Modeling and data analysis of the calcium activity in somatostatin interneurons from in vivo imaging on mice

M.Sc. Thesis of
Fabrizio Bernardi

Supervisor:
Prof. Riccardo Sacco

Co-supervisor:
Dr. Francesco Papaleo

Programme:
Mathematical Engineering

Department of Mathematics
Politecnico di Milano

Contents

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 2.
Finally, a quite modern and extremely relevant topic for this work is introduced: the synchronization among neural signals.

1.1 Neuronal circuits to describe behaviour

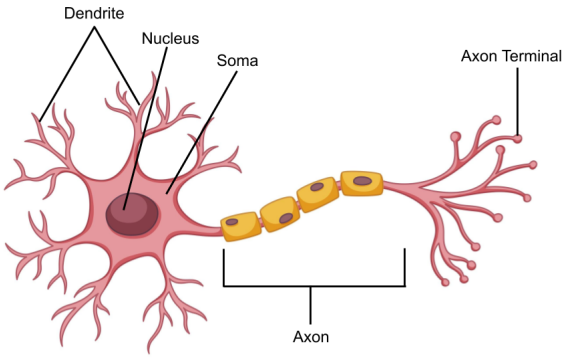

Figure 1: *Schematic representation of a neuron*

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. 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, which conurs 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.

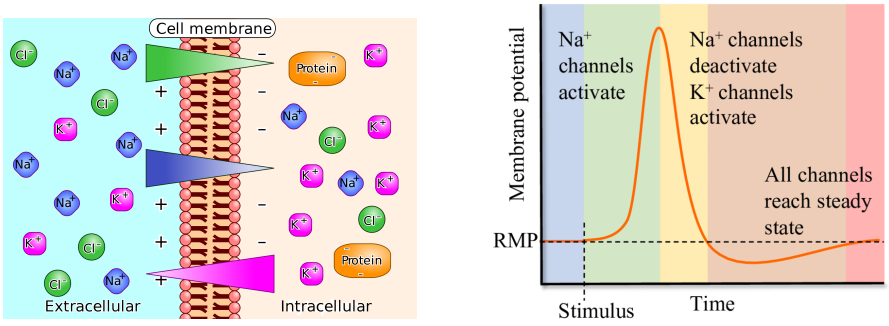

Figure 2: *Left: schematic representation of ionic species inside and outside the cellular membrane. Such species are allowed to pass through the membrane only in specific circumstances through adhibited ionic gates*
Right: example of action potential formation in reaction to a stimulus

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**, 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.

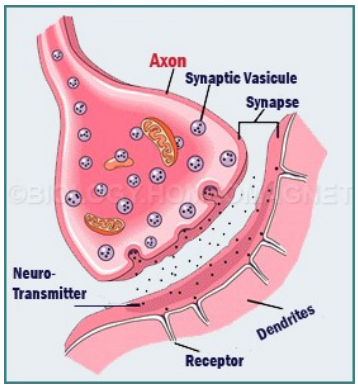

Figure 3: *Schematic representation of a chemical synapse*

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, 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.

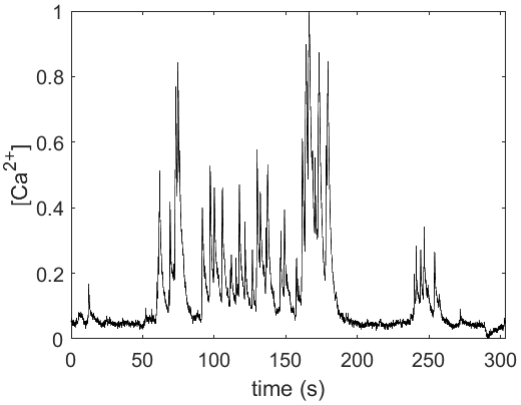

Figure 4: *Example of Ca^{2+} concentration recorded in a neuron, as function of time*

Having defined the meaning of neuronal 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?]

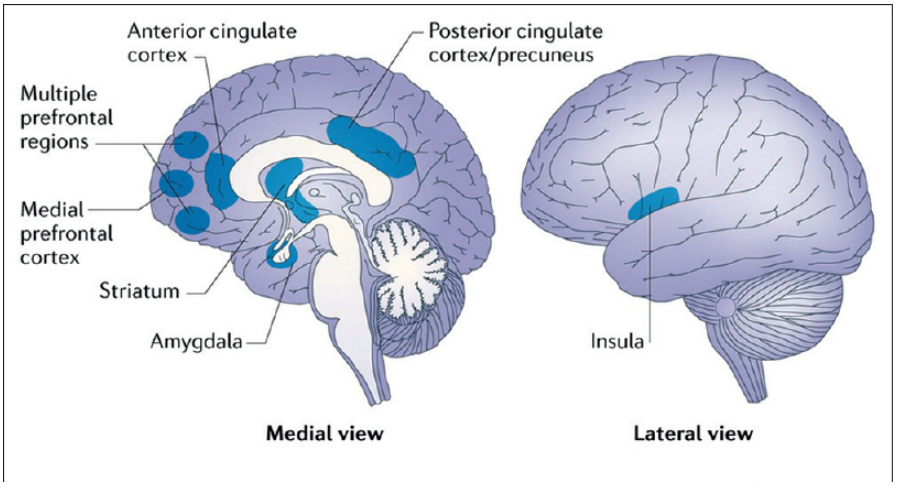

Figure 5: *Main areas of the brain involved in behaviour and emotion recognition*

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 measurments
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**. 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 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,

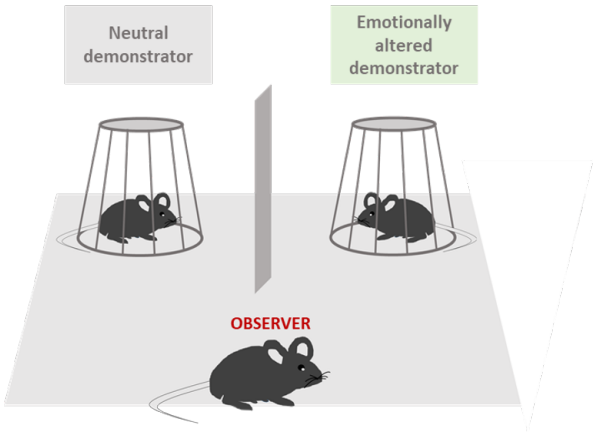

Figure 6: *Basic setting of the emotion discrimination task: one mouse (the observer) is free to move in the arena, while the others two (the demonstrators) remain inside a cage. One of the demonstrators has a neutral affective state, the other an altered affective state*

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
- 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 repetition of the task using only two neutral demonstrators no significant differences emerged
- By repeating the task elinitaing 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 optoenetic 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

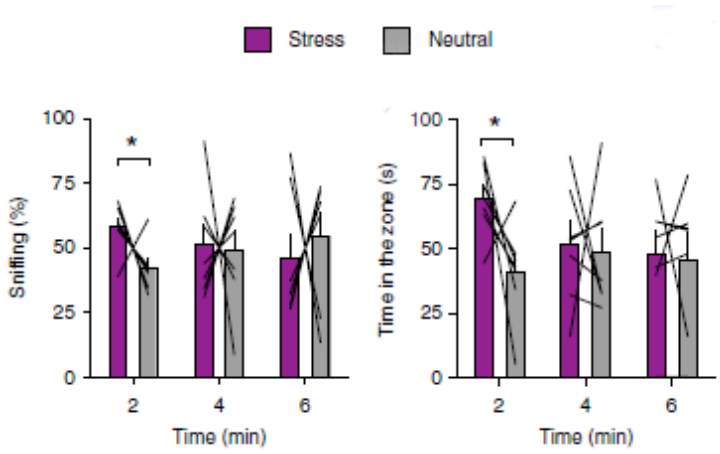

Figure 7: *Recording of sniffing times (left) and proximity times (right) between observer mouse and stressed demonstrator*

- 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 interneurons are 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 neuropep-

tide 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 (Chapter 2).

1.3 Calcium imaging through Inscopix

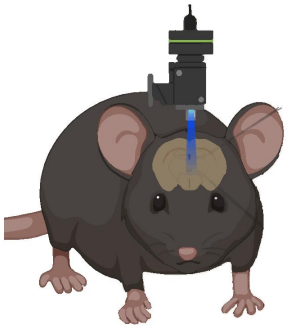

Figure 8: *The Inscopix miniscope*

In this section, a powerful tool for single neuron calcium imaging will be presented. 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 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. 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, however, the Inscopix software manages also the pre-processing part, which goes through the following steps:

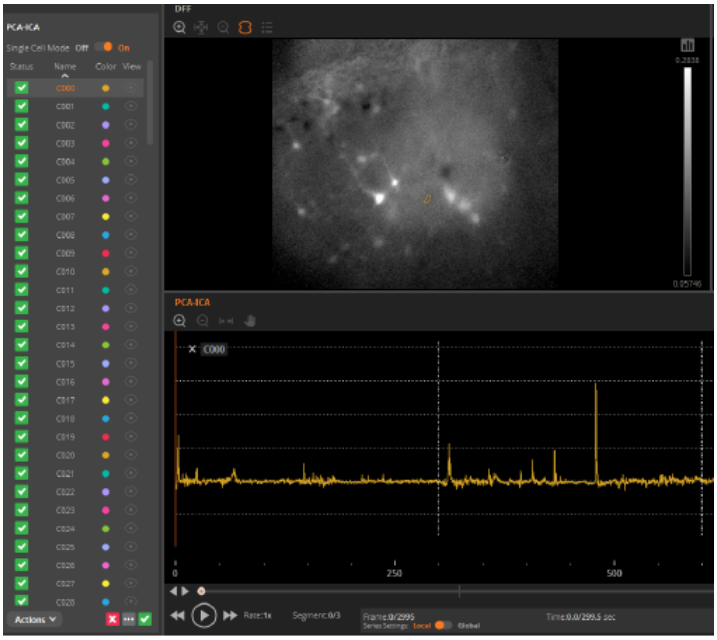

Figure 9: *Example of pre-processing work through Inscopix software. We can appreciate: the list of detected neurons (left), a segment of the video in the ROI (top-right) and an example of the calcium track recorded in a single neuron through time (bottom-right)*

1. If needed, more videos recorded from the test 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 $\sigma_i = \frac{2ln(2)}{2\pi\lambda_i}$ 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 adjustements**: 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

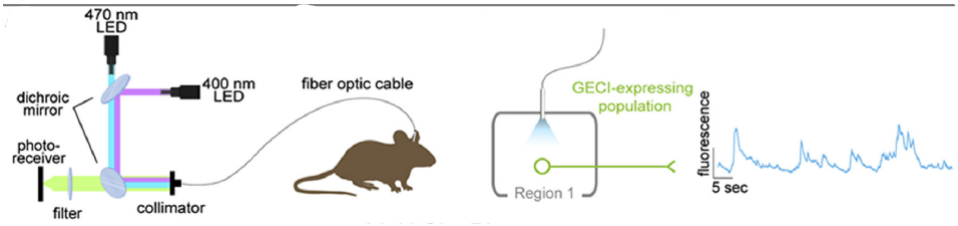

Figure 10: *Fiberphotometry experimental setup*

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 will be presented: **fiberphotometry**.
The main goal of this technique is the same as the Inscopix imaging one, however 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 ($470nm$), which is transmitted through the optical cannula implanted in the mouse, while an emission green light ($525nm$) 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. However, this technique is becomingincreasingly 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

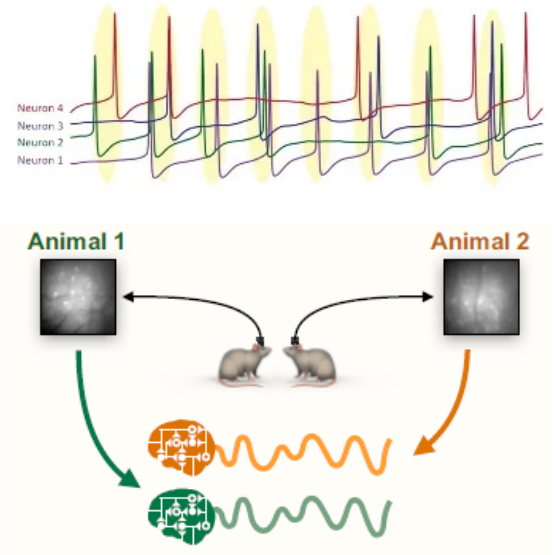

Figure 11: *Top: single neuron synchronization*
Bottom: overall activity synchronization

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 individuAL signals 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:

- **Intraneuronal synchronization**: the signals of single neurons show a correlated activity. Often (but not necessarily), the relevance of this phenomen 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.

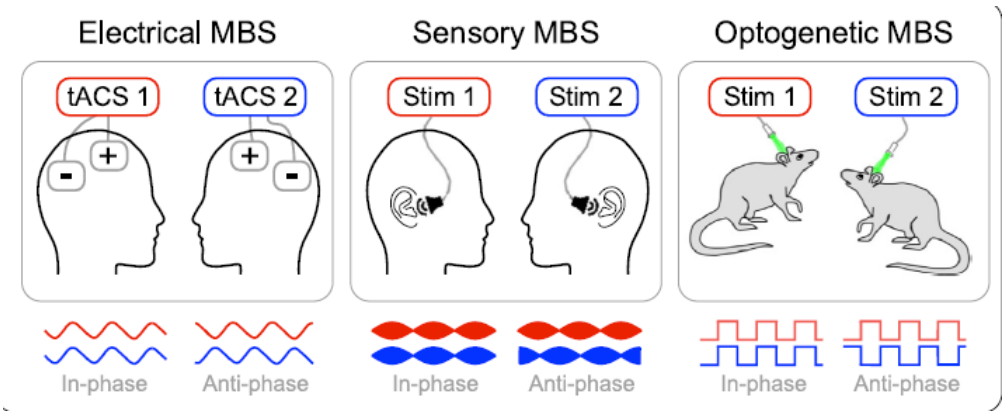

Figure 12: *Example of multi brain stimulation techniques*

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.

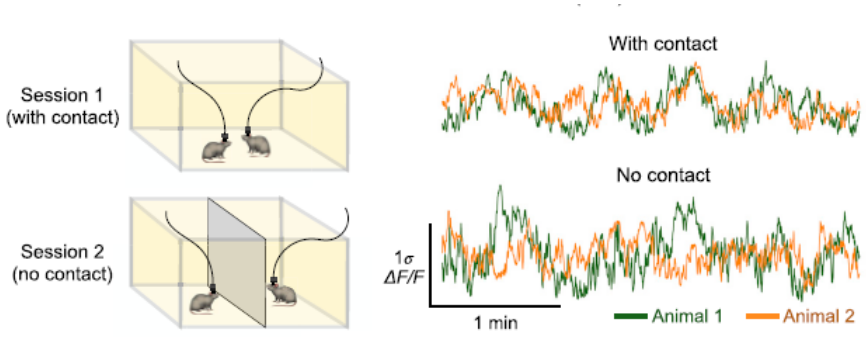

Figure 13: *Comparison of neuronal activities in the two mice with and without contact*

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.

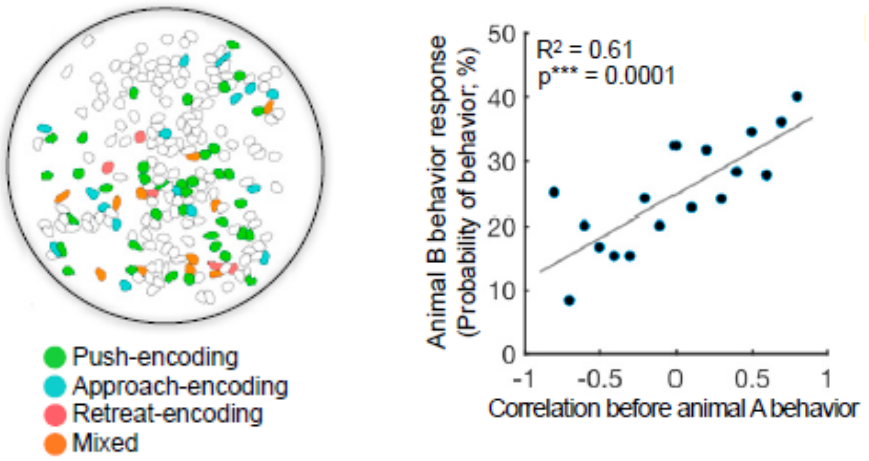

Figure 14: *Left: identification of behavioural neurons in the ROI.*
Right: correlation between the interbrain synchronization preceding behaviour in one animal and the response probability of the interacting partner

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 definitions and their role in the application of *Intebrain synchronization*.
After dealing with the problem of *what* synchronization actually means in this context and which measures *should not* be considered, several valid tools will be proposed: **cross-correlation**, **peak-synchronization**, **angular similarity**, L^2 **error analysis**. 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 dictating the behaviour of the other?". To answer such questions, a sophisticated tool will be object of study: the **Granger causality**.

2.1 Understanding synchronization

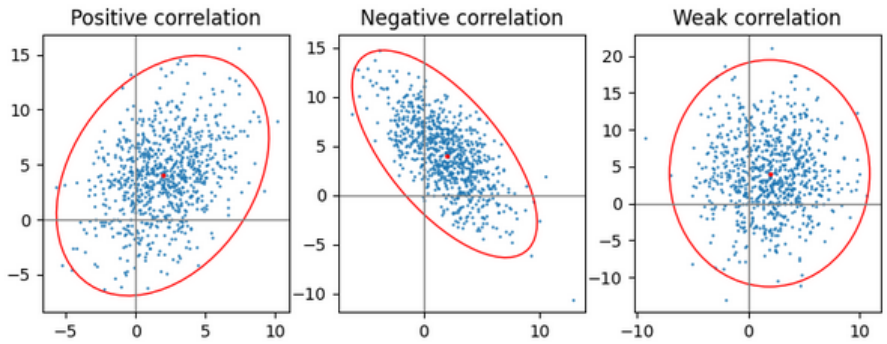

Figure 15: *Geometrical interpretation of the Pearson correlation: such coefficient gives a measure of how well the pairs obtained from the two samples distribute linearly along the diagonal of an ellipse covering the scatterplot.*

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, one first hint about what to look for or not can be found in the type of signal which is under analysis. In the case of this work, dealing with single

neuron measurements of the intracellular concentration of $[Ca^{2+}]$ means dealing with *strongly nonlinear and unstable signals*, which present sudden peaks and an often chaotic and difficult to predict behaviour .
For this reason, the measure of correlation which is probably the most famous and used, 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 them is defined as

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

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 which presents 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 hypoteses fail, and it follows that other correlation measures should be inspected instead.

2.2 Cross-correlation

When investigating the similarity between two signals depending on time, the first tool to be considered probably has to be the **cross-correlation**.
Formally, given two functions $f = f(t)$ and $g = g(t)$, we define the cross-correlation between them as

$$[f(t) \star g(t)](\tau) = \int_{-\infty}^{+\infty} f(\tau) g(t+\tau) dt \tag{1}$$

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

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

it follows that

$$[f(t) \star g(t)](t) = [f(-t) * g(t)](t) \tag{3}$$

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 (eq.1), the resulting quantity is not expressed as function of the time variable t, but as 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 functions, one can identify its maximum value and retrieve the correspondent lag-value, 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 happens

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) correspondent to the lag $\tau = 0$. Moreover, the shape of the cross-correlation will be always symmetrical (see Figure).

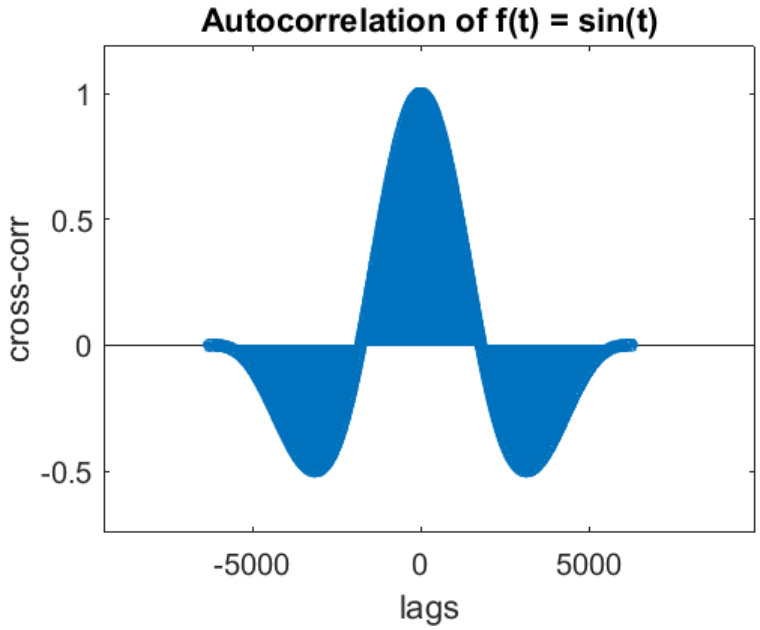

Figure 16: *Aurocorrelation of the function* $f(t) = \sin(t)$

While the autocorrelation represents an ideal case, when dealing with two different functions $f(t)$ and $g(t)$, the goodness of their cross-correlation will be given by the amount of shared properties to the case of autocorrelation; however, depending on the application in study, it could be normal to expect a peak value in correspondent 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 (eq.1) reads:

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

Where m is the 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, probably the most used and valid. With the following concept, the **peak-synchronization**, the focus of the analysis narrows to the specific type of signal characterizing this work: recording of intracellular calcium activity.
As it has been 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 will 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) will be synchronized, under the peak-synchronization point of view, if a *pattern* of simultaneous firing emerges.

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

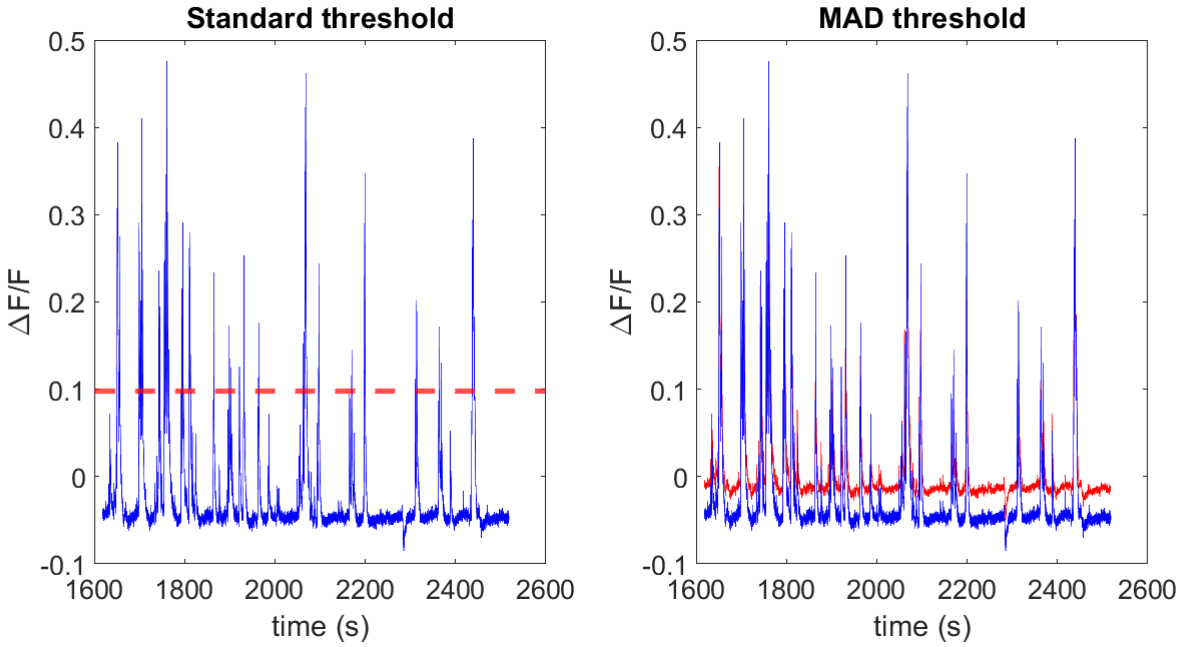

Figure 17: *Example of a calcium signal (blu). The red curves are the thresholds built from the two algorithms: standard threshold (left) and MAD threshold (right)*

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. Determine a way to quantify such correlation

The way to solve all of these three steps is not unique, but it is strictly dependent on the current biological application under study and by the choice of appropriate algorithms and formulas for the activity detection and for the synchronization quantification. As for this work, the following considerations and choices have been made:

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

   A naive way to create such 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 this fixed quantity will be considered active and vice versa. This approach is good enough when dealing with "well-behaving" signals, presenting a baseline low activity alternated with 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 their correspondent 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 dynamism.

Algorithm 2 MAD threshold algorithm

1: Start from a baseline threshold given by $MAD = median(X_i - median(X))$
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

2. A typical time interval in which a firing happens usually strongly depends on the specific case. However, in general it is safe to say that the peak of a neuron usually last at least $250 - 500$ 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 the simultaneous synchronization, a common tool in literature is the **Peak-correlation index** [Cutts and Eglen], defined as

$$i_{AB} = \frac{N_{AB}T}{2N_A N_B dT}$$

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 along every synchronization window.

The peak correlation index gives a representation of how well two signals (neurons) are peak-synchronized. However, it is necessary to notice 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 to compare different pairs.

2.4 Angular distance and L^2 error

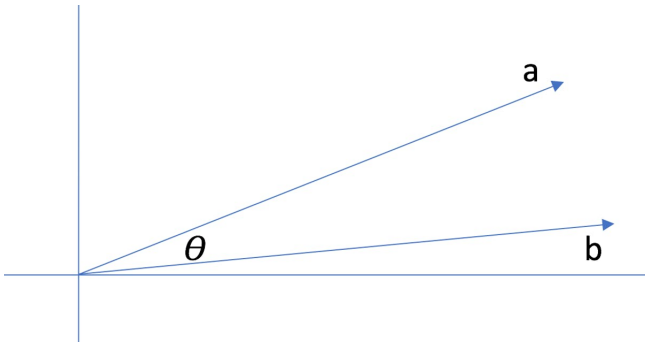

Figure 18: *Angle between two vectors in Euclidean space (2D case)*

After considering the two main tools for the synchronization analysis, cross-correlation and peak-synchronization, other concepts of correlation and similarity can be investigated as well. Here, two supplementary measures have been chosen: the **angular distance** and the L^2 **error**.

The purpose of the angular distance analysis is to determine the similarity between two signals in a *geometric sense*. The idea is indeed rather simple: project the signals on a Euclidean space and observe the angle between them. The two signals will be closer, and thus more similar, if the angle between them

will be small. A way to obtain the angle θ between two vectors is to compute first the **cosine similarity**

$$Sim_C(\mathbf{x}, \mathbf{y}) = \cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{||\mathbf{x}||_E ||\mathbf{y}||_E}$$

where $\mathbf{x} \cdot \mathbf{y}$ denotes the scalar product between $\mathbf{x}$ and $\mathbf{y}$ and $|| \cdot ||_E$ denotes the Euclidean norm.

Once computed the cosine of the angle, the angular distance can be retrieved as

$$d_\theta(\mathbf{x}, \mathbf{y}) = \frac{\arccos(Sim_C(\mathbf{x}, \mathbf{y}))}{\pi}$$

Where the actual angle is divided by a reference π angle.
As for the L^2 error, between two continuous functions f and g defined on an interval $[t_1, t_2]$, the L^2 error between them is defined as

$$||f - g||_2 = \int_{t_1}^{t_2} |f(t) - g(t)|^2 dt$$

When dealing with a discrete approximation, it actually coincides with the **mean squared error (MSE)**

$$MSE(\mathbf{x}, \mathbf{y}) = \frac{1}{N} \sum_{i=1}^{N} |x_i - y_i|^2$$

The MSE analysis performs a point-by-point comparison of the two signals, penalizing quadratically the differences between the two. It follows that in order to have a realistic estimate of this error, it is necessary to confront well-aligned signals. For this purpose, a previous analysis of cross-correlation can be helpful, since, as described in section (2.2), such analysis can also identify the *delay* between two time series, detected in correspondence to the cross-correlation peak. In conclusion, an alignment based on such value of lag between the two time signals should be performed before this L^2 error analysis.

2.5 Granger causality

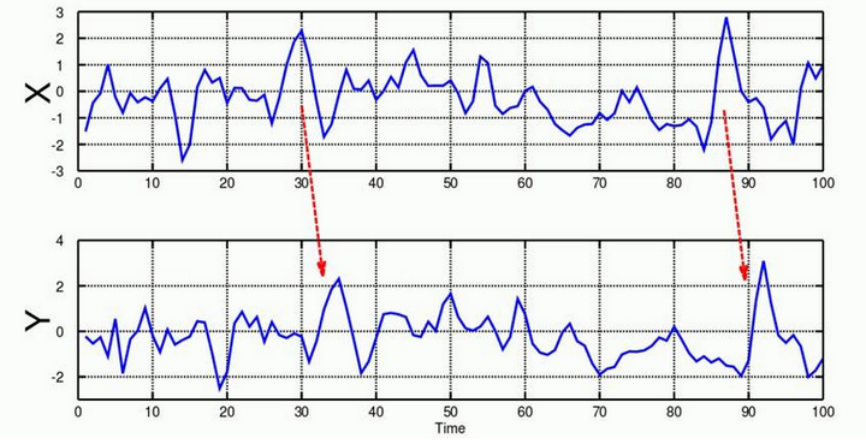

Figure 19: *Example of signal X Granger-predicting signal Y*

As already mentioned in section (1.5), observing a correlation between two time series is not enough to establish a *relationship* between them. Indeed, a parallel issue to the synchronization between signals is the presence of a *cause-effect* mechanism underlying it all. If such relationship is found, besides observing a synchronization, one can find as well 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 the previous chapter, such relationship had been investigated through a manipulation on an experimental level. In the case of the current work, once dealing with data already measured, a different and sophisticated approach is adopted, relying on straightforward statistical principles: the **Granger causality**.
The Granger causality (**G-causality**) is a method aimed to identify causal relationship between time series data. The adjective "casual" 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 main ideas of the G-causality method rely on the following scheme:

1. Creation of a **vector autoregression model (VAR)** of order p, integer to be determined, for one of the two time series, considering its previous values

$$\hat{x}_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \cdots + a_p x_{t-p} \tag{4}$$

2. Creation of a second model, in which the values of the second series are added to the previous model

$$\hat{x}_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \cdots + a_p x_{t-p} + b_1 y_{t-1} + \cdots + b_p y_{t-p} \tag{5}$$

3. Comparison of the two models: if model 2 is more significant (in a way which will be clarified in the following) than model 1, then signal **y** *Granger-predicts* signal **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-cause 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

$$\mathbf{U}_t = \sum_{k=1}^{p} A_k \mathbf{U}_{t-k} + \varepsilon_t \tag{6}$$

Where p is the **order** of the model, $\{A_k\}$ are the **regression coefficients** and ε the **residuals**, assumed normally and independently distributed. The **residual covariance matrix** of the model is defined as $\Sigma = Cov(\varepsilon_t)$ 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** $\{\Gamma_k\}$ is defined as $\Gamma_k = Cov(\mathbf{U}_t, \mathbf{U}_{t-k})$, and it is possible to relate this quantity to the autoregression coefficients $\{A_k\}$ thanks to the **Yule-Walker** equations [Anderson,1971]

$$\Gamma_k = \sum_{i=1}^{p} A_i \Gamma_{k-i} + \delta\Sigma \qquad \delta \in \mathbb{R} \tag{7}$$

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

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

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

$$\mathbf{U}_t = \sum_{k=1}^{p} \begin{bmatrix} A_{xx,k} & A_{xy,k} \\ A_{yx,k} & A_{yy,k} \end{bmatrix} \mathbf{U}_{t-k} + \begin{bmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \end{bmatrix} \tag{8}$$

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

$$\mathbf{X}_t = \sum_{k=1}^{p} A_{xx,k} \mathbf{X}_{t-k} + \sum_{k=1}^{p} A_{xy,k} \mathbf{Y}_{t-k} + \varepsilon_{x,t} \tag{9}$$

If the process Y does not G-cause the process X, it will follow that the coefficients $\{A_{xy,k}\}_{k=1}^{p}$ will be all null, and the model will take a form

$$\mathbf{X}_t = \sum_{k=1}^{p} A'_{xx,k}\mathbf{X}_{t-k} + \varepsilon'_{x,t} \tag{10}$$

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

$$H_0 : A_{xy,1} = A_{xy,2} = \cdots = A_{xy,p} = 0 \tag{11}$$

If $\Sigma'_{xx} = Cov(\varepsilon'_{x,t})$ is the residual covariance matrix of model (10), 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 $\mathcal{F}_{Y\to X}$):

$$\mathcal{F}_{Y\to X} = \ln \frac{|\Sigma'_{xx}|}{|\Sigma_{xx}|} \tag{12}$$

Since the determinant of a covariance matrix (i.e. the *generalized variance*) quantifies the *prediction error* of its regression model, the interpretation of (11) is that the G-causality statistics $\mathcal{F}_{Y\to X}$ is a measure of how much the prediction error is reduced when the process Y is included in the regression model as well. Clearly, this same procedure applies as well for the statistics $\mathcal{F}_{X\to Y}$, in which the directionality of the relationship is inverted. It can be proven [Wilks & Wald] that, under the null hypothesis, $(N-p)\mathcal{F}_{Y\to X} \sim \chi^2(d)$, where $d = pN^2$. To summarize, the typical workflow for G-causality estimation, consists in the following steps:

1. Estimation of the model order p via appropriate criterion (such as AIC and BIC)
2. Estimation of the autocovariance sequence Γ_k and the VAR coefficients (A_k, Σ) through the Yule-Walker equations (7), both for reduced and augmented models. Check for the satisfaction of the validity hypotheses
3. Computation of the G-statistics $\mathcal{F}_{Y\to X}$ and $\mathcal{F}_{X\to Y}$ through (12)
4. Testing of the significance of the statistical tests against the null hypotesis, 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 are presented.

3.1 Emotion discrimination task

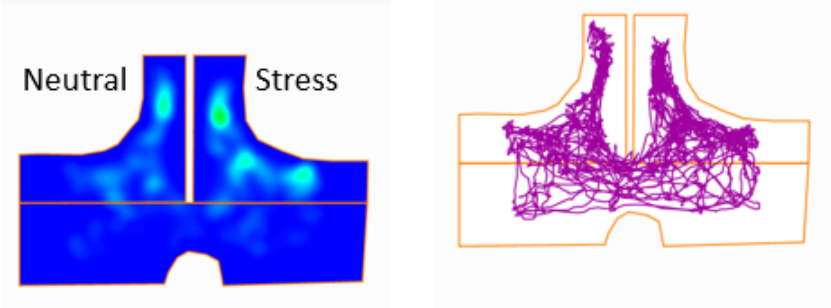

Figure 20: *Open arena of the test, along with positions of demonstrators and movement's trajectories of the observer*

The first task which has been analyzed is the *emotion discrimination task* described in chapter 1. Here, an observer mouse faces two demonstrators in an open arena. One demonstrator is in a *neutral* state, the other in a *stressed* state. In particular, all three mice are subjected to the following protocol:

1. **Homecage restrainment**: the three mice are kept separate in a cage in normal conditions.
   $\longrightarrow$ *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 stressing procedure
   $\longrightarrow$ *Duration*: 15 minutes
3. **Test**: the demonstrators join the observer in the arena and the test phase can begin
   $\longrightarrow$ *Duration*: 15 minutes

During all the three phases, neural recording are performing thorugh in vivo endoscopic calcium imaging as described in (1.3). Moreover, the position of the observer mouse along time is recorded as well. Finally, TTL signals of the reciprocal *sniffing* between one observer and a demonstrators have been measured as well.
The goals of the data analysis work on this experiment are severals and will cover the next sections:

- To compute the activities of the three mice (average or from single neurons) and look for differences in their values between the different phases of the experiment. Is it possible to explain the results based on the simultaneous protocol in action?

•