



Data analysis and modeling of calcium activity in mice somatostatin interneurons

M.Sc. Thesis of
Fabrizio Bernardi

Supervisor:
Prof. Riccardo Sacco

Co-supervisors:
Dr. Francesco Papaleo
Dr. Greta Chiaravalli

Programme:
Mathematical Engineering

Department of Mathematics
Politecnico di Milano

Contents

0.1 Neurobiology of behaviour	4
0.1.1 Neuronal circuits to describe behaviour	4
0.1.2 In vivo studies on mice	9
0.1.3 Calcium imaging through Inscopix	12
0.1.4 Calcium imaging through Fiberphotometry	16
0.1.5 Synchronization of neural activities	17
0.2 Main tools for synchronization analysis	22
0.2.1 Understanding synchronization	22
0.2.2 Cross-correlation	23
0.2.3 Peak-synchronization	25
0.2.4 Granger causality	28
0.3 Interbrain data analysis: emotion discrimination task	31
0.3.1 Emotion discrimination task	31
0.3.2 Normalization of the dataset	33
0.3.3 Behavioural analysis of the task	35
0.3.4 Cross-correlation analysis for the mean activity	37
0.3.5 Peak synchronization analysis	40
0.3.6 Granger causality analysis	41
0.3.7 Self-experience task	42
0.3.8 Conclusions on the emotion discrimination task	45
0.4 Activity of amygdala in altruistic decision making	46
0.4.1 Social decision making in mammals	46
0.4.2 Description of the altruism task	47
0.4.3 Behavioural results in the altruism task	48
0.4.4 The role of social hierarchy in altruism display	50
0.4.5 The role of BLA neurons in selfish and prosocial choices	52
0.5 Mathematical modeling in cellular electrophysiology	55
0.5.1 Electric activity in neurons	55
0.5.2 ODE local and global models	59
0.5.3 The LRM model	60
0.5.4 The GHK model	61
0.5.5 The Hodgkin-Huxley model	62
0.5.6 Cable model	65
0.5.7 Numerical solution of the cable equation	67
0.6 Modeling calcium patterns	70

1 Neurobiology of behaviour

In this first chapter, the biological and experimental foundations of this work will be presented. After a brief review of the main notions of neurobiology, such as the structure of a neuron, the propagation of an action potential in neuronal circuits and the intracellular calcium dynamics, there will be a closer focus on the areas of the brain that are involved in the following discussions.

Next, the experimental setup for such studies is presented, through the description of the behavioral tasks performed on mice and the experimental techniques for calcium imaging, from which it will be possible to extract the data analyzed in Chapter 3.

Finally, a quite modern and, for this work, extremely relevant topic is introduced: the synchronization among neural signals.

1.1 Neuronal circuits to describe behaviour

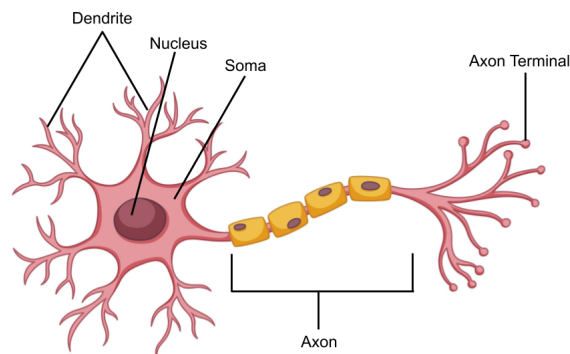


Figure 1: *Schematic representation of a neuron.*

The importance of the brain in mammals is by today extremely evident: this vital organ allows us to think, to capture the stimuli from the environment and elaborate them, enriching our memory and our learning. The brain is the control center of our movements and our actions, it allows us to speak, to understand other individuals and elaborate responses. Most importantly for this work, however, the brain, from its basic cellular units, can explain our behaviour and our emotions.

The brain itself is part of the **nervous system (NS)**, and its most important cell is the **neuron**. Neurons are one of the basic unit needed for the transmission of the electric signals, both within the same area and between different areas of the brain, allowing the NS to collect information and to react to stimuli, elaborating decisions based on them.

The neuron is composed of a central body called *soma*, of the *dendrites*, cellular extensions which collect stimuli from near areas, and of the *axon*, the biological cable connecting one neuron to another, in order to propagate information (Figure 1). Another type of cells present in the NS are the **glial cells** (like astrocytes), which perform functions of protection, sustainment and nutrition for the neurons. Although glial cells are not the main focus of this work, their contribution should not be neglected, as it has been shown to be relevant in many important processes of the nervous system [Semyanov, Henneberger 2020].

Neurons are **excitable** and **conductive**, i.e. they can generate an electrical impulse and transmit it to other neurons, forming neuronal microcircuits, often associated with a specific area and/or task of the brain.

Inside every neuron, in the cytoplasm, there is a coexistence of different ionic species (mostly Na^+ , Cl^- , K^+ , Ca^{2+}) which, in equilibrium conditions, assume a certain concentration, concurring to determine a difference between the electric potential assumed inside and outside the cell: we call this quantity the **membrane potential** of the cell (indeed it is the potential drop formed across the cell's membrane). In reaction to an external stimulus, the ionic concentrations change rapidly their values, provoking a heavy change in the membrane potential. As a consequence, the excitation of a neuron happens, as well as the formation of an **action potential**, which will propagate to other neurons across the axon (Figure 2).

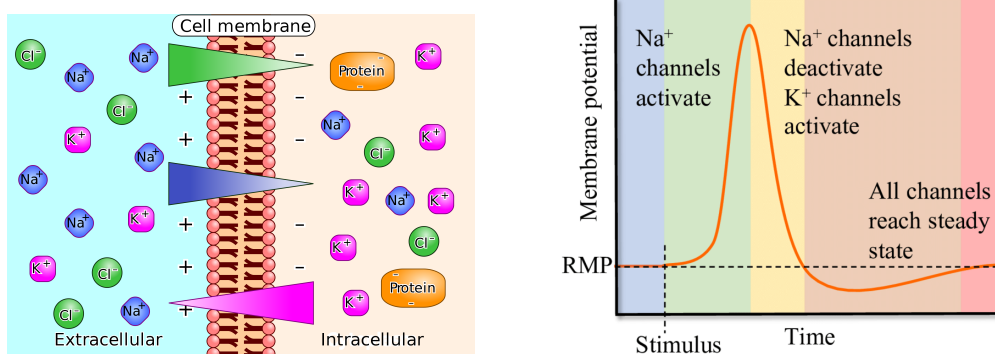


Figure 2: *Left: schematic representation of ionic species inside and outside the cellular membrane. Such species are allowed to pass through the membrane only in specific circumstances through adhibited ionic gates*
Right: example of action potential formation in reaction to a stimulus.

At the end of the axon, the link between a neuron and its neighbour, and the relative passage of the electrical impulse, happens thanks to a chemical **synapse** (Figure 3), a particular structure located in the terminal of the axon. Here, special molecules called *neurotransmitters* are syntethized, and the arrival of an action potential allows such molecules to travel towards the *intersynaptic space*, where they can bind to receptors located in the post-synaptic cell. The response of the post-synaptic cell, then, can be either excitatory or inhibitory, based on whether the impulse is preserved in the circuit or supressed.

The synapses (and thus their corresponding neuron) can be classified in four main groups, depending on the type of neurotransmitter which they release: glutamatergic, GABAergic, cholinergic, adrenergic. The most relevant for this work will be the one of GABAergic synapses, consisting of inhibitory neurotransmitters which reduce the excitability of the neurons.

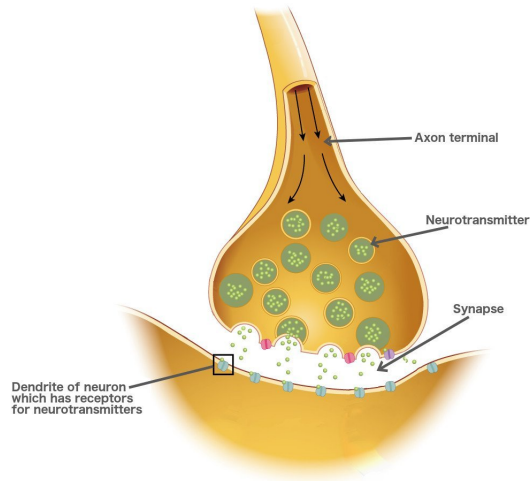


Figure 3: *Schematic representation of a chemical synapse.*

In neurobiology, the conditions under which a neuron can be considered **active** are not unique and matter of debate; in the context of neural activity, the literature usually means either the already discussed electrical activity, or the **calcium activity**. The intracellular dynamic of this particular element, in ionic form of Ca^{2+} , is known to be essential for the main cellular processes, and is strictly related to the formation of an action potential and subsequent propagation of the electrical impulse.

Indeed, we can usually observe *strong instabilities* in the intracellular calcium concentration (Figure 4), which often show fast oscillations and changes, through the formation of sudden peaks: we will define the neuron as *active* in correspondence to these peaks.

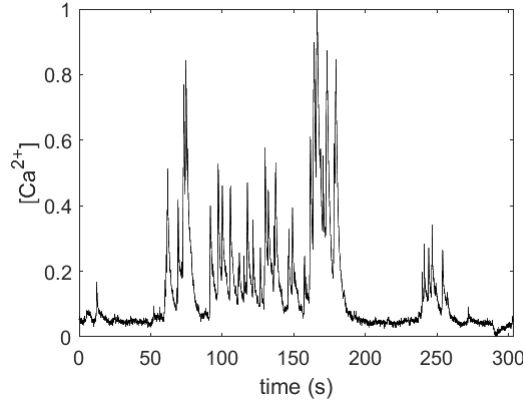


Figure 4: *Example of Ca^{2+} concentration recorded in a neuron, as function of time.*

Having defined the meaning of neural activity, we can ask ourselves whether this activity could be related to aspects like behaviour, emotion, body language, mood: in other words, to investigate whether there is a connection between the activation of some specific neurons, and what animals do.

This topic has fascinated neuroscientists for decades, and from what it has been discovered so far, it seems evident that **different brain circuits are responsible for different behaviours** (although these distinctions may not always be that marked, unique and easy to detect).

The neurobiology experiments on which this work relies on refer mainly to some particular areas of the brain, which previous studies found to be strictly connected to emotional and behavioral processes [Etkin et al. 2011]:

- **Medial prefrontal cortex (mPFC)**: part of the frontal cortex, the area of the brain located in the frontal lobe. This area is implicated in cognition processes, including socio cognitive abilities, with strong connections with decision making [Carlen 2017]
- **Anterior Cingulate Cortex (ACC)**: part of the cingulate cortex and situated in proximity of the mPFC, with which ACC shares a lot of functionalities. Indeed, experimental evidence of the connection between this area of the brain and emotions have been found [Zheng 2020]. This area seems also to be implicated in social aspects like morality or empathy [Carillo 2019], as reaction to interactions with another individual

- **Amygdala:** nuclear complex located in the medial part of the temporal lobe. It is responsible for the elaboration of the emotions, it collects stimuli from the thalamus and elaborate responses: in other words, it plays the role of emotional thermometer of the body and the decision maker for adequate responses [ARTICOLO?]

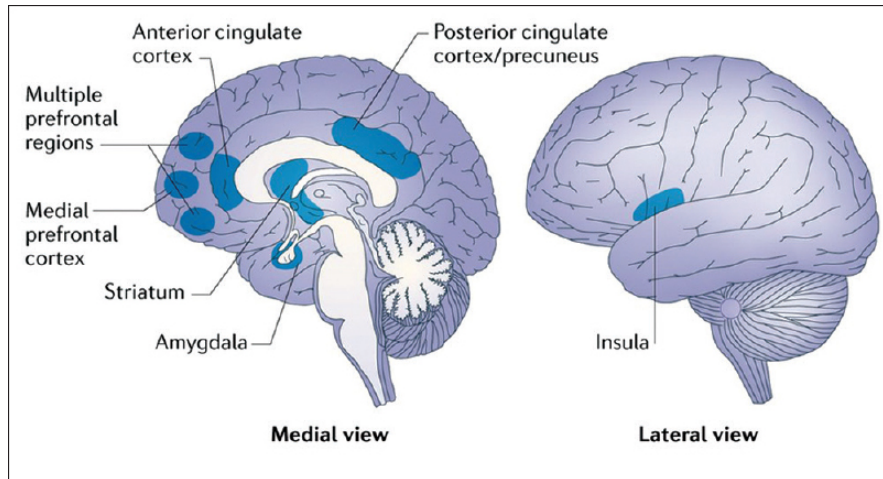


Figure 5: *Main areas of the brain involved in behaviour and emotion recognition.*

1.2 In vivo studies on mice

According to the *Foundation of Biomedical Research (FMR)*, approximately 95% of all lab animals are mice or rats, as in the case of this current work. Although very often the final goal of the research is an application useful for humans, the reasons why mice are the first choice for experiments are numerous. First of all, mice are mammals quite similar to humans when it concerns genetics, at the point that scientists have been able to reproduce genes in mice similar to the ones implied in human diseases ("*transgenic mice*" [Harari,Abramovich 2014]); rodents, indeed, are also easy to manipulate from a genetic point of view. Another reason is definitely the fact that mice are small, easy to control, quite cheap to buy and usually docile. Finally, being the most used animal in research, the largest information and literature which can be found is about them, their anatomy, their typical behaviours, and by today large populations of rodents have been created and used exclusively for experimental purposes, being almost all identical and therefore setting a uniform standard for study and validation of results.

Through a **behavioral task**, one or more mice are put in a particular situation designed to study their reactions, such as a shown behaviour or movement, and, in this context, the goal is to find a relationship with their correspondent neural activity. Therefore, the overall experiment consists in the following steps:

1. Preparation of the arena and equipment for the task
2. Preparation of the mice for the experiment, both in terms of behavioral conditioning, and for neural activity measurements
3. Performance of the test with simultaneous recordings of aspects of interest such as behaviour or neuronal activity
4. Pre-processing and analysis of the collected data

In this work, the behavioural task which has been studied is a particular realization of the **emotion discrimination task**, schematized in Figure 6. The first step of these types of tasks, performed in this way several times in the past, sees the presence of three mice: one, called the **observer**, is a mouse free to move in an arena, while the others two, called the **demonstrators**, remain in a cage. In the phase preceding the test, the *habituation phase*, one of the two demonstrators is *emotionally altered*: namely, it is subjected to specific conditions and procedures which provoke an alteration of its affective state. This alteration could be either negative (usually stress condition following a 15-minutes period of restrain) or positive (usually relief

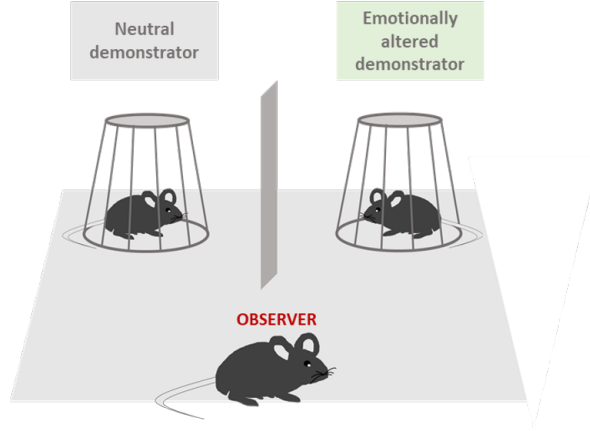


Figure 6: *Basic setting of the emotion discrimination task: one mouse (the observer) is free to move in the arena, while the others two (the demonstrators) remain inside a cage. One of the demonstrators has a neutral affective state, the other an altered affective state.*

condition, consisting in water *ad libitum* following water deprivation). The other demonstrator and the observer are instead in a *neutral* state.

Previous works on this setup [Scheggia-Managò] were performed using both positively and negatively affected demonstrators. The cellular target consisted in a specific subpopulation of neurons, which is thought to be involved in processes like *affective state discrimination*: the **somatostatin (SOM+)** interneurons in the medial prefrontal cortex.

These neurons, expressing the somatostatin neurotransmitter, form a small subgroup, distinct from others like pyramidal neurons or parvalbumin interneurons, but unlike these former ones, they seem to be directly implicated in the affective state discrimination of a conspecific mouse. The data analysis on the task showed some remarkable results:

- The time spent by the observer near the altered demonstrator was significantly higher than the time spent near the neutral demonstrator. A similar consideration holds for the duration of the moments when reciprocal sniffing between two close mice happened (Figure 7)
- This discrimination appears to be stronger in the first period of the test
- The discrimination is directly connected to the presence of a mouse in an altered state (both negative and positive), since from the repe-

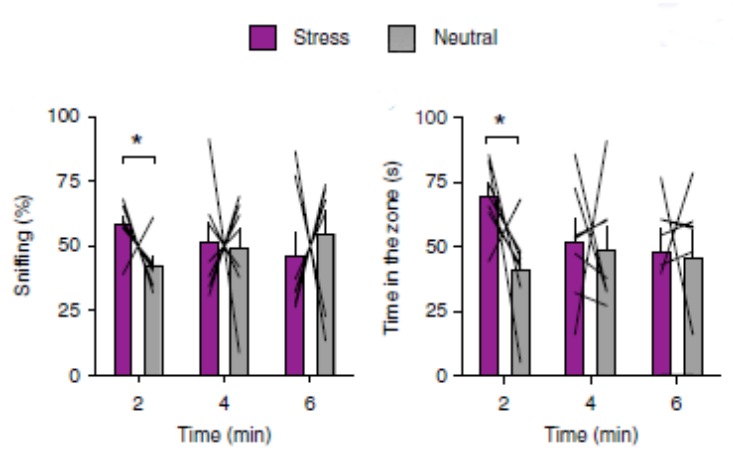


Figure 7: *Recording of sniffing times (left) and proximity times (right) between observer mouse and stressed demonstrator.*

tition of the task using only two neutral demonstrators no significant differences emerged

- By repeating the task eliminating one sense at a time, it has been shown that the main sense responsible for the discrimination is olfaction. However, in order to achieve the same results of the test, a combination of both olfactory and visual cues is needed
- The neural activity of the observer (both in terms of electrical impulse and calcium peaks) showed stronger values during the interactions with the altered mouse rather than the neutral one
- If optogenetic inhibition of SOM+ cells [REFERENCE?] is performed on the observer before the test, the discrimination disappears, namely no relevant differences between time spent and sniffing behaviour have been observed in relation to the altered demonstrator rather than the neutral one
- All the previous considerations do not apply to other neuronal categories like pyramidal or parvalbumin interneurons

Overall, these results seem to show a **key role of the somatostatin interneurons in affective state discrimination**.

Other examples of emotion discrimination in mice related to SOM+ interneurons can be found in [Mariotti et al 2018], with a focus on their effect

on cortical astrocytes: the main result shown in the study consist in a **crucial role of somatostatin** in the stimulation of responses by astrocytes. As a first result, it has been shown that the stimulation of SOM+ interneurons (via 10-pulses light stimulation of calcium indicators), increases the Ca^{2+} response from astrocytes. Also in this case, a comparison with parvalbumin (PV+) interneurons show the *critical sensitivity of astrocytes only to SOM+ interneurons, but not PV+*. Overall, this study reveals that a sustained activity in the SOM+ interneuron circuits is complemented by a sustained activity in the astrocytic network, thus confirming the importance of astrocytes and their link to inhibitory neuronal circuits.

Somatostatin is not the only neuropeptide that has been shown to be linked to emotion discrimination. For example, in [Ferretti-Maltese 2018], the in vivo task on mice targets the **oxytocin (OXT)** neurotransmitter, studying its release in the central amygdala. The behavioral task of this study has a similar structure to [Scheggia-Manago], in which neutral and altered demonstrators were contraposed to an observer mouse. Also this case shows evidence of discrimination caused by the altered mouse (for example, through an increased sniffing activity), as well as the importance of olfactory clues in the process of such discrimination. A reduction of OXT level in the central amygdala resulted in the *abolishment* of emotion discrimination, while a new increase of such level resulted in the rescue of the former conditions, thus remarking the central role of the OXT neuropeptide in the emotion discrimination.

When targeting a specific neuronal subpopulation, the hope is to always find more evidence on its connection with properly shown tendencies, emotions or behaviors that are common to known diseases. In the case of the emotion discrimination task targeting SOM+ interneurons in the mPFC, for example, the observed effects in the abolition of discrimination, following photoinhibition of such neurons, could be similar to the ones observed in neurodevelopmental disorders such as autism spectrum disorders (ASDs) or schizophrenia. Therefore, this approach aims to investigate the causes of a specific issue at a basic and detailed level such the one of single neuron precision, with the ultimate goal of acting on such level to perform a change on the macroscopic effects.

In order to do so, the paradigm of the task has to be defined in a way that it allows a collection of unbiased results, but an adequate way to perform an imaging of the neuronal activity of interest is necessary as well (and described in the following sections). As a final step, a proper data analysis on those results has to be performed (in the present case, using the methods

introduced in Chapter 2).

1.3 Calcium imaging through Inscopix

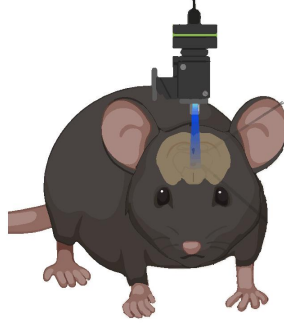


Figure 8: *The Inscopix miniscope*

In this section, a powerful tool for single neuron calcium imaging will be presented: **microendoscopic calcium imaging**. The Inscopix company [<https://www.inscopix.com/>] provides tools and software to extract calcium tracks as one photon measurements from in vivo experiments on mice.

To measure the intracellular Ca^{2+} concentration in a single neuron of a mouse during a behavioral task, the first step is the performance of a surgery on the mouse to implant a miniscope in the *region of interest (ROI)* of the brain to be investigated. On the same mouse, **genetically encoded calcium indicators (GECI)**, such as the **GCaMP** protein, are adopted. These types of proteins, when bound to Ca^{2+} , emit a fluorescent light, which intensity will be captured by the miniscope. At the end of the test, the miniscope will be able to give back a video of the ROI, in which the evolution of the fluorescence through time can be appreciated in the neurons, as shown in Figure 9. In this way, it is possible to obtain a representation of the intracellular calcium oscillations behaviour during the whole task. This information can be combined with behavioral or positional data of the mice in the arena, making it possible to start a data analysis and look for correlations between behaviour and neural activity.

In order to have the data ready to be analyzed, the Inscopix software manages also the pre-processing part, which goes through the following steps:



Figure 9: *Example of pre-processing work through Inscopix software. We can appreciate: the list of detected neurons (left), a segment of the video in the ROI (top-right) and an example of the calcium track recorded in a single neuron through time (bottom-right)*

1. If needed, more videos recorded from the task can be combined, in order to obtain a single final video of the ROI through time
2. **Spatial and temporal downsampling** are performed, selecting the appropriate scale factors of space and time which allow to capture all the important information
3. **Spatial filtering** of the image: it removes low and high frequencies with a bandpass filtering. The filter has the form

$$M_f^{band} = GB(M_f, \sigma_{high}) - GB(M_f, \sigma_{low})$$

Where M_f represents the frame f of movie M , GB is the *Gaussian Blur* function, and the standard deviations $\sigma_i = \frac{2\ln(2)}{2\pi\lambda_i}$ are computed from the cut-off values λ_{high} and λ_{low}

4. **Motion correction**: it accounts for the motions between different frames of the videos, applying a correction to let every pixel staying at the same place along the different frames of the movie

5. **Pixel normalization:** each pixel of a frame represents a different luminosity, caused by the fluorescence obtained from the reaction with the GCaMP. After the preprocessing, the value of the fluorescence (which estimates the calcium concentration) is expressed as

$$\frac{\Delta F}{F} = \frac{F(x, y, t) - F_b}{F_b}$$

where $F(x, y, t)$ represents the measured fluorescence at the point (x, y) at time t , while F_b is a baseline fluorescence value (usually the mean value of the movie, in some cases the minimum). After this step, the returned value for every pixel is an adimensional relative value of fluorescence

6. The **PCA-ICA algorithm** attempts to recognize the neurons of the ROI in the movie. Based on given information such as average cell's diameter, a principal component analysis (PCA) is performed, followed by an independent component analysis (ICA), until convergence. In particular, the frames of the movie, represented as matrices, are rasterized into 1D vectors and they are normalized on mean and standard deviation. At this point, a principal component analysis is performed, in order to reduce dimensionality. Every frame is then approximated by a weighted sum of the principal trace components. Finally, an ICA algorithm is launched. At the end of this process, every neuron will be identified, labeled, and gifted with a calcium activity value (which is actually $\frac{\Delta F}{F}$) for every time instant of the movie.
7. **Final adjustments:** appropriate algorithms or manual fixing are performed to take into account some imperfections, such as mistakes in neuron identification, inaccuracies due to overlapping neurons, excessive noise in the calcium traces for a neuron (discarding the corresponding cell from the analysis)
8. **Data import:** finally, data are imported in a csv file and are ready to be analyzed (in this work, using the software MATLAB)

At the end of all these steps, the final data available consist of a list of every neuron detected, in which for every one of them, at every time stamp is associated the corresponding value of $\frac{\Delta F}{F}$ recorded. The preprocessing step is now over, and the calcium traces of every neuron in the ROI are available for the data analysis process.

1.4 Calcium imaging through Fiberphotometry

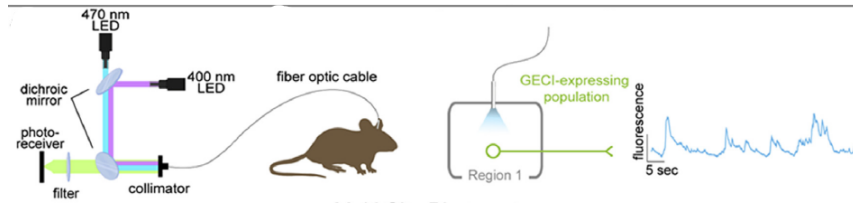


Figure 10: *Fiberphotometry experimental setup*

In the previous section, it has been shown a calcium imaging technique based on single neuron traces, measured with single photon miniscopes. In this section, another common technique for calcium imaging is under investigation: **fiberphotometry**.

Although the main goal of this technique is the same as the Inscopix imaging one, technical equipment, methodology and results present substantial differences.

Relying on genetically encoded calcium indicators such as GCaMP, with fiberphotometry the fluorescence signal is recorded through *optic fibers*, typically of length $300 - 400 \mu m$. An excitation of GCaMP is performed using a blu LED light at the appropriate wavelength ($470 nm$), which is transmitted through the optical cannula implanted in the mouse, while an emission green light ($525 nm$) is relayed to a photoreceiver. Often, a second excitation light (violet, $405nm$), is used as well, to take into account autofluorescence and produce an isosbestic (calcium-independent) control signal [TESI PHD GIULIA].

Finally, a software manages the output signal in order to perform filtering and return a collective raw signal, separating the two contributions from blu and violet lights. Such signal, unlike the miniscope neuronal imaging, can only be an *aggregate* signal of the observed area. This implies that the fiberphotometry technique produces signals at lower spatial resolution than the previously introduced technique of microendoscopic imaging. However, this technique is becoming increasingly popular because its smaller accuracy is compensated by several positive factors:

- The overall experimental setup for fiberphotometry results cheaper than the one for miniscope imaging
- The procedure operated in mice are less invasive, and the subjects are more suitable to long freely behaving experiments

- The implant of the cannula is relatively easy to perform, and a good signal is collected even with fiber placements in the neighborhood of GECI -expressing populations [Siciliano and Tie 2018]
- Multiple areas of the brain can be investigated at the same time
- Fiberphotometry can be used for other fluorescent indicators, such as norepinephrine (Feng et al., 2019), GABA (Marvin et al., 2019), glutamate (Liang et al., 2015), acetylcholine (Jing et al., 2018) and dopamine (Patriarchi et al., 2018)

1.5 Synchronization of neural activities

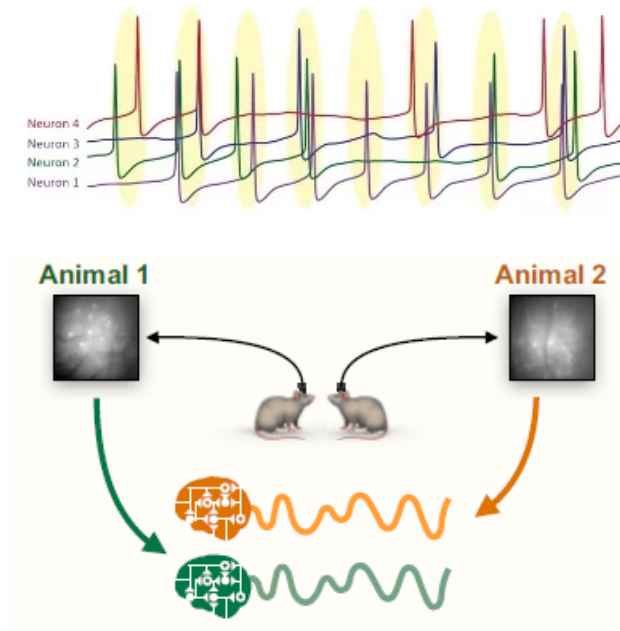


Figure 11: *Top: single neuron synchronization*
Bottom: overall activity synchronization

In the previous sections, the concept of *neural activity* has been introduced in its different shapes. Not only it can be identified with the electrical and calcium activity, but *activity* necessarily refers to a specific area: it can be the activity of a single neuron or of a group of neurons in a ROI (often, in this case, the activity is attributed to the whole animal for that specific context). In any case, the main object of interest of neural activity is a **signal evolving in time**.

While, historically, the study of neural activity has been mostly dealing with signals of single individuals treated one by one, the focus, in the most recent years, is extending to cover as well a more complex topic: the **synchronization between different brain activities**. The definition of synchronization between signals, not unique and assuming different shapes for different needs (Chapter 2), can involve two types of correlations (Figure 11):

- **Intraneuronal synchronization:** the signals of single neurons show a correlated activity. Often (but not necessarily), the relevance of this phenomenon is to look for correlations *within* the same region of interest (ROI), namely for correlations among the neurons of the same animal in a particular area of its brain. However, the correlation could be investigated as well between different areas of the same individual, such as between neurons of different mice
- **Interbrain synchronization:** defining the activity of one individual (for example through a mean of its neuronal activities), it is possible to study the synchronization between two individuals by investigating correlations between their overall signals

The first type of synchronization, i.e. the correlation between a subset of neurons in the same individual, is a property of a neural circuit: it has been observed that the chain of firing of neurons (both in electrophysiology and calcium imaging) often exhibits some **patterns**. This means that some neurons tend to show simultaneous peaks through time, or a similar order and timing of firing. The presence of correlated activity in a ROI has been shown to be linked with expressed behaviours [Frost et al. 2020], leading to the conclusion that **different behaviors can cause the synchronizations of different subpopulations of neurons**.

This phenomenon could be independent by a rise of neural activity as consequence of an observed behaviour, in the sense that, in principle, it may happen that a rise in a correlated activity does not correspond to a rise in neural activity and viceversa. One of the goals of the correlation analysis is the **identification of a group of neurons encoding a specific behaviour**. The neuronal ensemble in terms of synchronization may be different from the neuronal ensemble in terms of activity peaks (even if often the intersection of the two is quite significant [Scheggia-Managò]).

In the second type of synchronization, the **interbrain synchronization**, the correlation analysis is performed between signals of different individuals; this could be intended both for single neurons and for overall individual activities. Therefore, this type of synchronization is studied through an *interaction* between two subjects. Again, the goal is either to find an ensemble

of neurons which tend to correlate to the partner ones, or to show that in specific situations the correlation between subjects increases or decreases.

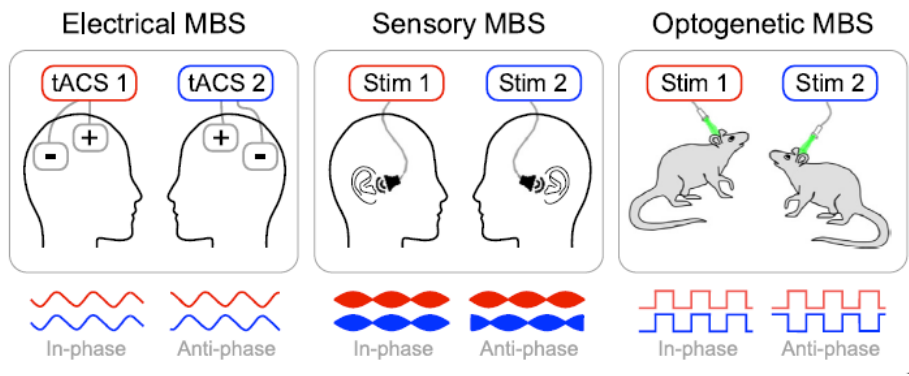


Figure 12: *Example of multi brain stimulation techniques*

Previous works ([Wass et al. 2020]) have been performed on humans in order to investigate the causes and effects of interbrain synchrony (often referred to as *hyperscanning*). Performing an **electroencephalography (EEG)** on two interacting subjects, a strong connection has been identified between a rise in synchronization and aspects like interpersonal coordination, cooperation and communication. In [Novembre et al. 2021], this relationship is further inspected: the simultaneous presence of synchronization and behaviour is not enough to establish a *causality* between the two, which, instead, can be shown only by studying the effects which the manipulation of the first causes on the second, or viceversa. This is achieved through **multi brain stimulation (MBS)**: stimulation processes (usually of non invasive type in humans, invasive on animals) are performed on the subjects, and the provoked effects measured. Therefore, a correct analysis should combine the two approaches (hyperscanning and MBS) to be able to show true causality between interbrain synchrony and behaviour.

As for interbrain analysis on mice based on microendoscopic calcium imaging, few works have been done and the topic is still largely unexplored. In one of the most significant papers, [Kingsbury et al. 2019], intracellular calcium has been recorded from neurons of the mPFC in two subjects interacting in an open arena. The mean activity (as $\frac{\Delta F}{F}$) has been computed for every mouse, and the correlation between the two corresponding signals has been computed as well (using mostly cross-correlation, see Chapter 2). The

first result was that animals engaged in social interaction showed higher interbrain synchronization, reporting higher values of correlation during social interaction rather than solitary periods. Moreover, the abolishment of physical interaction, using a barrier, resulted as well in the inhibition of interbrain correlation, meaning that the synchronization is not due to environmental factors, but instead strictly related to the direct interaction between conspecifics (Figure 13).

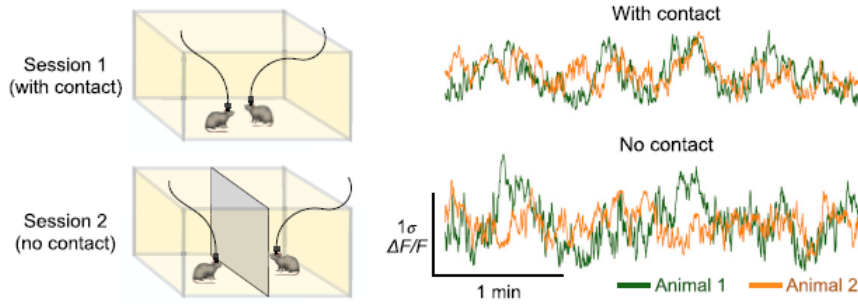


Figure 13: *Comparison of neuronal activities in the two mice with and without contact*

The second task performed in the experiment was the *tube test*, in which the two subjects were placed facing each other in a horizontal tube, allowing the passage of only one individual. In this configuration, the mice can assume three types of behaviors: *approach*, *push*, and *retreat*; here animals showed again synchronization during the competitive encounters. In particular, a classification of neurons as behavioral cells allowed to link subgroups of neurons with the three types of behaviors depending on their activity values. Interestingly, the removal of these behaviour cells resulted in a marked reduction of the correlated activity. Finally, between the two animals, it was evident that one would tend to assume the role of *dominant* (prevalence of push behaviour), while the other of *subordinate* (prevalence of retreat). Through generalized linear models (GLMs), it has been proposed a description of the dependence between the behaviour of one mouse and the neural activity of the other, observing that cells in subordinates responded more to the ones of the dominant than vice-versa. The consequence at synchronization level is that the dominant dictates the *rhythm* for the subordinate's signal, making it able to predict the expected behaviour of one mouse based on the activity of the other.

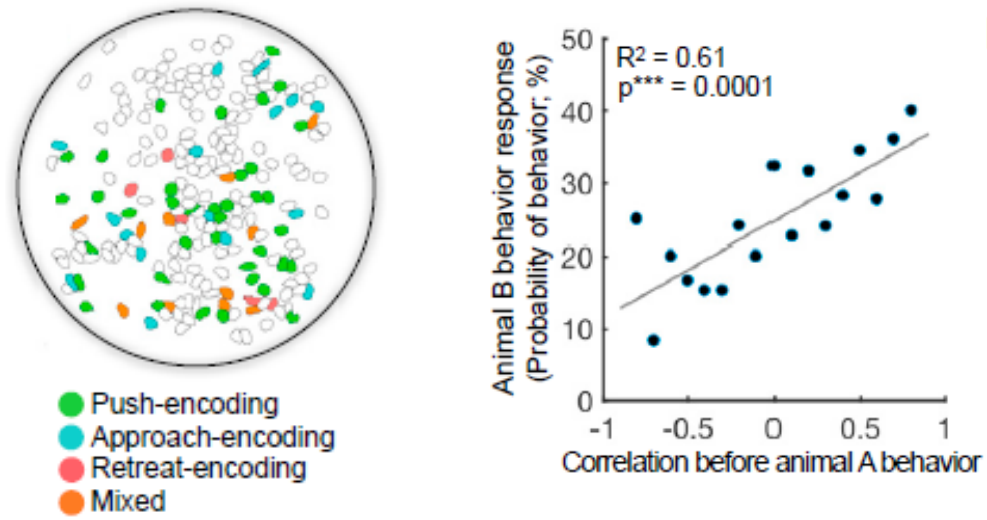


Figure 14: *Left: identification of behavioural neurons in the ROI. Right: correlation between the interbrain synchronization preceding behaviour in one animal and the response probability of the interacting partner*

2 Main tools for synchronization analysis

In this chapter, the main tools for the analysis of synchronization between signals will be presented, with a focus on their mathematical definition and their role in the application of *Interbrain synchronization*.

After dealing with the problem of *what* synchronization actually means in this context and which measures *should not* be considered, the two main tools corresponding to the two different synchronizations discussed in Section 0.1.5 will be presented : **cross-correlation**, to be computed considering overall activity signals, and **peak-synchronization**, to be computed between single neurons. Each one of these tools has a specific purpose and meaning, and they will all be considered in the analysis of this work, in order to give different shapes to the concept of synchronization and to try to study it under the largest possible perspective.

Finally, when observing the presence of correlation between two signals, a different (although strongly connected) question may arise: "which is the *relationship* of such connection?", or, in other words, "is one signal determining the behaviour of the other?". To answer such questions, a sophisticated tool will be object of study: the **Granger causality**.

2.1 Understanding synchronization

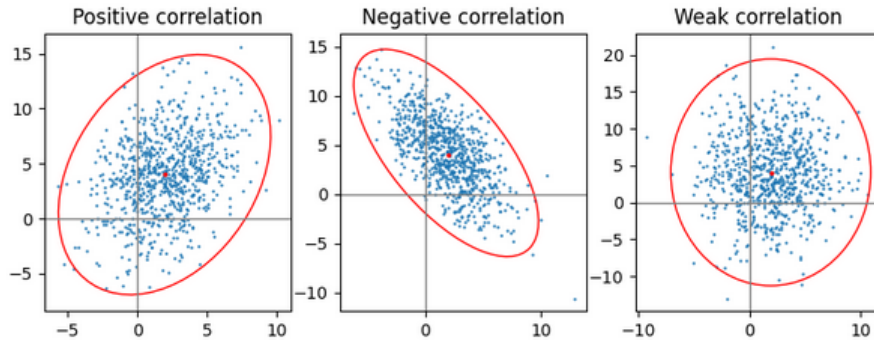


Figure 15: *Geometrical interpretation of the Pearson correlation: such coefficient gives a measure of how well the pairs obtained from the two samples distribute linearly along the diagonal of an ellipse covering the scatterplot.*

One of the main goals of this work will be to perform a data analysis on the Interbrain synchronization of neural signals, recorded in behavioural tasks.

In order to do so, first one has to clarify the meaning of the term *synchronization*.

In a statistical sense, the synchronization between two time series can be seen as a measure of the **correlation** between them: the more two series are correlated, the more similar and connected they will be. This leads inevitably to the fact that, when talking about synchronization, the definition can't be unique, since the similarity between two series can be assumed in different ways, depending on the particular aspect of interest.

However, a first hint about what to look for or not can be found in the type of signal which is under analysis. In the present case, dealing with single neuron measurements of the intracellular concentration of $[Ca^{2+}]$ means dealing with *strongly nonlinear and unstable signals*, which exhibit sudden peaks and are often chaotic and difficult to predict.

For this reason, the most commonly adopted measure of correlation, namely the **Pearson correlation**, is not suited to describe such signals. Given two random samples $\mathbf{x} = \{x_i\}_{i=1}^N$ and $\mathbf{y} = \{y_i\}_{i=1}^N$, the Pearson correlation (PC) between \mathbf{x} and \mathbf{y} is defined as

$$\rho(x, y) = \frac{Cov(x, y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2} \quad (1)$$

where:

- $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$ is the sample mean
- $Cov(x, y) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$ is the sample covariance
- $\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$ is the sample standard deviation

The PC measures the *linear correlation* between two variables. This implies that this type of correlation should be used to inspect a linear relationship between two variables, which have a distribution close to the Gaussian one, and a uniform variance (i.e. reduced presence of outliers) [Applied multivariate statistics book by Johnson]. Unfortunately, with the type of data incoming from the neural recordings, all these hypotheses fail, and it follows that other correlation measures should be inspected instead.

2.2 Cross-correlation

When investigating the similarity between two time-dependent signals, the first tool to be considered is the **cross-correlation**.

Formally, given two functions $f = f(t)$ and $g = g(t)$, we define the cross-correlation between them as

$$[f(t) \star g(t)](\tau) = \int_{-\infty}^{+\infty} f(\tau)g(t + \tau)d\tau \quad (2)$$

Given that the **convolution** between such functions is defined as

$$[f(t) * g(t)](t) = \int_{-\infty}^{+\infty} f(t)g(t - \tau)d\tau \quad (3)$$

it follows that

$$[f(t) \star g(t)](t) = [f(-t) * g(t)](t) \quad (4)$$

This means that the cross-correlation coincides with a convolution in which one function is considered backward in time. Moreover, in the common form of cross-correlation, as shown in (0.2), the resulting quantity is not expressed as a function of the time variable t , but as a function of τ , i.e. the *lag* or *delay* between the signals. The interpretation is straightforward: when, for a given value of the delay τ , simultaneous peaks of the two signals are both present, the contribution of their product in the integral will be more relevant for the computed value of cross-correlation correspondent to that specific lag.

As a consequence of this, once computing the cross-correlation between two time-dependent functions, one can identify its maximum value and retrieve the corresponding value of the lag τ , from which it is possible to obtain an estimation of the delay between the two functions. To summarize, a cross-correlation analysis allows to:

1. Compute the cross-correlation between the two signals as a function of the lag value τ
2. Estimate the real delay between such signals, observing when the peak of cross-correlation occurs

When the cross-correlation is computed between the same function, it takes the name of **autocorrelation**. Confronting two identical signals, there will always be a peak of correlation (equal to 1 using *normalized* cross-correlation) corresponding to the lag $\tau = 0$. Moreover, the shape of the cross-correlation will be always symmetrical (see Figure 16).

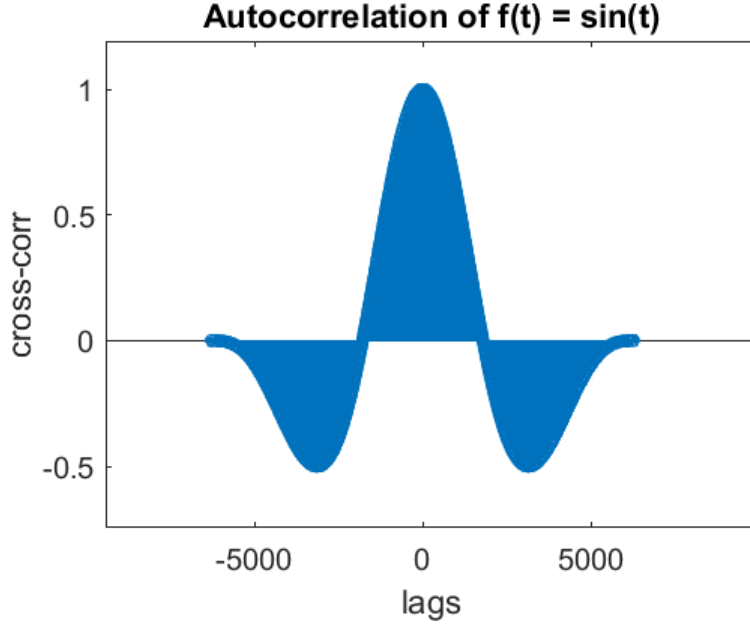


Figure 16: *Aurocorrelation of the function $f(t) = \sin(t)$*

While the autocorrelation represents an ideal case, when dealing with two different functions $f(t)$ and $g(t)$, the significance of their cross-correlation will be given by the amount of shared properties to the case of autocorrelation; however, depending on the current application, it can be reasonable to expect a peak value in correspondence to a nonzero value of the lag, as a representation of an actual physical delay.

Given two time series $\mathbf{x} = \{x_i\}_{i=1}^N$ and $\mathbf{y} = \{y_i\}_{i=1}^N$, the discrete approximation of (0.2) reads:

$$[\mathbf{x} \star \mathbf{y}](m) = \sum_{n=1}^{N-1} x_n y_{n+m} \quad m_{min} < m < m_{max} \quad (5)$$

where m is the approximate lag value, chosen in an appropriate interval.

2.3 Peak-synchronization

Cross-correlation is a general and widely used tool to quantify the similarity between two signals evolving in time. With the following concept of **peak-synchronization**, the focus of the analysis is restricted to the characteristic type of signal of this work: recording of intracellular calcium activity.

As shown in Chapter 1, a typical recording of the activity of a single neuron is characterized by the presence of rapid and intense *peaks*, which define the neuron as *active*. It follows that a way to intend synchronization between neurons could be related to the presence of *close* or *simultaneous* peaks observed at the same time. In other words, two signals (hence, neurons) are synchronized, under the peak-synchronization point of view, if a *pattern* of simultaneous firing occurs.

In order to quantify the peak-synchronization between two signals, the following steps have to be faced:

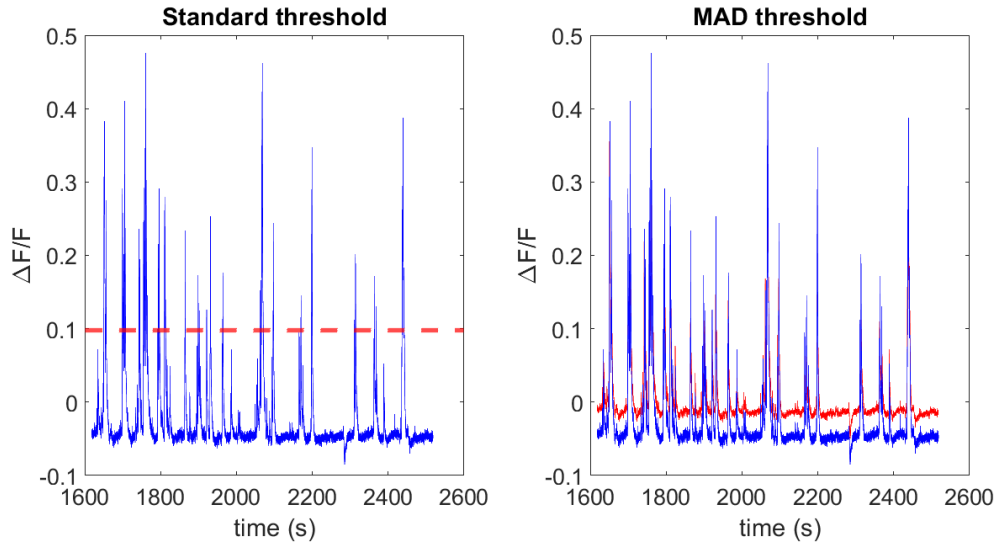


Figure 17: *Example of a calcium signal (blue). The red curves are the thresholds built from the two algorithms: standard threshold (left panel) and MAD threshold (right panel)*

1. Identify when a neuron is *active*, i.e. when we are in presence of a peak
2. Choose an appropriate time interval in which two neurons firing simultaneously can be considered as synchronized
3. Quantify the correlation between the two neurons

The way to solve all of these three steps is not unique, but it is strictly depending on the current biological application under study and by the choice of

appropriate algorithms for the activity detection and synchronization quantification. In this work, the following considerations and choices have been made:

1. Given a discrete time series, several algorithms are available or could be designed to detect, when a peak is occurring. A straightforward idea is to establish a threshold for the activity, in such a way that all the points above the threshold are active and all the ones below non active. Therefore, the output of such algorithm is of binary type (1 for activity, 0 for non-activity).

A naive way to define the threshold could be to consider a horizontal line, for example given by the equation $y = \mu + 2\sigma$, where μ and σ are the mean and the standard deviation of the signal. In this way, all the values higher than y will be considered active and vice versa. This approach is good enough when dealing with "well-behaving" signals, presenting a baseline low activity alternated by huge peaks, however it seems to fail when dealing with more complex and noisy signals.

Algorithm 1 Standard threshold algorithm

- 1: Consider a horizontal threshold given by $y = \mu + 2\sigma$
 - 2: Confront every point of the signal with each corresponding point of the threshold
 - 3: Every point above the threshold is labeled as 1, all the points below as 0
-

For this reason, a different algorithm can be considered as well [Inscopix manual]: the **MAD threshold** algorithm. The threshold established by this algorithm is not constant, but it "follows" the signal, in order to better capture its dynamics.

The two types of threshold resulting from the algorithms are schematized in Figure 17.

2. A typical time interval in which neuronal firing occurs usually strongly depends on the specific case. However, in general it is safe to say that the peak of a neuron usually has a duration of 500 – 1000 ms. Such value will define the reference time window.
3. Finally, once the binary vectors of activations are available, as well as a reference time window, to quantify their synchronization, a common tool adopted is the **Peak-correlation index** [Cutts and Eglén], defined as

Algorithm 2 MAD threshold algorithm

- 1: Start from a baseline threshold given by $MAD = \text{median}(X_i - \text{median}(X))$
 - 2: Identify the points where the slope changes from positive to negative (PN) and from negative to positive (NP)
 - 3: At every PN point, the threshold value is the MAD value plus the previous NP point's value
 - 4: The overall threshold is obtained from linear interpolation of the threshold points
 - 5: Every point above the threshold is labeled as 1, all the points below as 0
-

$$i_{AB} = \frac{N_{AB}T}{2N_A N_B dT} \quad (6)$$

Here T is the overall signal time window, dT is the synchronization time window, N_A is the number of peaks in signal A, N_B is the number of peaks in signal B and finally $N_{AB} = \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} I_{[-dT, dT]}(|a_i - b_j|)$ is the sum of simultaneous peaks within each synchronization window.

The peak correlation index gives a representation of how well two signals (neurons) are peak-synchronized. However, it should be noticed that such measure is not normalized, meaning that the value of one index considered by itself has no real meaning, and this measure should be used only as tool to compare the same pair of signals through different phases.

2.4 Granger causality

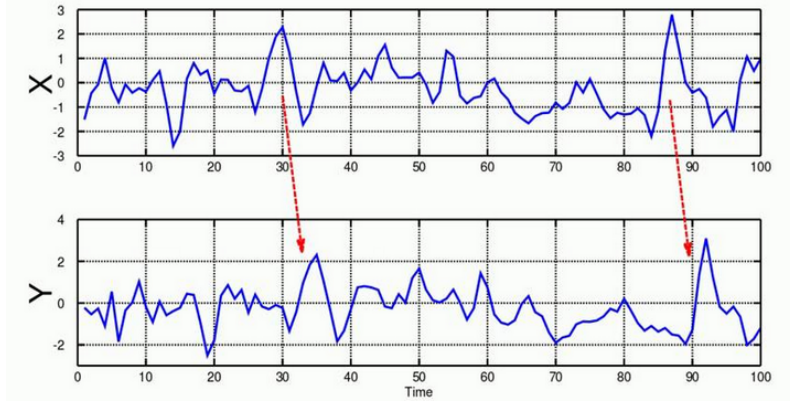


Figure 18: *Example of signal X Granger-predicting signal Y*

As already mentioned in Section 0.1.5, observing a correlation between two time series is not enough to establish a *relationship* between them. Indeed, a further indicator of the synchronization between signals is the presence of an underlying *cause-effect* mechanism. If such relationship is found, besides observing a synchronization, one can determine also which signal (and, in this case, mouse) is responsible for causing the opponent's one.

In situations like the EEG experiment in [Novembre et al.] discussed in Chapter 1, such relationship is investigated through a manipulations on the experimental level (via *Multi-brain stimulation*). In the current work, once dealing with data already measured, a different and sophisticated approach is adopted, based on statistical principles: the **Granger causality**.

The Granger causality (**G-causality**) is a method aimed to identify causal relationship between time series data. The word "casuality" is mainly due to historical reasons, since many debates about its correctness are still ongoing, and it would be probably more precise to refer to it as "Granger *prediction*".

Given two time series $\mathbf{x} = \{x_i\}_{i=1}^N$ and $\mathbf{y} = \{y_i\}_{i=1}^N$, the G-causality method is based on the following scheme:

1. Generation of a **vector autoregressive model (VAR)** of order p , where $p \geq 1$ is an integer to be determined, for one of the two time series, considering its previous values

$$\hat{x}_t = a_0 + a_1x_{t-1} + a_2x_{t-2} + \cdots + a_px_{t-p} \quad (7)$$

2. Generation of a second autoregressive model, in which the values of the second series \mathbf{y} are added to 0.7

$$\hat{x}_t = a_0 + a_1x_{t-1} + a_2x_{t-2} + \cdots + a_px_{t-p} + b_1y_{t-1} + \cdots + b_py_{t-p} \quad (8)$$

3. Comparison between the two models: if (0.8) is more significant than (0.7) (in a way to be clarified in the following), then signal \mathbf{y} *Granger-predicts* signal \mathbf{x}

More formally [Barnett-Seth], given two stochastic processes $\mathbf{X} = \{X\}_{i=1}^N$ and $\mathbf{Y} = \{Y\}_{i=1}^N$, process Y *does not* G-causes process X if X , conditional its past, is independent by the past of Y . A vector autoregressive model for a process U takes the form

$$\mathbf{U}_t = \sum_{k=1}^p A_k \mathbf{U}_{t-k} + \varepsilon_t \quad (9)$$

where p is the **order** of the model, $\{A_k\}_{k=1}^p$ are the **regression coefficients** and ε_t the **residuals**, assumed normally and independently distributed. The **residual covariance matrix** of the model is defined as $\Sigma = \text{Cov}(\varepsilon_t)$ and it is assumed to be stationary. The process U can then be identified both as X and Y . Given a VAR model of the form (6), the **autocovariance sequence** $\{\Gamma_k\}_{k=1}^p$ is defined as $\Gamma_k = \text{Cov}(\mathbf{U}_t, \mathbf{U}_{t-k})$, and it is possible to relate this quantity to the autoregression coefficients $\{A_k\}$ thanks to the **Yule-Walker** equations [Anderson,1971]

$$\Gamma_k = \sum_{i=1}^p A_i \Gamma_{k-i} + \delta \Sigma \quad k = 1, \dots, p \quad (10)$$

Standard VAR theory [Hamilton-Lutkepohl] requires the condition $\sum_{k=1}^N \|A_k\|^2 < \infty$. Moreover, defining the **characteristic polynomial** as

$$\phi_A(z) = \det \left(I - \sum_{k=1}^p A_k z^k \right)$$

it must be that the **spectral radius** $\rho(A) := \max_{\phi_A(z)=0} |z|^{-1}$ is strictly less than 1, as a *stability* condition.

Considering now a process in which $\mathbf{U}_t = \begin{bmatrix} \mathbf{X}_t \\ \mathbf{Y}_t \end{bmatrix}$, its VAR formulation reads

$$\mathbf{U}_t = \sum_{k=1}^p \begin{bmatrix} A_{xx,k} & A_{xy,k} \\ A_{yx,k} & A_{yy,k} \end{bmatrix} \mathbf{U}_{t-k} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \quad (11)$$

and its residual covariance is $\Sigma = Cov \left(\begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \right) = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}$.

This augmented formulation contains both the regression models for process X and Y . For example, its first component reads

$$\mathbf{X}_t = \sum_{k=1}^p A_{xx,k} \mathbf{X}_{t-k} + \sum_{k=1}^p A_{xy,k} \mathbf{Y}_{t-k} + \varepsilon_{x,t} \quad (12)$$

If the process Y does not G-cause the process X , it follows that the coefficients $\{A_{xy,k}\}_{k=1}^p$ are all equal to 0, and the model becomes

$$\mathbf{X}_t = \sum_{k=1}^p A'_{xx,k} \mathbf{X}_{t-k} + \varepsilon'_{x,t} \quad (13)$$

Therefore, a statistic test checking the null hypothesis $\{H_0 : Y \text{ does not G-predicts } X\}$ has the form

$$H_0 : A_{xy,1} = A_{xy,2} = \dots = A_{xy,p} = 0 \quad (14)$$

If $\Sigma'_{xx} = Cov(\varepsilon'_{x,t})$ is the residual covariance matrix of model (0.13), standard theory [Edwards, 1992] suggests the use of the **(log-)likelihood statistics** to obtain a *maximum-likelihood* estimator of the G-causality between Y and X (here referred as $\mathcal{F}_{Y \rightarrow X}$):

$$\mathcal{F}_{Y \rightarrow X} = \ln \frac{\det(\Sigma'_{xx})}{\det(\Sigma_{xx})} \quad (15)$$

Since the determinant of a covariance matrix (i.e. the *generalized variance*) quantifies the *prediction error* of its regression model, the interpretation of (0.15) is that the G-causality statistics $\mathcal{F}_{Y \rightarrow X}$ is a measure of how much the prediction error is reduced when also the process Y is included in the regression model. Clearly, this same procedure applies to the statistics $\mathcal{F}_{X \rightarrow Y}$, in which the directionality of the relationship is inverted. It can be proven [Wilks & Wald] that, under the null hypothesis, $(N - p)\mathcal{F}_{Y \rightarrow X} \sim \chi^2(d)$, where $d = pN^2$.

To summarize, the typical workflow for G-causality estimation, consists in the following steps:

1. Estimate the model order p via appropriate criterion (such as AIC and BIC)

2. Estimate the autocovariance sequence Γ_k and the VAR coefficients (A_k, Σ) through the Yule-Walker equations (0.10), both for reduced and augmented models. Verify that $\sum_{k=1}^N \|A_k\|^2 < \infty$ and $\rho(A) < 1$
3. Computation of the G-statistics $\mathcal{F}_{Y \rightarrow X}$ and $\mathcal{F}_{X \rightarrow Y}$ through (0.15)
4. Test the significance of the statistical tests against the null hypothesis, computing the p -values of the two tests

3 Interbrain data analysis: emotion discrimination task

In this second chapter, the main results on the data analysis for the emotion discrimination task are presented. After describing the phases of the task and trying to understand which could be an appropriate way to represent the data through normalizations, first considerations on the neural activity of mice, in relation to their behaviour, are taken into account. Then, the focus goes to the main part of the analysis: the *interbrain synchrony* among neural activities. The different concepts of synchrony introduced in Chapter 2 have been investigated, i.e. the cross-correlation between the mean activities of the mice and the peak synchronization among neurons. Finally, the topic of causality between signal is dealt using the tool of Granger prediction. All the results refer to two versions of the emotion discrimination task: a standard task, where the observer is in a *neutral condition*, and the **self-experience task**, in which the observer has been stressed before the task.

3.1 Emotion discrimination task

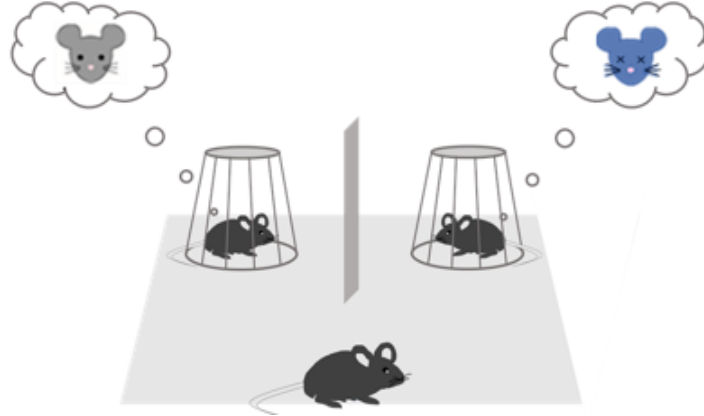


Figure 19: *Scheme of the emotion discrimination task.*

In the *emotion discrimination task*, an **observer** mouse is facing two **demonstrators** in an open arena. One demonstrator is in a *neutral* state, the other in a *stressed* state, meaning that it has been subjected to a *stress protocol* (forced restraint) before the test. The task consists of three main parts:

1. **Homecage restraint**: the three mice are kept separate in their

cages in normal conditions.

—→ *Duration*: 5 minutes

2. **Habituation**: the observer mouse is free to move in an empty open arena; in the meantime the neutral demonstrator is kept in the cage, while the stressed demonstrator is being subjected to the stress procedure.

—→ *Duration*: 15 minutes

3. **Test**: the demonstrators join the observer in the arena. Only the observer is free to move, since the two demonstrators are kept behind a cage allowing sight, sniffing but no direct contact.

—→ *Duration*: 15 minutes

During all the three phases, the neural activities of the three mice are recorded via microendoscopic calcium imaging, as described in Section 0.1.3, in which the target is the set of *somatostating-expressing* interneurons in the anterior cingulate cortex (ACC, Section 0.1). Moreover, the position of the observer mouse along time is recorded during habituation and test phases, and TTL signals (binary recordings of events) of the reciprocal sniffing between one observer and a demonstrators have been recorded during the test. Finally, spatial data of the neurons in the *region of interest (ROI)* captured by the miniscope, are available as well, such as their position and size.

These type of data can lead to several analyses:

- Inspecting if there exist correlations among the neural activity levels of one mouse and what is happening during the task, such as the stressing procedure, proximity or sniffing between two mice
- Testing the presence of *activity synchronization* among the neural signals of two mice, as described in Sections 2.1-2.4, in relation to what is happening in the test (proximity of two subjects, sniffing)
- Investigating whether the activity in one mouse is *predicting* the activity in another, for example via Granger causality analysis introduced in Section 2.5
- Investigating whether the neuronal firing in single mice seems to follow *patterns*, or if it is possible to predict mathematically such chain of events (problem dealt in Chapter 7)

The analyses involve two types of data: in the first one, a neutral observer interacts with the demonstrators, while the second one investigate the so

called **self-experience** setting, in which, before the test, also the observer is subjected to a stress protocol. For each of the two cases, two triplets of mice have been tested with the same procedure, and the results have been averaged among them.

It is worth to observe that the phases of homecage and habituation play the role of a *control* in the data analysis: given a quantity measured from the data of calcium concentrations (such as the level of synchronization), we can assume that a result is significant, and thus related to the interaction between mice, if it is present in the test phase but not in the control phases of homecage and habituation, in which the mice cannot interact and are located in different places.

3.2 Normalization of the dataset

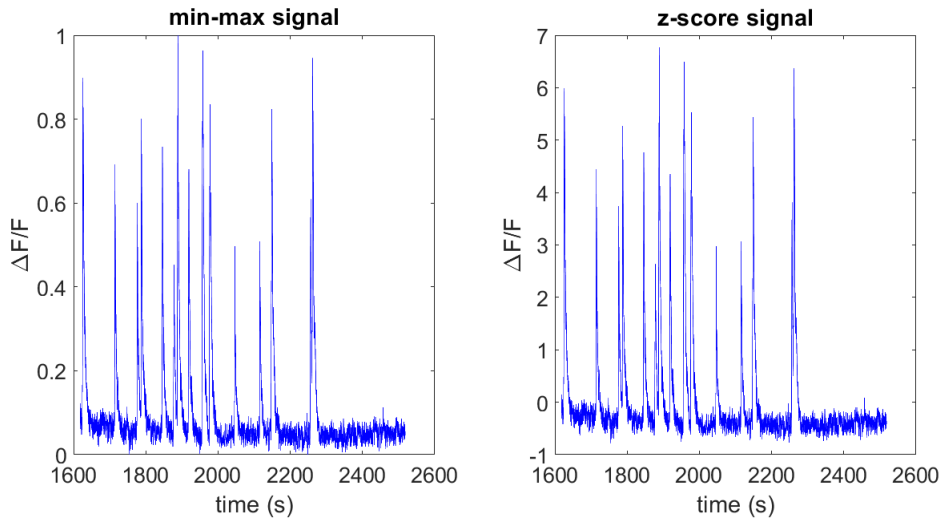


Figure 20: *Different normalizations of the signal. Left: min-max normalization, in which the signal assumes values in $[0, 1]$. Right: z-score normalization, with no constraint on the scale of the y-axis.*

The main object of the analysis is given by the calcium signals recorded in every single neurons, expressed in terms of relative fluorescence $\frac{\Delta F}{F}$. Such signals can be subjected to operations and comparisons, either among neurons of a single mouse, or among activities in different mice. based on this scenario, it necessarilly follows that it is fundamental to choose a proper *normalization* of the data which can allow such processes. The literature proposes several ways to treat data, (see Figure 20):

1. **Raw approach.** The raw data, coming from the Inscopix pre-processing described in Section 1.4, are usually not ready to be analyzed yet. Although porcesses of normalizations on the ROI videos, filtering and noise detection have been already performed, the single information of a fluorescence value $\frac{\Delta F}{F}$ needs to be treated carefully. Every single neuron is characterized by a baseline activity, alternated with rapid and huge spikes. Such baseline value, which biologically represents the *non activity* of a neuron, can take different values across different neurons. Moreover, the values reached by the spikes can appear in different magnitudes across different neurons, but this does not necessarily reflect what is really happening in the reality: a spike may exhibit smaller amplitudes for the simple fact that less GCaMP protein, which reacts with calcium and show flourescence, was present at that moment, and consequently the spiking of that neuron assumed smaller values.
2. **Min-max normalization.** With the min-max normalization, every value x_t of the signal $x = \{x_t\}_{t=1}^N$ at time t is normalized through

$$x_t^{mm} = \frac{x_t - \min(x)}{\max(x) - \min(x)} \quad (16)$$

With the min-max normalization, all the signals are bounded in the interval $[0, 1]$, where the values 0 and 1 occur in correspondence of the lowest and highest values of the signal, respectively. Therefore, using this normalization, all the signals are treated in the same way, since they are all in the same interval, whatever their original value and scale was in the raw dataset.

3. **Z-score normalization.** For a value x_t of the signal $x = \{x_t\}_{t=1}^N$, the z-score normalization reads.

$$x_t^z = \frac{x_t - \mu}{\sigma} \quad (17)$$

where μ and σ are the mean and standard deviation of the signal, respectively. The goal of this normalization is to obtain signals with 0 mean and unitary standard deviation, making them comparable without constrains on their values, as in the case of the min-max normalization. For this reason, z-score is the most used type of normalization in the literature.

All the different normalizations do not modify the shape of the signals, but their *scale*. As for this work, the choice went on a modifcacion on

the z -score normalization, defined as **z -score on baseline**, in which signals are aligned and then divided by their standard deviations. Therefore, the proposed approach belongs to the z -score case, but the alignment is not performed on the mean of each signal, but on their baseline activities. The reason for this is that two signals coming from neurons, when subtracted by their own mean, can still exhibit different baselines. Therefore, with the chosen normalization the alignment is done on such baselines, which are set to 0 for every neuron, and operations on signals such as sum or direct comparison, can become meaningful.

3.3 Behavioural analysis of the task

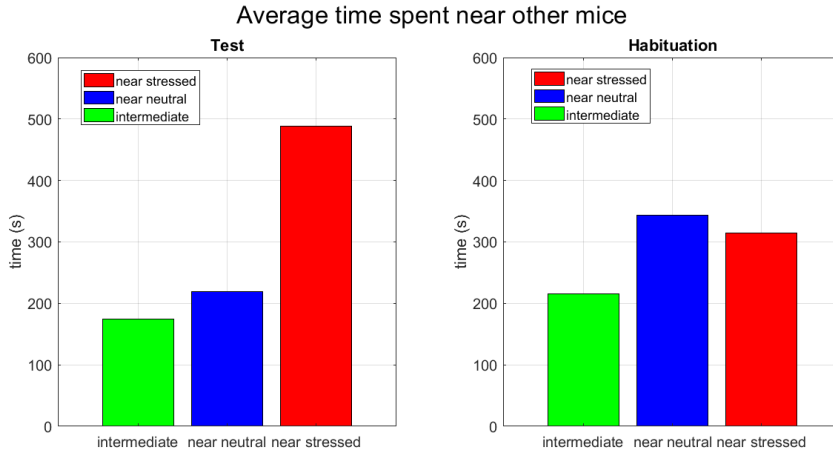


Figure 21: *Average time spent by the observer near the demonstrator cages, or in the intermediate area. Left: results during the test. Right: results during the habituation.*

In [Scheggia-Managò], the same neuronal population, namely the somatostin interneurons, was observed during a task following a protocol similar to the current emotion discrimination task. In that case, however, the area of interest was the medial prefrontal cortex, and the main focus on the electrophysiological activity. As results, it was evident that the observer mouse spent more time near the stressed demonstrators, and its neural activity (in terms of action potentials) showed higher values during those periods.

In this work, we obtained the following results:

- The observer spent a significantly higher time near the stressed demonstrators, rather than near the neutral one or in an intermediate zone

(see Figure 21). Moreover, in order to rule out the possibility that this effect was due to spatial factors present in the cage where the stressed mouse was located, a comparison with the habituation has been performed. During the habituation, the observer is free to move in the arena for the same time as in the test, but in the arena the two cages for the demonstrators are empty. The results show clearly that in this situation the observer doesn't show a particular tendency for proximity, and in particular spends considerably less time near the cage of the stressed mouse.

- In the analysis of the calcium activity, no evidence of correlation between activity and proximity to demonstrators seemed to emerge during the test: from the recorded neurons, the observer maintains an overall similar average activity through the different areas of the arena. As for the demonstrators, only the neutral appears to show an increase activity during the sniffing with the observer (Figure 22).

Next, a comparison on the average activity of mice during the phases of the task have been taken into account (see Figure 23). The results showed that, in general, the mean activity seems to be higher during the habituation. This happens in particular for the stressed mouse, coherently with the fact that during the habituation, it is subjected to the stress protocol (forced restraintment), for which it is expected to show higher activity. Overall, the neutral mouse shows smaller values of calcium activity, during both habituation and test, compared to the observer and the stressed conspecifics.

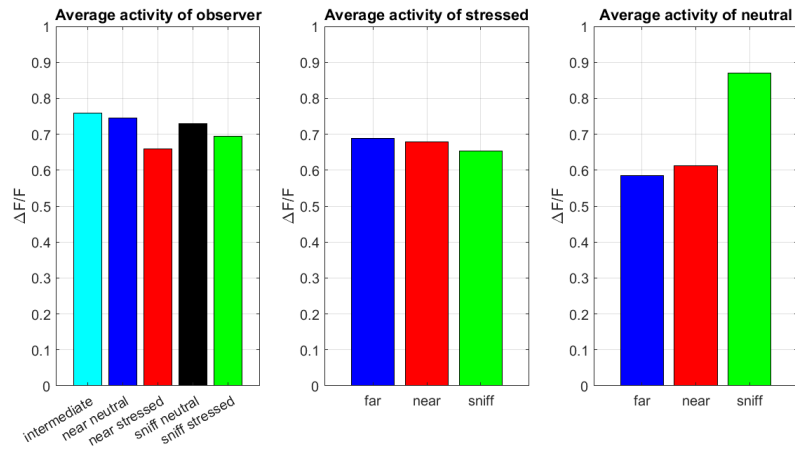


Figure 22: *Average activity showed by the mice during the test. Left: observer. No tendencies seem to emerge when it is in intermediate area, close to demonstrators or sniffing with demonstrators. Center: stressed. Right: neutral.*

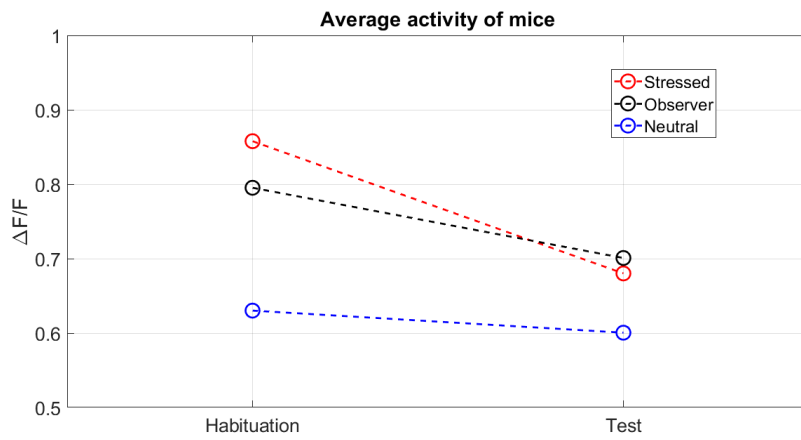


Figure 23: *Evolution of the average activity of the three mice from the habituation to the test*

3.4 Cross-correlation analysis for the mean activity

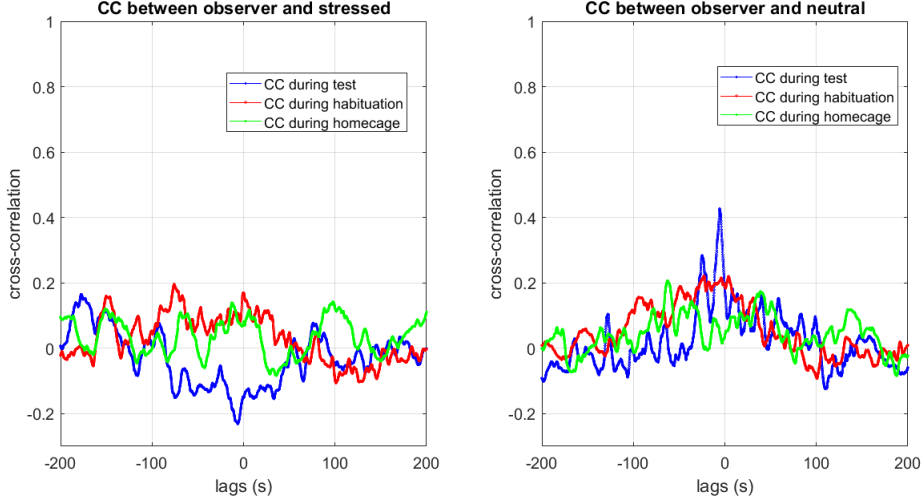


Figure 24: *Left: average cross-correlation between observer and stressed. Right: average cross-correlation between observer and neutral.*

With the *Interbrain analysis*, we study whether the interactions between two mice are followed by a synchronization of their neural activities. As explained in Section 2.2, one of the main tools for the synchronization analysis is the **cross-correlation**.

As first step, following the approach of [Kingsbury], every mouse has been gifted with its *overall activity*, consisting in a mean of the activities of all its recorded neurons. The spatial data on the position of the observer in the arena during habituation and test, made it possible to divide such mean activities in several parts::

1. Restriction to the times of proximity between the observer and the stressed
2. Restriction to the times of proximity between the observer and the neutral
3. Restriction to the times of reciprocal sniffing between two mice

The above subdivision allows us to analyze the correlation among the activities of two mice when they are actually close to each other. In this context, a cross-correlation analysis on the mean activity has been performed,

comparing what happens during the homecage, habituation and test. During homecage and habituation phases, the three mice are kept separate. Therefore, inspecting the correlation among the activities during these phases is a useful *control*, i.e. a case of comparison in order to see what to expect when correlation should not be present.

After computing the mean of all the $\frac{\Delta F}{F}$ signals in each neuron for every mouse, and applying the above classification 1. – 3., the cross correlation between observer and stressed and between observer and neutral mean activities has been computed, as described in (ref chap 2). As shown in Figure 24, analyzing the average cross correlation between the two experiments, one can notice the presence of a peak around $lag = 0$ for the pair observer-neutral in the test phase. This tendency is not present in the control cases of habituation and homecage, making it a proper feature of the test. On the other side, the cross correlation between observer and stressed does not show any particular difference with respect to the homecage and habituation phases. A further verification has been given by observing the cross correlation between the observer and one demonstrator considering the periods when the observer was visiting the opposite demonstrator. As a result, the cross correlation between observer and neutral is inhibited considering the instants when the observer was visiting the stressed, and the cross correlation between observer and stressed remains non significant when the observer is visiting the neutral (Figure 25).

Overall, these results may show the presence of a synchronization between neural activities of the observer and the neutral demonstrator, but not between the observer and the stressed one, and this effect is present only during the test and during the proximity of the two mice.

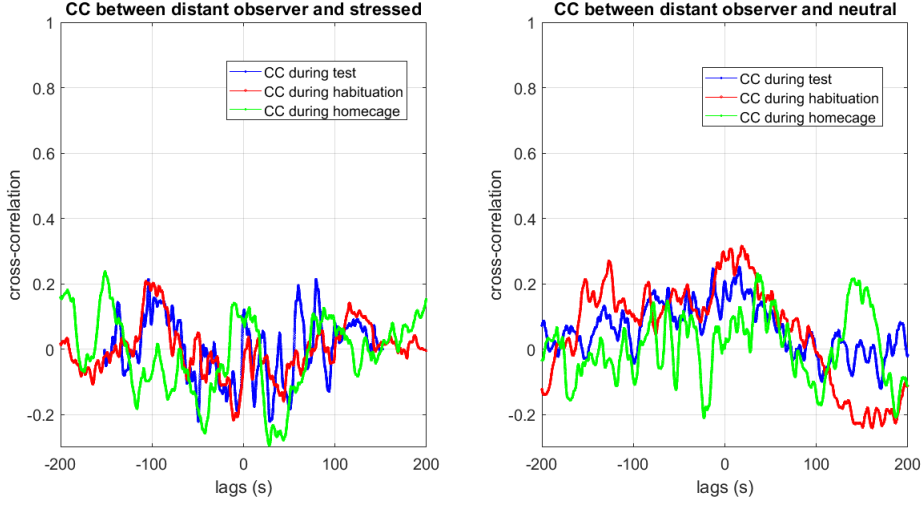


Figure 25: *Left: average cross-correlation peaks between observer and stressed when the observer is visiting the neutral. Right: average cross-correlation peaks between observer and neutral when the observer is visiting stressed.*

Next, two particular cases have been considered: the sniffing interactions and the restriction to the first part of test (first 2 minutes of the test). The sniffing is known to be one of the most direct interactions that two mice can display between each other, and therefore is of a particular interest for what could happen at synchronization level. The restriction to the first two minutes of the test is of interest based on previous results [Kingsbury], which show that the first period of interactions among mice is the one which determines some most relevant neural tendencies, such as the display of synchronization.

The obtained results allowed us to draw the following conclusions (Figure 26):

- Data measurements during the sniffing interactions are coherent with those obtained during the whole test: a peak around $lag = 0$ is present only in the cross correlation computed between observer and neutral, but not between observer and stressed
- Consistent with previous results, the peak of cross-correlation between observer and neutral is higher if we restrict the analysis to the first two minutes of interaction.

A summary of the cross-correlation results can be seen in Figure 27. For every case, the values refer to the cross-correlation peak recorded around

$lag = 0$. While for the observer-stressed case no significant differences arise, between observer and neutral the test phase shows higher correlations, in particular in the sniffing and at the beginning of the test.

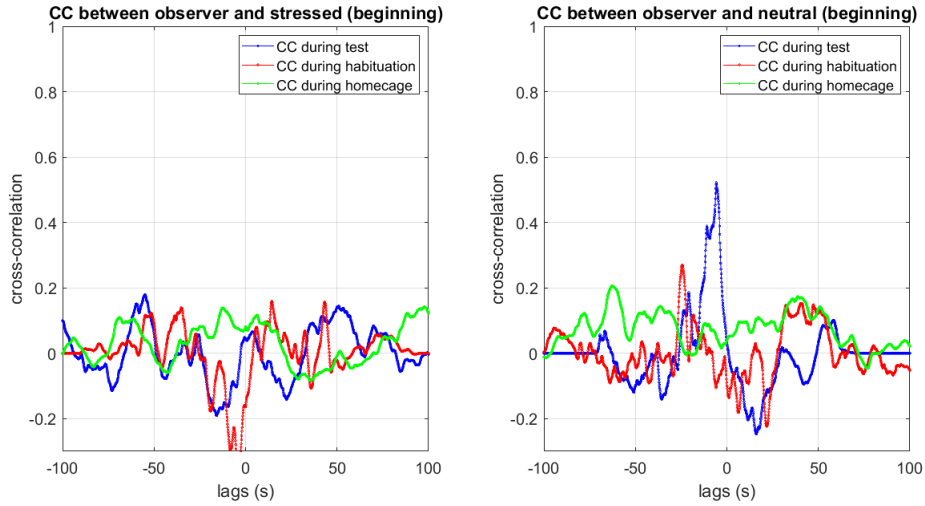


Figure 26: *Left: average cross-correlation between observer and stressed, during the first two minutes of test. Right: average cross-correlation between observer and neutral, during the first two minutes of test.*

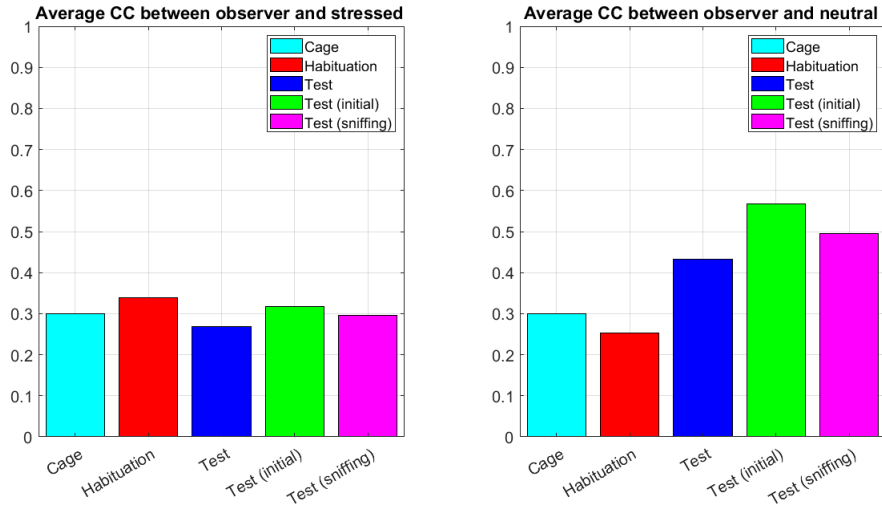


Figure 27: *Left: average cross-correlation peaks between observer and stressed. Right: average cross-correlation peaks between observer and neutral.*

3.5 Peak synchronization analysis

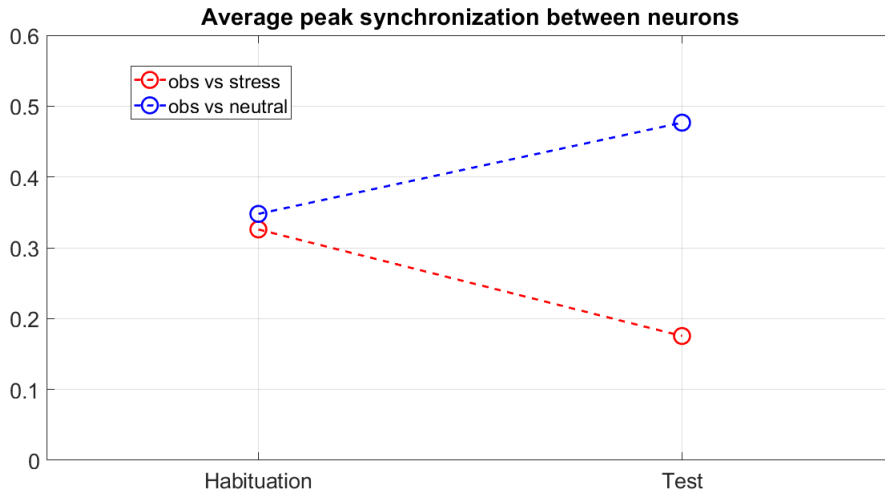


Figure 28: *Evolution of the average peak synchronization index from the habituation to the test, for the cases observer-stressed and observer-neutral.*

From the mean activity of the mice, the focus goes now to single neuron synchronization. As presented in Section 0.2.3, a useful tool to quantify the degree of simultaneous firing between two neurons is given by the **peak correlation index** 0.6, which takes into account the number of simultaneous peaks observed during a synchronization time window dT between two signals. As for the present work, the chosen time window, in which two peaks are considered simultaneous, has been set to $dT = 1s$. The analysis of peak correlation proceeds according the following steps:

1. Computaton of the peak correlation index for every pair of neurons between the observer and one demonstrator
2. Computation of the average of the indices for a couple of mice during the test, to be confronted with the same average obtained during the habituation. In this way we can compare the change in simultaneous firing between two neurons from one phase of the task to the other
3. The result on the evolution of the mean correlation index is accompanied by a *one-way ANOVA* test [Book] on the significance of the difference between the two groups of indices. The null hypothesis of the ANOVA states that there is no significance in the difference between the two pupulations, i.e. that $\mu_1 = \mu_2$, μ_1 and μ_2 being the means of the two populations. Such hypothesis is *rejected* for low p -values of the associated test.

The obtained results are showed in Figure 28. Consistent with the cross-correlation results for the mean activity, the average peak correlation index shows higher values in the test, rather than in the habituation, only for the observer-neutral pair, while for the observer-stressed pair it actually decreases. This average result is consistent between all the two investigated datasets, and in both cases the difference between the indices computed in the test and the ones computed in the habituation is significant (p -values of $p = 0.0025$ and $p = 0.0128$ for the two datasets from the ANOVA test)

3.6 Granger causality analysis

Granger analysis has been performed in accordance to Section 2.5, to investigate whether the activity of one mouse is Granger-predicting the activity in another. This analysis consists in two statistical tests with null hypothesis

$$H_0 : Y \text{ does not G-predicts } X$$

where in this case X and Y represent the mean activities of the observer and one demonstrator, respectively. In the two analyzed datasets, the results of the tests, in terms of p -values, are summarized in Table 29. Some considerations are in order:

- In the first dataset, there is significance to conclude that the observer activity is G-predicting the demonstrator ones (both stressed and neutral), while in the second dataset there is evidence only for the prediction from observer to stressed
- In all cases, it seems that, when the observer is visiting a demonstrator, the same demonstrator is influenced by the observer, which G-predicts its activity, i.e. the *directionality* of the prediction goes from the observer to the demonstrator
- The only coherent result is the one involving the stressed demonstrator, whose activity is always G-predicted by the observer one. However, more data should be available in order to draw more definitive conclusions

Dataset	Directionality	p-value
First	observer \rightarrow stressed	$p = 0.02$
First	observer \rightarrow neutral	$p = 0.005$
Second	observer \rightarrow stressed	$p = 0.01$

Figure 29: Results of the Granger analysis.

3.7 Self-experience task

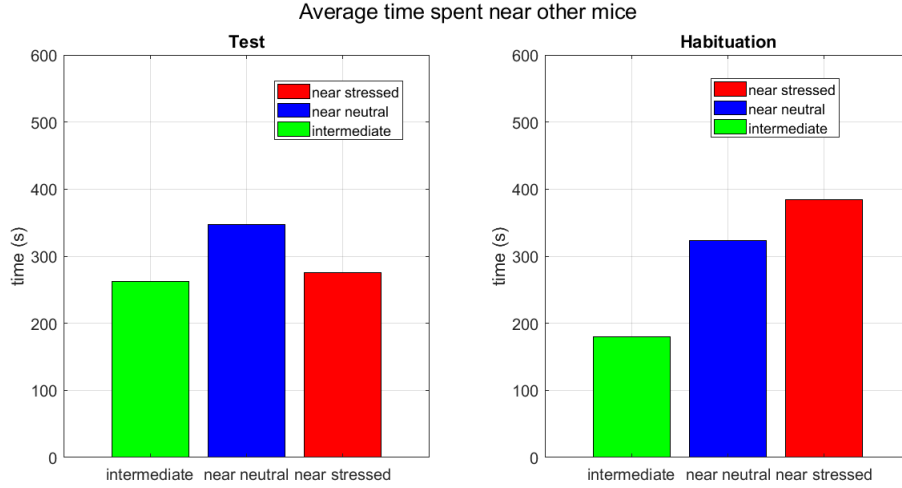


Figure 30: *Average time spent by the observer near the demonstrator cages, or in the intermediate area, for the self-experience task. Left: results during the test. Right: results during the habituation.*

Two more datasets of the emotion discrimination task have been analyzed. This time, however, they refer to the **self-experience** configuration, in which the observer has been stressed before the task. The same analyses with the same data normalization as in the standard configuration has been performed. The obtained results are the following:

- The observer spends, on average, less time near the stressed, compared to the standard case of non-stressed observer (see Figure 30). This is consistent with an *avoidance* behaviour between two stressed individuals, already observed in past studies [REF?]
- As for the average activity displayed by the three mice, no relevant differences seem to emerge with respect to the standard case. Once again, the strongest activity is observed during the sniffing between observer and neutral (Figure 31)

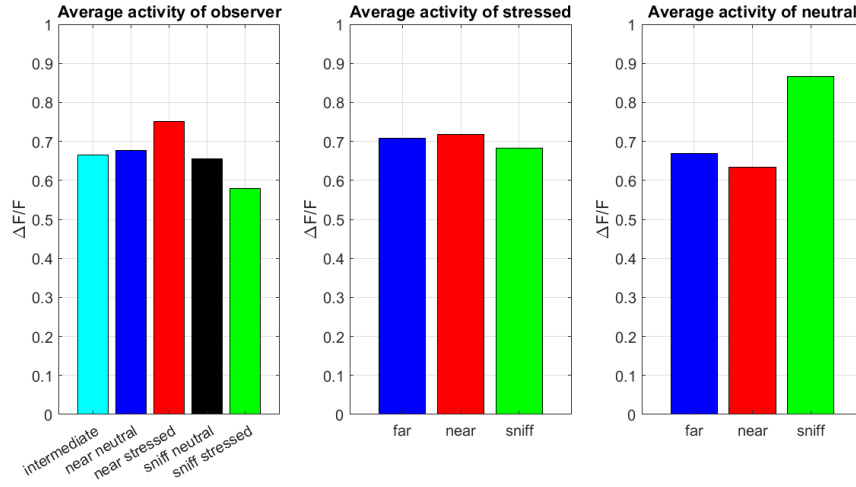


Figure 31: *Average activity showed by the mice in the self-experience task. Left: observer. Center: stressed. Right: neutral.*

- The cross-correlation peak, observed in the previous case between observer and neutral, now is not present. For both analyzed datasets of self-experience type, no correlation seems to emerge between any pair of mice, not even considering the first two minutes of the task (Figure 32)

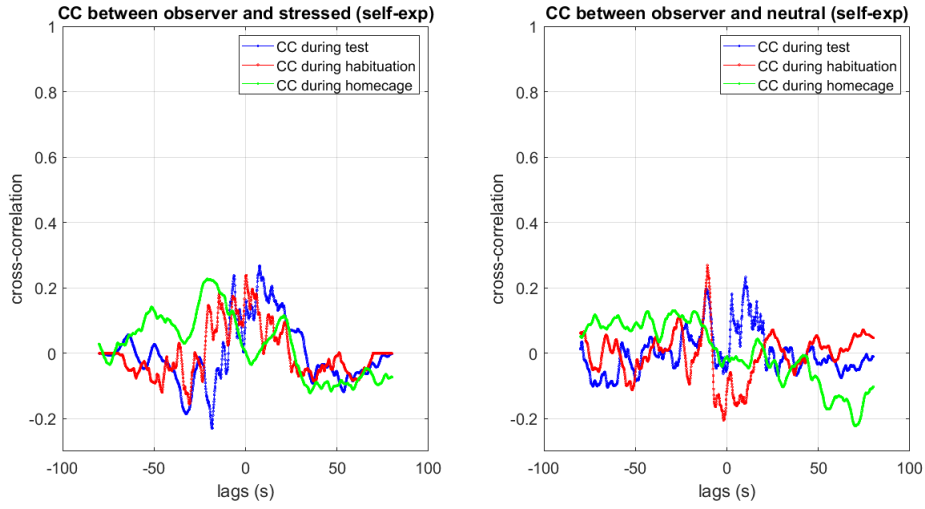


Figure 32: *Average cross-correlation, during the first two minutes, in the self-experience task. Left: between observer and stressed. Right: between observer and neutral.*

- The average correlation index slightly decreases in the observer-neutral case and increases in the observer-stressed. These results, however, do not show significance to the ANOVA test (Figure 33)

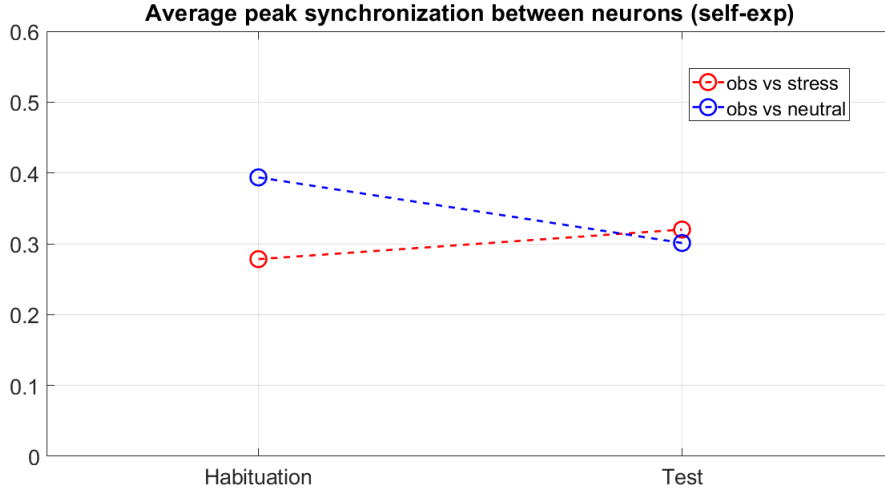


Figure 33: *Evolution of the average peak synchronization index from the habituation to the test, for the observer-stressed pairs and observer-neutral, in the self-experience task.*

- Concerning the Granger analysis, in both the analyzed datasets, no significance of Granger prediction in any directionality has been found, both for observer-neutral for observer-stressed pairs

3.8 Conclusions on the emotion discrimination task

The analyses on the available data (two experiments performed in the standard configuration, two in the self-experience configuration), suggest the manifestation of some unified tendency in the results, and allow us to make some conclusive considerations:

- The relevant factor explaining the onset or inhibition of the synchronization among neural activities of SOM+ interneurons in the ACC, seems to be the *stress condition*. Indeed, in the standard emotion discrimination task, a higher synchronization of mean neural activities appears between observer and neutral mice, but not observer and stressed, such as in the peak synchronization between neuron pairs. On the other side, when the observer itself is in a stress condition, also the results with the neutral demonstrator are not present, and the other analyses are less significant. This indicates that a stressed mouse may lose the ability to synchronize

- In order to have credibility, such conclusions need to be repeated in more experiments to be analyzed, since the consistency observed so far only refer to a small dataset
- In any event, the conclusions do not apply to the whole brain activity of mice. Rather, the conducted analyses consider a very specific neuronal population in a very specific area (SOM+ neurons in the ACC). It is possible, and actually plausible, that different areas and different neurons show a different response in activity synchronization

4 Activity of amygdala in altruistic decision making

In this chapter, a new aspect of the behavioural analysis from *in vivo* tasks on mice will be investigated: the ability of mice to display selfish or prosocial choices towards other individuals.

Through the *altruism task*, two mice are put in a situation in which they can choose whether to perform an altruistic or selfish choice, while having their neural activity recorded. In this context, however, the employed technique for calcium imaging is not microendoscopic calcium imaging, but *Fiberphotometry* (discussed in Section 0.1.4), from which an overall signal for the calcium activity is obtained.

As for the area of the brain under investigation, in this case the focus is on the **basolateral amygdala (BLA)**, which has been shown to be connected with prosocial choices in rodents (as anticipated in Section 0.1).

After describing the results of the altruism task in terms of behaviour, the focus will concern with the analysis of the calcium tracks from fiberphotometry, and in particular their relative change with respect to a baseline activity, when making prosocial or selfish choices.

4.1 Social decision making in mammals

In order to live as a group, mammals have to perform decisions based on their *social interactions*. Often, such decisions are strictly egoistic and selfish, in order to survive as a single individual and obtain the best for itself. However, evolution taught us that **altruistic** decisions, engaged with the purpose of creating a benefit for the entire community, are present as well, and are actually necessary in order to survive in social groups [Batson, C. D.].

The reasons behind whether a social choice may be of altruistic rather than selfish type are often nontrivial and may depend on several factors, such as the presence of a *dominance* relation inside a group, the degree of closeness between two individuals, or biological factors like age and gender.

In the first years of study of social interactions and decision making, the term *altruism* was used only for humans; however, along the years, the increase of discoveries of similar attributes in other animals changed the situation. For example, frequent habits of sharing food have been observed in parrots [Brucks, D. & von Bayern] and monkeys [Krupenye, C., Tan, J. & Hare]. Mice do not differ: rodents have been shown to display altruistic decision making such as consolatory and collaborative behaviours, or help to conspe-

cific in case of need [Bartal, I. B.-A., Decety, J. & Mason].

Moreover, a dysfunctional behaviour in social decision making, such as a lack of empathy, could be strictly related to many psycho-pathologic conditions such as schizophrenia, autism, Alzheimer’s disease or dementia.

Through the years, the interest in observation of social decision making in groups of mammals has been growing. However, very often the focus of such studies is limited to the *expressed* behaviour, rather than the characterization of the underlying neural causes. Recently, studies on primates [Dal Monte, O., Chu, C. C. J., Fagan] seem to have identified the neural circuits responsible for social interactions and decision making in the basolateral amygdala (BLA), and following studies on mice confirmed it as well [Felix-Ortiz, A. C., Burgos-Robles, A., Bhagat].

4.2 Description of the altruism task

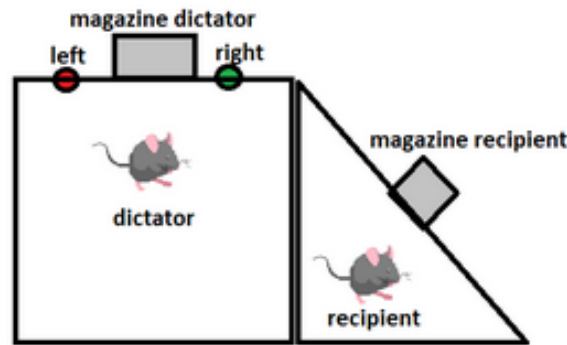


Figure 34: *Main setting of the altruism task: a dictator mouse (left) can choose between two pokes, one of which will deliver a food pellet to a recipient mouse (right)*

The *altruism task* is a task in which two mice are tested in their social decision making. It consists in two mice located in adjacent compartments, separated by a metal mesh which allows them to see and sniff each other, but not enter in direct contact. One mouse, called the *recipient*, assumes only a passive role: it stays in its space and, when a food delivery happens from an apposite box (called *magazine*) it can eat the delivered food pellet. In contrast, the other mouse, called *dictator*, has to perform a social choice. Indeed, in any case a food delivery to its own magazine will happen by poking

one of two buttons available, but even if both types of pokes will provoke a pellet delivery for the dictator, only one of them will give food to the recipient as well. It follows that the two choices of the dictator can be classified as *selfish* and *altruistic* decisions.

In order to stimulate the seeking for food, mice have been kept at 90% of their standard body weight before the test. The setup is organized as follows:

- Before the tasks were performed, both mice were kept together in same sex pairs for two weeks
- The animals were tested for 5 days. The first day involves a *learning* process for the dictator to understand the food delivery mechanism. The main test is considered to be during the following 3 days, and finally the process is considered well assimilated by the dictator during the final day
- The position of the altruistic and selfish pokes (i.e. of the two buttons providing or not the food delivery to the recipient) has been randomized through the different days of the test, in order to rule out spatial causes for the results
- A control test has been performed in parallel. Here, the recipient mice were replaced by inanimate objects, in order to determine the effective relevance of their presence in the task
- During the performance of the 5 days of the test, the calcium activity in the basolateral amygdala of the dictator has been recorded using the Fiberphotometry technique, giving as result a collective signal of the BLA activity (as discussed in Section 0.1.4)

4.3 Behavioural results in the altruism task

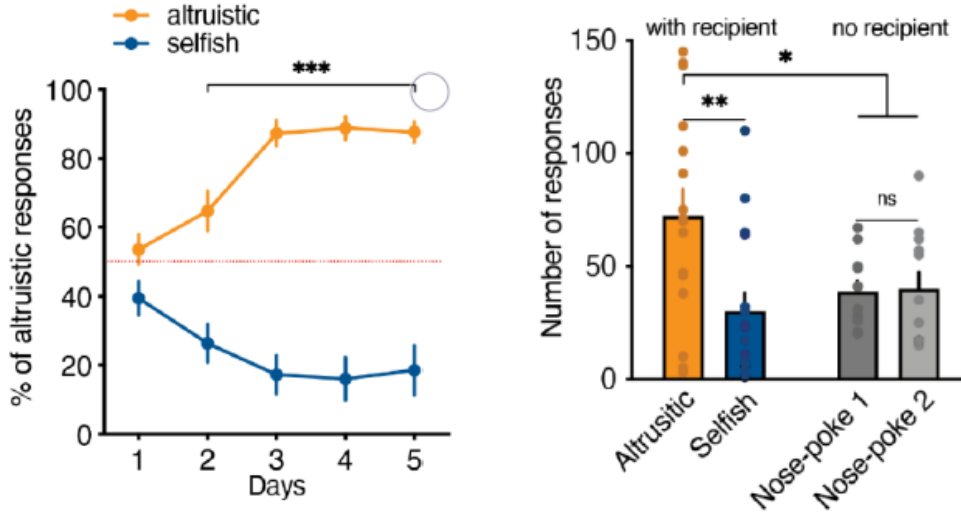


Figure 35: *Left: Percentage of change in altruistic response (i.e in the number of altruistic pokes) for altruistic and selfish mice. Right: difference in number of altruistic and selfish pokes by the dictator when the recipient is present or not*

As shown in Figure 35, the first evident result of the altruism task is that, when a recipient mouse was present and same sex pairs adopted, the dictator mice showed an increase in the number of altruistic pokes during the test days. In contrast, when recipient mice were replaced by inanimate objects, no appreciable difference between altruistic and selfish preferences have been detected. Such differences started to emerge from the second day of the test, as confirmation of a successful learning period on the first day.

An important difference has been observed between male and female mice: only male dictators showed altruism preference, while female ones did not show an appreciable difference between selfish and altruistic pokes.

Next, the spatial exploration of the mice has been under investigation. Results indicate that altruistic dictators (i.e. dictators which showed preference for altruistic pokes) tend to explore more the area in proximity to the recipient, in contrast with selfish dictators which did not show particular interest in their partners.

To test whether visual interaction plays a role determining the altruistic preferences, the task has been repeated replacing the metal mesh with an opaque screen, which still allowed olfactory and auditory stimuli, but not visual ones.

In this situation, mice showed a marked decrease in the altruistic responses, with respect to the previous case in which a metal mesh was used. This implies that the dictator mouse needs to be able to see the recipient partner getting the food delivery, in order to be able to select the altruistic alternative.

Finally, to further determine the ability of the dictator mice to *learn* from experience and change their habits in light of an altruistic choice, dictators were tested with no partner in adjacent cages, to trigger one nose poke over the other. Subsequently, the recipient mouse was put in the near cage, and the location of the altruistic poke, delivering food to it as well, has been assigned to the poke less pressed by the dictator. Despite this, the dictator showed through days a change in its preference towards the altruistic poke. Changing the recipient with an inanimate object, such preference decreased once again. The conclusion is straightforward: the presence of the recipient is the determining factor for the altruistic preference by the dictator.

Once, in normal conditions, the altruistic preference has been established, one may ask under which conditions such preference is maintained. To this purpose, the task was repeated changing the rules of food delivering. Now, two nose pokes were necessary for the food delivery in the altruistic poke, while only one in the selfish one. Even with this additional effort, male and female mice which previously showed altruistic preferences, increased the number of altruistic over selfish preferences also in this case. Increasing the necessary number of pokes to 4, male dictators kept their altruistic preferences, while female dictators did not show anymore a significant difference. Finally, requiring 6 nose-pokes for altruistic delivery still kept male mice in their altruistic preference, while made female ones prefer the selfish one, and only from 8 pokes, also males stopped to be altruistic. Overall, these results showed that *male mice tend to share food with their cagemates partners, even at costly conditions* (Figure 36).

Finally, mice were tested in two more conditions: absence of rewards for both dictator and absence of reward for dictator, but not for the recipient mouse. While, in the first case, no preference between the two pokes was expressed by the dictator, in the second case there was an altruistic preference. This means that the dictator can choose an altruistic decision even if there is no direct benefit for itself.

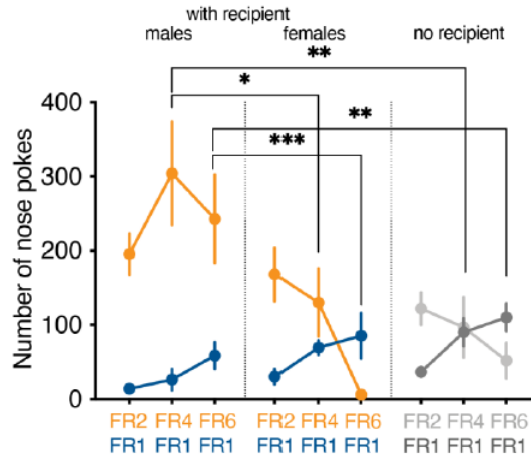


Figure 36: *Number of nose pokes of male and female mice, with and without recipients, changing the number of nose-pokes necessary for a food delivery.*

4.4 The role of social hierarchy in altruism display

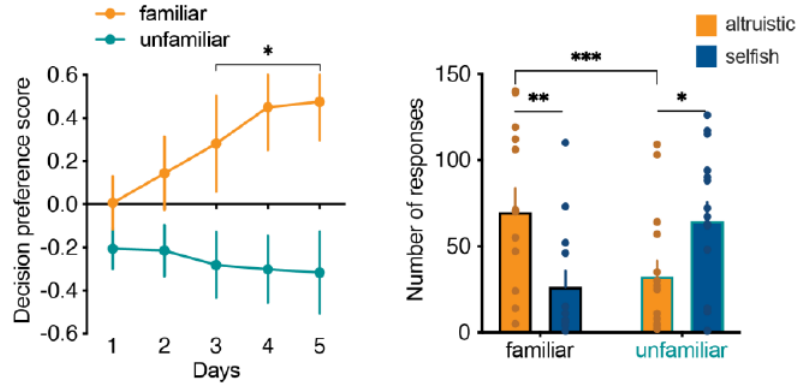


Figure 37: *Left: decision preference score with familiar and unfamiliar recipients. Right: Number of responses by selfish and altruistic dictators, with familiar and unfamiliar recipients*

As of today, it has been recognized that familiarity between individuals plays a key role in social relationships [De Waal, F. B. M. & Preston]. In the case of the altruism task, it is to be expected that the altruism shown by the dictator is strictly depending on the period preceding the test, in which dictator and

recipient have been cagemates for two weeks. To test this hypothesis, the task has been repeated using unfamiliar recipients, showing no preference for altruistic choices, where indeed most of dictator mice showed preference for the selfish pokes. This has been measured through the **decision preference score**

$$DPS = \frac{A - S}{A + S} \quad (18)$$

where A and S are the number of altruistic and selfish responses, respectively.

Next, a *dominant/follower* relationship has been established through the **tube test**, in which two mice have been paired in a tube allowing the passage of just one individual. This test helped to figure out which of the two was the *dominant* individual, i.e. the one pushing the other, and the *follower* individual, i.e. the one retreating. The dominance was quantified through the **Normalized David's score** [De Vries, H., Stevens, J. M. G. & Vervaecke]

$$NDS = \frac{1}{N} \left(DS + \frac{N(N-1)}{2} \right) \quad (19)$$

where N is the number of tested subjects and DS is the David's score, i.e. the ratio of wins.

The analysis of 9 pairs of mice ($N = 18$) yielded the main following results:

1. The dominance relationship is *transitive*: if mouse A is dominant with respect to mouse B, and mouse B is dominant with respect to mouse C, then mouse A dominates also mouse C
2. Dictator mice showing selfish preferences in the altruism task displayed low scores in the NDS compared to their recipients, which means that selfish dictators tend to be less dominant than their corresponding recipients
3. No significant differences have been detected between the dominance scores of altruistic dictators and their recipients
4. The more a dictator has a higher dominance value NDS , the more it is willing to perform altruistic preferences in the altruism task

Next, the study tested the hypothesis that *an increase in altruism is related to an increase in affective state matching between individuals*. To do so, the recipient has been subject to a manipulation of its emotional state,

in this case a fear conditioning paradigm through mild shocks. In the meantime, the dictator mouse (here assuming the role of *observer*) was allowed to see the recipient through a transparent separating wall. The main result was that altruistic mice displayed an increased freezing behaviour during this phase, compared to selfish ones. Since these latter are the dominant ones, this result lead to the conclusion that *dominant mice tend to show more empathic-like behaviours*, from altruism to emotional contagion.

4.5 The role of BLA neurons in selfish and prosocial choices

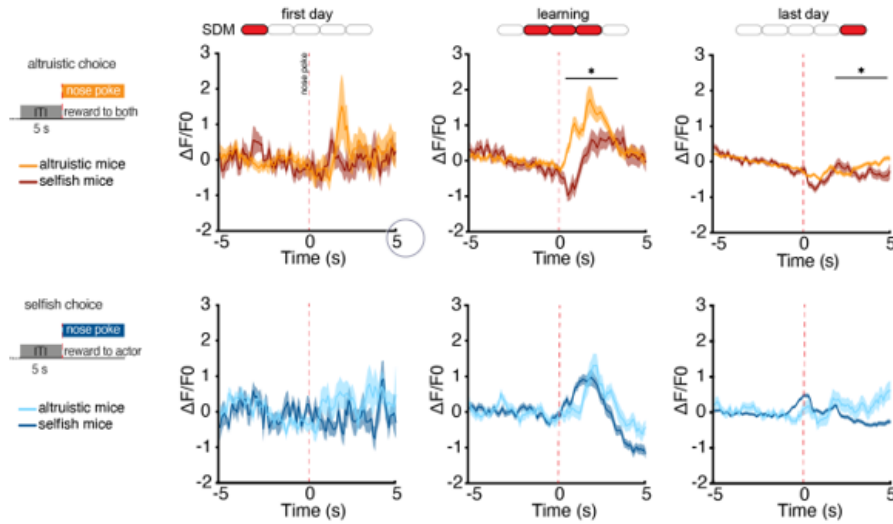


Figure 38: *Top row: PSTH of selfish and altruistic mice during nose pokes in altruistic pokes. Bottom row: PSTH of selfish and altruistic mice during nose pokes in selfish pokes. Results have been averaged on all nose pokes*

The fiberphotometry analysis of the calcium activity in the BLA allowed to establish a connection between the type of choice adopted by dictator mice and their corresponding neural activity. In order to analyze the relative change of such activities during altruistic or selfish pokes, the raw data from the photometry have been subject to a **peristimulus time histogram (PSTH)** analysis. Such analysis goes through the following steps:

1. Set a START and STOP time for the analysis, i.e. set the time window in which considering the variation of the neural activity (in this case $START = STOP = 5\text{ s}$ for a window of 10 s)

2. For every time point x_t of the overall activity given by Fiberphotometry, consider the interval $[x_t - START, x_t + STOP]$
3. Compute the mean μ_t and standard deviation σ_t of the activity in such interval
4. The value of the PSTH for the selected time point is then

$$P_t = \frac{x_t - \mu_t}{\sigma_t}$$

Next, in order to quantify the intensity of the activity, **areas under the curve (AUC)** were estimated numerically using the trapezoid quadrature rule [Quarteroni-Sacco].

The overall analysis on raw photometry data, PSTH tracks and AUC showed some remarkable results (Figure 38):

- Dictator mice showed an increased activity in the BLA when a recipient mouse was present, compared to the case where an inanimate object was placed in the near cage
- During the first and last day, in all cases no significant differences emerged in the activation of BLA
- In the intermediate days of learning, however, altruistic mice showed a strong activation during the nose poking in the altruistic pokes, but not in the selfish ones. Selfish mice, in contrast, did not show any particular activation of their neural activity during both pokes.
- The AUC analysis suggested that the levels of activity were higher during the intermediate days, rather than in the first and last day

To further inspect the role of BLA neurons in empathy display, such neurons have been silenced via a **chemogenetic procedure** [Qi-Gang Zhou, Ashley D Nemes]. In this approach, a viral vector, carrying specific receptors, is injected into the area of interest. Such receptors, called **hM4D receptors**, belong to the group of *inhibitory designer receptors exclusively activated by designer drugs (DREADD)*, which activate themselves only when in contact with a particular drug, the **Clozapine N-oxide (CNO)**. In this way, mice injected with hM4D-CNO will exhibit an *inhibition* of BLA-neurons. In contrast a control group of mice was injected with CNO only, without hM4D receptors.

The obtained results after these injections, displayed in Figure 39, showed a significant reduction of freezing behaviour in the dictators, while their recipient partners were subject to emotion conditioning as previously described. This confirms the conclusions of previous studies [Allsop, S. A. et al], namely *a critical role of the amygdala in emotional matching*. Silencing BLA neurons revealed to be correlated also to the amount of altruism preferences. Indeed, mice with silenced BLA neurons failed to show a marked preference for altruistic choices, in contrast with the control case. Indeed, while approximately 3 out of 4 mice showed altruistic preferences in control conditions, a similar percentage of BLA-silenced mice showed preference for selfish pokes. BLA silencing did not affect the number of responses or the latency of choices, rather it affected the type of the social decision. This is a further evidence of the primary role of BLA in altruism display. Finally, silencing of BLA neurons appeared to reduce also the dominance (Figure 40): a higher number of BLA-silenced mice assumed a subordinate role with respect to control mice. In particular, if transitive relationships of dominance are distributed in levels $\alpha \rightarrow \beta \rightarrow \gamma \rightarrow \delta$, silenced mice belong only to levels γ and δ .

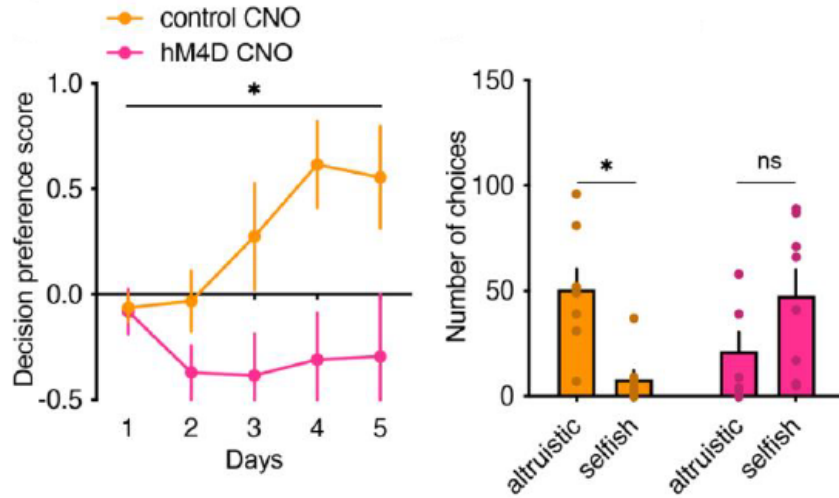


Figure 39: *Left: Decision preference score in BLA-silenced mice (hM4D CNO) and non-silenced mice (control CNO). Right: number of selfish and altruistic pokes in BLA-silenced mice (hM4D CNO) and non-silenced mice (control CNO)*

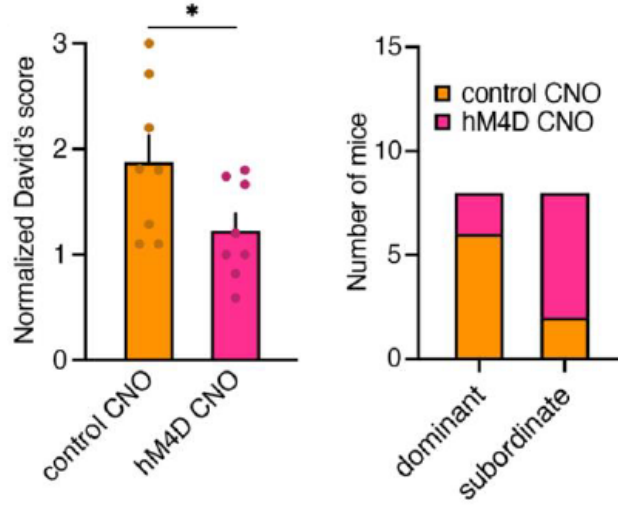


Figure 40: *Left: Normalized David's score in BLA-silenced mice (hM4D CNO) and non-silenced mice (control CNO). Right: number of dominant and subordinate mice in BLA-silenced mice (hM4D CNO) and non-silenced mice (control CNO)*

5 Mathematical modeling in cellular electrophysiology

In Section 0.1, the concept of *excitability* of neurons has been introduced: neurons are able to generate and transmit **electrical impulses** as a response to external stimuli. These impulses are a direct consequence of rapid changes in intracellular and extracellular *ionic concentrations* of main chemicals such as Na^+ , Cl^- , K^+ , Ca^{2+} . The change of such concentration values determines a change in the electric potential difference formed across the cell's membrane, and thus the formation of an **action potential**.

In this chapter, mathematical models to describe cellular electric activity are presented. After introducing the main assumptions of the models, the equivalent circuit formulations for cellular electrophysiology will be presented, along with different ways to describe its electrical components. Finally, the main model to describe the propagation of the action potential between two neurons will be considered, namely the **cable equation model**. All these models consist in either **ordinary differential equations (ODEs)** or **partial differential equations (PDEs)**, to be numerically solved through

appropriate techniques, briefly discussed in the final part of the chapter.

5.1 Electric activity in neurons

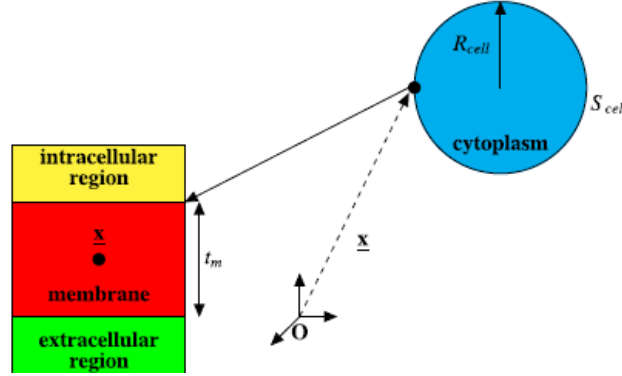


Figure 41: *Schematic model of a cell with a focus on a local point \mathbf{x} of the membrane*

In order to mathematically describe the main electric functionalities of a cell, we first need to understand the set of phenomena which characterize this complex component of our body, such as the processes leading to its condition of *dynamical equilibrium*, called the **cellular homeostasis**. The cellular homeostasis is achieved through the equilibrium of several forces acting on the cell, which can be of fluid, mechanical, chemical or electric nature, although the focus in this work will regard only the electrical part.

The *intracellular* environment of a cell consists mainly in a water solution containing vital nutrients and elements in ionic form such as Na^+ , Cl^- , K^+ , Ca^{2+} , making the solution behaving like a charged fluid. The intracellular environment is separated by the extracellular one by the **membrane** of the cell (see Figure 41). The membrane is formed by a bilayer of lipids (of thickness considerably smaller than the cell's radius) which does not allow passage of substances through it. The passage of some ions is still allowed through specific sites called **ionic channels**, each of one allowing, under certain conditions, the passage of a single ionic species. The movement of one ionic species from the intracellular to the extracellular environment, or viceversa, is caused by the different concentration of such quantity in the two regions, leading to an *electrochemical gradient* which forces positive charges to move from high to low potentials, and thus forming a potential difference across the membrane. We call this quantity the **membrane potential** of the cell.

The main role of the mathematical models for cellular electrophysiology is to describe the evolution of such quantity over time and space.

The particular biology of the cell makes it at the same time a *dielectric* and a *conductor*: the lipid layers of the membrane isolate the intracellular environment from the external environment, and cause the accumulation of charges of opposite sign around the internal and external surfaces of the membrane. This effect is typical of dielectric materials such as capacitors, and it will cause a *capacitive* current across the membrane. This current is not associated with motion of ions, but to the time rate of change of the surface charge on each side of the membrane. On the other side, when an ionic channel is open, there is a passage of ions which makes the cell behaving like a conductor. Both of these contributes need to be taken into account in the modeling of the cell, as done in the next Section.

In order to start modeling these electrical phenomena, let us consider a geometrical setting representing the relevant components of a cell in its electrical activity, with the following assumptions:

1. The cell is modeled as a sphere of radius R_c and surface $S_c = 4\pi R_c^2$
2. The intracellular environment Ω_{in} is separated by the extracellular region Ω_{out} by the membrane, with internal surface Σ_{in} and external surface Σ_{out} .
3. The thickness of the membrane t_m is such that $\eta_c := \frac{t_m}{R_c} \ll 1$. This allow us to neglect it and assume that $\Sigma_{in} \sim \Sigma_{out} \sim \Sigma$
4. The dielectric behaviour of the membrane is expressed by a **membrane capacitance** $c_m = \frac{\varepsilon_m}{t_m}$, where ε_m is the dielectric constant of the membrane

Let us consider now two points $\mathbf{x}' \in \Sigma_{in}$ and $\mathbf{x}'' \in \Sigma_{out}$, as represented in Figure (da mettere). The axis connecting \mathbf{x}' and \mathbf{x}'' is denoted by x and \mathbf{e} is the unit vector oriented from Ω_{in} to Ω_{out} . For any time $t \in [0, T]$, we introduce the electric potentials $\psi^{in}(\mathbf{x}', t)$ and $\psi^{out}(\mathbf{x}'', t)$. The **membrane potential** at $\mathbf{x} \in \Sigma$ is

$$\psi_m(\mathbf{x}, t) = \psi^{in}(\mathbf{x}', t) - \psi^{out}(\mathbf{x}'', t) \quad t \in [0, T], \quad \mathbf{x} \in \Sigma \quad (20)$$

Having introduced the above geometrical description of the cell's membrane, we define the **total current density** for the ion α as

$$\mathbf{e}J^{TOT}(\mathbf{x}, t) = \mathbf{J}^{TOT}(\mathbf{x}, t) = \mathbf{J}_{TOT}^{cond}(\mathbf{x}, t) + \mathbf{J}^{cap}(\mathbf{x}, t) \quad (21)$$

which consists in the following two contributions:

- **Conduction current density:**

$$\mathbf{J}_{TOT}^{cond}(\mathbf{x}, t) = \sum_{\alpha} \mathbf{J}_{\alpha}(\mathbf{x}, t) \quad (22)$$

where $\mathbf{J}_{\alpha}(\mathbf{x}, t) = \mathbf{e}J_{\alpha}(\mathbf{x}, t)$ is the current associated with the conduction of ion species α along the ionic channel.

- **Capacitive current**

$$\mathbf{J}^{cap}(\mathbf{x}, t) = \frac{\partial \mathbf{D}(\mathbf{x}, t)}{\partial t} \quad (23)$$

$$\mathbf{D}(\mathbf{x}, t) = \varepsilon_m \mathbf{E}(\mathbf{x}, t) \quad (24)$$

where $\mathbf{D}(\mathbf{x}, t)$ is the electric displacement vector at $\mathbf{x} \forall t \in [0, T]$ and $\mathbf{E}(\mathbf{x}, t)$ is the electric field. Assuming quasi-static conditions we have

$$\mathbf{E}(\mathbf{x}, t) = -\nabla \psi(\mathbf{x}, t) \quad (25)$$

so that

$$\mathbf{J}^{cap}(\mathbf{x}, t) = -\varepsilon_m \frac{\partial}{\partial t} \nabla \psi(\mathbf{x}, t) \quad (26)$$

Using assumption (3), we may write

$$-\nabla \psi(\mathbf{x}, t) \simeq -\mathbf{e} \frac{\psi(\mathbf{x}'', t) - \psi(\mathbf{x}', t)}{t_m} = \mathbf{e} \frac{\psi_m(\mathbf{x}, t)}{t_m}$$

Finally, using assumption (4), the capacitive current density takes the form

$$\mathbf{J}^{cap}(\mathbf{x}, t) = \mathbf{e}c_m \frac{\partial \psi_m(\mathbf{x}, t)}{\partial t} \quad (27)$$

In conclusion, the total current density flowing across the cell membrane at $\mathbf{x} \in \Sigma \forall t \in [0, T]$ is given by

$$\mathbf{J}^{TOT}(\mathbf{x}, t) = \sum_{\alpha} \mathbf{e} J_{\alpha}(\mathbf{x}, t) + \mathbf{e} c_m \frac{\partial \psi_m(\mathbf{x}, t)}{\partial t} \quad (28)$$

In the next sections, we are going to use Eq. (0.28) in the construction of local and global models for the electric activity of the cell. While the quantity c_m is typical of the cell and taken from experimental observations, the modeling will involve the conduction current.

Remark 1 *The quantity \mathbf{J}^{cap} is defined to as capacitive current density because it mathematically represents the effect of the time rate of change of the surface charge density*

$$\sigma(\mathbf{x}, t) = c_m \psi_m(\mathbf{x}, t) = c_m (\psi(\mathbf{x}', t) - \psi(\mathbf{x}'', t))$$

which accumulates with opposite sign on the internal (as σ_{in}) and external (as σ_{out}) surfaces Σ_{in} and Σ_{out} in such a way that

$$\sigma_{in} = -\sigma_{out}$$

In this interpretation, the cell membrane is interpretable, in the neighbourhood of $\mathbf{x} \in \Sigma$, as a capacitor with plane plates separated by a dielectric medium (the lipid bilayer) with specific capacitance $c_m = \frac{\epsilon_m}{t_m}$ and surface charge density $\sigma(\mathbf{x}, t)$.

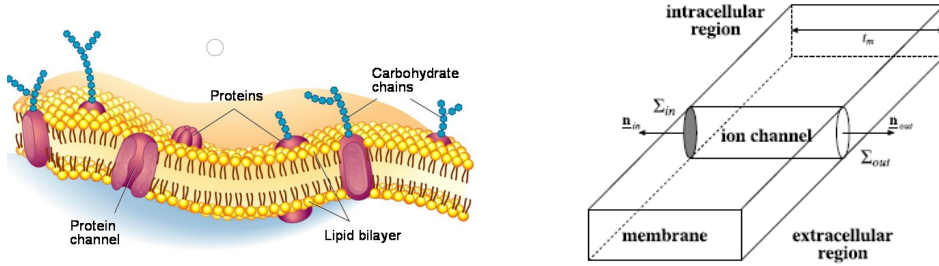


Figure 42: *Left: schematic representation of the cellular membrane. Right: mathematical description of the cellular membrane*

Electric quantity	Name	Unit measure
t_m	Membrane thickness	m
R_c	Cell radius	m
ψ_m	Membrane potential	V
J	Current density	A/m^2
I	Current	A
σ	Surface charge density	C/m^2
c_m	Specific membrane capacitance	F/m^2
C_m	Membrane capacitance	F
g_m	Specific membrane conductance	S/m^2
G_m	Membrane conductance	S

Figure 43: Main electrical quantities in cellular electrophysiology

5.2 ODE local and global models

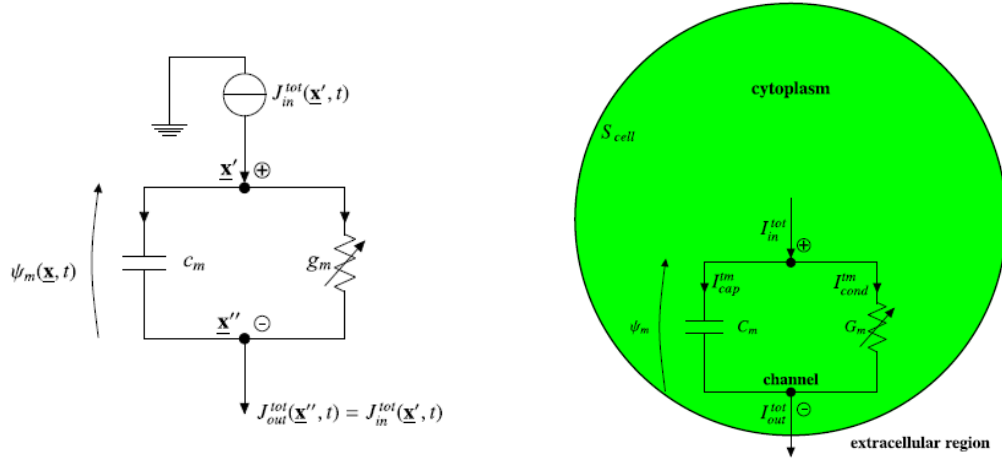


Figure 44: *Left: equivalent circuit for a fixed point $\mathbf{x} \in \Sigma$ (local model)*
Right: equivalent circuit for the whole membrane Σ (global model)

Considering a point $\mathbf{x} \in \Sigma$, equation (0.28) expresses a balance law for the current densities J^{TOT} , which can be divided in two contributions (capacitive and conduction), where the capacitive current assumes the form $J^{cap}(\mathbf{x}, t) = c_m \frac{\partial \psi_m(\mathbf{x}, t)}{\partial t}$.

The conduction current takes the general form

$$J_{\alpha}^{cond}(\mathbf{x}, t) = g_m(\mathbf{x}, t, \psi_m(\mathbf{x}, t))[\psi_m(\mathbf{x}, t) - E_{c,\alpha}] \quad (29)$$

where g_m is the **specific conductance**, depending on ψ_m , and $E_{c,\alpha}$ is the **Nernst potential**, which corresponds to the membrane potential for which the ion conduction current is equal to 0.

Such expressions for the current densities can be syntethized in a *circuitual form*, in which their two contributions (capacitive and conduction) are modeled respectively through a specific capacitance and a specific conductance (see Figure 44). Then, an application of the Kirchhoff current law gives the current balance for the circuit.

For any fixed $\mathbf{x} \in \Sigma$, equation (0.28) is an **ordinary differential equation (ODE)** in the unknown ψ_m . This model is a *local* model, since the focus has been on a specific point of the membrane \mathbf{x} .

If instead of a specific $\mathbf{x} \in \Sigma$, the aim is to write a current balance across the whole cell's membrane Σ , the resulting differential equation, which in this case expresses a *global* whole cell model, is obtained by integration of (0.28) over the surface Σ , which yields the conservation of the total current

$$I_{\alpha}^{TOT} = \int_{\Sigma} \mathbf{J}_{\alpha}^{TOT} \cdot \mathbf{n}_{out} dS \quad (30)$$

In this case, we need to define the whole cell membrane conductance G_m and capacitance C_m , so that the current balance takes the form

$$I_{\alpha}^{TOT}(t) = I_{\alpha}^{cap}(t) + I_{\alpha}^{cond}(t) \quad (31)$$

where $I_{\alpha}^{cap}(t) = C_m \frac{d\psi_m(t)}{dt}$ and $I_{\alpha}^{cond}(t) = G_m(t, \psi_m)[\psi_m(t) - E_{c,\alpha}]$.

While the (specific) capacitance is an intrinsic value of the cell, the conduction current, and in particular the (specific) conductance G_m , need to be modeled. In the next sections, three main models used in literature for this purpose will be presented: the **linear resistor model (LRM)**, the **Goldman-Hodgkin-Katz (GHK)** model and finally the **Hodgkin-Huxley (HH)** model.

5.3 The LRM model

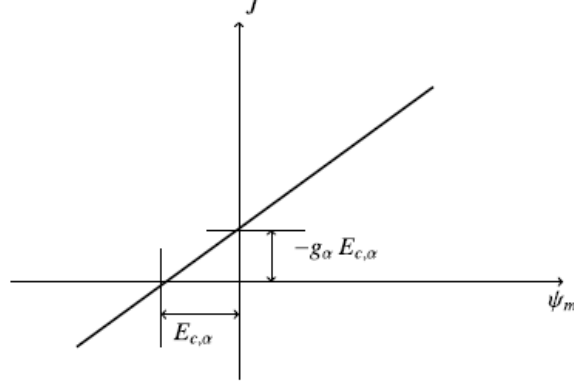


Figure 45: *Characteristic curve for the linear resistor model*

The **linear resistor model (LRM)** makes the simplest assumptions

$$g_{m,\alpha}(\mathbf{x}, t, \psi_m(\mathbf{x}, t)) = g_{m,\alpha}^0 = \text{const} \quad E_{c,\alpha}(\mathbf{x}, t) := E_{c,\alpha}^0 = \text{const} \quad \forall \alpha \quad (32)$$

Therefore, this model assumes the conductance and the Nernst potential as constant. The resulting expression for the conduction current density is

$$J_G(t) = g_{m,\alpha}^0 [\psi_m - E_{c,\alpha}^0] \quad (33)$$

This law defines a straight line in the plane J_G vs ψ_m (Figure 45), such that there exists a unique value of membrane potential for which $J_G = 0$. This value is exactly the Nernst potential $E_{c,\alpha}^0$.

The Nernst potential $E_{c,\alpha}^0$ is related to the concept of **thermodynamic equilibrium**. Micorscopic theory assumes that ion particles in the channel are subjected to two different contributions: diffusion process caused by the chemical gradient, and a transport caused by the presence of an *electrical field* $\mathbf{E}(\mathbf{x}, t)$ [Sacco-Guidoboni]. This phenomenological perspective leads to the **drift-diffusion (DD)** model for the conduction current density, which reads

$$\mathbf{J}_\alpha^{DD} = \mathbf{J}_\alpha^{DIFF} + \mathbf{J}_\alpha^{DRIFT} = -F z_\alpha D_\alpha \nabla C_\alpha + F |z_\alpha| C_\alpha \mu_\alpha \mathbf{E} \quad (34)$$

where

- F is the **Faraday constant**, such that $F = qN_{av}$, being q the elementary charge and N_{av} the Avogadro's constant
- z_α is the **ion number**, positive for cations and negative for anions
- μ_α is the **ion electrical mobility**
- C_α is the molar density of ion α
- D_α is the **ion diffusivity**. this value can be retrieved from the *Einstein-Smoluchowski* relation [Md. Akhtarul Islam] $D_\alpha = V_{TH} \frac{\mu_\alpha}{|z_\alpha|}$, where $V_{TH} \sim 26 \text{ mV}$ is the *thermal voltage*

Theorem 5.1 *Let us assume one dimensional geometry. Then, the thermodynamic equilibrium, for which $\mathbf{J}_\alpha^{DD} = 0$ and the drift current density balances the diffusion current density, is obtained at the value of potential*

$$E_{c,\alpha}^0 = \frac{V_{TH}}{z_\alpha} \ln \left(\frac{C_\alpha^{OUT}}{C_\alpha^{IN}} \right) \quad (35)$$

5.4 The GHK model

With the **Goldman-Hodgkin-Katz (GHK)** model, the modeling of the conduction current becomes nonlinear. Here, the following assumptions are adopted:

- The electric field in the ionic channels assumes a constant value of $E = \frac{\psi_m}{t_m}$, i.e. the ratio between the membrane potential and the membrane thickness
- The geometrical setting is one dimensional
- The conduction current \mathbf{J}_α^{cond} varies only in time, i.e. is constant in space and $\frac{\partial J_\alpha^{cond}}{\partial x} = 0$. This means that there is no production or consumption of ions along the channel

Applying these hypothesis to eq.0.34, it can be shown that the conduction current density takes the form

$$J_\alpha^{GHK}(\psi_m, t) = -qz_\alpha P_\alpha [n_\alpha^{out} \mathcal{B}(\beta_m) - n_\alpha^{in} \mathcal{B}(-\beta_m)] \quad (36)$$

where $P_\alpha = \frac{D_\alpha}{t_m}$ is the **membrane permeability**, $\beta_m := \frac{z_\alpha \psi_m}{V_{th}}$ and $\mathcal{B}(x) = \frac{x}{e^x - 1}$ is the **inverse Bernoulli function**.

It can be proven [Sacco-Guidoboni] that (0.36) can be written in general form (0.29) by setting:

$$g_\alpha(\psi_m, t) = P_\alpha \frac{q z_\alpha^2 N_{av} C_\alpha^{in}(t)}{V_{TH}} \quad (37)$$

$$E_{c,\alpha}(\psi_m, t) = P_\alpha \frac{V_{TH}}{z_\alpha} \left(\frac{C_\alpha^{out}(t)}{C_\alpha^{in}(t)} - 1 \right) \mathcal{B}(\beta_m) \quad (38)$$

5.5 The Hodgkin-Huxley model

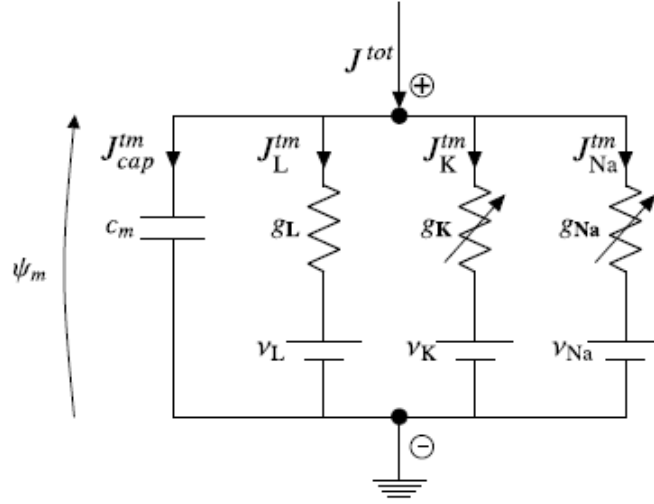


Figure 46: *Equivalent circuit for the Hodgkin-Huxley model*

In 1963, Hodgkin and Huxley (HH) published in a paper a revolutionary work, which awarded them the Nobel prize, on mathematical modeling of the electrophysiological activity of cells, with the focus on the neurons of the giant squid [Hodgkin-Huxley]. The equivalent circuit of the HH model is shown in Figure 46. The main assumptions of the model can be resumed as follows:

- The total current density J^{tot} can be split in two contributions: a capacitive transmembrane current $J_{cap}^{tm} = c_m \frac{\partial \psi_m}{\partial t}$, where the capacitance is assumed to be a constant quantity, and, for every ion considered, a conduction transmembrane current J_{cond}^{tm} to be modeled. The main ions considered are *potassium* (K^+) and *sodium* (Na^+). The remaining

contribution is mostly due to *chlorine* (Cl^-) and is summarized in a *leakage* current J_L^{tm}

- With every ion α is associated its Nernst potential $E_{c,\alpha}$, representing the value for which thermodynamic equilibrium is achieved for each ion current density
- In order to better describe what happens at a biological level of the cell membrane, the ionic channel have a *state*, i.e. an associated probability value which tells the percentage of their opening. Indeed, the behaviour of ionic exchange between the cytoplasm and the extracellular environment, strongly depends on the closing and opening of the ionic channel. With the HH model, the attempt is to mathematically describe the opening status of the channel along time

In the original form of the paper, a change of variable is performed: let us define the equilibrium potential of the whole cell as E_m , and define for every ion α the quantities

$$v_\alpha := E_{c,\alpha} - E_m \quad v := \psi_m - E_m \quad (39)$$

In this way, the form of the current density of ion α becomes

$$J_\alpha^{cond} = g_{m,\alpha}(v - v_\alpha) \quad (40)$$

and the Kirchhoff current law for the circuit, expressed for the new variable $v = v(t)$, is

$$\begin{cases} J^{tot} = c_m \frac{\partial v}{\partial t} + g_{Na}(v - v_{Na}) + g_K(v - v_K) + g_L(v - v_L) \\ v(0) = 0 \end{cases} \quad (41)$$

having assumed the membrane potential at equilibrium at time $t = 0$. For the whole equilibrium potential E_m , the **Goldman potential** can be used [Ermantraut-Terman]

$$E_m = V_{TH} \ln \left[\frac{P_{Na}C_{Na}^{out} + P_KC_K^{out} + P_{Cl}C_{Cl}^{in}}{P_{Na}C_{Na}^{in} + P_KC_K^{in} + P_{Cl}C_{Cl}^{out}} \right] \quad (42)$$

where P_α and C_α denote the permeability and molar density of ion α , respectively.

As for the modeling of conductances, let us introduce the **gating variables** n, m, h for the three considered ion species K^+, Na^+, Cl^- . They represent the probability of opening of the ionic channel. For example, if the cell is in a situation in which $C_{Na}^{out} > C_{Na}^{in}$, the opening of Na^+ channel will result

in a current flow from outside to inside, and to a corresponding rise of the potential ψ_m . When the membrane potential reaches the value of equilibrium E_{Na} , the sodium channel will close. For this reason, the gating variables are expressed as the solution of ordinary differential equation of the type

$$\frac{dy}{dt} = \gamma(1 - y) - \delta y \quad \gamma, \delta > 0 \quad y = m, n, h \quad (43)$$

so that there is a balance between a production rate γ and a consumption rate δ . The solution of an ODE in this form is

$$y(t) = y_\infty \left[1 - \exp\left(\frac{-t}{\tau}\right) \right] \quad (44)$$

where

$$y_\infty = \frac{\gamma}{\gamma + \delta} \quad \tau = \frac{1}{\gamma + \delta} \quad (45)$$

For the potassium conductance, the chosen dependence on its gating variable (in agreement with experimental data) is $g_K = \bar{g}_K n^4$. For the sodium variable, it has been observed that also a second process is involved, so that the proposed form is $g_{Na} = \bar{g}_{Na} m^3 h$. The values \bar{g}_K and \bar{g}_{Na} are constant and determined to fitting of experimental data.

To summarize, the HH model consists in the following sytem of nonlinear ordinary differential equations:

$$\left\{ \begin{array}{ll} J^{tot} = c_m \frac{dv}{dt} + J_{Na} + J_K + J_L & \text{(Kirchoff current law)} \\ J_\alpha = g_\alpha (v - v_\alpha) \quad \alpha = Na, K, Cl & \text{(Conduction current densities)} \\ \frac{dn}{dt} = \gamma_n(1 - n) - \delta_n n & \text{(ODE for gating variable n)} \\ \frac{dm}{dt} = \gamma_m(1 - m) - \delta_m m & \text{(ODE for gating variable m)} \\ \frac{dh}{dt} = \gamma_h(1 - h) - \delta_h h & \text{(ODE for gating variable h)} \\ g_K = \bar{g}_K n^4 & \text{(Modeling conductance for K)} \\ g_{Na} = \bar{g}_{Na} m^3 h & \text{(Modeling conductance for Na)} \end{array} \right. \quad (46)$$

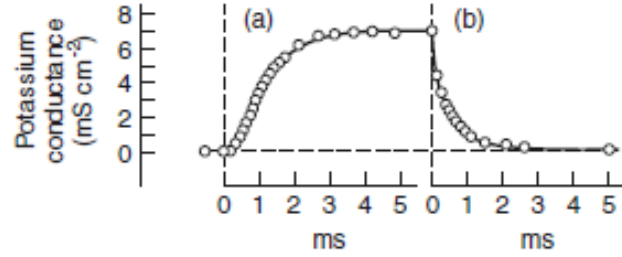


Figure 47: *Example of behaviour of the potassium conductance g_K : a step increase is followed by a decrease. The dots are experimental data*

5.6 Cable model

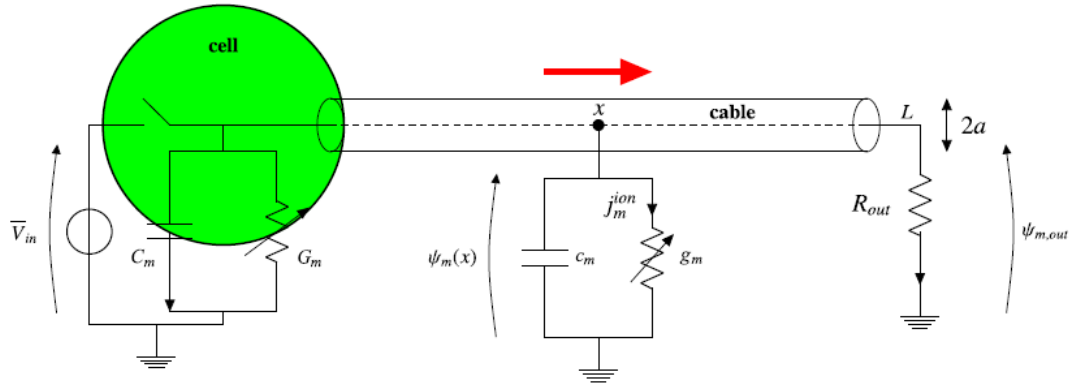


Figure 48: *Schematic representation of the cable model*

The models studied so far described the dynamic evolution of the membrane potential for a single cell. In order to go further, we need to keep into account also the *communication* between two cells, i.e. the propagation of the action potential from one neuron to another. Biologically, this propagation is mediated by the *axon* connecting the two units. The axon is a part of the neuron, and at its every point we may assume a communication with the surrounding environment through electrical circuits as described in the previous sections.

The main assumptions for the cable model, schematized in Figure 48, are the following:

- The biological cable, i.e. the axon, has length L and radius a (thus a cross-sectional area $S = \pi a^2$), with $\frac{a}{L} \ll 1$, allowing us to represent the cable as a 1D object along a longitudinal axis $x \in [0, L]$
- The electric potential outside the cell is constant, as well as the concentrations inside and outside the cell for every ion α :

$$\psi^{out}(x, t) = \text{const} \quad C_{\alpha}^{in}(x, t) = \text{const} \quad C_{\alpha}^{out}(x, t) = \text{const}$$

- The **resistivity** of the axon ρ_{ax} is constant. The resistivity is the reciprocal of the *conductivity* σ , and is related to the *resistance* R_{ax} by

$$R_{ax} = \frac{\rho_{ax} L}{S}$$

If we consider an element of length dx of the cable, its contribution to resistance is $dR(x) = \frac{\rho_{ax} dx}{\pi a^2}$. If a current I_{in} enters the element, the consequent potential difference across the element is

$$d\psi_m(x, t) = -(\psi_m(x + dx, t) - \psi_m(x, t)) = dR(x)I_{in}(x, t) = \frac{\rho_{ax} dx}{\pi a^2} I_{in}(x, t)$$

Therefore, letting $dx \rightarrow 0$, this implies that

$$I_{in}(x, t) = -\frac{\pi a^2}{\rho_{ax}} \frac{\partial \psi_m(x, t)}{\partial x} \quad (47)$$

Next, to express a current balance at each point x , we need to consider again an infinitesimal element dx , in which a current $I_{in}(x, t)$ is entering the element, a current $I_{in}(x + dx, t)$ is leaving the element, and a transmembrane current $I_{TM}(x + \frac{dx}{2}, t)$ is flowing from the axon to the extracellular environment, with a capacitive and a conduction contributions, as shown in Figure 49.

The application of Kirchhoff current law at the node $x + \frac{dx}{2}$ gives

$$-I_{in}(x, t) + I_{cap}(x + \frac{dx}{2}, t) + I_{cond}(x + \frac{dx}{2}, t) + I_{in}(x + dx, t) = 0 \quad (48)$$

The capacitive and conductive contributions to the transmembrane current are

$$I_{cap}(x + \frac{dx}{2}, t) = J_{cap}(x + \frac{dx}{2}, t) 2\pi a dx = c_m \frac{\partial \psi_m(x, t)}{\partial t} 2\pi a dx \quad (49)$$

$$I_{cond}(x + \frac{dx}{2}, t) = J_{cond}(x + \frac{dx}{2}, t)2\pi a dx \quad (50)$$

where J_{cond} can be mathematically described using any of the LR, GHK or HH models. Dividing equation (0.48) by $2\pi a dx$ and letting $dx \rightarrow 0$ gives rise to the equation system

$$\begin{cases} \frac{\partial \psi_m(x, t)}{\partial t} + \frac{1}{2\pi a} \frac{\partial I_{in}(x, t)}{\partial x} + J_{cond}(x, t) = 0 \\ I_{in}(x, t) = -\frac{\pi a^2}{\rho_{ax}} \frac{\partial \psi_m(x, t)}{\partial x} \end{cases} \quad (51)$$

The balance law (0.51) is a *parabolic partial differential equation* for the unknown membrane potential $\psi_m(x, t)$, varying along space and time. To be solvable, (0.51) needs to be completed with an **initial condition** $\psi_m(x, 0) = \psi_m^0(x)$ and appropriate **boundary conditions**. The nature of the cable equation will be linear or nonlinear depending on the chosen model for J_{cond} .

As for the boundary conditions, a common choice is to consider *Robin boundary conditions*. Defining the potential at $x = 0$ as $\bar{V}_{in}(t)$ and modeling the right-hand terminal of the axon through a resistance R_{out} , the resulting boundary conditions are

$$\begin{cases} \psi_m(0, t) = \bar{V}_{in}(t) \\ -I_{in}(L, t) + \frac{1}{R_{out}}\psi_m(L, t) = 0 \end{cases} \quad (52)$$

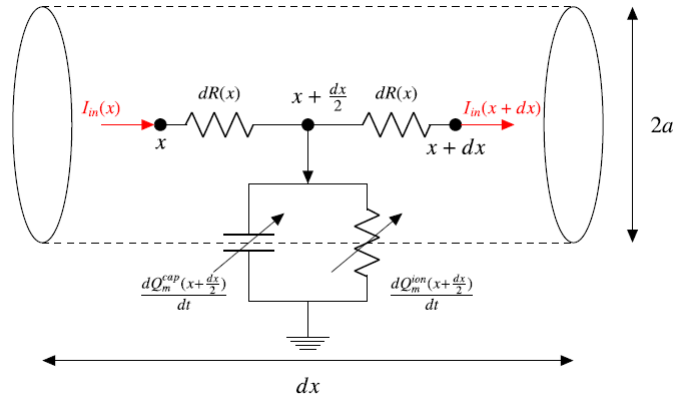


Figure 49: *Current balance in an element dx of the axon*

5.7 Numerical solution of the cable equation

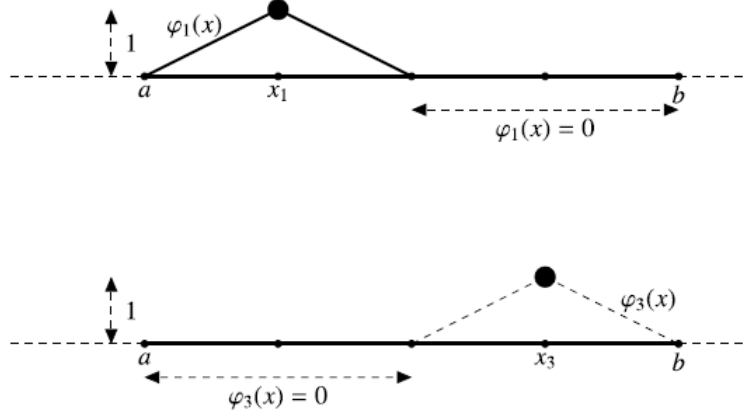


Figure 50: *Lagrangian basis functions*

The cable equation model is an example of the following one-dimensional *diffusion-reaction* boundary value problem for unknowns $u(x, t)$ and $\mathbf{J} = J(x, t)\mathbf{e}$:

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial J}{\partial x} = \mathcal{P} & \forall (x, t) \in (0, L) \times (0, T) & \text{(Conservation law)} \\ J = -\mu \frac{\partial u}{\partial x} & \forall (x, t) \in (0, L) \times (0, T) & \text{(Flux law)} \\ u(x, 0) = u^0(x) & \forall x \in (0, L) & \text{(Initial condition)} \\ \gamma \mathbf{J} \cdot \mathbf{n} = \alpha u - \beta & \forall (x, t) \in \partial\Omega \times (0, T) & \text{(Boundary conditions)} \end{cases} \quad (53)$$

Here, \mathbf{J} represents a *flux* taking into account the diffusion, while \mathcal{P} takes into account the net production rate. The boundary coefficients α, β, γ , depending on their values, can give rise to the main types of boundary conditions (Dirichlet, Neumann, Robin).

In order to numerically solve equation (0.53), the following two main steps need to be performed:

1. **Time semidiscretization.** The time derivative is approximated with a finite-difference scheme. In the present case, the **Backward Euler method (BE)** is chosen, because of its unconditional stability [Quarteroni-Sacco]. Dividing the interval $[0, T]$ in N_T subintervals of the form $[t^k, t^{k+1}]$ with $k = 0, \dots, N_T - 1$, and setting $\Delta t = \frac{T}{N_T}$, the semidiscretization of (0.53) using BE reads

$$\begin{cases} \frac{u^{k+1}-u^k}{\Delta t} + \frac{\partial J^{k+1}}{\partial x} = \mathcal{P}^{k+1} \\ J^{k+1} = -\mu \frac{\partial u^{k+1}}{\partial x} \end{cases} \quad (54)$$

where, for a scalar quantity $\phi(x, t)$, it has been used the notation $\phi^k := \phi(x, t^k)$. The first equation of (0.54) can be rewritten as

$$\sigma^{k+1} u^{k+1} + \frac{\partial J^{k+1}}{\partial x} = f^{k+1} \quad (55)$$

2. **Spatial discretization.** A Galerkin-FEM strategy is adopted to solve (0.54). Assuming, without loss of generality, $\gamma = \beta = 0$ and $\alpha = 1$, and setting $V = H_0^1(0, L)$, problem (0.54) can be rewritten in its *variational form* [Quarteroni]

$$\text{Find } u \in V \text{ s.t. } B(u, \varphi) = F(\varphi) \quad \forall \varphi \in V \quad (56)$$

in which we have defined the forms

$$B(u, \varphi) = \int_{\Omega} \left[\mu \frac{\partial u}{\partial x} \frac{\partial \varphi}{\partial x} + \sigma u \varphi \right] dx \quad (57)$$

$$F(\varphi) = \int_{\Omega} [f \varphi] dx \quad (58)$$

Lax-Milgram theorem [Quarteroni] ensures existence, uniqueness and stability for the solution of (0.56).

With the **Finite element method (FEM)**, the infinite dimensional setting of (0.56) is approximated by the introduction of a finite dimensional space V_h s.t. $\dim(V_h) = N_h < \infty$. For the present case, the chosen space is the linear combination of piecewise linear polynomials. Partitioning the interval $[0, L]$ into M_h subintervals of uniform length $h = \frac{L}{M_h}$ and indicating with τ_h the collection of subintervals, we define the space

$$V_h := \{w_h \in C^0(\Omega) \quad \text{s.t.} \quad w_h|_{K_i} \in \mathbb{P}^1(K_i) \quad \forall K_i \in \tau_h, \quad w_h(0) = w_h(L) = 0\}$$

where $\mathbb{P}^1(K_i)$ is the set of linear polynomials defined on the interval K_i . The space V_h is therefore a finite dimensional approximation of the space $H_0^1(0, L)$. Using for V_h the *Lagrangian basis* (Figure 50) $\{\varphi_j\}_{j=1}^{N_h}$, where $\varphi_j(x_i) = \delta_{ij}$, every element $w_h \in V_h$ can be written as

$$w_h = \sum_{j=1}^{N_h} w_j \varphi_j(x) \quad (59)$$

In light of this, the FE discrete formulation becomes

$$\text{Find } u_h = \sum_{j=1}^{N_h} u_j \varphi_j(x) \in V_h \text{ s.t. } B(u_h, \varphi_i) = F(\varphi_i) \quad \forall i = 1, \dots, N_h \quad (60)$$

resulting in the *algebraic system*

$$B\mathbf{u} = \mathbf{F} \quad (61)$$

where $\mathbf{u} = [u_j]_{j=1}^{N_h}$, $B_{ij} = B(\varphi_j, \varphi_i)$ and $F_i = F(\varphi_i)$. The following theorem expresses the convergence of the FEM solution to the exact solution of (ref 36) [Quarteroni]:

Theorem 5.2 (Convergence) *Let $u \in H^2(\Omega) \cap H_0^1(\Omega)$. Then there exist two constants $c_1, c_2 > 0$, independent of h , such that*

$$\|u - u_h\|_{H_0^1} \leq c_1 h \|u\|_{H^2}$$

$$\|u - u_h\|_{L^2} \leq c_2 h^2 \|u\|_{H^2}$$

The above error estimates express the property that u_h converges to u with order 1, with respect to h , in the H^1 norm, and with order 2 in the L^2 norm, respectively.

The application of this discretization procedure to the cable equation (0.51) will results in a different algebraic system of type (0.61) depending on the adopted modeling of J^{cond} . If a linear resistor model is adopted, the PDE, and consequently the algebraic system, will be linear. On the opposite side, if a nonlinear model is adopted, the nonlinearity will reflect also on the algebraic formulation. To this purpose, functional iterations can be adopted to iteratively solve the obtained nonlinear system, such as in the *Newton method* or the *Alternative lagging method* described in [Sacco-Guidoboni].

6 Modeling calcium patterns