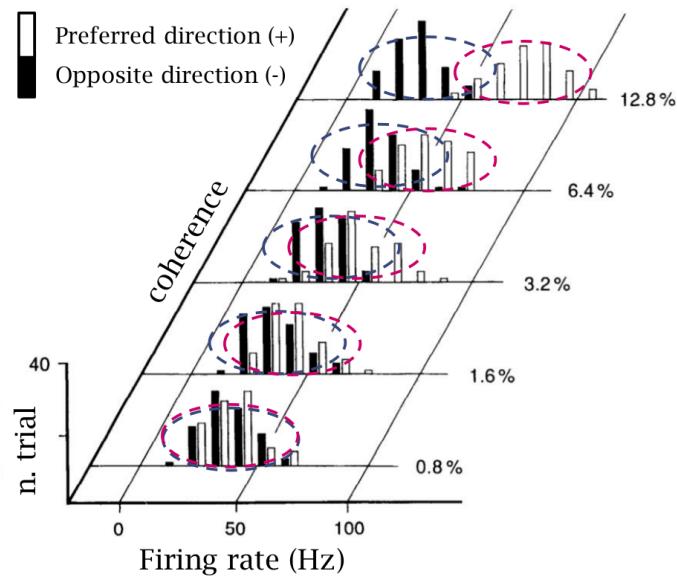
Histogram has been built on 60 trials for each combination of the two factors, considering the choice made by the subject.

The previous figure essentially shows **encoding** process performed by "magenta" neuron when it receives one of test images. We can note that the neuron reaches only to movements in its preferred direction (+)

NEURAL DECODING

Once collected neural data we can build our **neural decoding** model. Assuming we know coherency level and firing rate of "magenta" neuron, we want to imitate subject classification choice, namely imitate the neural decoding process made by subject's brain.

For each coherence level we can represent the classification choice as two Gaussians $P[r|+]$ and $P[r|-]$ with the same variance σ_r^2 and different means $\langle r \rangle_+$ and $\langle r \rangle_-$.



We can note that for high levels of coherency it is easy to make the discrimination and imitate subject's decision. For instance, consider 12.8% coherence level, we can choose a threshold $z = 50\text{Hz}$, if we measure a firing rate higher than z then we choose (+) otherwise (-). For lower levels of coherence the discrimination is more difficult though.

We can formalize this concept with **discriminability** d' :

$$d' = \frac{\langle r \rangle_+ - \langle r \rangle_-}{\sigma_r}$$

So given a threshold z between the two distributions we make our classification as:

- If $r \geq z$ the inferred direction is (+)
- If $r < z$ the inferred direction is (-)

The probability to discriminate between the two directions (classification performance) is increased with increased coherency because the two distributions are more separate. Performance depends also on the choice of z .

EVALUATION OF THE CLASSIFICATION

Now we are interested in the **evaluation of the classification** performed by our neural decoding model that essentially uses a threshold z .

Recall basic quantities

If +:

$$P[r \geq z|+] = \beta(z) \text{ (true positive)}$$

$$P[r < z|+] = 1 - \beta(z) \text{ (false negative)}$$

If -:

$$P[r \geq z|-] = \alpha(z) \text{ (false positive)}$$

$$P[r < z|-] = 1 - \alpha(z) \text{ (true negative)}$$

Probability	Stimulus	
	+	-
$r \geq z$	$\beta(z)$	$\alpha(z)$
$r < z$	$1 - \beta(z)$	$1 - \alpha(z)$
	$\beta(z) + [1 - \beta(z)] = 1$	$\alpha(z) + [1 - \alpha(z)] = 1$

Ideally:

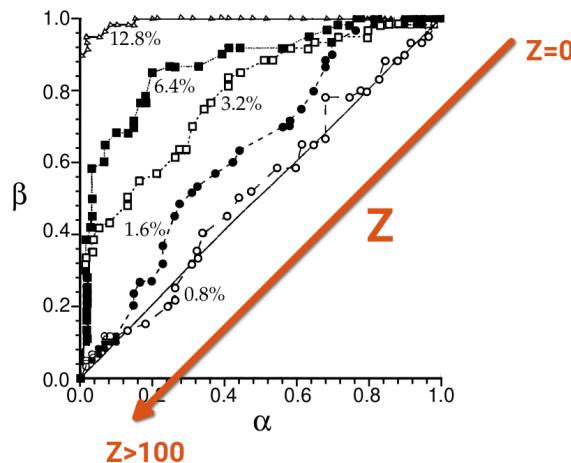
$$\begin{cases} \beta(z) = 1 \\ \alpha(z) = 0 \end{cases}$$

Receiver Operating Characteristic (ROC) curves

We represent variation of classification performance according to the threshold z and the level of coherence using **ROC Curves** in which:

- A curve for each coherence level
- A dot for each threshold z with its false positives $\alpha(z)$ and true positive $\beta(z)$

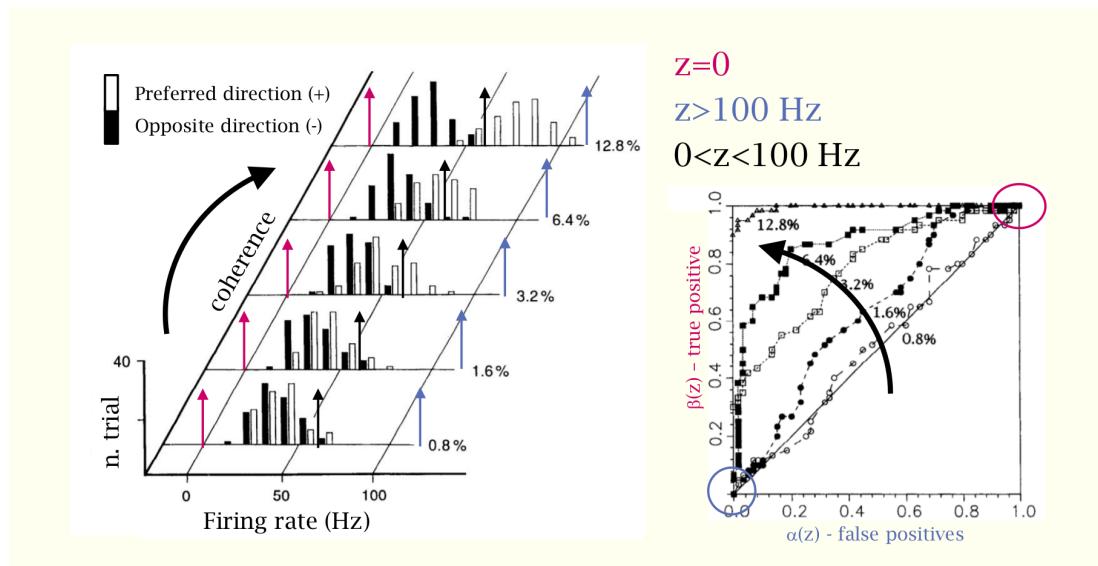
For each curve, we start with $z = 0$ (always (+) prediction) at the upper right corner and we end up with $z > 100$ (always (-)) at the bottom left corner.



So the curves share the same starting and ending points:

- $z = 0 \rightarrow \alpha(z) = 1, \beta(z) = 1$
- $z > 100 \rightarrow \alpha(z) = 0, \beta(z) = 0$

We can understand better the effect of z on the performance by looking also to the neural data collected:



Each ROC summarizes the classification performances:

- Curve close to the diagonal \rightarrow random classification (50%)
- Curve close to the upper left corner (ideal conditions) \rightarrow maximum accuracy (close to 100%)

Area Under the Curve (AUC)

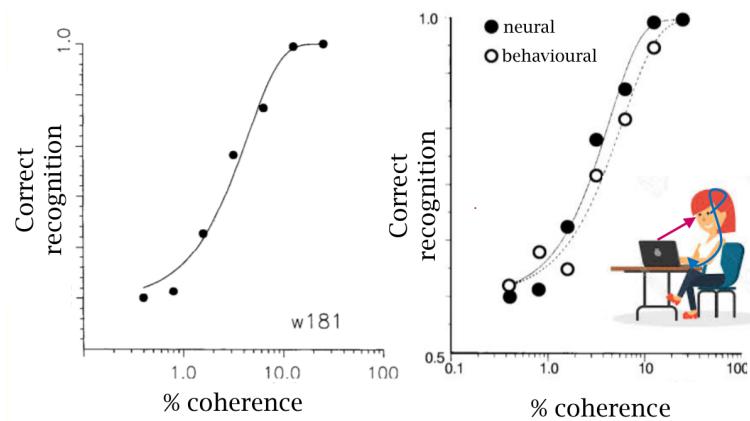
An useful quantity that we compute for each curve is **Area Under the Curve (AUC)** that indicates the percentage of correct classifications between 0.5 (chance level) and 1 (perfect classification)

NOTE: (No unique z)

There is not any unique ideal threshold z that allows to obtain the maximum performance for each level of coherence.

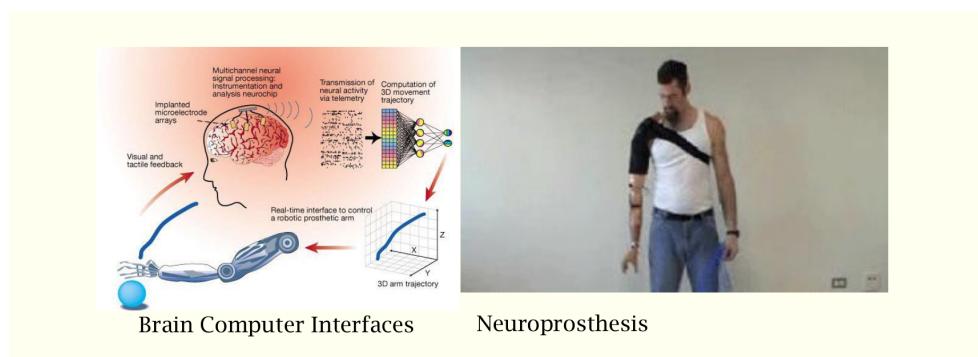
.SUBJECT AND CLASSIFICATION PERFORMANCES

Now we are finally ready to compare performance of our Z threshold based neural decoding model with the performance obtained by subject:



The classifier performances are close to the subject performances therefore for this family of neurons, we can hypothesize that a threshold mechanism can correctly describe the neuronal information processing

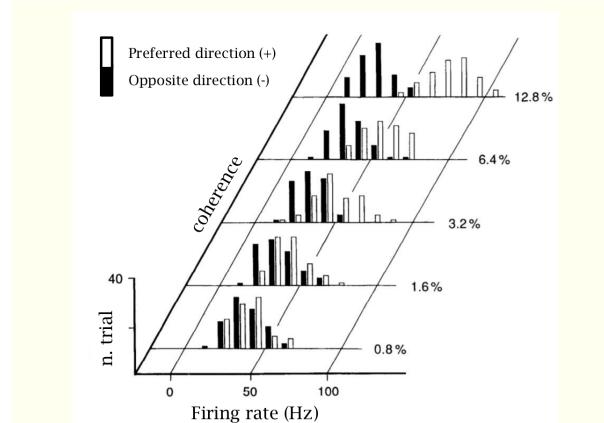
APPLICATIONS OF NEURAL DECODING



Self-evaluation test

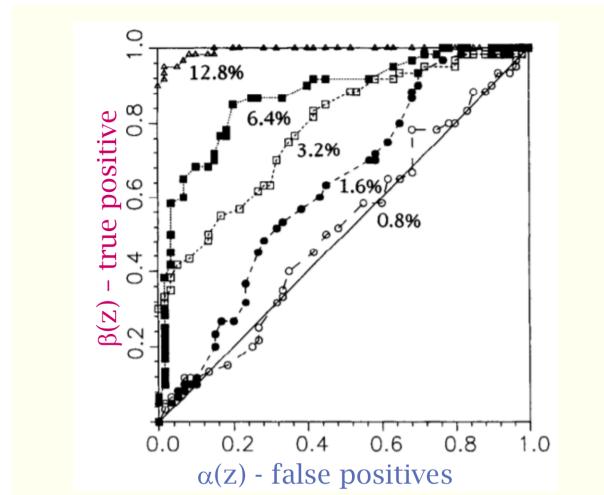
1. Given the distribution of firing rates in the figure:

- A. The discriminability d' is higher when the coherence is =1.6 than =6.4 (T/F)
- B. There is an optimal value of z that can be used for all coherence levels (T/F)
- C. r_+ distribution is more affected by the coherence level than r_- (T/F)



2. Given the ROC curves in the figure:

- A. The related AUC is between [0,1] (T/F)
- B. The classification performances do not depend on z (T/F)
- C. The classification performances depend on the experimental conditions (T/F)
- D. The ideal curve is the one closer to the diagonal (T/F)



References

- Dayan & Abbott:
 - Chapter 3.1 (Encoding and decoding)
 - Chapter 3.2 (Discrimination; ROC curves; ROC Analysis of Motion Discrimination)

07 - BRAIN NETWORKS - PART 1

Learning objectives of the lesson

1. **Understand** the need for the multivariate analysis of brain activity
2. **List** the different definitions of brain connectivity and their meaning
3. **Define** the spectral matrix and the spectral index called ordinary coherence
4. **Describe** its advantages and limitations
5. **Explain** the main limitations of the correlation analysis and the problem of the common source

INTRODUCTION

.MULTIVARIATE ANALYSIS OF BIOLOGICAL SIGNALS

When we analyze biological signals we can make two different types of analysis:

- **Univariate Analysis:** We analyze each signal (physiological correlate) independently from the others
- **Multivariate Analysis:** In complex systems, we need to analyze the **signals AND their interdependency**



NOTE: (*Interdependency*)

The key point that distinguishes univariate and multivariate analysis is that in this last we are also interested to study the interaction between the signals, so not only their individual values.

.NETWORK NEUROSCIENCE

Through an integrative perspective, **network neuroscience** aims to map, record, analyze and model the elements and interactions of neurobiological systems. In a nutshell, network neuroscience models parts of the brain as a network.

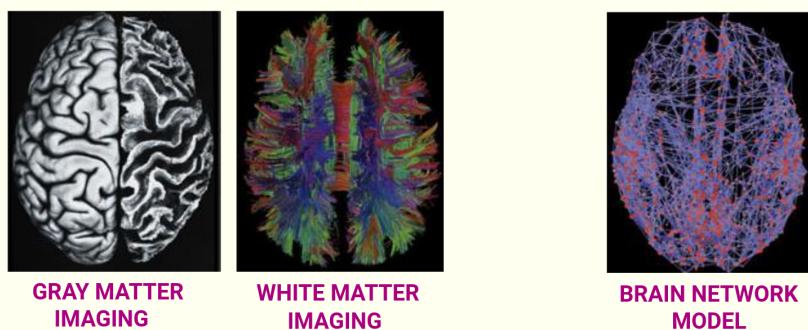
Objective: to estimate **dynamic** brain networks at all levels (among molecules, neurons, brain areas and social systems). Convergence of empirical and computational advancements, to study network dynamics and integrate network processes across spatiotemporal domains.

We used the term **dynamic** for the brain networks since it cannot be static, if the network represents information exchanges then the network changes very quickly, for instance, it changes when the subject moves a finger. Instead, if the network represents brain structure, the changes happen more slowly (Brain plasticity).

Example brain network

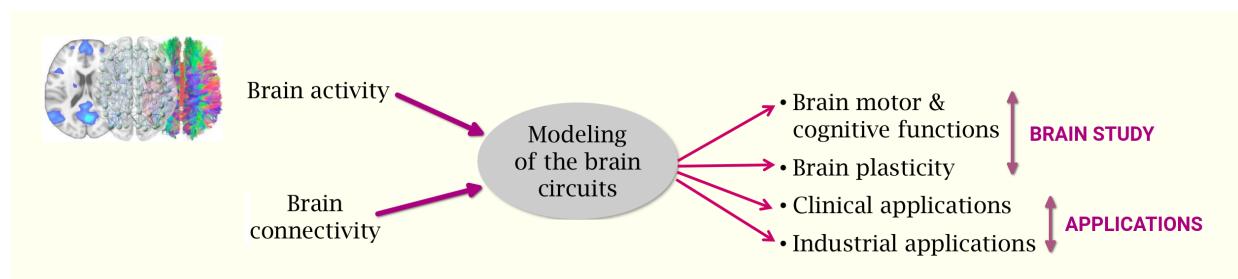
Neural populations (functionally specialized regions) are physically connected (anatomical connectivity) and interact within and between themselves (brain networks)

Adapted from: Sporns, Olaf, and Patric Hagmann. 2008. *The Human Connectome*.



UNDERSTANDING THE BRAIN FUNCTIONS

In order to **understand the brain functions** and develop clinical and industrial applications, we build a model of the brain circuit using **brain activity** information and **brain connectivity**:



There are different types of brain connectivity:

- **Anatomical connectivity:** Physical connection
- **Functional connectivity:** Two or more regions have similar behaviors
- **Effective connectivity:** A brain region influences the behavior of another one

BRAIN CONNECTIVITY

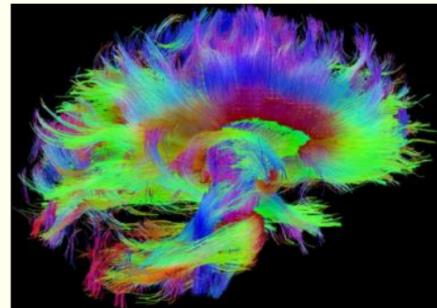
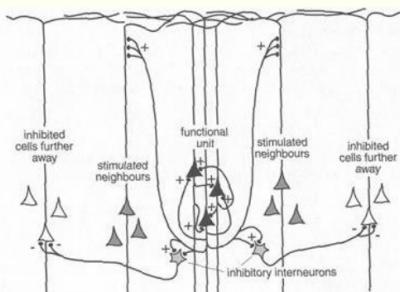
Let's see the different types of **brain connectivity** in more details.

Anatomical connectivity

Physical or structural connections linking sets of neurons or neuronal elements.

Anatomical connectivity is relatively stable at shorter time scales (seconds to minutes), at longer time scales (hours to days), structural connectivity patterns are likely to be subject to significant morphological change and plasticity

It can be acquired by invasive tracing studies or by noninvasive diffusion weighted imaging techniques. The following image on the right shows a tractography that represent essentially neurons axons. It is important to highlight that the anatomical connectivity is independent by what the subject was doing during the imaging acquisition.



NOTE: (*Limitation*)

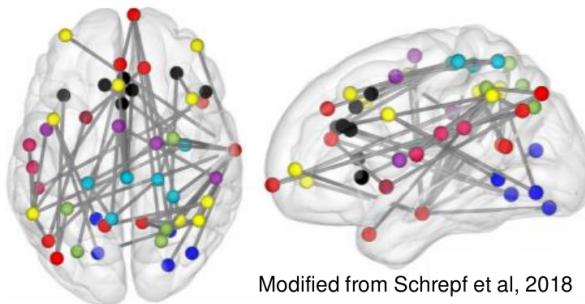
Anatomical connectivity is at the basis of functional/effective connectivity, but **cannot explain it**

Study of anatomical connectivity in clinical setting has two main applications:

- Monitor evolution of a brain disease
- Monitor outcome and progress during neurorehabilitation (e.g. after a stroke)

Functional connectivity

Functional connectivity represents how much two or more brain regions are related between them. Functional connectivity is essentially descriptive.



Modified from Schrepf et al, 2018

Relation between regions can be quantified with different measures such as:

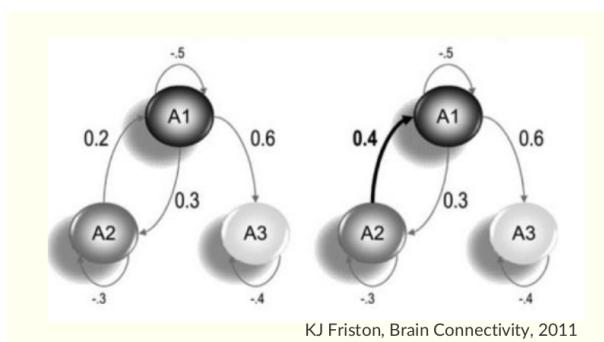
- **Correlation:** It measures how much two signals are synchronized or similar in time
- **Coherence:** It is the same of *correlation* but in frequency domain
- **Transfer entropy:** It measures more general changes (entropy changes during interactions of regions)

NOTE: (*Dependency on what subject is doing*)

Differently from anatomical connectivity, functional connectivity depends on what the subject is doing, for instance we have two different networks when the subject is solving a problem and when he is at rest.

Effective connectivity

Effective connectivity refers explicitly to the influence that one neural system exerts over another, either at a synaptic or population level.



KJ Friston, Brain Connectivity, 2011

In this case we need to use a **parameterized model**. Indeed, effective connectivity corresponds to the parameters of the model that tries to explain observed dependencies. The model can be based on coupling or directed causal influence, it can use Hypothesis-led or large model spaces.

NOTE: (Functional VS Effective connectivity)

The main differences between the previous connectivity are:

- **Model based:** Functional connectivity does not rest on any model, effective one instead, is model based.
- **Causality:** Functional connectivity uses a non-oriented graph therefore it does not give any information about **causality**. Effective instead, gives information about causality.

In a nutshell, let consider two regions A and B:

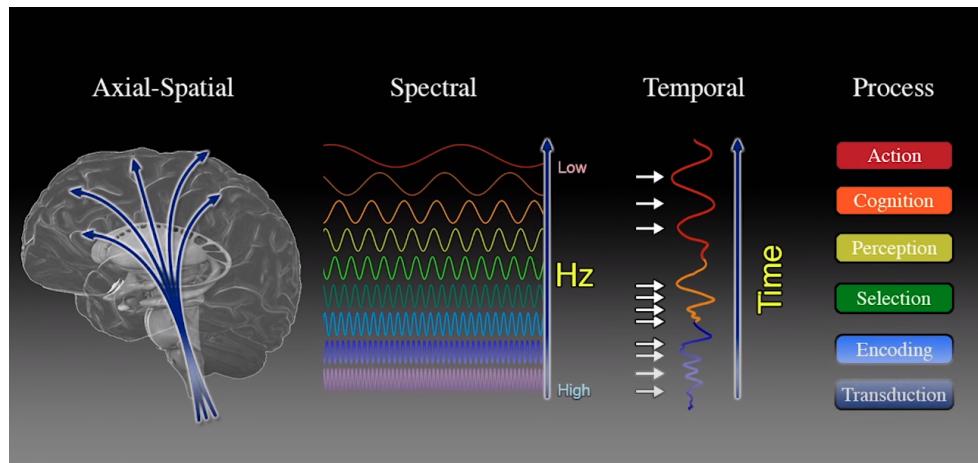
- **Functional connectivity** answers to the question: *Are A and B similar or related?*
- **Effective connectivity** answers to: *Does A cause what happens in B?*

CORRELATION

.SYNCHRONICITY IN THE BRAIN

“Cells that fire together wire together” (Hebbian Theory)

We have already seen that EEG signal is produced by synchronization between several neurons. In the brain there are other kinds of synchronicity, in this section we have interested into synchronicity between brain regions, that is at the base of **brain rhythmic**s.



https://www.youtube.com/watch?v=OCpYdSN_kts&feature=youtu.be

Due to their excellent temporal resolution, neuroelectrical measures (from invasive to scalp-recorded EEG) keep the information coded in the brain rhythms, therefore they are suitable for study them.

In this section we will introduce the concept of **correlation**.

.AUTO-CORRELATION

Given a *complex values* time series $x[n]$ with sampling interval T , the **auto-correlation** function is defined as:

$$r_{xx}[k] = \sum_{n=-\infty}^{\infty} x^*[n] x[n+k]$$

Where:

- $x^*[n]$ is the conjugate complex number of $x[n]$

OBS: (Time series)

In our case, the time series $x[n]$ represents our discrete EEG signal.

AUTO-CORRELATION AND SPECTRAL ANALYSIS

Fourier Transform

We can pass from time domain to the frequency domain by computing the **Fourier Transform** of the time series $x[n]$:

$$X(f) = \sum_{n=-\infty}^{\infty} x[n] \exp(-j 2\pi f n T)$$

NOTE: (*Continuous function in frequency domain*)

Applying Fourier transformation we passed from a time-discrete signal $x[n]$ to a continuous function $X(f)$.

Power Spectral Density

Then we can compute the **Power Spectral Density** (PSD) as:

$$S_{xx}(f) = |X(f)|^2$$

Wiener-Khinchin Theorem:

Power Spectral Density can be computed by Fourier-transforming the auto-correlation function:

$$S_{xx}(f) = \sum_{n=-\infty}^{\infty} r_{xx}[n] \exp(-j 2\pi f k T)$$

NOTE: (*Importance of this theorem*)

For us, Fourier transformation $X(f)$ of the time series $x[n]$ has not a physical meaning and it is not useful, in the contrary, the PSD has a physical meaning (signal strength) and we will be used by us later on. Wiener-Khinchin Theorem is very important since allows us to always compute the PSD of $x[n]$, indeed not all signals can be Fourier-transformed but the auto-correlation can be always transformed.

CROSS-CORRELATION

Given two *complex values* time series $x[n]$ and $y[n]$ with sampling interval T , the **cross-correlation** function is defined as:

$$r_{xy}[k] = \sum_{n=-\infty}^{\infty} x^*[n] y[n+k]$$

Where:

- $x^*[n]$ is the conjugate complex number of $x[n]$

CROSS-CORRELATION AND BIVARIATE SPECTRAL ANALYSIS

Given two time series $x[n]$ and $y[n]$ the **Mutual Power Spectral Density** are computed as:

$$S_{xy}(f) = X(f) Y^*(f)$$

$$S_{yx}(f) = Y(f) X^*(f)$$

It's symmetrical:

$$S_{xy}(f) = S_{yx}(f)$$

NOTE: (*Physical meaning*)

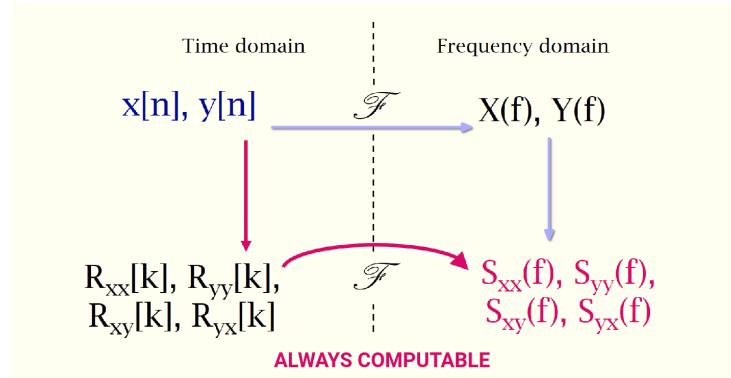
Mutual Power Spectral Density represents the quantity of power that both signals "share" in a frequency f .

Wiener-Khinchin Theorem:

Power Spectral Density can be computed by Fourier-transforming the cross-correlation function:

$$S_{xy}(f) = \sum_{n=-\infty}^{\infty} r_{xy}[n] \exp(-j 2\pi f k T)$$

.CONSEQUENCES W-K THEOREM



.SPECTRAL MATRIX

We can collect all information about coherency between two signals $x[n]$ and $y[n]$ in the **spectral matrix**:

$$S(f) = \begin{pmatrix} S_{xx}(f) & S_{xy}(f) \\ S_{yx}(f) & S_{yy}(f) \end{pmatrix}$$

Where:

- S_{xx}, S_{yy} = Power Spectral Density of x and y , respectively. (How the power of the signal is distributed at different frequencies)
- S_{xy}, S_{yx} = Mutual Power Spectral Density of x and y . (How the two signals share power at different frequencies)

OBS: (Symmetrical matrix)

$S(f)$ is symmetrical since $S_{xy}(f) = S_{yx}(f)$

COHERENCE

Coherence is essentially the same concept of *correlation* but defined in frequency domain using spectral power density function defined before.

There are several levels of coherence, we will see only the first order (**ordinary coherence**).

.ORDINARY COHERENCE

Ordinary Coherence is the linear correlation between two signals at a given frequency, it is defined as the Mutual Power Spectral Density divided by the product of their individual Power Spectral Densities. These last are used to normalize and take into account the ratio of exchanged power on the total power of the signals:

$$C_{xy}(f) = \frac{|S_{xy}(f)|^2}{|S_{xx}(f)| |S_{yy}(f)|}$$

PROPERTIES

Properties of ordinary coherence are:

- **Spectral:** It's spectral (=a function of frequency)
- **Normalized:** It's normalized between 0 and 1

For a given frequency f_0 :

- $C_{xy}(f_0) = 0 \rightarrow$ The two signals are **independent** at that frequency
- $C_{xy}(f_0) = 1 \rightarrow$ the two signals are **maximally correlated** at that frequency

Advantages

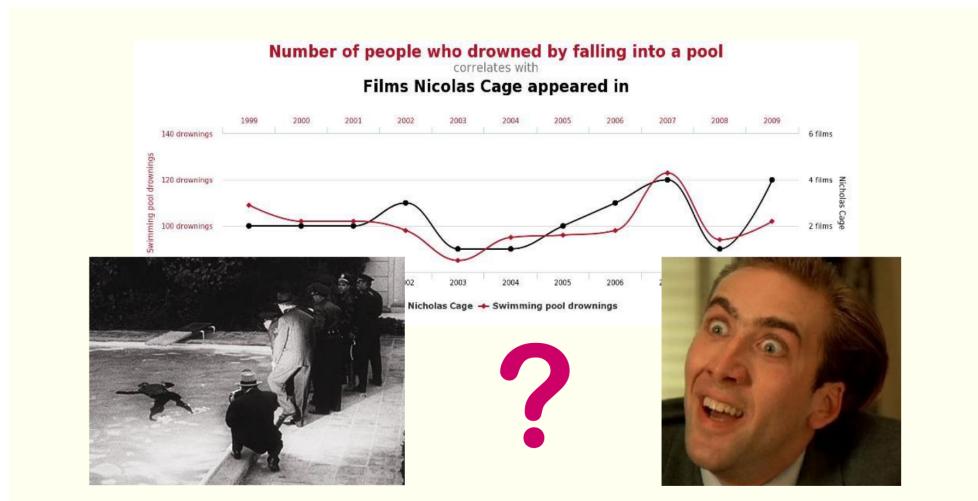
- **Easy to compute:** All the computation is based on well-known Fourier transformation
- **Useful to build a graph:** For each frequency, we can build a coherence network in which nodes are EEG signals and arcs are weighted with coherence

Limitations

- **Non-oriented graph:** The graph that we can build cannot be oriented since ordinary coherence is symmetrical ($C_{xy}(f) = C_{yx}(f)$), therefore it does not give any information about the direction of the interaction
- **No causality:** A consequence of the previous property is that coherence measures synchronicity but not causality
- **Bivariate:** It's bivariate (what if the time series are more than 2?)

CORRELATION VS CAUSATION

A very important concept to understand is that correlation is not causation, namely two signals can be synchronized but they can be independent:



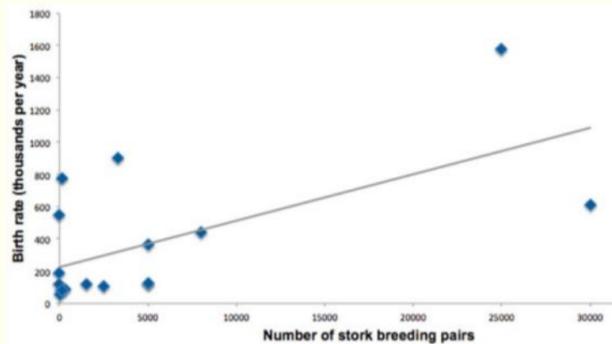
PROBLEM OF THE COMMON SOURCE

There are other cases in which two signals are synchronized and there is not causation between them but there is a third signal (common source) that influences both of them. This can be very common in neuroscience when we use a bivariate method like ordinary coherence.

Example (Common source)

Do storks deliver babies?

The number of storks in a country bears a correlation to the number of human births in that country:



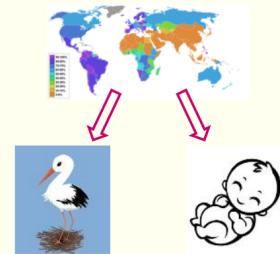
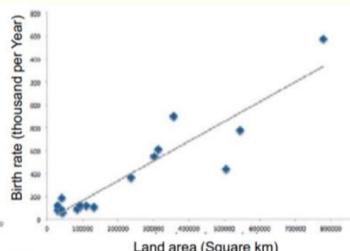
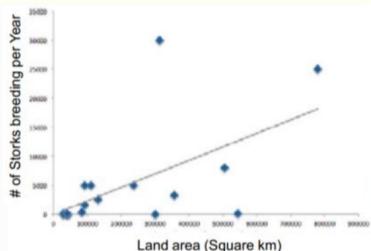
Is correlation due to a cause-effect relation or is it due to a third factor (unknown)?



?



Common effect of the country extension on the number of strokes and on the human birth rate:



Matthews, 2001

Self-evaluation test

1. Explain the difference between univariate and multivariate analysis of the brain
2. Explain the difference between anatomical, functional and effective connectivity
3. If $C_{xy}(f)$ is the ordinary coherence between x and y, indicate, for each of the following sentences, if they are true or false:
 - a. C_{xy} is a function of frequency
 - b. $C_{xy} \in [0, 1]$
 - c. $C_{xy} = C_{yx}$
 - d. C_{xy} can be computed also if the Fourier transform of x and y does not exist
4. Describe at least 2 advantages and 2 limitations of the ordinary coherence
5. Make an example of the problem of the common source

References

- Hari & Puce, Chapter 9 (Coherence and other measures of association, Some issues with coherence calculations)
- M.X. Cohen, Chapter 25

08 - BRAIN NETWORKS - PART 2

Learning objectives of the lesson

1. **Understand** the two main definitions of causality and their differences
2. **Remember** the definition of causality in the statistical sense
3. **Describe** the AR model (and its bivariate version) and its use as a linear predictor
4. **Compute** the Granger causality index from (BI)AR models
5. **Illustrate** its values range
6. **List** its advantages and its limitations

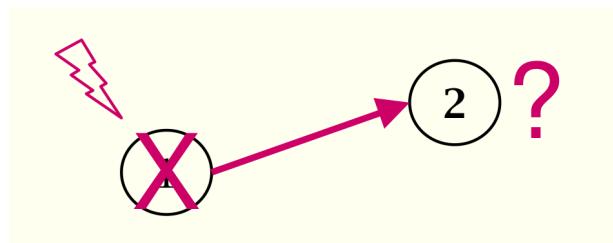
DEFINITIONS OF CAUSALITY

We can define the concept of causality between two quantities in two different ways. Let's introduce both of them although In this course we will use only the second one (**temporal precedence**).

PHYSICAL INFLUENCE

The first way to define causality is by studying their **physical influence**. A quantity x physically influences another quantity y if changing the **cause** (or **control**) x also the **consequence** y changes.

In order to verify if there is a physical influence between two brain regions, we have to perform experimentally controlled interventions in which we physically stimulate a brain region x (1) and we check if an other region y (2) changes. Physically acting upon brain activity effectively removes any other physical influence this node receives.



Main limitations

When we are interested into check a physical influence between two brain regions we have to perform a laboratory experiment with a subject, the procedure can be done with invasive methods based on electrodes implant, or with non-invasive one using for instance a magnet field. However in any cases we influence the brain activity by injecting current so we always physical act on the brain's subject. This is the first limitation.

The second limitation is that we have to do a separate experiment for each hypothesis we want to test:

- Hyp 1: Does x influence y ? -> 1' EXPERIMENT

- Hyp 2: Does y influence x ? -> 2' EXPERIMENT

Therefore we have a huge time limitation about the number of studies we can make about physical influence.

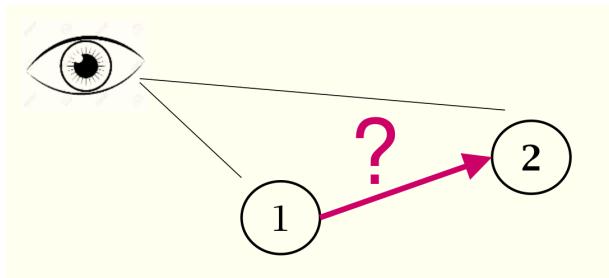
Due to these two main limitations, physical influence causal definition is not used a lot in neuroscience.

.TEMPORAL PRECEDENCE

Another way to define causality between two signals is based on the study of their **temporal precedence**, namely we require that the cause x precedes the consequence y , however this is not enough to say that x causes y .

We will see better later on how to establish if x causes y or not, in a nutshell we will test whether knowing past of x helps a better prediction of y future.

So this method is based on testing for improvement in predictive capacity between temporally distinct neural events and it is totally **observational**, namely we do not influence brain activity directly (non interventional).



Properties

Advantages

The main advantage of this approach is that we can collect the data in a single experiment and it is not strictly required to start with an hypothesis, namely we can use the same signals collected for several hypothesis tests. The other advantages are its simplicity and the fact that does not require to influence subject's brain activity:

- **Single experiment, multiple hypothesis tests**
- **Easy to use**
- **Non interventional**

Disadvantages

The main disadvantage is that we do not check physically for causality, for instance we can fall in the *common source problem*.

Due to its advantages, in neuroscience, in most of the cases the definition of causality is based on the temporal precedence.

CAUSALITY IN STATISTICAL SENSE

In order to define formally a relation of causality between two signals we use the **definition of causality is the statistical sense**.

The first definition of causality in a statistical framework has been provided by Norbert Wiener (1956):

*Given two **simultaneously measured** signals y, x , if one can **predict** the first signal y better by incorporating the past **information** from the second signal x than using only information from the first one y , then the second signal x can be called **causal** to the first one y (Wiener, 1956).*

Wiener provides to us only a general theoretical concept about statistic causality that can be summarized in the following three steps:

- **Step 1:** Predict y future with only y past values
- **Step 2:** Predict y future with y and x past values
- **Step 3:** If we obtained better prediction performance in step 2 then x causes y

OBS: (*How to predict y*)

Wiener's definition does not specify how to perform the predictions of a signal y by using its past and another signal x . The method/model to use has been introduced by Granger in 1969 (**Wiener-Granger causality**)

.WIENER-GRANGER CAUSALITY

Granger added a model (autoregressive model) to be used in the Wiener causality, the result is called **Wiener-Granger causality**:

*An observed time series $a[n]$ is said to **Granger-cause** another series $b[n]$ if knowledge of $a[n]$'s past significantly improves prediction of $b[n]$ by an **autoregressive modeling** (Granger, 1969)*

In theory we can use any other prediction model, however an autoregressive model is very simple and works pretty well with EEG signals.

Let's a first example:

Example ($a[t] \rightarrow b[t]$)

For instance, if we want to check if $a[t] \rightarrow b[t]$, we try to predict $b[n]$ in two ways:

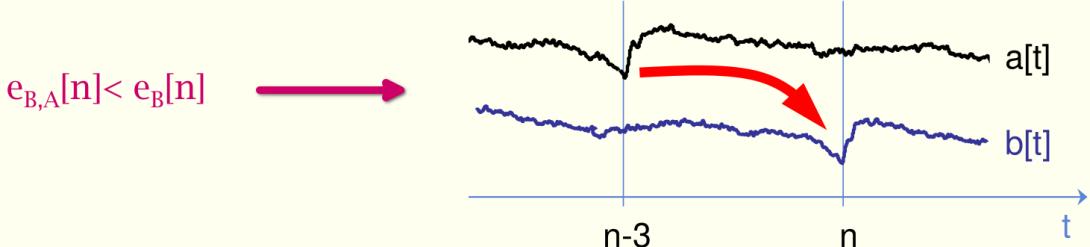
- **Using only b history:**

$$b[n] = B_1 b[n - 1] + \dots + B_N b[n - N] + e_B[n]$$

- Using a and b history:

$$b[n] = B_1 b[n-1] + \dots + B_N b[n-N] + A_1 a[n-1] + A_2 a[n-2] + \dots + A_m a[n-M] + e_{B,A}[n]$$

$e_B[n]$ and $e_{B,A}[n]$ are called **residual** and represent the prediction errors, therefore if $e_{B,A}[n] < e_B[n]$ it means that we obtain better prediction performance using also a past.



NOTE: (No information about $b[t] \rightarrow a[t]$)

In this example we proved only $a[t] \rightarrow b[t]$, in general It can be $a[t] \rightarrow b[t]$ without necessarily being $b[t] \rightarrow a[t]$ namely here we have directionality information. However, in neuroscience the relation $a[t] \leftrightarrow b[t]$ is very common (mutual influence)

We will see better later how to check Granger causality, now let's have a look to **autoregressive modeling**.

Linear autoregressive (AR) model

Using a **linear autoregressive model** we can express a sample at time $t = n$ of a time series $x[t]$ in terms of its previous values ($x[n-1], \dots, x[n-p]$ where p represents the model order):

$$x[n] = - \sum_{k=1}^p a[k] x[n-k] + e[n] = -a[1] x[n-1] - a[2] x[n-2] - \dots - a[p] x[n-p] + e[n]$$

Where:

- $x[n]$ = time series value at $t = n$
- $a[k]$ = autoregressive parameter, lag k
- p = model order
- $e[n]$ = model residual for $x[n]$

p is the model order and it essentially indicates the number of previous samples used to write $x[n]$, usually for EEG signal: $p = 10$ or $p = 20|25$.

Hypothesis:

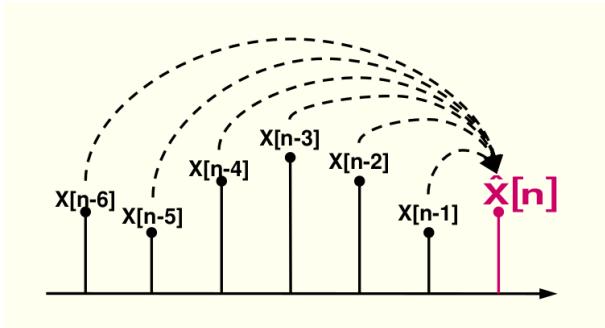
In order to properly represent a sample $x[n]$ using an AR model we have to satisfy the following **hypothesis**:

- $x[n]$: It is a **wide-sense stationary** signal (its mean and autocorrelation function are time invariant)
- $e[n]$: All residuals must belong to a **zero mean, uncorrelated white noise** ($e[n]$ must never contain relevant signal information about amplitude and frequency)

Autoregressive linear prediction

Assuming previous hypothesis satisfied, an AR model can be used as a linear predictor:

$$\hat{x}[n] = - \sum_{k=1}^p a[k] x[n-k]$$



The residual is here equal to the prediction error:

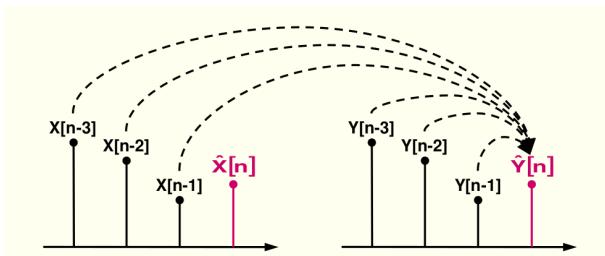
$$e[n] = x[n] - \hat{x}[n]$$

We must determine the coefficients $a[k]$, by minimizing the squared error $e[n]^2$

Bivariate autoregressive prediction

The autoregressive prediction of a signal $y[n]$ can be made by including information its own past and about the past samples of another signal $x[t]$, it is called **bivariate autoregressive prediction**:

$$x[n] = \sum_{k=1}^p a_{xy}[k] x[n-k] + \sum_{k=1}^p b_{xy}[k] y[n-k] + e_{xy}[n]$$



OBS: (Hypothesis)

Both signals $x[t]$ $y[t]$ and the residual must satisfy the hypothesis described previously.

We can do the same with $y[t]$:

$$y[n] = \sum_{k=1}^p a_{yx}[k] x[n-k] + \sum_{k=1}^p b_{yx}[k] y[n-k] + e_{yx}[n]$$

GRANGER CAUSALITY

We are finally ready to introduce the formal method that we will use to test causality between two signals. The name of the method is **Granger causality test** and it is based on computation and comparison of univariate and bivariate AR predictions.

.USE OF AUTOREGRESSION MODEL

All equation required to test both $y[t] \rightarrow x[t]$ $x[t] \rightarrow y[t]$ are the following:

- **Univariate AR predictions:** Here, the prediction error $e[n]$ depends only on the past values of the **own** signal.

$$x[n] = \sum_{k=1}^p a_x[k] x[n-k] + e_x[n] \quad (1)$$

$$y[n] = \sum_{k=1}^p a_y[k] y[n-k] + e_y[n] \quad (2)$$

- **Bivariate AR predictions:** Here, the prediction error for each individual signal depends on the past values of **both** signals.

$$x[n] = \sum_{k=1}^p a_{xy}[k] x[n-k] + \sum_{k=1}^p b_{xy}[k] y[n-k] + e_{xy}[n] \quad (3)$$

$$y[n] = \sum_{k=1}^p a_{yx}[k] x[n-k] + \sum_{k=1}^p b_{yx}[k] y[n-k] + e_{yx}[n] \quad (4)$$

OBS: (Use of equations)

We will not use all the equations at the same time but we will use them in pairs:

- To test $y[t] \rightarrow x[t]$: We use (1) (3)
- To test $x[t] \rightarrow y[t]$: We use (2) (4)

.GRANGER CAUSALITY TEST

The prediction performances for both models can be assessed by the **variances** of the prediction errors where $X|X$ and $X|X,Y$ indicate predicting X by its past values alone and by past values of X and Y, respectively:

- For **univariate** models:

$$V_{x|x} = \text{var}(e_x)$$

$$V_{y|y} = \text{var}(e_y)$$

- For **bivariate** models:

$$V_{x|x,y} = \text{var}(e_{xy})$$

$$V_{y|y,x} = \text{var}(e_{yx})$$

Suppose we are testing $y[t] \rightarrow x[t]$, if $V_{x|x,y} < V_{x|x}$ then y causes x in the sense of Granger causality, and a measure of that is given by:

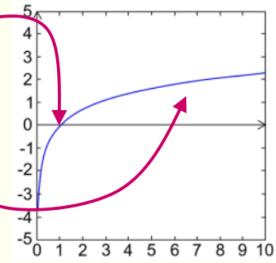
$$G_{y \rightarrow x} = \ln \left(\frac{V_{x|x}}{V_{x|x,y}} \right)$$

From a mathematical point of view, we can have three different scenario:

- $V_{x|x,y} = V_{x|x} \implies G_{y \rightarrow x} = 0$: The past of y does not improve the prediction of x
- $V_{x|x,y} < V_{x|x} \implies G_{y \rightarrow x} > 0$: Improvement in prediction of x by the inclusion of y
- $V_{x|x,y} > V_{x|x} \implies G_{y \rightarrow x} < 0$: In our case it does not have any sense since adding information cannot decrease the model accuracy, namely $V_{x|x,y}$ should not be greater than $V_{x|x}$, if it does, there's probably something wrong with the modeling

$$G_{y \rightarrow x} = \ln \left(\frac{V_{x|x}}{V_{x|x,y}} \right)$$

- ① $V_{x|x,y} = V_{x|x} \rightarrow G_{y \rightarrow x} = \ln(1) = 0$
- ② $V_{x|x,y} < V_{x|x} \rightarrow G_{y \rightarrow x} = \ln(> 1) = \uparrow$
- ③ $V_{x|x,y} > V_{x|x} \rightarrow \text{ERROR IN MODEL}$



Therefore, for our purpose $G_{y \rightarrow x}$ can assume only non-negative values:

$$G_{y \rightarrow x} \in [0, +\infty]$$

NOTE: (Directionality)

In the example above we tested causality in the direction $y \rightarrow x$ computing $G_{y \rightarrow x}$, this last does not give any information about the opposite direction $x \rightarrow y$. In order to study it we compute $G_{x \rightarrow y}$ that in principle is different with respect to $G_{y \rightarrow x}$:

$$G_{y \rightarrow x} = \ln \left(\frac{V_{x|x}}{V_{x|x,y}} \right) \neq G_{x \rightarrow y} = \ln \left(\frac{V_{y|y}}{V_{y|y,x}} \right)$$

.PROPERTIES

Let's summarize advantages and limitations of this approach.

Advantages

The main **advantages** are:

- **Directionality:** $G_{y \rightarrow x} \neq G_{x \rightarrow y}$
- **No interventional:** This method has a statistical meaning and does not require any intervention on subject's brain activity

Limitations

With respect to *coherence*, Granger test is based on a model and therefore it is a bit more complex to manage since we have to find the best parameters for it, the other main difference is that in this approach we are working in time domain and we don not have any information about frequencies:

- **Model required**
- **Defined in time domain:** Defined in the time domain (in the time window we used to identify the model) provided that the signals are stationary in that window
- **Non Normalized index:** $G \in [0, +\infty]$ is difficult to be interpreter
- **No true causality is guaranteed:** True causality can only be assessed if the set of two time series contains all possible relevant information and sources of activities for the problem (Granger, 1980).
- **Pair wise defined:** As coherence, also Granger test is defined only between two signals, therefore we can fall in the *common (hidden) source problem*

Self-evaluation test

1. Testing causality as temporal precedence is more practical than testing the physical influence (T/F)
2. The difference between the Wiener's and Granger's definitions of causality in the statistical sense is in the definition of a model to be used for the prediction (T/F)
3. Given two time series x and y :
 - a) $G_{x \rightarrow y}$ is always equal to $G_{y \rightarrow x}$ (T/F)
 - b) $G_{x \rightarrow y} \in [-\infty, +\infty]$ (T/F)
 - c) A negative value of $G_{x \rightarrow y}$ means an inverse precedence between the two time series (T/F)
4. List two advantages and two limitations of the Granger test

08 - BRAIN NETWORKS - PART 3

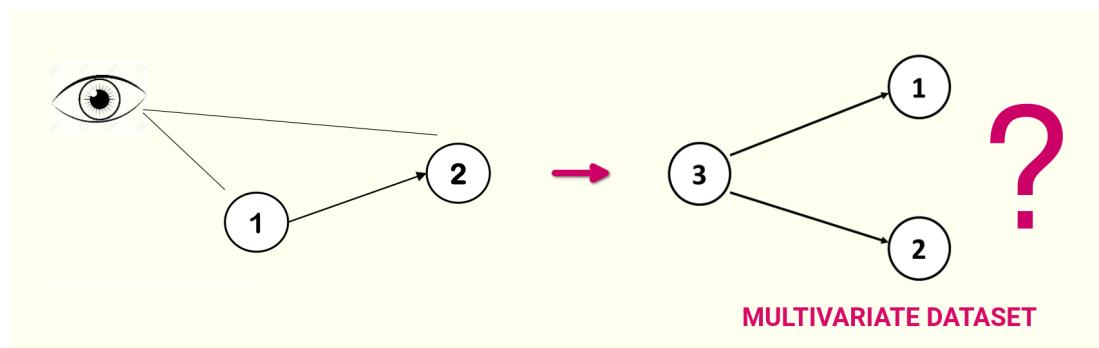
Learning objectives of the lesson

1. **Describe** the pairwise and the multivariate approaches
2. **Define** the PDC index, based on a multivariate, spectral AR model
3. **Illustrate** how it can be used to build directed brain networks at different frequencies
4. **Illustrate** its values range
5. **Compare** the advantages and limitations of pairwise vs multivariate approaches

INTRODUCTION

In the previous sections we have introduced the **hidden source problem** in which by analyzing two time series one could be identified as cause of the other when actually it is not true. This is due to a third *hidden* time series that influences both of them.

In this section we will try to solve, at least theoretically, the hidden source problem by introducing **multivariate datasets problem** in which we are interested to study relation between more than two time series:



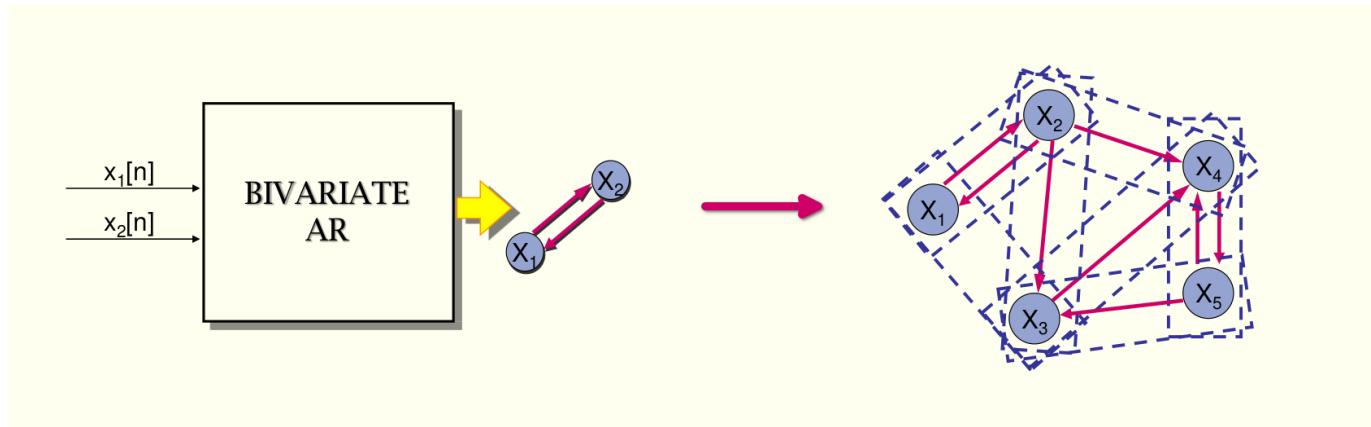
.MULTIVARIATE DATASETS PROBLEM

Let's see how we can manage the study of more than two time series. There are two main different approaches:

- **Pairwise approach:** Model studies every pairs of signals individually
- **Multivariate approach:** Model studies all signals together

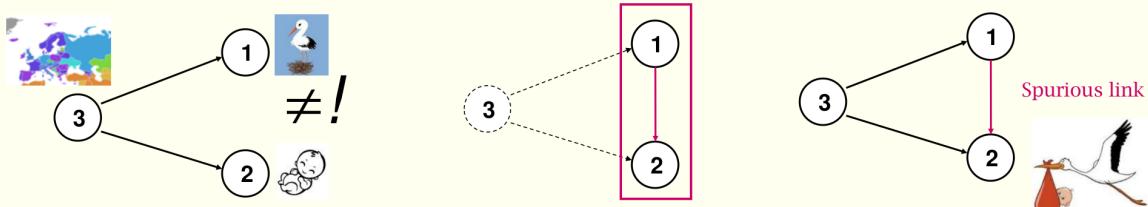
Pairwise approach

In the **pairwise approach** all signals in the dataset are taken pairwise and causality is assessed:



Limitations

The issue with this approach is that the set of two time series does not contain all possible relevant information, this is a limitation of the Granger definition (Granger, 1980). Therefore, any pairwise model analyzing the relation between two signals 1 and 2 cannot recognize that the statistical link between them is due to the common effect of a signal 3, which is not included in the model:

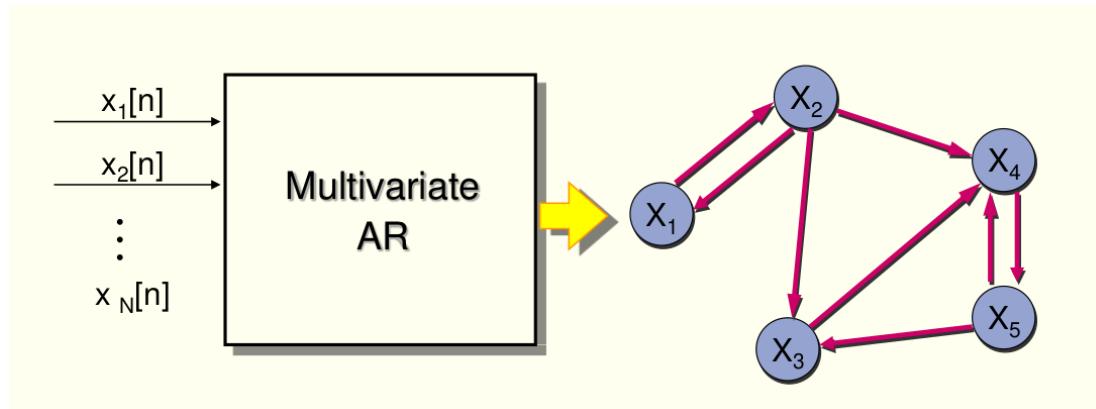


NOTE: (Pairwise approach in neuroscience)

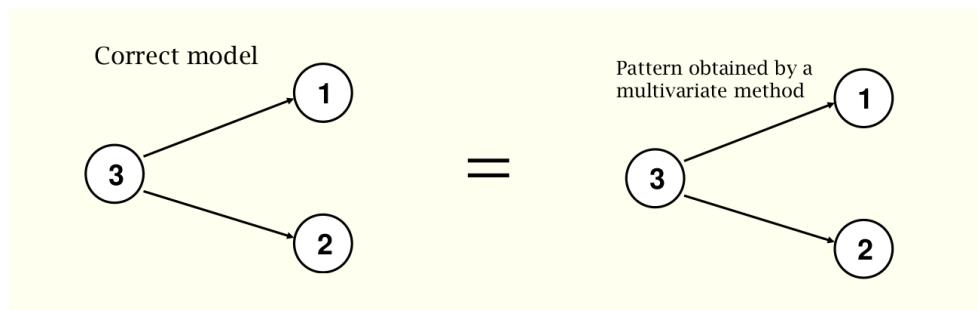
Although the previous limitation, pairwise approach still widely used in neuroscience especially when we deal with very large brain network, indeed as we will see later, the multivariate approach is more complex and requires a lot of data to work properly.

Multivariate approach

In the **multivariate approach** all the signals are studied together, so the connectivity pattern is obtained by a unique generation model (multivariate autoregressive) estimated on the entire set of data and takes into account all their interactions:



Multivariate methods, by building a unique model based on all the signals, use all the information at disposal and thus allow a better comprehension of the relationship between the signals:



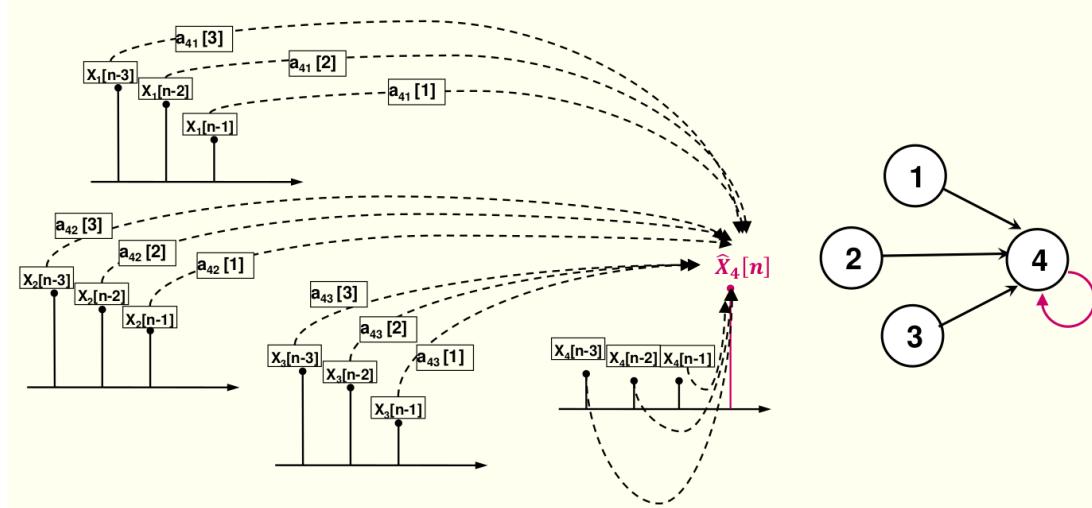
NOTE: (Spurious link in neuroscience)

Theoretically, using a multivariate approach we solve the hidden source problem and we avoid spurious links since we include all the available signals in the same model. However, in practice in neuroscience not all the signals can be acquired, indeed through an EEG some brain regions cannot be reached, therefore we still have hidden source problem.

So, a multivariate model can solve the hidden source problem only if we can record all the signals in a network and unfortunately this is not the case in neuroscience when we use an EEG.

MULTIVARIATE AUTOREGRESSIVE MODELS (MV-AR)

A **multivariate autoregressive model** is a generalization of the bivariate AR model seen in the previous section.



Also in this case, each signal and the model must satisfy the same hypothesis introduced last time:

- $x_i[t]$: It is a **wide-sense stationary** signal (hypothesis on the signals)
- $e_i[t]$: All residuals must belong to a **zero mean, uncorrelated white noise** (hypothesis on the model)

Equations

Given a set of N signals: $\tilde{X} = [x_1[1] \ x_2[1] \ \dots \ x_N[1]]^T$, a multivariate AR model of dimensions N is:

$$\begin{aligned} x_1[n] &= -\sum_{k=1}^p a_{11}[k]x_1[n-k] - \sum_{k=1}^p a_{12}[k]x_2[n-k] - \dots - \sum_{k=1}^p a_{1N}[k]x_N[n-k] + e_1[n] \\ x_2[n] &= -\sum_{k=1}^p a_{21}[k]x_1[n-k] - \sum_{k=1}^p a_{22}[k]x_2[n-k] - \dots - \sum_{k=1}^p a_{2N}[k]x_N[n-k] + e_2[n] \\ &\vdots \\ x_N[n] &= -\sum_{k=1}^p a_{N1}[k]x_1[n-k] - \sum_{k=1}^p a_{N2}[k]x_2[n-k] - \dots - \sum_{k=1}^p a_{NN}[k]x_N[n-k] + e_N[n] \end{aligned}$$

We recall that this model includes directionality information indeed parameters a_{12} and a_{21} are different:

- a_{12} is related to $2 \rightarrow 1$ (Past of 2 improves prediction of 1)
- a_{21} is related to $1 \rightarrow 2$ (Past of 1 improves prediction of 2)

The model parameters are $N * N * P$ where N is the number of signals and P is the model order:

$$\bar{a}[1] = \begin{bmatrix} a_{11}[1] & \cdots & a_{1N}[1] \\ \vdots & \ddots & \vdots \\ a_{N1}[1] & \cdots & a_{NN}[1] \end{bmatrix} \quad \bar{a}[2] = \begin{bmatrix} a_{11}[2] & \cdots & a_{1N}[2] \\ \vdots & \ddots & \vdots \\ a_{N1}[2] & \cdots & a_{NN}[2] \end{bmatrix} \quad \dots \quad \bar{a}[p] = \begin{bmatrix} a_{11}[p] & \cdots & a_{1N}[p] \\ \vdots & \ddots & \vdots \\ a_{N1}[p] & \cdots & a_{NN}[p] \end{bmatrix}$$

Including the N variances of the residuals $S_E = [\sigma_1, \sigma_2, \dots, \sigma_N]^T$, the total number of parameters to be estimated is:

$$N * N * P * N = N(N * P + 1)$$

MV-AR in frequency domain

In the previous sections we have seen two methods to study the relation between time series both of them have a pros and a cons:

- **Ordinary coherence:**

- PROS: Defined in frequency domain
- CONS: No directionality

- **Granger Test:**

- PROS: Directionality
- CONS: Defined in time domain

In the new method that we will introduce in this section (**partial directed coherence**) we take the best of both methods by computing the Fourier transformation of the MV-AR model:

$$\sum_{k=1}^p A[k]X[n-k] = E[n] \quad a_{ij}[k] = a_{ij}[1], a_{ij}[2], \dots, a_{ij}[p]$$

\mathcal{F}

$$\bar{A}(f)\bar{X}(f) = \bar{E}(f) \quad A_{ij}(f) = \sum_{k=0}^p a_{ij}[k]e^{-j2\pi fTk}$$

Where:

$$\bar{A}(f) = \begin{bmatrix} A_{11}(f) & \cdots & A_{1N}(f) \\ \vdots & \ddots & \vdots \\ A_{N1}(f) & \cdots & A_{NN}(f) \end{bmatrix} \quad \bar{X}(f) = \begin{bmatrix} X_1(f) \\ \vdots \\ X_N(f) \end{bmatrix} \quad \bar{E}(f) = \begin{bmatrix} E_1(f) \\ \vdots \\ E_N(f) \end{bmatrix}$$

The new method is based on the non-symmetric matrix of parameters $\tilde{A}(f)$

PARTIAL DIRECTED COHERENCE (PDC)

The new method is called **Partial Directed Coherence** and the name has the following meanings:

- **Partial**: It is a multivariate method (like *partial* derivative)
- **Directed**: It gives information related to the interaction between a source j and a destination i
- **Coherence**: This method is an upgrade of *ordinary* coherence

Let's see two different formulations of PDC:

PDC (Non-normalized)

Partial Directed Coherence (PDC) from j to i is defined on the basis of previous A matrix (*Baccalà and Sameshima, 2001*):

$$\pi_{ij}(f) = |A_{ij}(f)|^2$$

Recall: π_{ij} means $j \rightarrow i$ (column \rightarrow row)

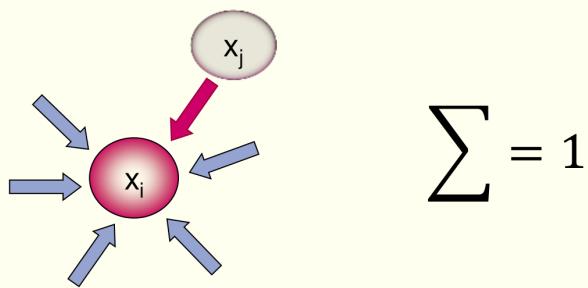
PDC Normalized

The non-normalized formulation is usually unstable and difficult to be interpreted therefore, different normalizations of PDC are provided, for instance (*Astolfi et al, 2007*):

$$\pi_{ij}(f) = \frac{|A_{ij}(f)|^2}{\sum_{m=1}^N |A_{im}(f)|^2}$$

Where:

- $\sum_{n=1}^N \pi_{in}(f) = 1$



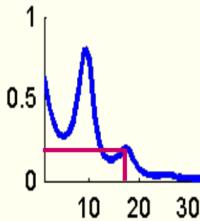
So at the numerator we have the influence of the signal j on the signal i and at denominator the sum of the influence of all the other signals $1 \dots N$ on i , therefore values assumed by π_{ij} have the following meanings:

- $\pi_{ij} = 0$: $\rightarrow j$ does not influence i
- $\pi_{ij} = 0.2$: $\rightarrow 20\%$ of i is influenced by j
- $\pi_{ij} = 1$: $\rightarrow 100\%$ of i is influenced by j

The normalized formulation is the most used one. We recall both formulations provided directionality information since

$$A_{ij}(f) \neq A_{ji}(f) \implies \pi_{ij}(f) \neq \pi_{ji}(f)$$

PDC can be computed at different frequencies, therefore for each frequency we can build a different network. The value of PDC_{ij} at a certain frequency f_0 represents the existence of a causality link directed from j to i .



.EXAMPLE

Let's see an example of how we can pass from spectral indices to brain networks.

- **Step 1:** Compute matrix $A(k)$ $k = 1 \dots p$ so as to satisfy work hypothesis on the residual $E(t)$

$$\sum_{k=0}^p A(k)X(t-k) = E(t)$$

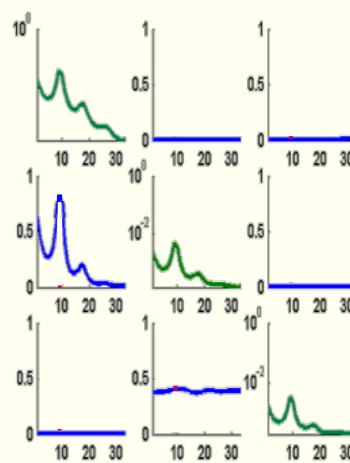
- **Step 2:** Compute Fourier transformation:

$$\sum_{k=0}^p A(k)X(t-k) = E(t) \quad \xrightarrow{\mathcal{F}} \quad A(f)X(f) = E(f)$$

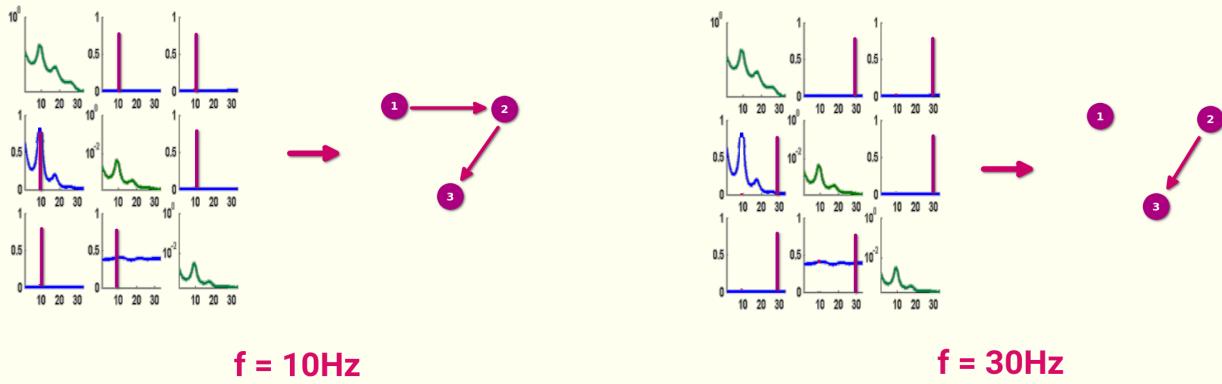
- **Step 3:** Compute PDC_{ij} for each pair of signals i, j . (PDC_{ij} estimates the influence of the region j toward the region i at a given frequency f)

$$\pi_{ij}(f) = \frac{|A_{ij}(f)|^2}{\sum_{m=1}^N |A_{im}(f)|^2}$$

- **Step 4:** Assuming N signals, from the previous step we obtain a square matrix of functions $N \times N$. We usually ignore PDC_{ii} , namely the "auto-influence", therefore in the main diagonal we put the spectrum density function of the signal i instead of $\pi_{ii}(f)$



- **Step 5:** Build network for each frequency. For instance, let consider two frequencies, 10Hz and 30Hz:



Recall: column \rightarrow row

.PROPERTIES

PDC methods provides several information about the relation between N signals:

- **Directed causality**
- **Strength of influence:** (Normalized PDC $\in [0, 1]$)
- **Spectral information:** Causality relation depend on the frequency
- **(Theoretically) No spurious links**

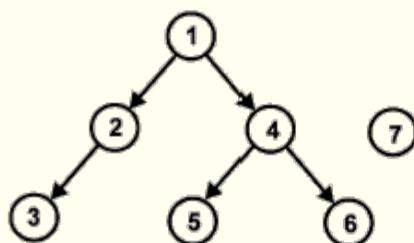
COMPARISON BETWEEN ESTIMATORS

.COMPARISON ON SIMULATED DATA

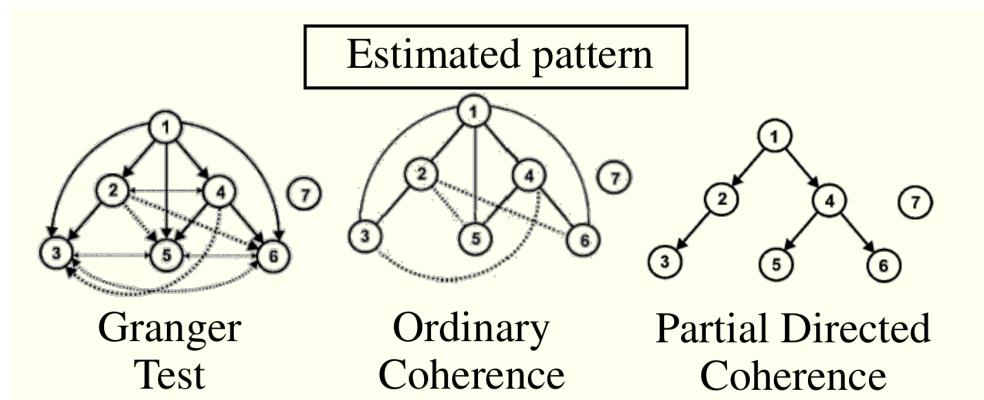
Let's compare all the methods seen in the course by using *simulation data*.

Given the following *imposed pattern* (target), we are interested in seeing how each method estimates it.

Imposed Pattern



The methods estimate the target pattern in the following ways, the **spurious links** are indicated with a dotted line.



NOTE: (Cascade effect)

We can notice only PDC is able to estimate the target correctly, the other two methods have several spurious links like 2->5 in ordinary coherence. However not all this excess links are spurious links, for instance the link 1-> 6 in ordinary coherency is not a spurious link since 1 affects indirectly (in cascade) the node 6 (1->4->6).

Source: Kus R, Kaminski M, Blinowska KJ, Determination of EEG activity propagation: pairwise versus multichannel estimate. *IEEE Trans Biomed Eng*, 2004.

. PAIRWISE VS MULTIVARIATE ESTIMATORS

Bivariate approach

Advantages

- No limit to the number of signals
- To be used when short data segments are available

Limitations

- Reduced accuracy

Multivariate approach

Advantages

- Better estimation performances
- Allows for inserting all data sources in the model

Limitations

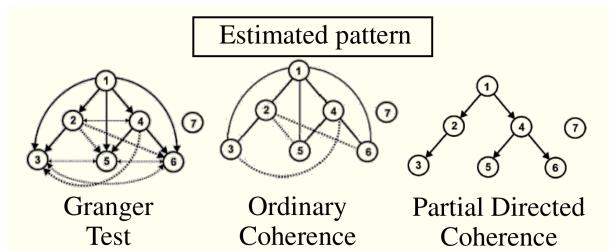
- Limitation in the number of channels/signal that can be modeled → **more data required**

. HOW TO CHOOSE THE METHOD

In order to choose the method to use, we should always answer to the following questions:

- Are we interested in *causality* or *synchronicity*?
- Is *frequency domain* required?
- How many data we have? (PDC requires many compared to other methods)

.SUMMARY OF CONNECTIVITY ESTIMATORS



Self-evaluation test

1. Show an example of network for which a pairwise approach is less accurate than a multivariate one
2. Given the PDC estimator:
 - a) $PDC_{i \rightarrow j}(f)$ is always equal to $PDC_{j \rightarrow i}(f)$ (T/F)
 - b) The normalized PDC $\in [-\infty, \infty]$ (T/F)
 - c) PDC can always avoid the problem of the "hidden source" (T/F)
3. List two advantages and a limitation of the pairwise and of the multivariate approach

References

- Cohen, Chapter 28

09 - GRAPH THEORY AND NEUROSCIENCE - PART 1

Learning objectives of the lesson

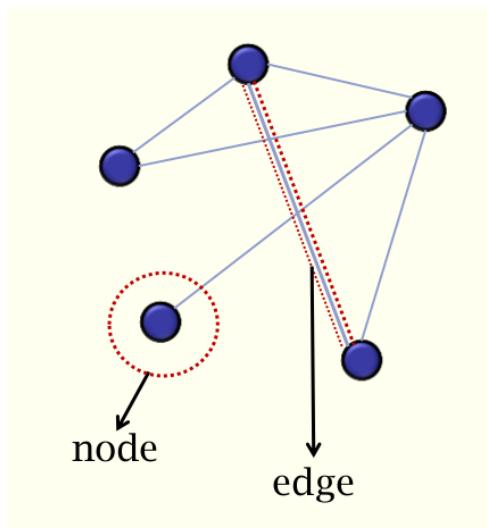
1. **Understand** what a graph is, its graphical and mathematical representations
2. **Understand** the need for graph theory in Neuroscience
3. **Remember** the differences between binary, weighted, directed and undirected graphs
4. **Describe** the properties of adjacency matrices for different graph types
5. **Remember** the definition and meaning of the main graph indices at the local scale used in neuroscience
6. Be able to **compute** such indices given a brain network

GRAPH THEORY

Networks can be represented as graphs or as adjacency matrices

Graphs have vertices (also called nodes), and edges that represent connections among vertices

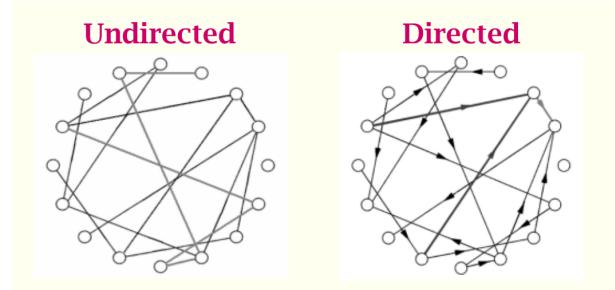
In network neuroscience, vertices correspond to brain regions or electrodes and edges correspond to some measure of connectivity



.GRAPH PROPERTIES

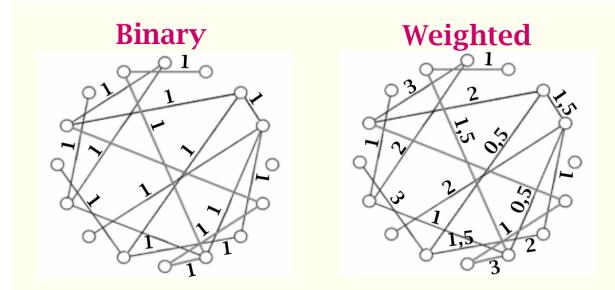
Directionality

- Undirected graphs
- Directed graphs



Weight

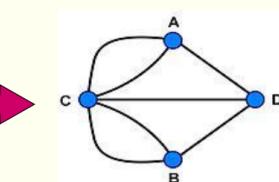
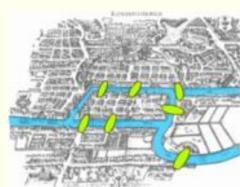
- Binary graphs
- Weighted graphs



.BIRTH OF GRAPH THEORY



1736: Euler solved the problem of «the seven bridges of Königsberg»: *is there a walk through the city that would cross each of the bridges once and only once?*



Euler proved that the problem has no solution and laid the foundations of graph theory and topology.

.EXAMPLES OF GRAPHS

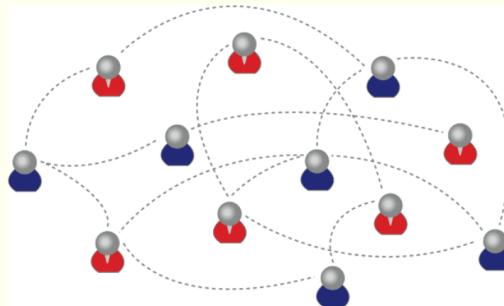
Social networks

Social networks

nodes: people

edges: relations

Binary,
undirected
graph



Highway networks

Highway networks

nodes: towns

edges: distances

Weighted,
undirected graph



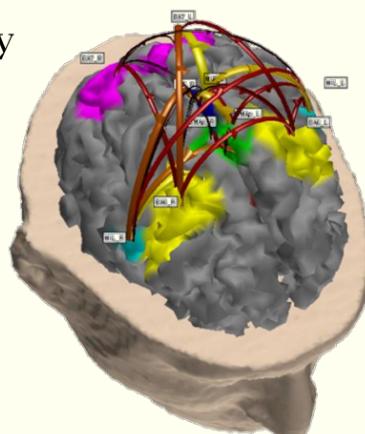
Brain networks

Brain networks

nodes: brain regions

edges: correlation or causality

Weighted,
(un)directed graph

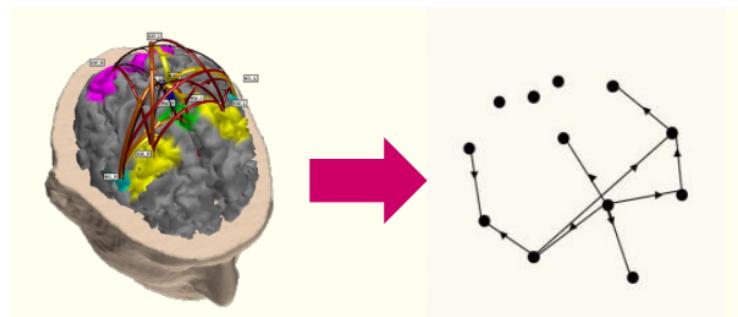


The type of graph adopted depends a lot on the method used. In the most simple cases, when we are interested only in the topology of the network, we can use **binary graph** (undirected).

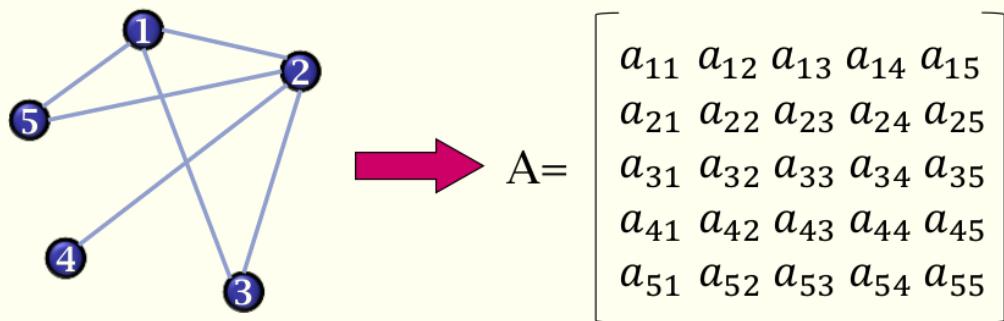
Need for graph analysis of brain networks

There are several motivations behind the use of graph theory in neuroscience, the main ones are the following:

- **Understand** the network structure and properties at the local, global and meso-scales
- **Extract quantifiable, objective, measurable and concise indices** to be used:
 - For a statistical comparison between conditions, time points, groups of subjects...
 - As markers of specific pathological conditions (to support diagnosis, prognosis and evaluation of the effects of a clinical intervention) (Monitoring brain plasticity)
 - As features for a classification aimed to decode the brain activity



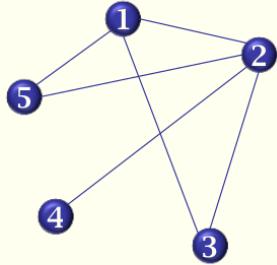
. ADJACENCY MATRIX



- $a_{ij} = 0$ if there is **no link** between i and j
- $a_{ij} \neq 0$ if there is **a link** between i and j
 - **Binary graphs** $\rightarrow a_{ij} = 1$
 - **Weighted graphs** $\rightarrow a_{ij} = w$ (w = weight of the edge)

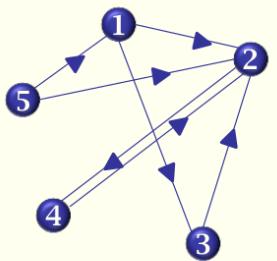
Binary graph

Binary undirected graph



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix} \rightarrow \text{Binary, symmetrical matrix}$$

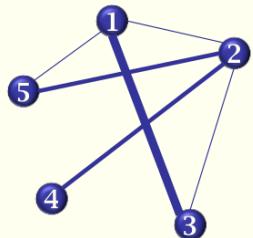
Binary directed graph



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \text{Binary, asymmetrical matrix}$$

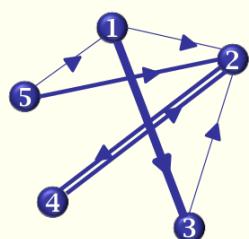
Weighted graph

Weighted undirected graph



$$A = \begin{bmatrix} 0 & 0.1 & 0.9 & 0 & 0.1 \\ 0.1 & 0 & 0.1 & 0.4 & 0.5 \\ 0.9 & 0.1 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 & 0 \\ 0.1 & 0.5 & 0 & 0 & 0 \end{bmatrix} \rightarrow \text{Weighted, symmetrical matrix}$$

Weighted directed graph



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.1 \\ 0.1 & 0 & 0.1 & 0.4 & 0.5 \\ 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow \text{Weighted, asymmetrical matrix}$$

GRAPH INDICES IN NEUROSCIENCE

Let's now see the graph indices commonly used in neuroscience, here we will see the definitions only for binary (unweighted) graph but they can be easily translated for a weighted graph.

Basic measures:

- Density
- Degree
- Distance

Global measures:

- Global Efficiency
- Local Efficiency

Mesoscale measures:

- Modularity
- Divisibility

With the term **mesoscale** we refer to a metric focused on a group of nodes

In the following part of this chapter and the next, we will see all of these indices in detail.

BASICS MEASURES

.DENSITY

Density k of a binary graph is the ratio between the **number L of edges** in the graph and the **maximum possible number of edges** L_{tot}

$$k = \frac{L}{L_{tot}} \quad k \in [0, 1]$$

- $k = 0 \rightarrow$ no edges in the network
- $k = 1 \rightarrow$ fully connected network

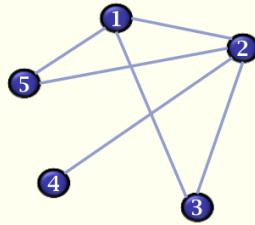
Density is a **global index** and is useful to compare two brain networks

Undirected graph

Given an N -nodes graph, for undirected graphs, L_{tot} is given by:

$$L_{tot} = N * \frac{N - 1}{2}$$

Example (Density)



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$N = 5$$

$$L = 6$$

$$L_{tot} = N(N-1)/2 = 10$$

$$k = L/L_{tot} = \frac{6}{10}$$

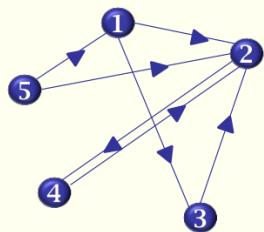
The number of existing arcs for an undirected graph can be determined from **half** of the matrix

Directed graph

Given an N -nodes graph, for directed graphs, L_{tot} is given by:

$$L_{tot} = N * (N - 1)$$

Example (Density)



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$N = 5$$

$$L = 7$$

$$L_{tot} = N(N - 1) = 20$$

$$k = L/L_{tot} = 7/20 = 0,35$$

For a directed graph we need the **full** matrix

.DEGREE

This is a **local index**, namely refers to a specific brain region. It is defined slightly different for undirected and directed graphs.

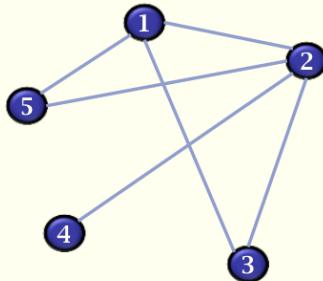
Undirected graph

The degree g of a node i ($i \in [1, N]$, N = number of nodes of the graph) is the **number of edges connected** to that node:

$$g_i = \sum_{j=1, i \neq j}^N a_{ij} \quad g_i \in [0, N - 1]$$

A node degree quantifies the role of a node in a network. The higher the degree, the more involved the node.

Example



$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$\rightarrow g_1 = 3$
 $\rightarrow g_2 = 4$
 $\rightarrow g_3 = 2$
 $\rightarrow g_4 = 1$
 $\rightarrow g_5 = 2$

Directed graph

If the graph is directed, we can define:

- **In-degree** of a node i ($i \in [1, N]$) as the number of edges **directed to** node i

$$g_i^{in} = \sum_{j=1, i \neq j}^N a_{ij} \quad g_i^{in} \in [0, N - 1]$$

- **Out-degree** of a node i ($i \in [1, N]$) as the number of edges **originated from** node i

$$g_i^{out} = \sum_{j=1, i \neq j}^N a_{ji} \quad g_i^{out} \in [0, N - 1]$$

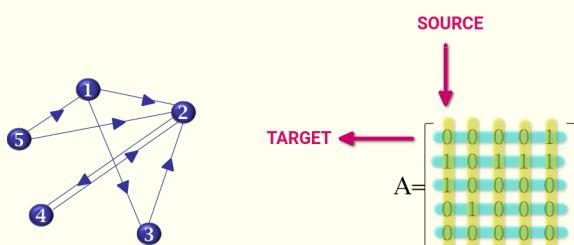
- **Degree** of a node i ($i \in [1, N]$) as the total number of edges **originated from and directed to** node i

$$g_i = g_i^{in} + g_i^{out} \quad g_i \in [0, 2 * (N - 1)]$$

OBS: (Meaning)

- $g_i^{in} < g_i^{out}$: The node is a target of information
- $g_i^{in} > g_i^{out}$: The node is a source of information
- $g_i^{in} = g_i^{out}$: The node is a hub of information

Example

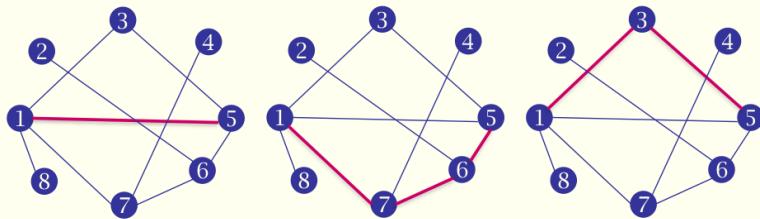


out-degree	in-degree	degree
$g_1^{out} = 2$	$g_1^{in} = 1$	$g_1 = 3$
$g_2^{out} = 1$	$g_2^{in} = 4$	$g_2 = 5$
$g_3^{out} = 1$	$g_3^{in} = 1$	$g_3 = 2$
$g_4^{out} = 1$	$g_4^{in} = 1$	$g_4 = 2$
$g_5^{out} = 2$	$g_5^{in} = 0$	$g_5 = 2$

DISTANCE

Path

Any possible sequence of edges linking two nodes, e.g.:



3 paths linking 1 and 5:

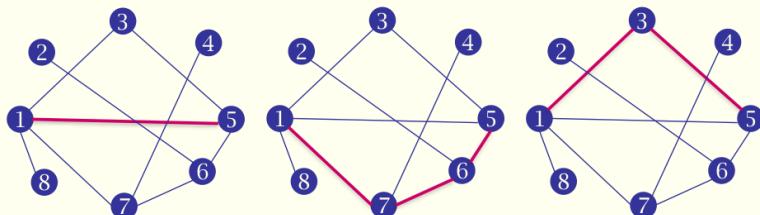
- $1 \rightarrow 5$ (1 step)
- $1 \rightarrow 7 \rightarrow 6 \rightarrow 5$ (3 steps)
- $1 \rightarrow 3 \rightarrow 5$ (2 steps)

The path length is given by the number of edges (not by the physical distance)

Distance $d(i, j)$ between nodes i and j is the shortest path between them.

Distance measures how efficient the interaction between two nodes is. The shorter the distance, the more efficient the interaction.

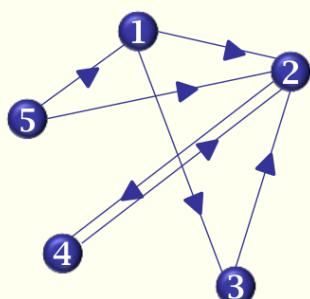
Example (Undirected graphs)



3 paths linking 1 and 5:

- $1 \rightarrow 5$ (1 step)
- $1 \rightarrow 7 \rightarrow 6 \rightarrow 5$ (3 steps)
- $1 \rightarrow 3 \rightarrow 5$ (2 steps)

Example (Directed graphs)



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow D = \begin{bmatrix} 0 & \infty & \infty & \infty & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & \infty & 0 & \infty & 2 \\ 2 & 1 & 2 & 0 & 2 \\ \infty & \infty & \infty & \infty & 0 \end{bmatrix}$$

When no path links two nodes, the distance is infinite

References

- Cohen, Chapter 31

10 - GRAPH THEORY AND NEUROSCIENCE - PART 2

Learning objectives of the lesson

1. **Remember** the definition and meaning of the main graph indices at the global and meso-scale used in neuroscience
2. **Understand** the concept of network integration and segregation
3. Be able to **compute** such indices given a brain network, for directed and undirected graphs
4. **Describe** the reference networks (regular or lattice, random and small-world) and their main properties
5. **Compare** a real (brain) network with the reference networks

GLOBAL MEASURES

In this section we will see graph indexes that refer to the all network. Despite the name of the second one, both methods are global, namely they are computed by looking at all nodes.

GLOBAL EFFICIENCY

The Global Efficiency E_g of a graph is the average of the reciprocal of the distances between any pair of nodes (*Latora e Marchiori, 2001*):

$$E_g = \frac{1}{L_{tot}} * \text{Sum inverse reciprocal distances}$$

Global Efficiency measures how efficiently the information is exchanged in the network. It varies in $[0, 1]$:

- $E_g = 1 \rightarrow$ graph fully connected
- $E_g = 0 \rightarrow$ void graph

This efficiency index provides a measure of **integration** of the network, namely the "communication degree of freedom" available in the graph.

This index is useful when monitored over time in order to understand brain changes.

Directed graph

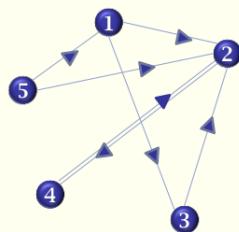
For a directed graph, global efficiency is defined as follows:

$$E_g = \frac{1}{N(N-1)} * \sum_{i,j}^N \frac{1}{d_{ij}} \quad E_g \in [0, 1]$$

Where:

- d_{ij} is the element at $i - th$ row and $j - th$ column of matrices of distances D

Example (Directed graph)



$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \rightarrow D = \begin{bmatrix} 0 & \infty & \infty & \infty & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & \infty & 0 & \infty & 2 \\ 2 & 1 & 2 & 0 & 2 \\ \infty & \infty & \infty & \infty & 0 \end{bmatrix}$$

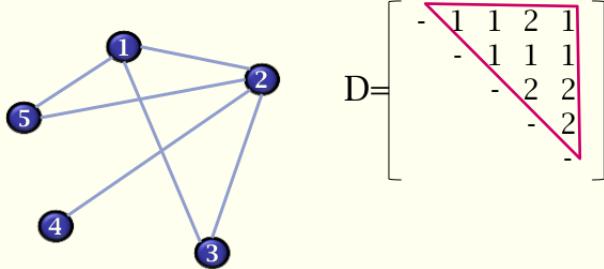
$$E_g = \frac{1}{N(N-1)} \sum_{i,j=1, i \neq j}^N \frac{1}{d_{ij}} = \frac{1}{5(5-1)} \left[1 + 1 + \frac{1}{2} + 1 + 1 + \frac{1}{2} + 1 + 1 + 1 + \frac{1}{2} + \frac{1}{2} \right] = 0,45$$

Undirected graph

For an undirected graph:

$$E_g = \frac{2}{N(N-1)} * \sum_{i,j}^N \frac{1}{d_{ij}} \quad E_g \in [0, 1]$$

Example (Undirected graph)



$$D = \begin{bmatrix} 0 & 1 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

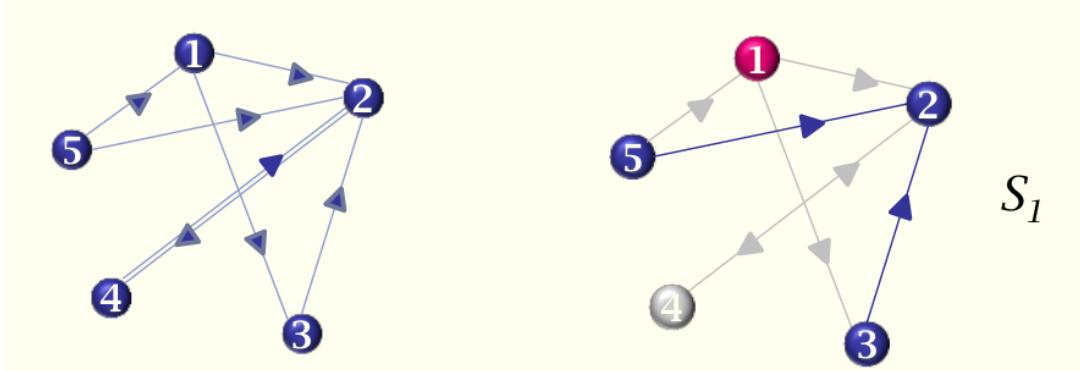
The distances in an undirected graph can be determined from half of the matrix. The number of connections for the normalization is $N(N-1)/2$ (maximum possible number of arcs)

$$E_g = \frac{2}{N(N-1)} \sum_{i,j=1, i \neq j}^N \frac{1}{d_{ij}} = \frac{2}{20} \left(1 + 1 + \frac{1}{2} + 1 + 1 + 1 + 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \right) = 0.8$$

LOCAL EFFICIENCY

The idea behind **local efficiency** is understand how the graph changes when we remove a node i . To do that we essentially remove the node i and compute the global efficiency over the resulting subnetwork.

For each node i we can extract a subnetwork S_i made of all the nodes directly connected to i (but not including i itself) and compute its Global Efficiency $E_g(S_i)$.



(Average) Local Efficiency

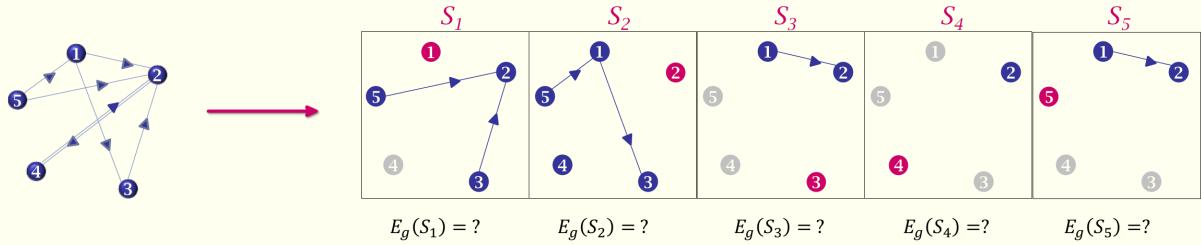
The average of $E_g(S_i)$ across all the nodes is called (average) Local Efficiency (*Latora e Marchiori, 2001*)

$$E_l = \frac{1}{N} \sum_{i=1}^N E_g(S_i)$$

Local Efficiency measures the tendency of the network to form strongly connected subgroups. It is also a measure of the robustness of the network to the removal of node i

Example (Directed graph)

A graph with N nodes has N subgraphs, each one associated with a node. Each subgraph S_i is obtained by removing the related node i and considering the subgraph formed by the nodes adjacent (= at distance 1, in either direction) to the removed node

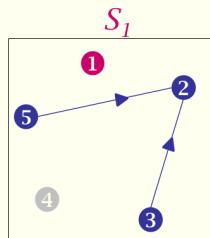


Let's compute all of them:

NOTE: (Number of nodes)

Note that the number of nodes N here is referred to the subgraph S_i (not to the original graph)

- S_1

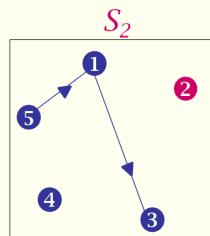


$$D = \begin{bmatrix} 2 & 3 & 5 \\ - & 1 & 1 \\ \infty & - & \infty \\ \infty & \infty & - \end{bmatrix}_{3 \times 3}$$

$$E_g(S_1) = \frac{1}{3(3-1)}(1+1) = \frac{1}{3}$$

Note that the number of nodes here is referred to the subgraph (not to the original graph)

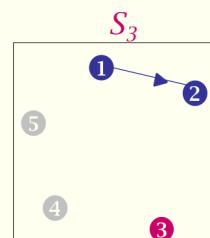
- S_2



$$D = \begin{bmatrix} 1 & 3 & 4 & 5 \\ - & \infty & \infty & 1 \\ 1 & - & \infty & 2 \\ \infty & \infty & - & \infty \\ \infty & \infty & \infty & - \end{bmatrix}_{4 \times 5}$$

$$E_g(S_2) = \frac{1}{4(4-1)}\left(1+1+\frac{1}{2}\right) = \frac{5}{24}$$

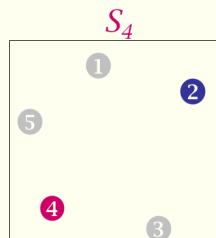
- S_3



$$D = \begin{bmatrix} 1 & 2 \\ - & \infty \\ 1 & - \end{bmatrix}_{2 \times 2}$$

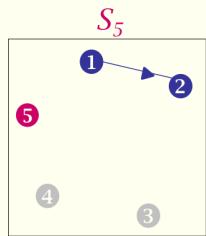
$$E_g(S_3) = \frac{1}{2(2-1)}(1) = \frac{1}{2}$$

- S_4



$$E_g(S_4) = 0$$

- S_5



$$D = \begin{bmatrix} 1 & 2 \\ - & \infty \\ 1 & - \end{bmatrix}$$

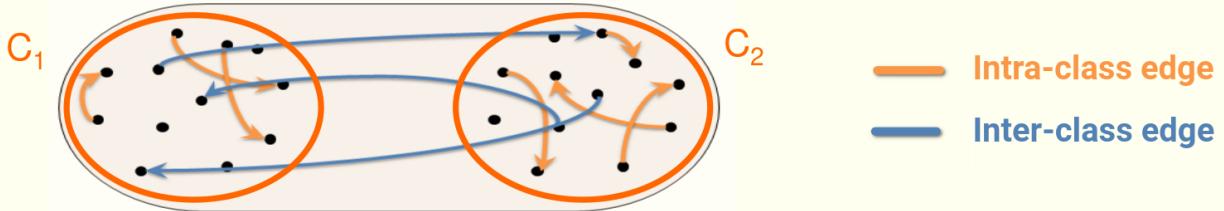
$$E_g(S_5) = \frac{1}{2(2-1)}(1) = \frac{1}{2}$$

- Average Local Efficiency

$$E_l = \frac{1}{N} \sum_{i=1}^N E_g(S_i) = \frac{1}{5} \left(\frac{1}{3} + \frac{5}{24} + \frac{1}{2} + 0 + \frac{1}{2} \right) = 0.308$$

MESOSCALE MEASURES

With mesoscale measures we study groups of brain regions. For some applications it may be interesting to measure how well a network can be divided into **communities** (a.k.a. *clusters*):



A subnetwork is a **community** if it shows an organized structure with respect to the entire network. Measures useful to determine if two or more subnetworks act as communities are **Divisibility** and **Modularity**.

.DIVISIBILITY

Divisibility D is a measure of the **segregation** between two communities, namely an **high divisibility** means is easy to remove the communication between the two communities. A **low divisibility** instead, corresponds to a very good connection between the two classes that cannot be easily interrupted.

Divisibility D is "conceptually" defined as follows:

$$D = \frac{(2^*)\#\text{Edges in the graph}}{k + \#\text{ inter-classes edges}}$$

Where:

- k : is a constant used to avoid divergence (division by zero), usually it's equal to the total number of edges in the network (#Edges in the graph)

OBS (Low divisibility)

When the divisibility is low that means the two classes are connected through many links, therefore it could be the case that they are not actually two classes (clusters) but all the nodes belong to the same cluster. We will see better this concept later.

NOTE:

Divisibility focuses on **inter**-community links

Undirected graph

For **undirected graphs**, divisibility is defined in the following way:

$$D = \frac{2L}{2L + \sum_{i,j=1}^N a_{ij}[1 - \delta(C_i, C_j)]}$$

Where:

- L : Number of edges in the network
- a_{ij} : Elements (i, j) of the adjacency matrix
- C_i : Community to which node i belongs
- $\delta(C_i, C_j)$: Same class?

$$\delta(C_i, C_j) = \begin{cases} 1 & \text{if } C_i = C_j \\ 0 & \text{if otherwise} \end{cases}$$

Directed graph

For **directed graphs** instead, divisibility is defined in the following way:

$$D = \frac{L}{L + \sum_{i,j=1}^N a_{ij}[1 - \delta(C_i, C_j)]}$$

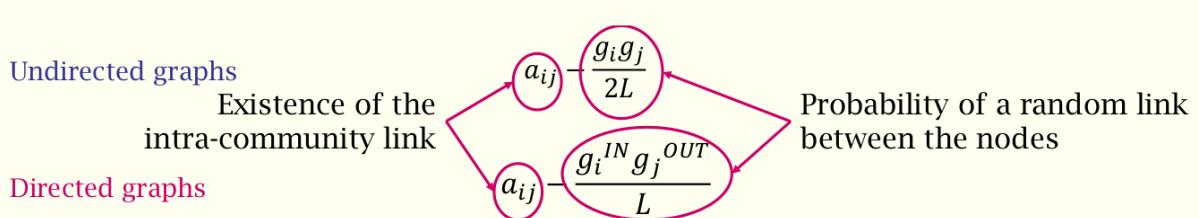
To summarize the meaning of D :

$$D = \begin{cases} 0.5 & \text{Classes not easy to separate} \\ 1 & \text{Easy to separate} \end{cases}$$

MODULARITY

Modularity Q measures the tendency of the subnetworks to form communities.

For each pair of nodes belonging to the same community, modularity Q compares the existence (or not) of a link between them with the probability of having a connection between them in a random distribution of the links:



NOTE: Modularity focuses on **intra**-community links

Undirected graph

$$Q = \frac{1}{2L} \sum_{i,j=1}^N \left(a_{ij} - \frac{g_i g_j}{2L} \right) \delta(C_i, C_j)$$

Where

- L : number of edges in the network
- a_{ij} : elements of the adjacency matrix
- g_i : degree of node i
- C_i : community to which node i belongs
- $\delta(C_i, C_j)$: Same class?

$$\delta(C_i, C_j) = \begin{cases} 1 & \text{if } C_i = C_j \\ 0 & \text{if otherwise} \end{cases}$$

Directed graph

$$Q = \frac{1}{L} \sum_{i,j=1}^N \left(a_{ij} - \frac{g_i^{IN} g_j^{OUT}}{L} \right) \delta(C_i, C_j)$$

Where

- L : number of edges in the network
- a_{ij} : elements of the adjacency matrix
- g_i : degree of node i (IN or OUT)
- C_i : community to which node i belongs
- $\delta(C_i, C_j)$: Same class?

$$\delta(C_i, C_j) = \begin{cases} 1 & \text{if } C_i = C_j \\ 0 & \text{if otherwise} \end{cases}$$

To summarize the meaning of Q :

$$Q = \begin{cases} > 0 & \text{there are more intra-community links than in a random division} \\ \leq 0 & \text{otherwise} \end{cases}$$

NETWORK INTEGRATION AND SEGREGATION

All the measures seen so far can be used to characterize a network or a network subdivision, using the terms **integration** and **segregation**:

NETWORK CHARACTERIZATION

Two opposite characterization for a network are the following:

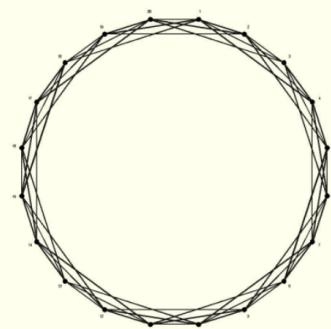
- **Integrated** The network is not divided into subgroups but each node can easily communicate with the others:
 - High ↑ Global Efficiency
 - Low ↓ Local Efficiency
- **Segregated** The network is composed by subgroups, each node can easily communicate with its neighbor but not with other nodes:
 - Low ↓ Global Efficiency
 - High ↑ Local Efficiency

Reference networks

According to the previous characterization, there are two **reference networks**

Regular networks (a.k.a. Lattice networks)

- Each node is linked to a small number of other nodes.
- The degree is the same for all nodes.
- Efficient communications between small groups
- Inefficient communications at the entire network level

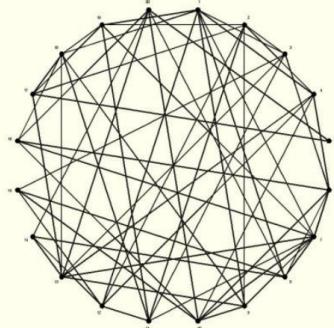


Lattice structure

- LOW Global Efficiency
- HIGH Local Efficiency

Random networks (Erdős Rènyi, 1959)

- Each node is linked to the others randomly.
- There are no small groups with a strong internal communication.



Random structure

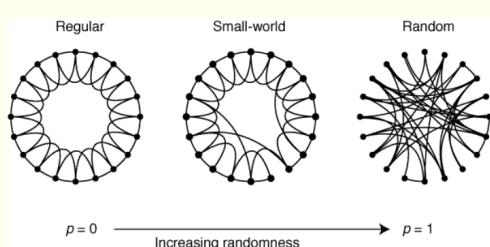
- HIGH Global Efficiency
- LOW Local Efficiency

Small-world networks (Watts and Strogatz, Nature, 1998)

Real networks are neither like the random nor like the regular graphs, both of them have pros and cons:

- **Regular network:**
 - Low Energy consumption
 - Slow communication
- **Random network:**
 - Fast communication
 - High energy consumption

Real networks usually have a structure called **small-world network** that correspond to a tradeoff between the two reference networks, however, it is more similar to a regular than a random network:

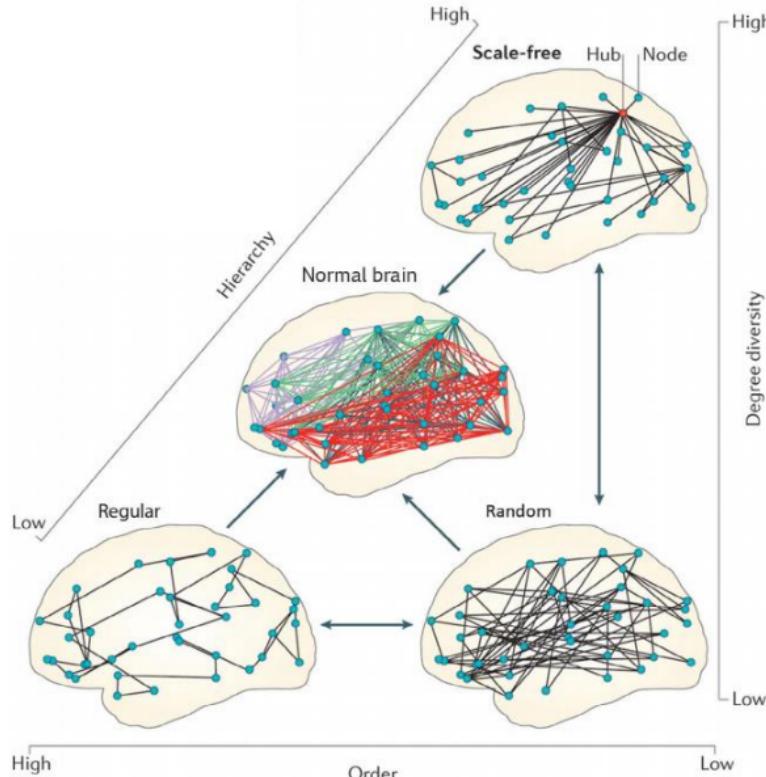


Eg regular < Eg < Eg random

El random << El < El regular

Six degree of separation (now 3.57)

The name **small-world network** came from the idea of **six degree of separation** for which all people are six or fewer social connections away from each other. With social networks now it is reduced to 3.57.#



SUBDIVISION CHARACTERIZATION

By using *divisibility* and *modularity* we are able to understand if a network subdivision is correct or not, namely if the two classes really act as two clusters.

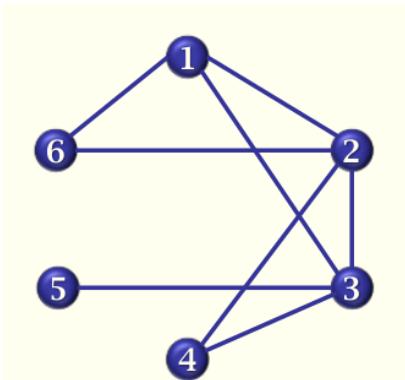
To do that we use the following two opposite characterization for the subdivision:

- **Integrated** The two classes have a lot of inter-communities connections (low divisibility) and the intra-class links are less than a random subdivision of the graph (low modularity) -> They are **NOT** really two classes
 - Low ↓ Divisibility
 - Low ↓ Modularity
- **Segregated** The two communities are connected by a small number of inter-classes links (high divisibility) and the number of intra-class links is higher than a random subdivision. -> They are really two classes
 - High ↑ Divisibility
 - High ↑ Modularity

Let's see two exercises in which we use divisibility and modularity

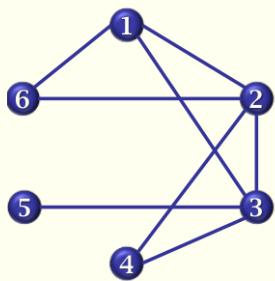
Exercise 1 (Undirected graph)

Given the graph in figure:



1. Write down the corresponding binary adjacency matrix A
2. Compute the degree of each node
3. Compute Divisibility and Modularity according to the following two classes: C=[1 1 1 2 2 2]
4. Compute Divisibility and Modularity according to the following two classes: C=[1 1 2 2 2 1]

1 (binary adjacency matrix A)

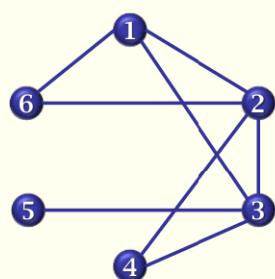


$$N = 6 \\ L = 8$$

A=

0	1	1	0	0	1
1	0	1	1	0	1
1	1	0	1	1	0
0	1	1	0	0	0
0	0	1	0	0	0
1	1	0	0	0	0

2 (degree of each node)

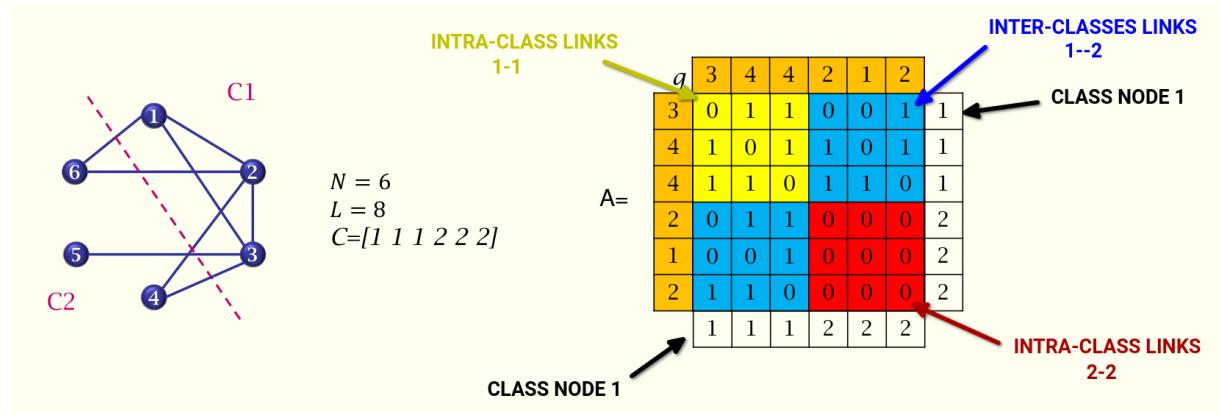


$$N = 6 \\ L = 8$$

A=

g	3	4	3	1	1	2
3	0	1	1	0	0	1
4	1	0	1	1	0	1
3	1	1	0	1	1	0
1	0	1	1	0	0	0
1	0	0	1	0	0	0
2	1	1	0	0	0	0

3 ($C=[1\ 1\ 1\ 2\ 2\ 2]$)



$$D = \frac{2L}{2L + \sum_{i,j=1}^N a_{ij} [1 - \delta(C_i, C_j)]} = \frac{16}{16 + 10} = 0.62 \quad D \in [0.5, 1]$$

Equal to 1 if i, j belong to 1-2

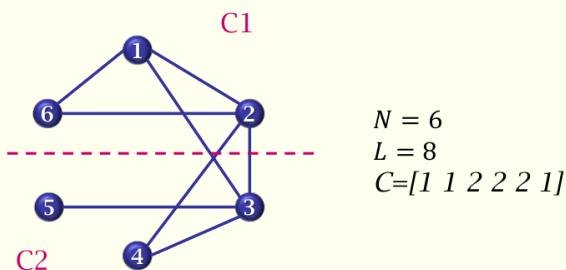
$$Q = \frac{1}{2L} \sum_{i,j=1}^N (a_{ij} - \frac{g_i g_j}{2L}) \delta(C_i, C_j) = \frac{1}{16} \left[6 - \frac{(g_1 * g_1) + (g_1 * g_2) + (g_1 * g_3) + (g_2 * g_1) + (g_2 * g_2) + (g_2 * g_3) + (g_3 * g_1) + (g_3 * g_2) + (g_3 * g_3)}{16} + \right. \\ \left. 0 - \frac{(g_4 * g_4) + (g_4 * g_5) + (g_4 * g_6) + (g_5 * g_4) + (g_5 * g_5) + (g_5 * g_6) + (g_6 * g_4) + (g_6 * g_5) + (g_6 * g_6)}{16} \right]$$

Equal to 1 if i, j belong to 1-1 or 2-2

$$Q = \frac{1}{16} \left[6 - \frac{(g_1 + g_2 + g_3)^2}{16} + 0 - \frac{(g_4 + g_5 + g_6)^2}{16} \right] = \frac{1}{16} \left[6 - \frac{(3+4+4)^2}{16} + 0 - \frac{(2+1+2)^2}{16} \right] = -0.20$$

Existence of intra-community links Probability of random links between the nodes

4 ($C=[1\ 1\ 2\ 2\ 2\ 1]$)



g	3	4	4	2	1	2
3	0	1	1	0	0	1
4	1	0	1	1	1	0
4	1	1	0	1	1	0
2	0	1	1	1	0	0
1	0	0	1	0	0	0
2	1	1	0	0	0	0
	1	1	2	2	2	1

$$D = \frac{2L}{2L + \sum_{i,j=1}^N a_{ij} [1 - \delta(C_i, C_j)]} = \frac{16}{16 + 6} = 0.72 \quad D \in [0.5, 1]$$

Equal to 1 if i, j belong to

1--2

$$Q = \frac{1}{2L} \sum_{i,j=1}^N (a_{ij} - \frac{g_i g_j}{2L}) \delta(C_i, C_j)$$

Equal to 1 if i, j belong to 1--1 or 2--2

$$Q = \frac{1}{16} \left[6 - \frac{(g_1+g_2+g_6)^2}{16} + 2 - \frac{(g_3+g_4+g_5)^2}{16} \right] = \frac{1}{16} \left[6 - \frac{(3+4+2)^2}{16} + 4 - \frac{(4+2+1)^2}{16} \right] = \frac{1}{16} [6 - 5.06 + 4 - 3.06] = 0.12$$

Existence of intra-community links

Probability of random links between the nodes

So to summarize:

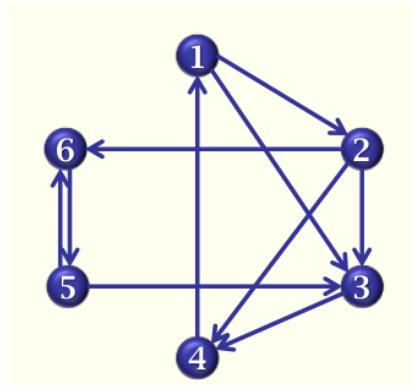
	Divisibility	Modularity	
C=[1 1 1 2 2 2]	0.62	-0.20	Lower D, lower Q: higher integration
C=[1 1 2 2 2 1]	0.72	0.12	Higher D, higher Q: higher segregation

← **BAD SUBDIVISION**

← **GOOD SUBDIVISION**

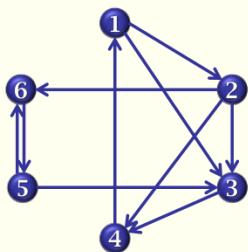
Exercise 2 (Directed graph)

Given the graph in figure:



1. Write down the corresponding binary adjacency matrix A
2. Compute the degree of each node
3. Compute Divisibility and Modularity according to the following two classes: $C=[1 \ 1 \ 1 \ 2 \ 2 \ 2]$
4. Compute Divisibility and Modularity according to the following two classes: $C=[1 \ 1 \ 1 \ 1 \ 2 \ 2]$

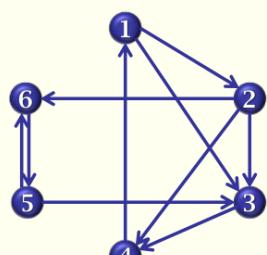
1 (binary adjacency matrix A)



$$N = 6 \\ L = 10$$

$1 \rightarrow 2$ $A =$	<table border="1" style="border-collapse: collapse; text-align: center;"> <tbody> <tr><td>0</td><td>0</td><td>0</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>1</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>0</td><td>1</td><td>0</td></tr> </tbody> </table>	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0
0	0	0	1	0	0																																
1	0	0	0	0	0																																
1	1	0	0	1	0																																
0	1	1	0	0	0																																
0	0	0	0	0	1																																
0	1	0	0	1	0																																

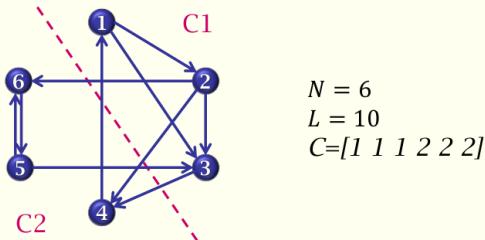
2 (degree of each node)



$$N = 6 \\ L = 10$$

g_{IN}	g_{OUT}					
	2	3	1	1	2	1
1	0	0	0	1	0	0
1	1	0	0	0	0	0
3	1	1	0	0	1	0
2	0	1	1	0	0	0
1	0	0	0	0	0	1
2	0	1	0	0	1	0

3 ($C=[1\ 1\ 1\ 2\ 2\ 2]$)



g_{IN}	g_{OUT}						
	2	3	1	1	2	1	
1	0	0	0	1	0	0	1
1	1	0	0	0	0	0	1
3	1	1	0	0	1	0	1
2	0	1	1	0	0	0	2
1	0	0	0	0	0	1	2
2	0	1	0	0	1	0	2
	1	1	1	2	2	2	

$$D = \frac{L}{L + \sum_{i,j=1}^N a_{ij} [1 - \delta(C_i, C_j)]}$$

$$D = \frac{10}{10 + 5} = 0.66 \quad D \in [0.5, 1]$$

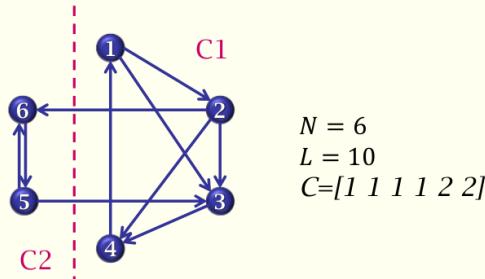
Equal to 1 if i, j belong to 1--2

$$Q = \frac{1}{L} \sum_{i,j=1}^N (a_{ij} - \frac{g_i^{OUT} g_j^{IN}}{L}) \delta(C_i, C_j) = \frac{1}{10} \left[3 - \frac{(g_1^{OUT} g_1^{IN}) + (g_1^{OUT} g_2^{IN}) + (g_1^{OUT} g_3^{IN}) + (g_2^{OUT} g_1^{IN}) + (g_2^{OUT} g_2^{IN}) + (g_2^{OUT} g_3^{IN}) + (g_3^{OUT} g_1^{IN}) + (g_3^{OUT} g_2^{IN}) + (g_3^{OUT} g_3^{IN}) + (g_4^{OUT} g_4^{IN}) + (g_4^{OUT} g_5^{IN}) + (g_4^{OUT} g_6^{IN}) + (g_5^{OUT} g_4^{IN}) + (g_5^{OUT} g_5^{IN}) + (g_5^{OUT} g_6^{IN}) + (g_6^{OUT} g_4^{IN}) + (g_6^{OUT} g_5^{IN}) + (g_6^{OUT} g_6^{IN})}{10} + \right.$$

Equal to 1 if i, j belong to 1--1 or 2--2

$$\begin{aligned} & \left. 2 - \frac{(g_4^{OUT} + g_5^{OUT} + g_6^{OUT})(g_4^{IN} + g_5^{IN} + g_6^{IN})}{10} \right] = \\ & = \frac{1}{10} \left[3 - \frac{(2+3+1)(1+1+3)}{10} + 2 - \frac{(1+2+1)(2+1+2)}{10} \right] = \frac{1}{10} [3 - 3 + 2 - 2] = 0 \end{aligned}$$

4 ($C=[1\ 1\ 1\ 1\ 2\ 2]$)



g_{IN}	g_{OUT}						
	2	3	1	1	2	1	
1	0	0	0	1	0	0	1
1	1	0	0	0	0	0	1
3	1	1	0	0	1	0	1
2	0	1	1	0	0	0	1
1	0	0	0	0	0	1	2
2	0	1	0	0	1	0	2
	1	1	1	1	2	2	

$$D = \frac{L}{L + \sum_{i,j=1}^N a_{ij} [1 - \delta(C_i, C_j)]}$$

$$D = \frac{10}{10 + 2} = 0.83 \quad D \in [0.5, 1]$$

Equal to 1 if i, j belong to 1--2

$$Q = \frac{1}{L} \sum_{i,j=1}^N (a_{ij} - \frac{g_i^{OUT} g_j^{IN}}{L}) \delta(C_i, C_j) = \frac{1}{10} \left[6 - \frac{(g_1^{OUT} + g_2^{OUT} + g_3^{OUT} + g_4^{OUT})(g_1^{IN} + g_2^{IN} + g_3^{IN} + g_4^{IN})}{10} + 2 - \frac{(g_5^{OUT} + g_6^{OUT})(g_5^{IN} + g_6^{IN})}{10} \right]$$

Equal to 1 if i, j belong to
1-1 or 2-2

$$Q = \frac{1}{10} \left[6 - \frac{(2+3+1+1)(1+1+3+2)}{10} + 2 - \frac{(2+1)(1+2)}{10} \right] = \frac{1}{10} [6 - 4.9 + 2 - 0.9] = 0.22$$

1-1 2-2

Existence of intra-community links Probability of random links between the nodes

So to summarize:

	Divisibility	Modularity	
C=[1 1 1 2 2 2]	0.66	0	Lower D, lower Q: higher integration
C=[1 1 1 1 2 2]	0.83	0.22	Higher D, higher Q: higher segregation

BAD SUBDIVISION ← GOOD SUBDIVISION

References

- Cohen, Chapter 31