MASTER'S THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

Title: Oscillatory dynamics with applications to cognitive tasks

Author: Andrea Suárez Segarra

Advisor: Gemma Huguet Casades

Department: Department of Mathematics

Academic year: 2021-2022

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Master in Advanced Mathematics and Mathematical Engineering
Master's thesis

Oscillatory dynamics with applications to cognitive tasks

Andrea Suárez Segarra

Supervised by Gemma Huguet

September, 2022

I wish to express my gratitude to my supervisor Gemma Huguet for her guidance and positive energy in every step of this project. It has been a rewarding experience and I have learnt a lot during these months.

I would also like to thank my family and friends, especially my father Carlos and my partner Basile who have thoroughly read my work and given valuable advice. Finally, I sincerely thank my mother Marta for her unconditional support.

Abstract

Oscillations are ubiquitous in the brain and robustly correlate with distinct cognitive tasks. A specific type of oscillatory signals allows robust switching between states in networks involved in memorizing tasks. In particular, slow oscillations lead to an activation of the neuronal populations whereas oscillations in the beta range are effective in clearing the memory states.

In this master thesis, previous works are revisited in order to provide a detailed analysis of the mechanisms underlying the states' switching and their dependence on the network parameters. The model studied is a macroscopic description of the network recently derived due to mean-field theory advances. The role of spiking synchrony in the "switching off" of the active states is identified by means of bifurcation analysis and the study of the fixed points under the stroboscopic map. Finally, we propose an application of the effect of oscillations in a context of working memory.

Keywords

Computational neuroscience, dynamical systems, working memory, mean field equations, bifurcation analysis, stroboscopic map.

Contents

1. Introduction

When recording the brain electrical activity using an electroencephalogram, the presence of oscillatory patterns is indisputable. During the second half of the 20th century several scientifics evidenced the role of neuronal circuits self producing oscillations ("central pattern generators") in motor functions such as walking or breathing [1]. However, the possible contribution of brain waves in cognitive tasks such as memorizing has not been completely explored yet.

Recent advances in mean-field theory [2] provide low-dimensional models for the macroscopic activity of complex neuronal networks. It is possible then to study the effect of oscillatory signals in the mean behaviour of the network. It has been observed that oscillatory inputs in particular frequency bands can generate robust switches between different neuronal states. In particular, low external input frequencies (around 1.6Hz) are effective in "switching on" the neuronal activity, whereas frequency values in the beta range [13Hz-30Hz] can serve to clear memory states. The described results have been presented and analysed in the research article *Network mechanisms underlying the role of oscillations in cognitive tasks*, 2008, by Schmidt et al. [3].

The initial goal of this master thesis is to retrieve the results obtained in [3] and to understand the drawn conclusions. First, we provide a detailed analysis of the mechanisms underlying the transitions between memory states. Specifically, a bistable neuronal network is considered as a canonical model for a memory circuit to which an oscillatory external current is applied (Section 2). The macroscopic model that describes the network behaviour is analysed using bifurcation theory, which is the mathematical study of changes in the qualitative behaviour of the system due to changes in the parameters. The first analysis is done for the non-forced network and allows to find the suitable parameters so that the network presents bistability and is appropriate for the working memory setting (Section 3). Then the response of the network to periodic forcing of different frequencies *freq* is studied (Sections 4, 5). This is done in particular by means of the stroboscopic map (also known as the Poincaré phase map), which sends each point of the system to its position after time $\frac{1}{freq}$ (i.e. the period of the external input).

Finally, to go beyond previous works, in section 8 we propose an application of the studied model in a specific context of a simple working memory task inspired by the results of the doctoral thesis [4]. In particular, the items to memorize are defined by the combination of one colour and one location, encoded by two networks connected via excitation. In the models that we present, the binding between colour and location is accomplished through the synchronization of neuronal activity of the two connected networks. The possible role of surrounding oscillations is explored. In particular, it is investigated if oscillations with frequency in a specific range can help to decode the color-location pairs that were memorized. The dynamical models used are inspired by the ones described in the extension of the model for multiple interconnected neuronal networks in Schmidt et al. [3].

2. A reduced model for a QIF neuronal network

The information received by a neural system is processed and transmitted by billions of neurons that are connected through synaptic connections. The propagation of a signal along each neuron is provoked by a rapid change in its membrane potential, also called an *action potential* or a *spike*. The information arrives to the axon of a neuron (pre-synaptic structure) and is then transmitted via a synapse to other neurons' dendrites (post-synaptic structure). The chemical synapses can be excitatory or inhibitory depending on whether an action potential in the pre-synaptic neuron increases or, respectively, decreases the probability that an action potential occurs in the post-synaptic neuron.

Neuronal activity has been extensively studied through different computational models. The Hodgkin-Huxley model provides a detailed and biologically accurate description of the initiation and propagation of actions potentials in individual neurons. However, it is a computationally expensive model when studying large networks. Another widely used family of models for single neurons are the integrate-and-fire models, significantly simpler to compute. In this family of models the spikes are described as events: their shape and particular characteristics are not taken into account, only their existence [5]. Thus, the model consists in a single differential equation, describing the membrane potential evolution, and a mechanism to generate spikes when the voltage reaches some threshold v_{th}. A type of integrate-and-fire model often used is the *quadratic integrate and fire* (QIF) model. The dynamics of a neuron's voltage v are described by an autonomous differential equation and a reset condition:

$$\dot{v} = v^2 + I, \qquad \text{if } v \geq v_{th}, \quad \text{then } v \leftarrow v_{reset},$$

where I is a constant representing the external input and v_{reset} is the value given to the voltage when it has reached the threshold v_{th} that determines the presence of a spike. The dot means a differentiation with respect to time t and, for easier readability, this dot notation will be used for almost all equations in this work. Similarly, note that the time dependence is assumed for variables x that have explicit dynamics (i.e. an equation for $\dot{x}$).

Another approach for the analysis of large neuronal networks arises from macroscopic observations of the neurons' mean *firing rate* (frequency at which spikes occur). The derived models, characterized by the firing-rate equations (FREs) [6], have been proved effective to understand specific network functions such as memory or decision making. In 2015 [2], researchers derived a set of macroscopic FREs for networks of heterogeneous, all-to-all coupled QIF neurons, which is exact in the thermodynamic limit (that is, for large numbers of neurons). This is the model that will be studied through this thesis.

2.1 Mean-field equations

The network considered in this work consists on N excitatory heterogeneous QIF neurons, all-to-all coupled. The individual voltage state description [2] for each neuron $j \in N$ of the network is given by the equations :

$$\dot{v}_j = v_j^2 + I_j(t), \qquad \text{if } v_j \geq v_{th}, \quad \text{then } v_j \leftarrow v_{reset} \tag{1}$$

$$I_j(t) = \eta_j + J \cdot s(t) + I(t). \tag{2}$$

Equation (1) is the ordinary differential equation describing the dynamics of the membrane potential $\{v_j\}_{j=1,\dots,N}$ and the reset condition. Without loss of generality, the time and voltage have been rescaled to simplify the equation. The input currents are described in equation (2). The recurrent input is the synaptic weight J multiplied by the mean-synaptic activation $s(t)$ (for more details see [2]). The external input has a time-varying component $I(t)$. Also, heterogeneity is introduced in the network via the constant input current η_j. It models the fact that not every neuron is at the same state of excitability or has the same probability to exhibit a spike. The constants $\{\eta_j\}_{j=1,\dots,N}$ are distributed following a Lorentzian distribution L of half-width Δ and centered at η:

$$L(x) = \frac{1}{\pi}\frac{\Delta}{(x-\eta)^2+\Delta^2}.$$

The Lorentzian distribution was validated with several numerical simulations and it is chosen because of its mathematically convenience for a dimensionality reduction [2]. Indeed, when N is very large, the microscopic model (1) -(2) is computationally expensive to study. Therefore, we will consider instead a system of FREs that provides an precise description of the macroscopic activity of the presented network of QIF neurons [3], assuming that N tends to infinity. The mean field equations, derived in [2], are:

$$\begin{cases} \tau^2\dot{r} = \frac{\Delta}{\pi} + 2\tau v r, \\ \tau\dot{v} = v^2 + J\tau r + \eta - \pi^2\tau^2 r^2 + I(t), \end{cases} \tag{3}$$

where $r(t)$ and $v(t)$ are respectively the average firing rate (spikes per second) and the average membrane potential (voltage) of the network. The strength of the synaptic weights is represented by J, and $I(t)$ is the external input. Parameters η and Δ are, respectively, the center and width of the Lorentzian distribution of inputs, which brings heterogeneity in the network (as explained in the previous paragraph). Finally τ is the membrane time constant of the individual neurons. Through all this work we fix the parameters $\Delta = 2$ and $\tau = 20$ ms, as it was done on the studied paper [3].

The external periodic input $I(t)$ models the existing oscillations in the brain that might be involved in the process of memorization and are the core of this work.

3. Analysis of the network model without forcing

In this section we study the network model (3) without external input. This will bring information about the possible final states of the network for different initial conditions of the mean firing rate and the mean voltage. For example, for some given parameters, a goal is to know whether the network initialized with a high mean firing rate will continue active as time goes, or its mean firing rate will decrease instead. The analysis is done by means of the bifurcation theory, a type of study that we introduce in the next paragraph.

3.1 Some concepts of bifurcation analysis

When analysing a dynamical system, we are interested in studying the invariant objects of the system, since they organize the dynamics around them. Amongst the invariant objects one can look for the fixed points, which are a type of steady states. For this thesis, the fixed points of interest are: the stable node, the saddle point and the stable focus. The type of fixed point for any x can be determined by the eigenvalues of the Jacobian matrix of system (3) evaluated in x. For instance, for $x \in \mathbb{R}^2$, x would be a stable (or attracting) node if both eigenvalues are negative and real, a saddle point if the eigenvalues are real and have different signs and, finally, a stable (or attracting) focus if both eigenvalues are complex with negative real part.

Once the steady states of the system are known, the aim is to understand their dependence on the parameters. How changes on one parameter can impact the existence or stability of the fixed points?

This question is studied through bifurcation analysis, which is the mathematical study of qualitative changes in the system response when one or more parameters vary. The type of bifurcation particularly relevant for the studied two dimensional model (3) is the *Saddle-node* bifurcation: as the parameter of interest changes, two fixed points of the system appear, merge or disappear.

Consider for instance a function $f_t : \mathbb{R}x\mathbb{R}^2 \to \mathbb{R}^2$ and a dynamical system **S** governed by the equation $\dot{x} = f_t(x)$, for $x \in \mathbb{R}^2$. The flow of f_t is a map $\varphi^t(x_0) : \mathbb{R}x\mathbb{R}^2 \to \mathbb{R}^2$ that is solution of the dynamical system in the sense that:

$$\begin{cases} \frac{d\varphi^t(x_0)}{dt} = f_t(\varphi^t(x_0)), \\ \varphi^0(x_0) = x_0. \end{cases}$$

Now let s^* be a saddle point of this system. The stable and unstable *manifolds* of s^*, designated respectively by C_{stable} and $C_{unstable}$, are curves defined as follows:

$$C_{stable} := \{x \in \mathbb{R}^2 | lim_{t\to\infty} \;\; \varphi^t(x) = s^*\},$$
$$C_{unstable} := \{x \in \mathbb{R}^2 | lim_{t\to-\infty} \;\; \varphi^t(x) = s^*\}.$$

In words, the stable manifold (resp. unstable) of s^* is the set of points such that solutions of **S** initiated at these points asymptotically approach the fixed point s^* as t goes to $+\infty$ (resp. $-\infty$). Moreover, the tangent vector to the unstable (resp. stable) manifold at the point s^* is an eigenvector associated to the positive (resp. negative) eigenvalue of the Jacobian matrix of f_t evaluated at s^*.

For a non linear system, it is useful to compute the invariant manifolds associated to a saddle node in order to understand the trajectories and behaviour of the system before, during and after the bifurcation.

In particular, the stable manifolds of a saddle point delimit the basins of attraction of the other attracting fixed points (See Fig.1).

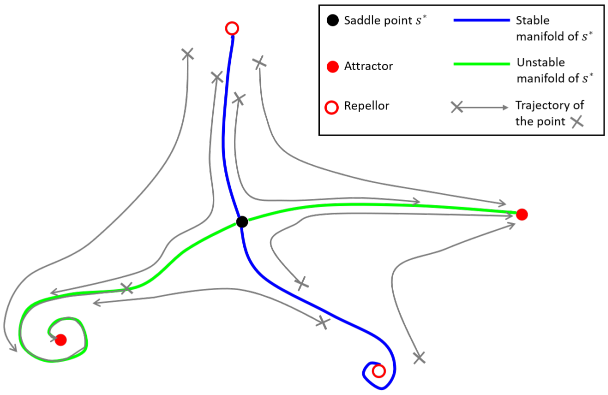


Figure 1: Example of stable and unstable manifolds of a saddle point s^*. Some trajectories converging to the attractors are drawn. It can be seen how the stable manifolds establish the limit between two basins of attraction.

3.2 Steady states and bistability region of the unforced network

In this section we analyse the stationary states of the macroscopic model (3) without external input, as parameters J and η vary. It is intended to find parameters that make the non forced network bistable, so that it can be later approached as a model for neuronal populations implicated in memorization tasks. In particular, the system should have two attractors: one resting state with a firing rate close to zero and one active state with a high firing rate.

For a fixed and sufficiently strong J it is possible to identify a suitable range of η values such that the network presents bistability (See Fig.2). Further investigation reveals that, in the region of bistability, the stable resting state (r_N, v_N) is a node whereas the stable active state (r_F, v_F) is a focus. Between those two stable states there is a saddle point (r_{SP}, v_{SP}). The parameters η_{c1}, η_{c2} for which a saddle node bifurcation occurs can be found by computing the solutions of equations:

$$\frac{\Delta}{\pi \cdot \tau^2} + \frac{2}{\tau} \cdot r \cdot v = 0 \tag{4}$$

$$\frac{v^2}{\tau} + J \cdot r + \frac{\eta}{\tau} - \pi^2 \cdot \tau \cdot r^2 = 0 \tag{5}$$

$$(\frac{2v}{\tau})^2 - \frac{2r}{\tau} \cdot (J - 2\pi^2 \tau \cdot r) = 0, \tag{6}$$

where r, v and η are variables. First, equations (4) and (5) impose (r, v) to be a steady state of the system (note the simplification by τ and τ^2 in the mean-field equations (3)). Then, equation (6) sets to 0 the determinant of the system's Jacobian matrix evaluated at the fixed point (r, v), which forces the solution to be a bifurcation point. The system (4)-(5)-(6) can be solved numerically, for example using the function *fsolve()* in Matlab. When J is strong enough, it is possible to use different initial seeds until achieving convergence to two solutions that have as third coordinate the wanted values η_{c1} and η_{c2}.

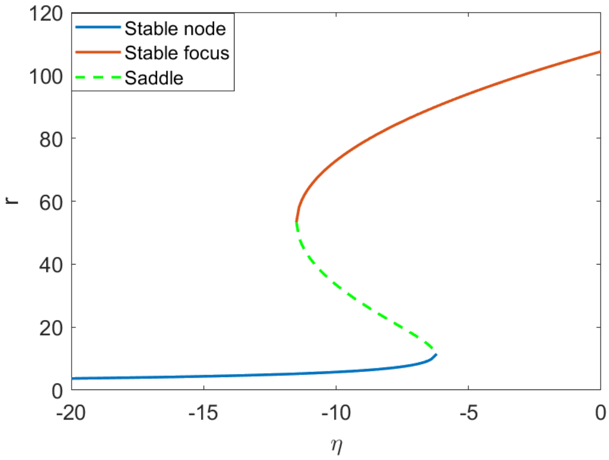


Figure 2: Bifurcation diagram for fixed $J = 15\sqrt{2}$ computed with a numerical procedure in Matlab. For a particular range of η values, approximately from -11.5 to -6.5, the three steady states (node, saddle and focus) coexist.

It is also interesting to vary the parameter J and compute the two dimensional bifurcation diagram with respect to the parameters J and η (Figure 3). This diagram is also computed numerically using Newton-Raphson method. (See the Annex A.2).

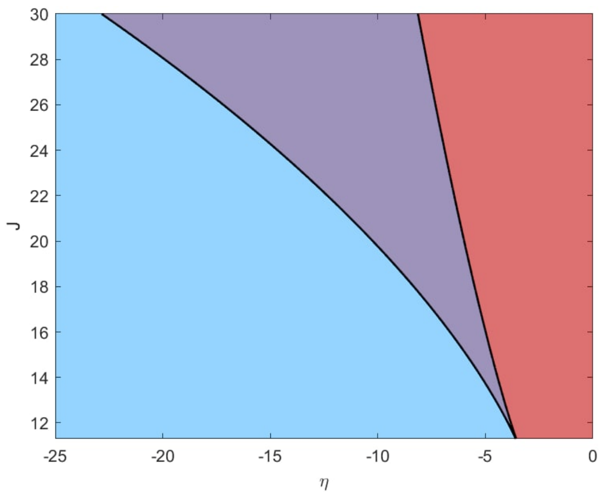


Figure 3: Bifurcation diagram with respect to J and η. We can identify the regions where the system has only one steady state: a low activity state (light blue region) or a high activity state (red region). The bistable region (purple) is delimited by the bistability boundary (black). Parameters: $\tau = 0.02, \Delta = 2$. Parameters for Newton-Raphson method: $\epsilon = 0.1$, tol$= 10^{-5}$ (See Annex A.2 for the numerical details)

It can be seen that for several well chosen and fixed J, there is a range of η values for which the system is bistable. The theory given in subsection 3.1 indicates that it is relevant to numerically compute the unstable and stable manifolds related to the saddle point (r_{SP}, v_{SP}) (See Annex A.1). This was done for different η values in the bistable range and helps to visualize how the fixed points appear, evolve and disappear as η varies (See Figures 4 and 5).

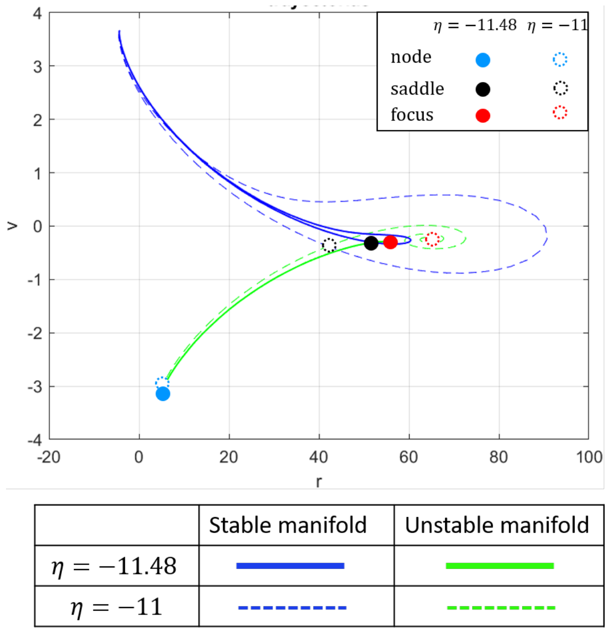


Figure 4: Superposition of the invariant manifolds of the saddle as η increases and moves away from the first bifurcation value η_{c1}. The figure shows how the saddle point (red) moves away from the stable focus (pink), two fixed points that have just been created. $J = 15\sqrt{\Delta}, \Delta = 2$.

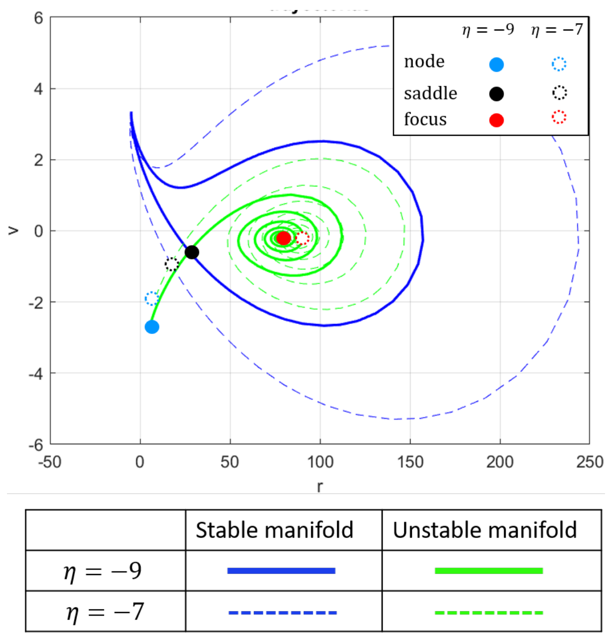


Figure 5: Superposition of the invariant manifolds of the saddle as η increases and approaches the second bifurcation value η_{c2}. The figure shows how the saddle point (red) gets closer to the stable node (black). When η reaches η_{c2}, these two fixed points will merge and disappear. $J = 15\sqrt{\Delta}, \Delta = 2$.

4. Behaviour of the network with periodic forcing

In this section we examine the response of the system (3) to a periodic external input of the form:

$$I(t) = A(\gamma \cdot \sin(\pi \cdot freq \cdot t)^{20} - 1), \tag{7}$$

where A is the amplitude, *freq* the frequency and γ is a constant (independent from A and *freq*) such that $\int_0^T I(t)\,dt = 0$. The input $I(t)$ oscillates between values $-A$ and $A(\gamma - 1)$. For the numerical simulations done in this thesis, we used a numerical approximation of γ.

We apply to the network this non-sinusoidal oscillatory input because the authors of the studied research article [3] proved that this type of forcing was more effective than pure sinusoidal inputs in allowing flexible switching between network states.

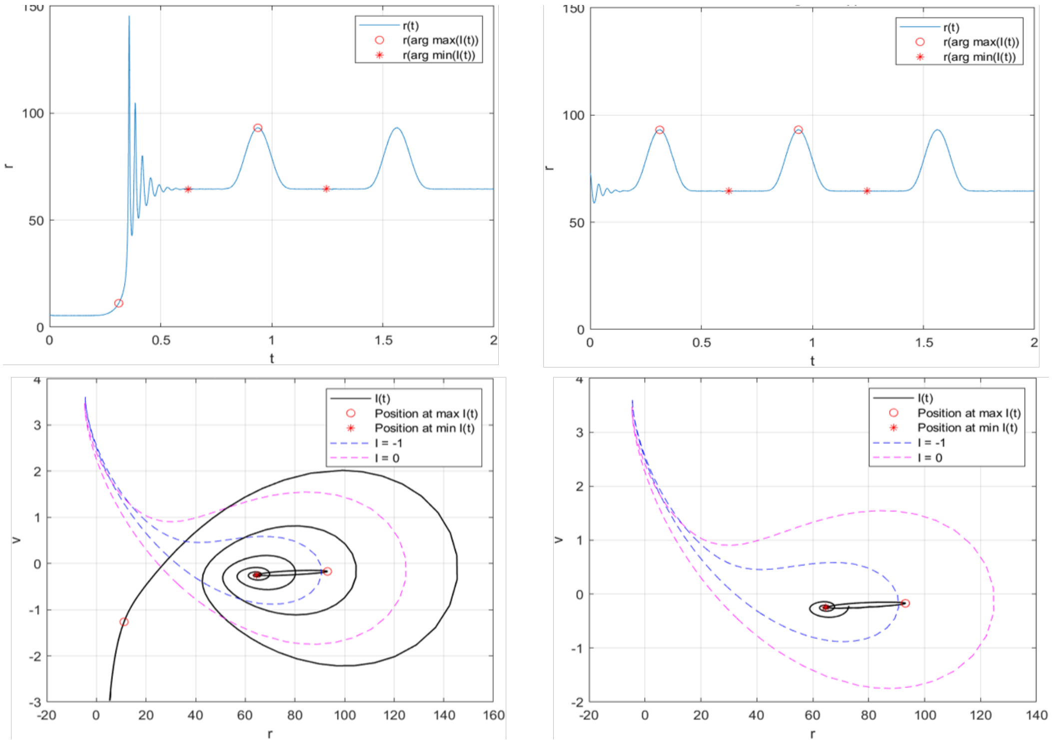
4.1 Numerical results

Fixed parameters J and η are chosen such that the network is bistable. It can be seen that the applied oscillatory current $I(t)$ (equation 7) has an important effect on the behaviour of the system (firing rates and trajectories in Fig.6). Depending on the input frequency, we can observe transitions between the different dynamical states. In the context of a network involved in a working memory task, the transition from a low firing rate to a high firing rate is called *Recall*. The opposite transition, when the activated network is "switched off", is called *Clearance*. We will later study both phenomena.

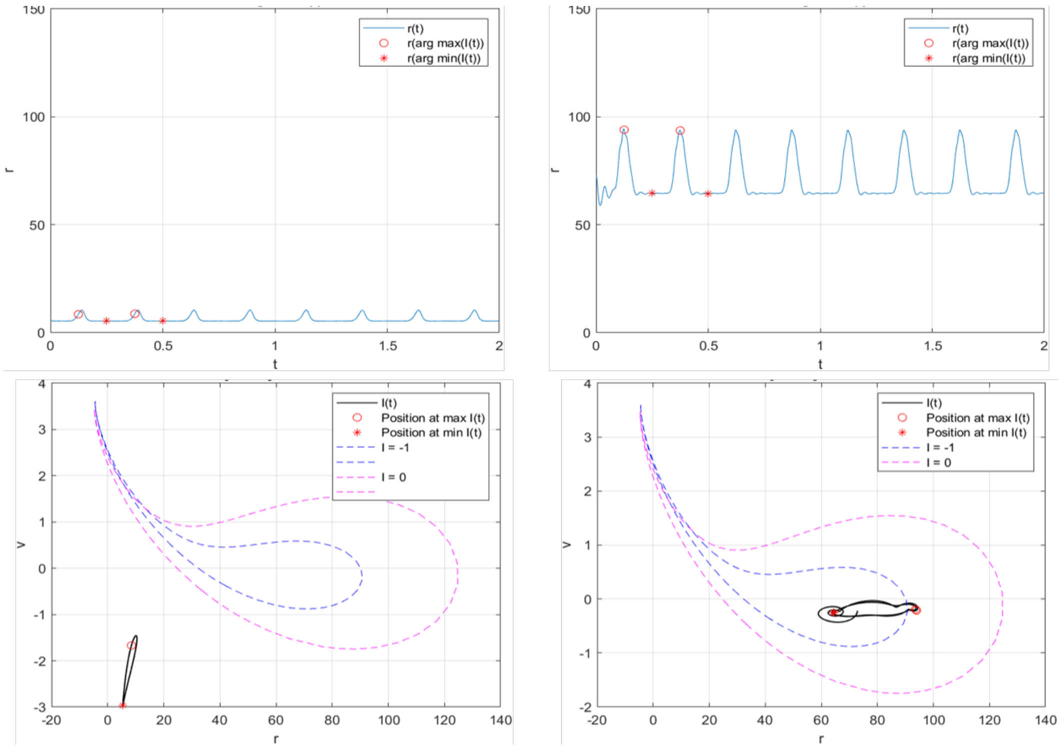
In order to visualize these transitions and the surrounding dynamics, Figure 6 shows the trajectories of the system starting at the resting and the active state, under the forcing of a periodic current, for different input frequencies. The plots also show the two stable invariant manifolds of the saddle point (dotted curves in Figure 6) associated to two particular values the input current takes. Only the stable invariant manifolds are computed because they define the limit between the basin of attraction of the stable focus and the one of the stable node. These basins change and even disappear as $I(t)$ oscillates. We wanted to plot the stable manifolds corresponding to the minimum and the maximum values of $I(t)$ (7). However, with the chosen parameters, the input varies between -1 and $\gamma - 1 \approx 4.68$, and it was found that the saddle and the stable node do not exist for fixed inputs larger than ≈ 3.73. Therefore, it was arbitrarily chosen to plot the stable manifolds for fixed $I = -1$ and $I = 0$ in order to get a feeling of the variations of the system phase portrait. Note that, given the governing equations (3), it is the same to consider the system with a fixed input $I = I^*$ than studying the non-forced system with a new η parameter $\eta_{new} := \eta + I^*$.

Therefore, Figure 6 shows the different behaviours of the bistable network depending on the frequency of the input and the initial state. It can be seen in plot (a) that for low enough frequencies the system is pushed from the state of low activity to the state of high activity. This behaviour corresponds to the *Recall* in the context of cognitive tasks because if the neurons are activated, then some information is being retained. The system undergoes damped oscillations when approaching the active state, what was expected as the attractor is a focus. This type of attraction reflects the transient spike synchrony occurring in the microscopic network when the neurons are globally activated and their firing rates increase. At slightly higher frequencies (b), both steady states persist under the forcing. At some range of intermediate frequencies (c), the high-activity state is driven to the low-activity state. In the context of working memory this is the *Clearance* behaviour: the neurons that were active because of the memorizing task become inactive and the information is forgotten. Finally, at high frequencies (d), both stable states persist under the forcing.

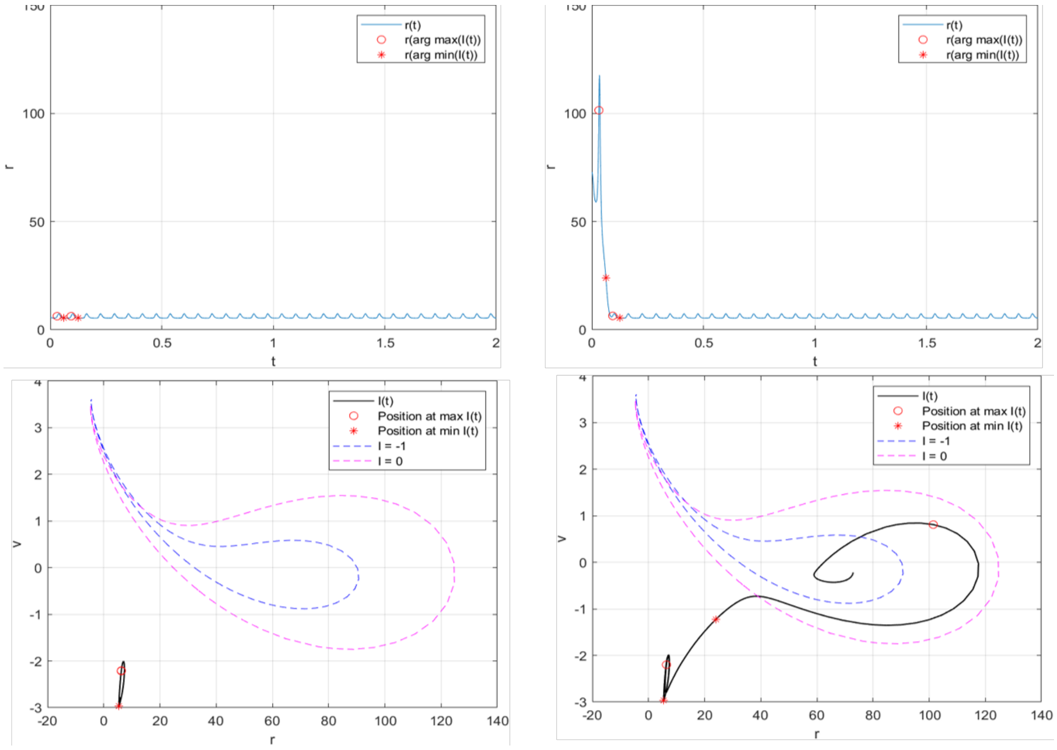
(a) freq=1.6Hz, *Recall*



(b) freq=4Hz



(c) freq=16Hz, *Clearance*



(d) freq=80Hz

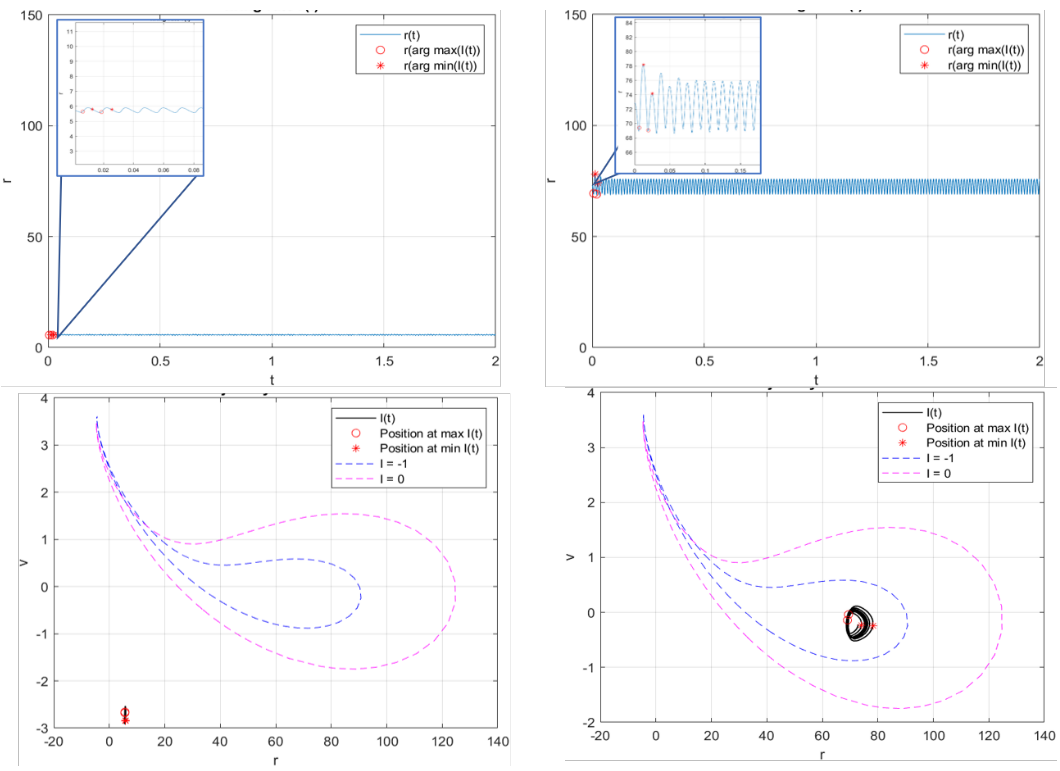


Figure 6: **Response of the network to periodic forcing.** For different input frequencies the figure shows the firing rate (top) and the trajectory described (bottom) when taking as initial condition the resting state (left) or the active state (right). The parameters are chosen so that the system presents bistability in the absence of forcing: $\tau = 0.02$s, $\eta = -10, \Delta = 2, J = 15\sqrt{\Delta}, A = 1$.

The numerical simulations of the mean firing rate model with the chosen parameters (specified in the Figure 6's caption) indicate that the *Recall* behaviour occurs for input frequencies in the range $]0, 1.75]$ Hz. Likewise, it was found that when the input frequency is between 13.5Hz and 33Hz the *Clearance* phenomenon is observed. This frequency band corresponds significantly with the $[13-30]$Hz beta range [7].

4.2 Importance of subthreshold dynamics

In general, the FREs models consist on a single equation describing the dynamics of the network's mean firing rate [6], and follow the assumption that spiking neuronal activity is uncorrelated. However, the authors of the research article [3] found that spike synchrony could be observed for the studied microscopic model (equations (1), (2)). This synchronization phenomenon is then considered in the mean-field model (3) by means of the voltage dynamics. Indeed, this equation describes the average activity of the neurons that are not firing (subthreshold dynamics), hence their predisposition to spike.

In order to appreciate the importance of spike synchrony in the model for obtaining the behaviour described in the previous paragraph, the mean-field equations are studied omitting the voltage equation. A $1-$dimensional equation can be derived by forcing $v(t)$ to be constant over time in the system given by (3), when $I(t) = 0$.

The first equation in system (3) gives an expression for r in terms of v:

$$r = \frac{-\Delta}{\pi} \cdot \frac{1}{2\tau v}.$$

Now the terms are lifted to the square and v^2 is replaced by its definition given in the second equation in system (3), which results in:

$$r^2 = \frac{\Delta^2}{4\tau^2\pi^2} \cdot \frac{1}{\pi^2\tau^2 r^2 - J\tau r - \eta}$$

$$\Leftrightarrow 0 = \pi^2\tau^2 r^4 - J\tau r^3 - \eta r^2 - \frac{\Delta^2}{4\tau^2\pi^2}$$

$$\Leftrightarrow 0 = \pi^2\tau^2 r^4 + (-J\tau r - \eta) r^2 + \frac{-\Delta^2}{4\tau^2\pi^2}.$$

This is an equation of the form $ar^4 + br^2 + c = 0$ and the corresponding generic solutions are:

$$r_{1,2,3,4} = \pm\sqrt{\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}}.$$

In the studied case, r is a firing rate so it cannot be negative nor an imaginary number. Thus, there is a unique biologically plausible solution:

$$\begin{aligned} r_0 &= \sqrt{\frac{-(-J\tau r_0 - \eta) + \sqrt{(-J\tau r_0 - \eta)^2 - 4\pi^2\tau^2(\frac{-\Delta^2}{4\tau^2\pi^2})}}{2\pi^2\tau^2}} \\ &= \frac{1}{\sqrt{2}\pi\tau}\sqrt{J\tau r_0 + \eta + \sqrt{(J\tau r_0 + \eta)^2 + \Delta^2}}. \end{aligned}$$

Therefore, the set of the first coordinates of the system (3)'s fixed points corresponds to the set of fixed points of the equation:

$$r = \Phi(J\tau r + \eta), \tag{8}$$

where $\Phi(x) = \frac{1}{\sqrt{2}\pi\tau}\sqrt{x + \sqrt{x^2 + \Delta^2}}$.

Now it is possible to construct a dynamical system in one dimension that has the same steady states for the firing rate as the original model when $I(t) = 0$, but that depends only on r, as it is standard in classical mean field models [6]. The heuristic model for this new system is:

$$\tau \dot{r} = -r + \Phi(J\tau r + \eta + I(t)). \tag{9}$$

However, there are not exactly the same dynamics for the firing rate $r(t)$ in the heuristic 1-dimensional as there are in the original 2-dimensional model. Linear stability analysis reveals that the stable focus in the complete model corresponds to an attracting node in the 1-dimensional model. This difference results in the disappearance of the *Clearance* behaviour [3] of the forced system, for any amplitude and frequency of the external input. Therefore the subthreshold dynamics are necessary to model the behaviour of the studied network.

Indeed, there is a relation between spiking synchrony, and thus an attracting focus, and the switch from the active to the inactive state. Consider the system without forcing and with parameters J, η such that there are two stable states (r_1, v_1) and (r_2, v_2), and the saddle point (r_{SP}, v_{SP}).

The phase-plane dynamics in the 1D model are:

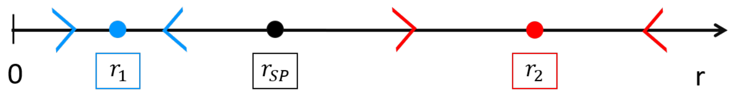


Figure 7: Phase portrait of the 1-dimensional model without external input

Whereas the phase-plane in the full model are:

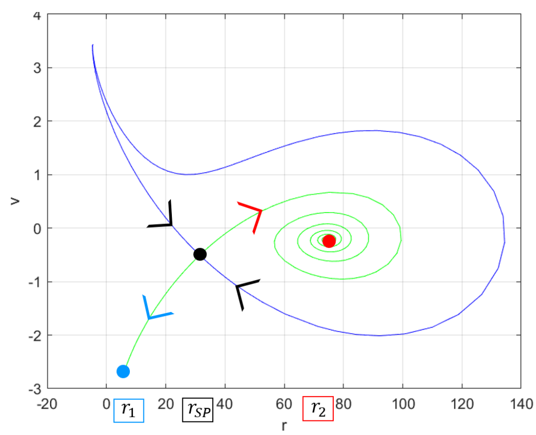


Figure 8: Phase portrait in the full model without external input

From these portraits it can be deduced that no excitatory perturbation on the 1D model (Fig. 7) can force it to change from the active to the inactive state. On the contrary, in the 2D- model (Fig. 8), the stable invariant manifolds of the saddle (in blue) have a particular form that allows and causes the switch. Indeed, if an excitatory perturbation is strong enough, the solution with high firing rate could leave the region of attraction of the stable focus and then be driven to the stable node, which corresponds to the inactive state. This is why the *Clearance* phenomenon can occur in the full model, while it never happens in the 1 dimensional simplification.

This dynamical analysis is useful to describe the different behaviours observed in the numerical results. Applying a constant excitatory input ($I(t)$ 7 is mostly positive), has the same effect on the invariant manifolds as if a higher η parameter is fixed in the non forced system. If the frequency of the input is low enough, the system will always tend to the high-activity state because, even when taking as initial condition the resting state, the maximal value of input will force the disappearance of this stability point and it will move into the region of attraction of the focus before the resting state becomes stable again (*Recall*, Fig.6,(a) low). This phenomenon will be further studied in Section 7.

However, when the frequency of the input increases, the system does not have time to get close enough to the focus and gets attracted by the node again. On the other side, the frequency of the manifolds transformation is still too low to make the high-activity state sufficiently leave the proximity of the focus and move to the stable node (Fig.6, (b), low).

Then, when the frequency of the input is further increased, the high-activity state is pushed outside the region of attraction of the focus and moves to the low-activity state before the zone of attraction of the focus gets bigger again (*Clearance*, Fig.6, (c) low). This phenomenon will be discussed in Sections 5 and 6.

Finally, for high frequencies, the invariant manifolds change too fast to have an impact on the trajectories (Fig.6, (d), low).

5. Linearized response of the network under periodic forcing

It has been explained why the *Clearance* behaviour can occur in the studied model, the aim is now to understand its relation and sensitivity to the input parameters. We have mentioned that this phenomenon is linked to the subthreshold dynamics and thus, to the spiking synchrony in the high-activity state. Therefore, this section and the following one (Section 6) will provide a further analysis of the dynamics organized by the stable focus.

Thanks to the simplicity mean-field equations (3), it is possible to compute analytically the linear response of the system under forcing. In particular, it is possible to find a resonant frequency of the high-activity state (i.e. the focus), already noticeable for low amplitude external inputs.

5.1 Dynamics near the stable focus

As explained in paragraph 3.1, a stable focus is a fixed point of the dynamical system for which the Jacobian matrix has complex eigenvalues with negative real part. The magnitude of the real part describes how "fast" the trajectories approach the fixed point: the trajectories will describe turns around the focus while getting closer to it at a rate that depends on the real part. And the value of the imaginary part corresponds to the angular frequency of the focus' inherent oscillations (i.e. the damped oscillations that appear when approaching the high-activity steady state). This frequency proper to the focus is in fact a resonant frequency for the system and we will now compute it.

Consider the system in equation (3) without forcing (that is, $I(t) = 0$). The Jacobian matrix M evaluated at the focus (r_F, v_F) is:

$$M(r_F, v_F) := \begin{pmatrix} \frac{2}{\tau} v_F & \frac{2}{\tau} r_F \\ J - 2\pi^2 r_F \tau & \frac{2}{\tau} v_F \end{pmatrix}, \tag{10}$$

and the corresponding eigenvalues are given by:

$$det(M(r_F, v_F) - \lambda Id) = \left(\frac{2}{\tau} v_F - \lambda\right)^2 - \frac{2}{\tau} r_F (J - 2\pi^2 r_F \tau) = \lambda^2 - \frac{4 v_F}{\tau}\lambda + \frac{4 v_F^2}{\tau^2} - \frac{2 J r_F}{\tau} + 4\pi^2 r_F^2$$

$$\Rightarrow \lambda_\pm = \frac{2}{\tau} v_F \pm \sqrt{-2 r_F (2\pi^2 r_F - \frac{J}{\tau})}.$$

As (r_F, v_F) corresponds to a stable focus, the obtained eigenvalues have non zero imaginary part. Therefore it is assumed that $2\pi^2 r_F > \frac{J}{\tau}$. The resonant angular frequency is then $\omega_{res} = \sqrt{2 r_F (2\pi^2 r_F - \frac{J}{\tau})}$, and the eigenvalues are $\lambda_\pm = \frac{2}{\tau} v_F \pm i \cdot \omega_{res}$. (A small error was found in the computation of the resonant frequency in the studied paper [3]). With the numerical parameters given in Fig. 6, and without forcing, the focus has coordinates $(r_F, v_F) \approx (72.87, -0.22)$. Thus, the linear resonant frequency for the focus in this setting is $freq_{res} = \frac{\omega_{res}}{2\pi} \approx 37.35$ Hz.

5.2 Fourier decomposition of the linearized system

The study of the system's linear response provides a first description of the dynamics and behaviour of the forced model, near the steady states. We consider the Fourier decomposition of the T-periodic external input with frequency *freq* $(= \frac{1}{T})$ and angular frequency $\omega := 2\pi \cdot \mathit{freq}$:

$$I(t) = I_0 + \sum_{k=1}^{\infty} I_k e^{ik\omega t}.$$

Forced by this type of periodic input (at least for small amplitudes) the response of the system is T-periodic as well. One can then give expressions for the firing rate $r(t)$ and the membrane potential $v(t)$ as Fourier series around an initial fixed point (r_0, v_0):

$$r(t) = r_0 + \sum_{k=1}^{\infty} r_k e^{ik\omega t}, \qquad v(t) = v_0 + \sum_{k=1}^{\infty} v_k e^{ik\omega t}.$$

In order to find an approximation of the Fourier coefficients in a close neighbourhood of a fixed point, it is sufficient to only consider the linear response of the model equations (and ignore the quadratic terms). We are interested here in the linear response of the system in the vicinity of the focus (i.e. $(r_0, v_0) = (r_F, v_F)$).

On one side the linearized system is:

$$\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} \frac{2}{\tau} v_F & \frac{2}{\tau} r_F \\ J - 2\pi^2 r_F \tau & \frac{2}{\tau} v_F \end{pmatrix} \cdot \begin{pmatrix} r - r_F \\ v - v_F \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{I(t)}{\tau} \end{pmatrix}.$$

$$\Leftrightarrow \begin{cases} \dot{r} = \frac{2v_F}{\tau} r + \frac{2r_F}{\tau} v - \frac{4v_F r_F}{\tau} \\ \dot{v} = (J - 2\pi^2 r_F \tau) r + \frac{2v_F}{\tau} v - (J - 2\pi^2 r_F \tau) r_F - \frac{2v_F^2}{\tau} + \frac{I(t)}{\tau} \end{cases}.$$

On the other side we consider the Fourier series for the linearized $r(t)$, $v(t)$ with initial point the focus (r_F, v_F), and we differentiate them with respect to time :

$$\begin{cases} \dot{r} = \sum_{k=1}^{\infty} ik\omega r_k e^{ik\omega t} \\ \dot{v} = \sum_{k=1}^{\infty} ik\omega v_k e^{ik\omega t} \end{cases}.$$

We proceed coefficient by coefficient and we obtain that, for $k \geq 1$, the coefficients r_k, v_k must satisfy:

$$\begin{cases} ik\omega r_k = \frac{2v_F}{\tau} r_k + \frac{2r_F}{\tau} v_k \\ ik\omega v_k = (J - 2\pi^2 r_F \tau) r_k + \frac{2v_F}{\tau} v_k + \frac{I_k}{\tau} \end{cases}.$$

We solve these equations and obtain that, for $k \geq 1$:

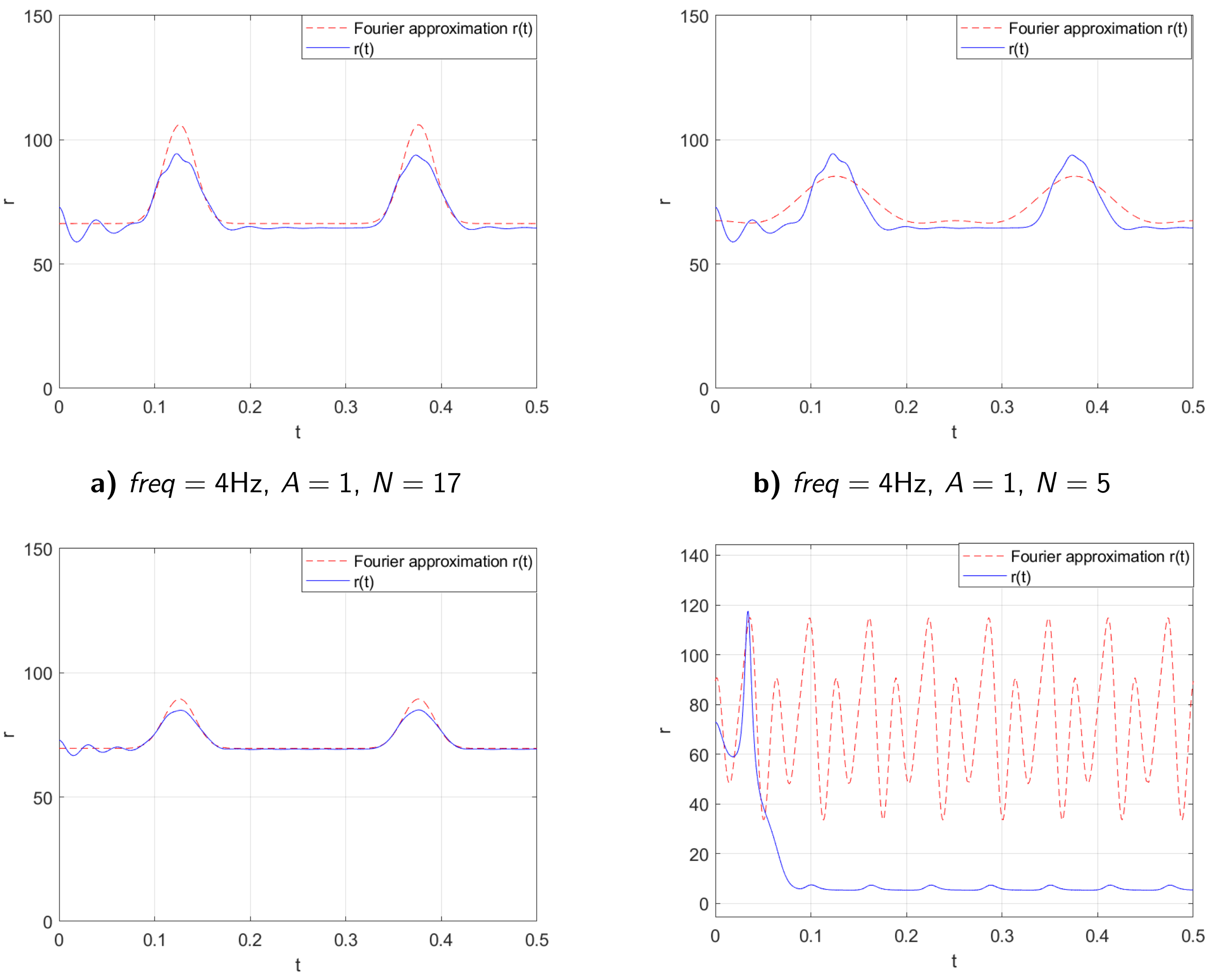
$$r_k = 2r_F I_k \Omega_k^{-1}, \qquad v_k = (\tau ik\omega - 2v_F) I_k \Omega_k^{-1},$$

with

$$\Omega_k = (\tau ik\omega - 2v_F)^2 + \tau^2 \omega_{res}^2.$$

We can easily recognize the presence of the focus resonant frequency in all terms. Finally, the full description of the linearized focus' response as Fourier series is obtained:

$$r(t) = r_F + \sum_{k=1}^{\infty} 2r_F I_k \Omega_k^{-1} e^{i\omega t}, \qquad v(t) = v_F + \sum_{k=1}^{\infty} (\tau ik\omega - 2v_F) I_k \Omega_k^{-1} e^{i\omega t}. \tag{11}$$



a) $freq = 4\text{Hz},\ A = 1,\ N = 17$ **b)** $freq = 4\text{Hz},\ A = 1,\ N = 5$

c) $freq = 4\text{Hz},\ A = 0.5,\ N = 17$ **d)** $freq = 16\text{Hz},\ A = 1,\ N = 17$

Figure 9: Different plots of the firing rate $r(t)$ (blue) and Fourier approximation of the linearized response of $r(t)$ (red) with the stable focus as initial condition and under the forcing of the external input $I(t)$ (7) with frequency *freq* and amplitude A. We considered N Fourier coefficients to compute the Fourier approximations of the firing rate, voltage and external current.

Figure 9 presents different superposed plots of the firing rate from the forced system in equation (3) and the Fourier approximation of the linearized firing rate response, when taking as initial condition the stable focus. As expected, using a large number of Fourier coefficients allows to have a better approximation for

$r(t)$, as confirms plot **b)**. A comparison of the two left plots (**a), c)**) suggests that the smaller the input amplitude, the better the Fourier approximation. This hypothesis is confirmed by the results in [3] which show that under weak forcing, the system's response is qualitatively identical to the linearized system's response.

Finally, plot **d)** invites to conjecture that the *Clearance* phenomenon does not occur when considering uniquely the linear response of the system. Indeed, Figure 10 (extracted from [3]) corroborates that for all input frequencies, beta range included, there exists a high-activity attracting orbit. This observation indicates that the *Clearance* behaviour cannot be explained by the linear response of the model, and thus the non linear effects have to be taken into account. Nevertheless, the linearized system already provides information on the resonant frequency for the high activity state (See Fig.10), which indeed seems to coincide with the frequence of the focus' inherent oscillations ($freq_{res}$) computed in paragraph 5.1. Moreover, sub-harmonic resonances of the linearized system in range $[12, 38]$Hz can also be observed.

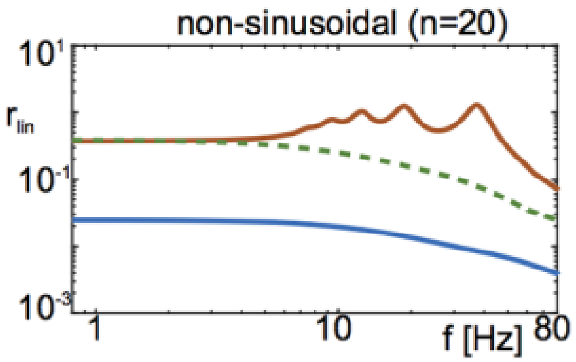


Figure 10: (Figure from Schmidt et al. [3]) Linear response of focus (red), saddle (green) and node (blue) to input $I(t)$ (7). The focus exhibits a characteristic resonant response at approximately 38Hz as well as sub-harmonic resonances. The function $r_{lin}(freq)$ is derived in [3] from the Fourier approximation of the firing rate in (11):

$$r_{lin}(2\pi \cdot freq) = \frac{\max_t r(t) - \min_t r(t)}{2}.$$

6. Analysis of the forced system by means of the stroboscopic map

In the previous section we advanced that, to explain phenomenon *Clearance*, the nonlinear resonance of the focus needs to be considered. This section analyses the complete response of the network to forcing for different frequency values by means of the stroboscopic map (also known as Poincaré map).

From this section ahead, the model equations (3) will be referred by the vector field f_t. That is $\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = f_t(r, v)$ with

$$f_t(r, v) := \begin{pmatrix} \frac{\Delta}{\pi \cdot \tau^2} + \frac{2}{\tau} rv \\ \frac{v^2}{\tau} + Jr + \frac{\eta}{\tau} - \pi^2 \tau \cdot r^2 + I(t) \end{pmatrix}. \tag{12}$$

6.1 Fixed points of the stroboscopic map

When an oscillatory input $I(t)$ of period $T = \frac{1}{freq}$ is applied to the system, there are no longer fixed points. However, we still want to identify invariant objects in order to understand the different dynamics. In particular, the aim of this section is to analyse further the conditions leading to the "disappearance" of a high activity attracting state, that is, the *Clearance* behaviour.

The stroboscopic map is a useful tool for finding periodic orbits of a system forced by a periodic input. Let $\varphi^t(r_0, v_0)$ be the flow of f_t (i.e. $\varphi^t(r_0, v_0) := (r(t), v(t))$ solution of $\frac{d(r,v)}{dt} = f_t(r, v)$ with initial condition $(r(0), v(0)) = (r_0, v_0)$). Given T the period of the external input, the stroboscopic map $P_T : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $P_T(x_0, v_0) = \varphi^T(r_0, v_0)$. In words, this map outputs the state of the system (3) under forcing after time T, for a given initial condition (r_0, v_0).

The aim now is to find fixed points of the map P_T. Indeed, if they exist, these points correspond to periodic orbits of the forced system. Thus, we are looking for points (r, v) for which $G(r, v) := P_T(r, v) - (r, v)$ vanishes. This will be done by means of the Newton-Raphson method, which involves the differentiation of map P_T that is only known numerically.

Suppose $G(r_i, v_i) \neq 0$ for some (r_i, v_i). The goal is to find some $(r_{i+1}, v_{i+1}) = (r_i, v_i) + (\Delta r_i, \Delta v_i)$ such that $G(r_{i+1}, v_{i+1}) = 0$. Let DG be the Jacobian matrix of G. If the quadratic terms are omitted in the Taylor expansion, $G(r_{i+1}, v_{i+1})$ can be approximated by $G(r_i, v_i) + DG(r_i, v_i) \cdot (\Delta r_i, \Delta v_i)^T$. Observe that $DG(r_i, v_i) = DP_T(r_i, v_i) - Id$, where DP_T is the Jacobian matrix of the map P_T and Id is the identity map in dimension 2.

Now follows the computation of $DP_T(r, v)$. We denote by $Df_t(r, v)$ the Jacobian matrix of the vector field f_t (defined as in equation (10)). The variational equations [8] are:

$$\frac{dDP_t(r, v)}{dt} = Df_t(r, v) \cdot DP_t(r, v). \tag{13}$$

Let us define $DP_t(r, v) =: \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}$. Then we obtain the following set of equations:

$$\begin{cases} \dot{r} = \frac{\Delta}{\pi\tau^2} + \frac{2}{\tau}vr, \\ \dot{v} = \frac{v^2}{\tau} + Jr + \frac{\eta}{\tau} - \pi^2\tau r^2 + \frac{I(t)}{\tau}, \\ \dot{a} = \frac{2}{\tau}va + \frac{2}{\tau}rc, \\ \dot{b} = \frac{2}{\tau}vb + \frac{2}{\tau}rd, \\ \dot{c} = (J - 2\pi^2 r\tau)a + \frac{2}{\tau}vc, \\ \dot{d} = (J - 2\pi^2 r\tau)b + \frac{2}{\tau}vd. \end{cases}$$

The two first equations are the mean-field equations (3) whereas the other four equations correspond to the variational equations (13). Taking as initial conditions (r_i, v_i) and defining $DP_0(r_i, v_i) = Id$, that is $(r(0), v(0)) = (r_i, v_i)$ and $(a(0), b(0), c(0), d(0)) = (1, 0, 0, 1)$, it is possible to integrate numerically the equations from $t = 0$ to $t = T$ and obtain $DP_T(r_i, v_i)$.

It remains to solve $G(r_i, v_i) + (DP_T(r_i, v_i) - Id) \cdot (\Delta r_i, \Delta v_i)^T = 0$. That is, finding $(\Delta r_i, \Delta v_i)$ such that:

$$\begin{pmatrix} a(T) - 1 & b(T) \\ c(T) & d(T) - 1 \end{pmatrix} \begin{pmatrix} \Delta r_i \\ \Delta v_i \end{pmatrix} + G(r_i, v_i) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Finally, if $G(r_{i+1}, v_{i+1})$ is small enough (up to some tolerance), (r_{i+1}, v_{i+1}) is considered a fixed point of the stroboscopic map. Otherwise, the procedure is iterated until the error is smaller than the chosen tolerance.

6.2 Periodic orbits of the forced system

Two oscillatory systems, one forcing the other, are in $p : q$ *phase* $-$ *locking* relation if, when the forcing system turns q times, the other completes p cycles. In the context of this work, the stimulating oscillator corresponds to the input $I(t)$ and the second system is the one governed by the vector field f_t (12).

As previously mentioned, if they exist, the fixed points of the stroboscopic map P_T correspond to periodic orbits in the forced system (3) and thus provide information on its $1 : 1$ phase-locking properties. Indeed, as these points are bound to return to their initial state, the trajectory of any fixed point of P_T after time T will be a periodic orbit in the forced system.

It is well known [9] that if a system has attracting fixed points and an oscillatory input with sufficiently small amplitude is applied, then a periodic orbit will appear around each one of those points. Or, equivalently, the stroboscopic map will have fixed points close to the non-forced system's fixed points. For the interest of the model studied in this work, we plot the periodic orbits around the focus (r_F, v_F) for increasing external input frequency. The previous numerical results (Fig. 6) suggested that, after some critical frequency between 4Hz and 16Hz, no periodic orbit exists around the point for the given parameters.

Figures 11 and 12 seem to indicate that, when the external input has amplitude $A = 1$ and its frequency gets higher than 13 Hz, the only fixed points for the stroboscopic map are located near the stable node as the drawn orbits appear only around the low-activity state. Then, when the frequency increases up to 34Hz, periodic orbits reappear around the high activity state (but the low activity state still has periodic orbits around).

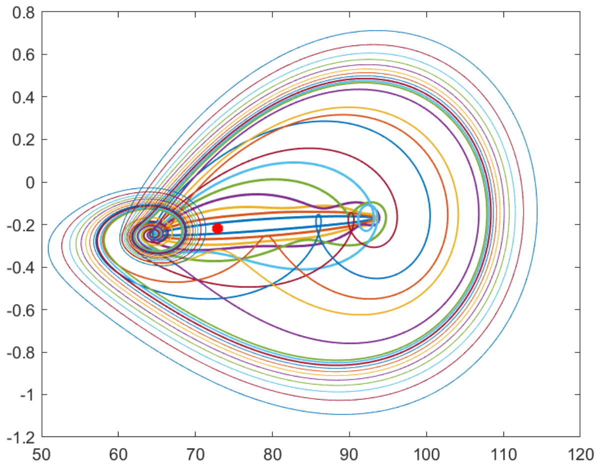


Figure 11: Periodic orbits near the focus (red point) computed using the stroboscopic map P_T for different external input frequencies and fixed amplitude $A = 1$. Frequencies increase from 1Hz (thicker line) to 13Hz (thinner line).

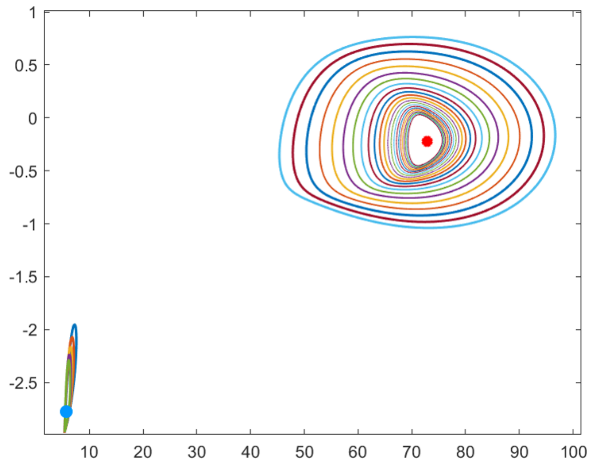


Figure 12: Periodic orbits computed from the fixed points of the stroboscopic map P_T for different external input frequencies and fixed amplitude $A = 1$. Frequencies increase from 15Hz (thicker line) to 80Hz (thinner line). First, we only observe orbits around the stable node (light blue point) and then they reappear also around the stable focus (red point) of the non-forced system.

To obtain further information on the system's bifurcations, it is useful to keep track of the eigenvalues λ_1, λ_2 of the matrix DP_T associated to each fixed point of P_T. The eigenvalues vary between complex and real values, depending on the attraction type of the trajectory under the stroboscopic map. Figure 13 shows the absolute values of these eigenvalues and, whenever they are complex, $|\lambda_1| = |\lambda_2|$. However, when the eigenvalues are real, their absolute values are different and this explains why both curves do not coincide in the figure for some frequency ranges.

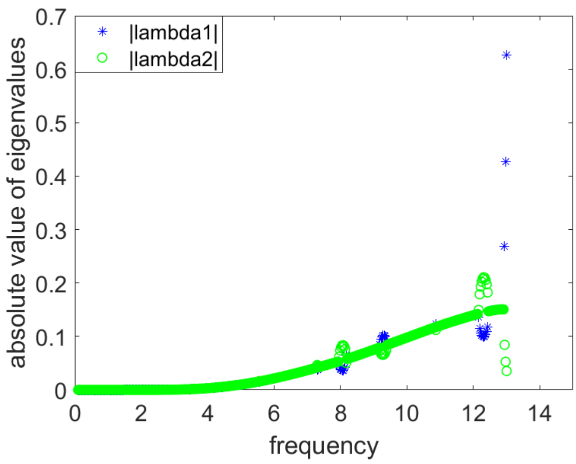


Figure 13: Absolute value of the eigenvalues for the matrix DP_T evaluated in the stroboscopic map's fixed point near the focus (r_F, v_F) for increasing input frequencies.

It can be noticed that, until reaching an input frequency of 13Hz, both values $|\lambda_1|, |\lambda_2|$ remain below 0.2. Then, between 13Hz and 13.1Hz, the absolute value of one eigenvalue increases very fast approaching the value of 1. This indicates the presence of a saddle-node bifurcation of fixed points of the stroboscopic map, which corresponds to a saddle node bifurcation of periodic orbits for the forced system (3).

According to the results of [3], the bifurcation structure of the forced system is indeed governed by SN bifurcations of periodic orbits, but also period-doubling (PD) bifurcations (See Fig.14). For large enough input amplitudes, these bifurcations lead to a frequency range where all periodic orbits in the high activity state are unstable. As a result, the only attracting orbit of the system is near the low activity attracting node, which explains the *Clearance* behaviour for this input frequency band.

It should also be remarked that this frequency range corresponds to the one causing a resonant response in the linearized system (Fig. 10). Therefore, the frequencies effective in switching off the active states depend on the resonant frequency of the focus' inherent oscillations, which themselves are determined by the subthreshold dynamics of the system. The *Clearance* frequency range depends then on the model parameters (J, τ, Δ, η), as they are all involved in the definition of the resonant frequency.

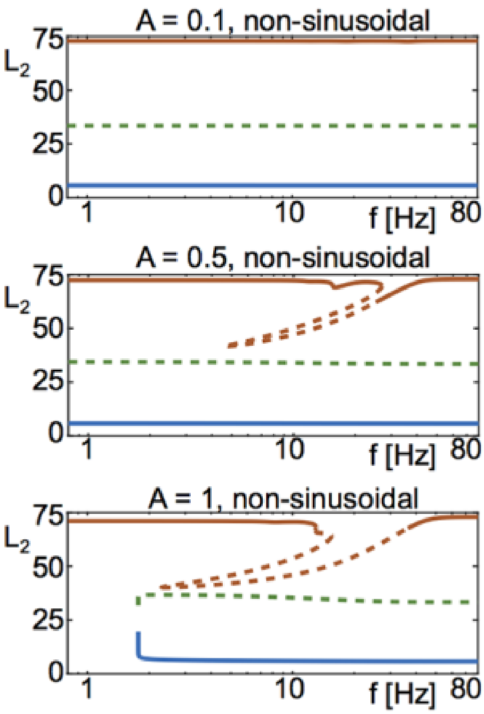


Figure 14: (Figure from Schmidt et al. [3]) Non-linear response of focus (red), saddle (green) and node (blue) by means of bifurcation analysis to input $I(t)$ (7) for different amplitudes. For each input frequency and each steady state, the figure shows the L_2-norm of the firing rate along the corresponding periodic orbit. The solid (resp. dashed) lines indicate that the periodic orbit is stable (resp. unstable).
We can see that, for small amplitudes ($A = 0.1$ in this figure), the network response is similar to the linear response: there is a stable periodic orbit near each non forced fixed point at any input frequency. As A increases, a richer bifurcation structure appears. In particular, for $A = 1$ several bifurcations occur and are responsible for the *Recall* and *Clearance* behaviours.

7. Mechanism underlying the "switching on" of the neuronal network

The previous sections were dedicated to the study of the "switching off" of the activated states. The current section studies the mechanisms involved in the "switching on" of the network, observed in Figure 6 (a). This is the switch corresponding to the *Recall* phenomenon in the context of working memory tasks.

When the system is forced with a low-frequency input, it can be studied as an almost stationary response. As time goes and $I(t)$ slowly varies, the time evolution of the firing rate almost tracks the system's fixed point for the current value of $I(t)$. As the switch to the active state happens when the input current has a low frequency (Fig. 6 (a)), this phenomenon can be studied by analysing the steady states of the system for different fixed values of I.

Consider the bifurcation diagram of the model with a chosen parameter J such that the system is in a bistable regime. This range is delimited by two saddle-node bifurcations, occurring at η_{c1} and η_{c2} (See boundary of bistability in Fig. 3). In order to understand the mechanisms that push the system from the vicinity of one stable state to the other, we must consider the system with a chosen parameter η_0 with $\eta_{c1} < \eta_0 < \eta_{c2}$ (Fig. 15), to which the external current $I(t)$ is applied. It will be shown that the switching depends on the amplitude A of the external input.

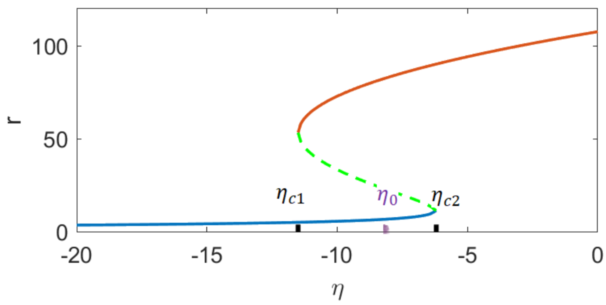


Figure 15: Bifurcation diagram of stationary states with critical values for saddle node bifurcation (η_{c1}, η_{c2}) and a chosen parameter η_0. Parameters: $J = 15\sqrt{\Delta}, \Delta = 2$.

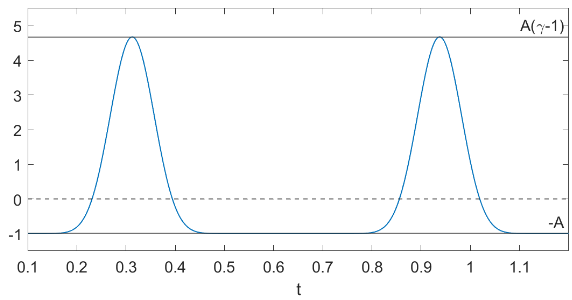


Figure 16: Non-sinusoidal external current (7) with indicated minimum and maximum values. Parameters: $A = 1$, $freq = 1.6$Hz, $\gamma \approx 5.68$.

The studied input current (7) oscillates between values $-A$ and $A(\gamma - 1)$ (see Fig. 16). Then, the minimal amplitude required to push the inactive steady-state of the system past the saddle-node bifurcation (η_{c2}) (force it to go the upper branch) and change it to an active state is A_1 such that:

$$\eta_0 + A_1(\gamma - 1) = \eta_{c2}$$
$$\Leftrightarrow A_1 = \frac{\eta_{c2} - \eta_0}{\gamma - 1}.$$

And the maximum amplitude up to which the system stays in the high activity state is A_2 such that:

$$\eta_0 - A_2 = \eta_{c1}$$
$$\Leftrightarrow A_2 = \eta_0 - \eta_{c1}.$$

For amplitudes below A_1 the system is unable to "jump" from the inactive to the active state and will just present oscillations of low firing rate. When the external input frequency A is greater than A_1, but also greater than A_2, the system will first jump from the inactive state to the active state but, as the oscillatory external input decreases to its lowest value, the system will switch back to the inactive state ("fall" to the lower branch). In this case the system constantly varies between the steady and the active state.

In conclusion, the *Recall* behaviour occurs for an external input with amplitude between A_1 and A_2. With this condition, the switch from inactive to active state happens and then the firing rate keeps oscillating around high values. Finally, if the parameters of the model are such that $A_2 < A_1$, then there is no range of amplitude values at which the robust "switching on" occurs and increasing the input amplitude directly leads the system to a constant switching between states. Thus, the range of suitable input amplitudes A allowing a *Recall* behaviour must verify the condition $A_1 \leq A \leq A_2$, so it is dependant on the network parameters.

Explanation of the Recall behaviour in numerical computations for Fig. 6:
In Figure 6 the plots were realized with chosen parameters $\eta_0 = -10$, $\tau = 0.02$, $J = 15\sqrt{2}$ and $A = 1$ (as indicated in the caption). With this setting it is possible to numerically compute the coordinates of the two saddle node bifurcations by solving the set of equations (4), (5) and (6). The obtained result is $\eta_{c1} \approx -11.49, \eta_{c2} \approx -6.27$. Then, the current $I(t) = A(\gamma \cdot \sin(\pi \cdot freq \cdot t)^{20} - 1)$ is applied, where $\gamma \approx 5.68$ (independent of A) was approximated numerically. Thus, applying the previous definitions, $A_1 \approx 0.80$ and $A_2 \approx 1.49$ in this setting. As the amplitude used in the plot was $A = 1$ and $A_1 < 1 < A_2$, the phenomenon of *Recall* for low input frequencies was to be expected, and indeed happens (Fig. 6 (a)).

In Figure 14, extracted from [3], it can be seen that for lower input amplitudes (0.1 and 0.5 Hz) there is always an existing attracting orbit near the low activity state. This implies that a "switching on" of the resting state would not occur in those cases, even for low input frequencies.

8. Application of the model in a setting of working memory

In the previous sections (Sections 3 to 7) the model (3) has been extensively analysed in order to understand the mechanisms underlying the role of an oscillatory input $I(t)$ in activating and clearing neuronal states of a single neuronal population. In particular it has been revealed that the amplitude of the input is important for the existence of the *Recall* and the *Clearance* behaviours, and that the second one happens only for well chosen input frequencies (close to the resonant frequency of the model's stable focus).

The aim of this last section is to reproduce the studied mechanisms in a setting of multiple interconnected neuronal networks involved in the realization of a working memory task.

8.1 Working memory and its relation with the model

8.1.1 What is the working memory?

Long-term memory is the memory we use to hold in our mind information for a long period, such as our sisters' name or our parents' birthdays, whereas we use the short-term memory or *working memory* (WM) to memorize things for a short period of time (usually less than a minute) [10]. The second type of memory is the system we use, for example, to remember the random number sequence we were assigned at the doctor's waiting room. We keep the sequence in our memory, expecting to be called by the physician, but as soon as we enter in the office we will undoubtedly forget the number.

This section is focused in the process of working memory and its different phases. As it is a temporary storage of information, the neuronal network implicated is first in a resting state, then, when the information to memorize is presented, the neurons spike and the firing rate of the population increases. This activation is the *Recall* behaviour we have been studying. While the subject is memorizing (usually a few seconds) the population remains active but, as soon as the information held is no longer useful (when the subject enters the physician's office in the previous example), the network firing rate drops back to a low state. This "switching off" phenomenon is *Clearance*, as we mentioned in previous sections.

8.1.2 The theory of the model in higher-dimensional networks

So far, this work has only been focused in the model of a single neuronal population. In the studied paper [3], the authors explore as well the effect of the periodic input $I(t)$ (7) in systems of multiple neuronal networks. They consider neuronal populations that, when isolated, follow the behaviour modelled by the studied mean-field equations (3) and are bistable. The authors propose a model for interacting neuronal populations, inspired by the single-population model, that takes into account excitatory or inhibitory synaptic connections among them. The simulations performed for different external input frequencies revealed that, in this higher-dimensional context, the robust switching between low and active states also happens. It was found that the application of an external current of low frequency (2Hz) in the presence of weak noise leads to the random activation and deactivation of some populations, whereas fast oscillations (40Hz) drive all populations to the inactive state.

These results, together with the fact that the behaviour of WM networks is accurately described by the mean-field model (3), motivated us to create a working memory model with several interacting populations.

8.2 A working memory model

Inspired by the PhD thesis of João Moura Barbosa [4], we propose a study of the effect of oscillations in a simple setting of working memory. The model proposed represents the neuronal behaviour in a subject realizing a simple short-term memorization task (Fig. 17). The experiment is as follows: the subject sees an image with two circles, one blue and one green, in different locations (Left and Right for instance). Then the image is removed. A short-time later, the subject is shown the image with one of the colors and must respond what was the position of that color. Alternatively, the subject might be shown a position and must indicate the color that was in that place.

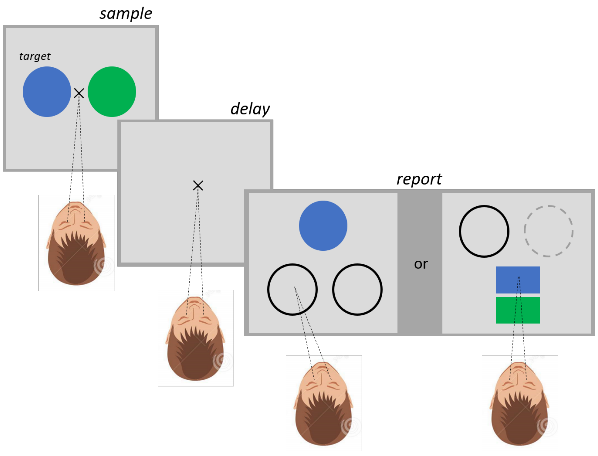


Figure 17: Working memory task. During a short period of time, the subject is shown two coloured circles: one blue and one green. The circles are located one at the right and one at the left of the sample image. After the delay period either i) the target is revealed by showing its colour and the subject must indicate its location, or ii) the location of the target is shown and the subject must report its colour. (The image was inspired by Figure 1.1.c) in [4])

To model the experiment, a simplification of the neuronal architecture considered in the PhD thesis[4] is done. We model four neuronal populations, two coding for the position (Left and Right) and two coding the color (Blue and Green). The two populations coding for the same type of information inhibit each other, while any two populations coding for a different feature mutually excite each other. Moreover, each neuronal population receives recurrent excitation (See Fig. 18). Finally, all populations have two stable states when uncoupled: either they have a low firing rate or a high firing rate.

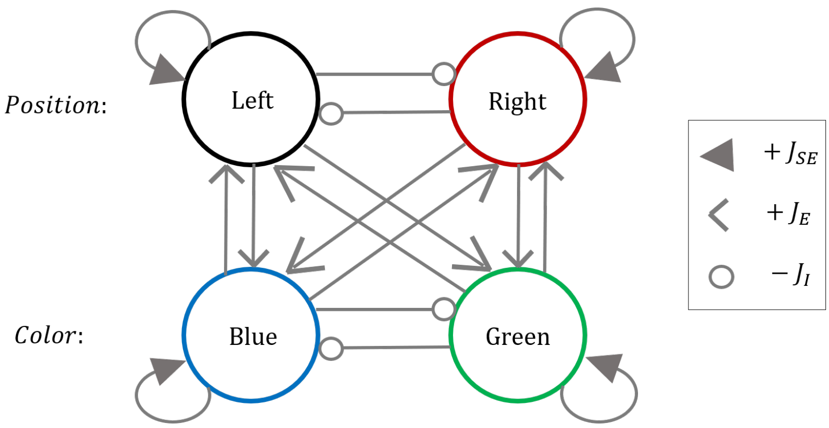


Figure 18: Four populations with self excitation (J_{SE}) and inter-connected with excitation (J_E) or inhibition (J_I)

The equations for the mean-activity of the Left population and the Blue population are:

$$\begin{aligned} \dot{r_L} &= \frac{\Delta}{\pi \cdot \tau^2} + \frac{2}{\tau} \cdot r_L \cdot v_L \\ \dot{v_L} &= \frac{v_L^2}{\tau} + J_{SE} \cdot r_L - J_I \cdot r_R + J_E \cdot r_B + J_E \cdot r_G + \frac{\eta}{\tau} + \frac{I(t)}{\tau} - \pi^2 \cdot \tau \cdot r_L^2 \end{aligned} \tag{14}$$

$$\begin{aligned} \dot{r_B} &= \frac{\Delta}{\pi \cdot \tau^2} + \frac{2}{\tau} \cdot r_B \cdot v_B \\ \dot{v_B} &= \frac{v_B^2}{\tau} + J_{SE} \cdot r_B - J_I \cdot r_G + J_E \cdot r_L + J_E \cdot r_R + \frac{\eta}{\tau} + \frac{I(t)}{\tau} - \pi^2 \cdot \tau \cdot r_B^2 \end{aligned} \tag{15}$$

The equations for the Right populations are found by exchanging L and R in system (14). Similarly, by exchanging indices B and G in system (15), we obtain the equations for the Green population. This system of four equations is a simplification of the set of equations proposed for multiple interconnected neuronal populations in the main studied research article [3].

Without loss of generality, we decided to consider only the experiment where the report consists in giving the location of the clued colour, and Blue was chosen to be the target. We explored several approaches to model the neuronal activity during the realization of the task. In all our simulations we assumed that the complete experiment had a duration of approximately 1 second, as this period was sufficient to see all the phases of the neuronal behaviour for the working memory task. Nevertheless, depending on the experiment we want to simulate, the models can be generalised for a longer period of time.

The two models presented in this section follow the same approach. In fact, the second one provides a further development of the first one. The common idea is to model the memorization part (for instance associating Left with Blue and Right with Green) by forcing each pair to exhibit oscillations that are synchronized and in opposite phase with respect to the other pair. The decoding part is modelled in different ways but in both models an external input is applied to one of the color populations, and the aim is to decode the location associated to the Blue population.

8.2.1 First model: phase synchronization between associated populations

Theory:

At the beginning of the experiment, all the populations are in a resting state, with oscillations of the form:

$$A(\gamma \cdot \sin(\pi \cdot 80 \cdot (t + phase))^{20} - 1).$$

Thanks to the previous sections we know that these fast oscillations ($freq = 80$Hz) do not provoke switches between the active and inactive state (Fig 6 (d)). They model here inherent oscillations that the neurons have (as they are never completely inactive) and will be present until the decoding is finished (i.e. the task is accomplished). The four *phases* are randomly chosen between 0 and $\frac{1}{freq}$ (which is the period of the rapid oscillations).

When the experiment starts, the image is presented to the subject and the four populations involved must be activated. First, this was modelled by adding the oscillating input that causes *Recall* in the single population model (Fig 6 (a)). But it turns out that a transient step current $I_1 = \gamma - 1$ is sufficient to activate the four populations. Then, to model the pairing in the memorization task, we synchronized two by two the phases of the fast oscillations (Blue with Left and Green with Right). When the image is removed, the neuronal activity is still color-position synchronised.

The decoding in this model is done by injecting an external transient input I_2 to the queried color neuronal population. Then, as this color was in the same phase than the associated location firing rate, and it exerts excitation on it, the wanted location is expected to follow the Blue firing rate increase, making possible its decoding. Finally, all four populations are sent to 0. Once the subject has answered the question, the task is finished.

Results:

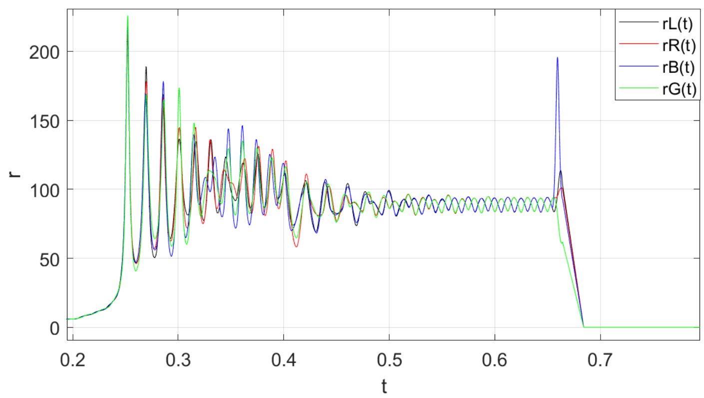
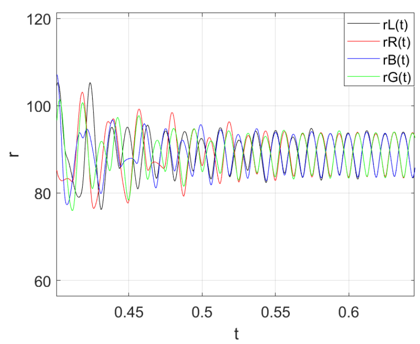
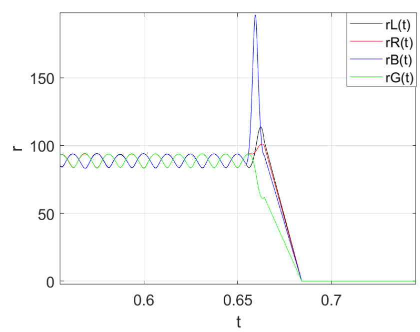


Figure 19: **Results First Model WM**. In this plot we can identify the different phases for the neuronal populations involved in the WM task: recall, memorization of the color-location pairs, correct report of the Blue location and clearance of all states. Parameters used: $\tau = 0.02, \eta = -10, \Delta = 2, A = 1, \gamma \approx 5.68$, $J_{SE} = 15 \cdot \sqrt{\Delta}, J_E = 3, J_I = 4$. Transient input for activation: $I_1 = \gamma - 1$ at $t_1 = 0.2$ s, transient input for decoding: $I_2 = 4(\gamma - 1)$ from $t_3 = t_2 + 20/freq + 1/(3 \cdot freq)$ s to $t_4 = t_3 + 0.8/freq$ s. Phase synchronization at $t_2 = 0.4$.



a) Synchronization: The color population gets synchronised with the corresponding location population (in this case Blue with Left and Green with Right).



b) Decoding: an external current is applied to the Blue population. The location population (Left here) that has an immediate increase of its firing rate is the location that was memorized for the Blue circle.

Figure 20: First model WM: zoom synchronization and decoding

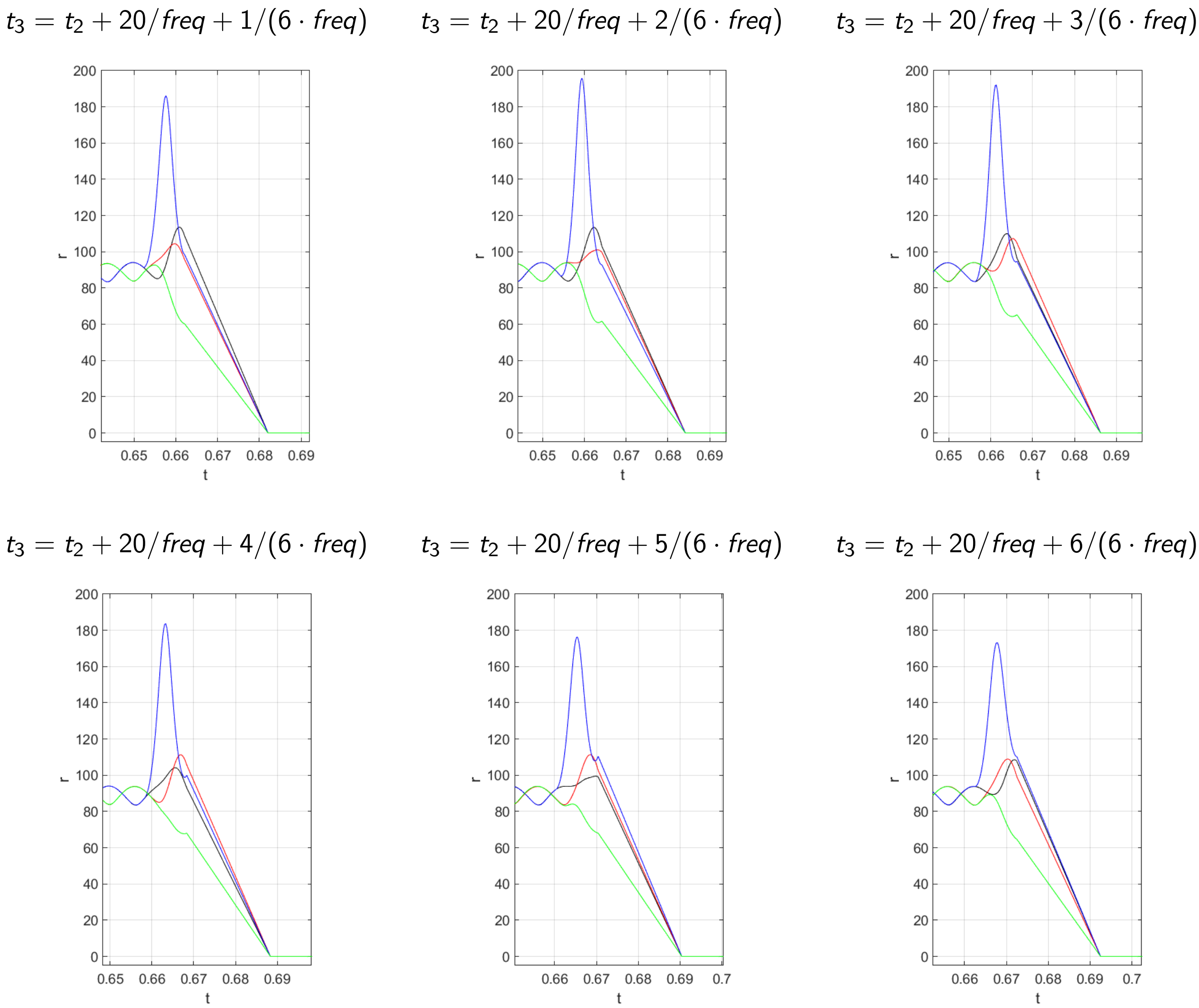


Figure 21: First model WM: decoding at different times t_3. If the decoding step current is applied when the Blue mean firing goes from its maximum to its minimum ($1/(6 \cdot freq)$ to $3/(6 \cdot freq)$), the Left population increases more than the Right population, so the decoding can be done correctly. On the contrary, the decoding is incorrect if the input is applied when the Blue mean firing rate varies from its minimal value to its maximal value ($4/(6 \cdot freq)$ to $6/(6 \cdot freq)$).

Observations:

After several plots, it was found that the outcome of this model depends considerably on the moment of decoding t_3 (See Fig. 21).

For a range of well chosen t_3, the model seems to be robust. As the Blue and Left populations are in phase and excite each other, the Left population firing rate also increases just after the Blue population receives the excitatory input. The Green population firing rate (r_G) also increases but, as it is in phase with the Right population activity which is being strongly inhibited by the Blue one, its increment is reduced. Therefore, a correct decoding is possible (See example in Figures 19 and 20). Moreover, the amplitude of the decoding step current can be moderately modified and the output will still be correct.

However, for t_3 that are not in the particular range, it is common to observe an increase on the Right population firing rate (r_R) following the Blue large spike, which would lead to an incorrect decoding. This happens because the Blue population excites both the Left and Right populations with the same intensity and, as they both experience oscillations, one of them will be more likely than the other to follow the Blue increase depending on their current oscillation phase. Unfortunately, this sensitivity to the decoding time implies the impossibility to change the target without changing the time parameter t_3 in the code. Indeed, a time t_3 that guarantees a correct decoding for the Blue location (Left in this example) will give a wrong outcome when asking the Green location. This is the consequence of having the Blue and the Green populations oscillating in anti-phase.

Possible improvements:
The presented model could be improved in several ways. First, it might be possible to define the decoding time t_3 as dependant on the phase of the target population. This could be implemented by means of trigonometric functions and it would allow us to use the same model for any chosen target.

Also, it is coherent to have small oscillations in all populations during the complete task, considering that the neuronal oscillations are widely observed in the brain. Thus, it would be more realistic if those oscillations persisted after the decoding, instead of imposing a 0 value for the four mean firing rates. In fact, it would be interesting to find a more natural way to make the populations go back to the resting state (*Clearance*). A way that does not involve manually changing their mean firing rate for a lower one.

8.2.2 Second model: synaptic facilitation

One of the improvements to do in the first model is to induce *Clearance* in a more natural way. In the previous sections and the higher-dimension model studied in [3], it was found that an oscillatory input with frequency in a specific range could drive all the targeted populations to their resting state. Unfortunately, the application of this type of current for the decoding and *Clearance* in a similar model to the previous one turned it even more dependant on the time for decoding and lead usually to incorrect results (See Annex B.1).

A new idea to overcome the phase dependency for decoding could be to explore the concept of synaptic facilitation. The brain is in constant activity: new neuronal connections arise, some old connections become weaker and other get stronger. Synaptic facilitation is a form of short-term brain plasticity that strengthens the connection between some neurons. It has the opposite effect to the usual synaptic fatigue, or short-term synaptic depression. The role of synaptic facilitation was studied in the scientific paper [11].

In the framework of this project, for the working memory problem, the connection between the memorized pairs of location and position could be strengthen as a consequence of synaptic facilitation. Further studies of the molecules implicated should be done to adequately justify the modifications in the model. But, in the meantime, to have a motivating result, we applied linear synaptic facilitation and linear synaptic fatigue to the first model (Paragraph 8.2.1). The synaptic strength parameters when the memorization starts are: $J_{SE} = 15 \cdot \sqrt{\Delta}$, $J_I = 4$, $J_{strongE}(t) = 3 + t$, $J_{weakE}(t) = 3 - t$. Then, for the example case where the Blue circle is on the Left, the diagram in Figure 22 is obtained.

The equations describing the dynamics of each population are the same as in the first model (14) - (15) with the exception of the synaptic parameters, that follow here the configuration given in Fig. 22.

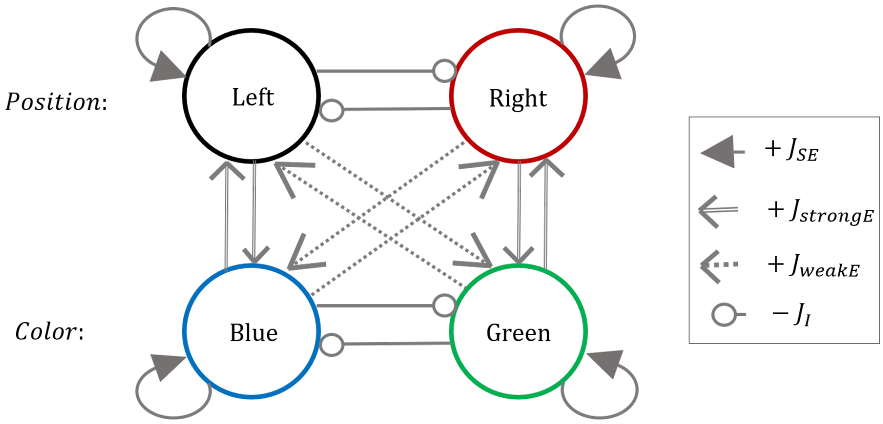
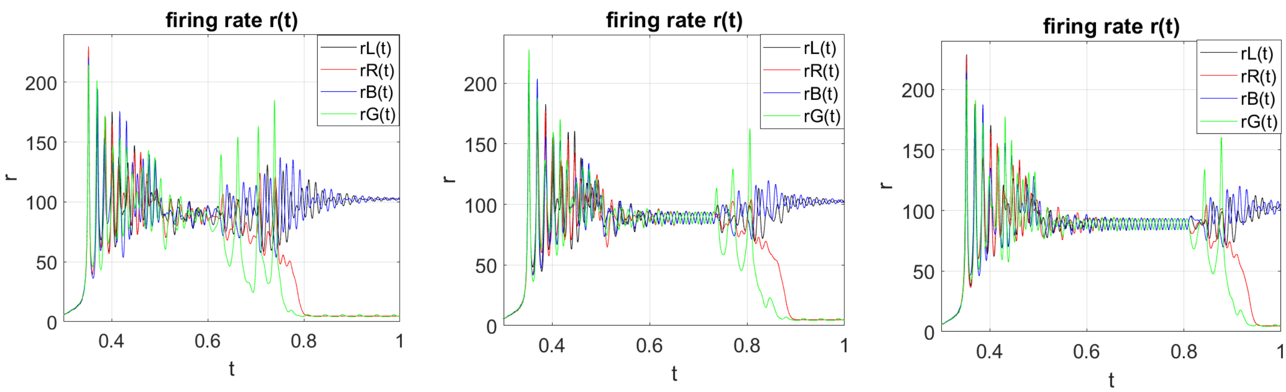


Figure 22: WM colour-position diagram with synaptic facilitation and fatigue

This model follows the same steps as the previous one but, instead of targeting the Blue population and clearing all states at the end, an external input with frequency in the clearing range is applied to the Green population so that the only color population remaining is the Blue one. As in the previous model, due to the fact that the firing rate of the Green and Right population are in phase, it is intended that the Right population follows the Green one and shifts to the inactive state. If this happens, the only active position population to still be active would be the Left population and there will be a correct decoding. Moreover, as this model considers stronger Green-Right and a weaker Green-Left synaptic connections, it is expected to be more robust than previous trials (See Annex B).

Results:



a) Decoding at t=0.6s

b) Decoding at t=0.7321s

c) Decoding at t=0.81s

Figure 23: Model applying linear synaptic facilitation and fatigue.

Observations:

The obtained plots show indeed correct outcomes for different decoding times, which is an improvement from the previous model. In this case the decoding and partial *Clearance* are a consequence of the same oscillating input, and not forced by the program.

This results are only a motivation to investigate further in this direction. As a matter of fact, it is not neurologically coherent that the synaptic facilitation and fatigue are linearly dependant of time. Probably, the synaptic strength should follow other better justified specific function [11].

8.3 Model discussion and further ideas

Both presented models proved to be very sensitive to changes in the synaptic strength J_{SE}, J_E, J_I parameters. If the inhibition is too strong or not strong enough, some or all populations never get activated. The same happens for the excitation parameter. Surely, there is a relation between the values of the synaptic weight parameters and the input amplitude necessary for activation. This dependency could be a subject for future studies.

In addition, we could consider neuronal networks for which the fast oscillations spontaneously happen. In the studied paper [3] the authors show that simple circuits of excitatory and inhibitory populations can generate oscillations of the desired form (7). It could be possible to use two populations, one excitatory and one inhibitory, to model each one of the four neuronal populations (Blue, Green, Left and Right). Then we should investigate how to induce phase synchronization to the oscillatory associated color-location pairs.

With respect to the synchronization, we also considered another approach: frequency-synchronization. As the first model revealed that a phase-synchronization of the inherent oscillations was not sufficient to have a correct decoding, the idea arose that synchronizing the pairs with respect to their oscillating frequency could give better results. Inspired by the previous sections and, particularly, by the Figure 6 (a) and (d), we decided to activate all four populations and then establish slow oscillations for the Blue-Left pair and fast oscillations for the Green-Right pair. The details of this trial are given in the Annex B.2. Unfortunately, the strength of the synaptic connections between the populations, which was necessary for their activation, was an obstacle for the frequency synchronization. However, on this line, other synchronization techniques could be explored.

9. Conclusion

In this work we analysed the role of oscillations in switching or maintaining specific neuronal states in the brain. Following the approach of the research article by Schmidt et al. [3], we studied a mean field model for a network of QIF neurons. First we considered a single network governed by the given mean field equations (3). Using tools from bifurcation analysis we could describe the different steady states and bifurcations parameters (see Fig. 2) in absence of forcing. This procedure allowed us to find the range of (J, η) values for which the system has two steady states (Fig. 3), and is likely to endure a switch from one to the other when an external current is applied.

Then, we analysed the behaviour of the system under periodic forcing $I(t)$ (7). The results presented in [3] were corroborated: for low input frequencies ($\approx$ 1.6Hz) the system switches from the low-activity to the high activity state and, for input frequencies in a band near the beta range (around 13-30 Hz), the active state is cleared (driven to the inactive state). Those two phenomena can be related to the *Recall* and *Clearance* phases in a context of working memory. In fact, the studied model was used later to describe the behaviour of neuronal populations implicated in a working memory task.

Some further analysis were proposed in order to achieve a better understanding of the two types of robust switching and their dependence on the model parameters. First, an inspection of the linearized system provided some information on the frequency band leading to a *Clearance* behaviour. Then, the study of the stroboscopic map and its fixed points revealed the presence of a saddle-node bifurcation of periodic orbits in the forced system. This bifurcation causes the disappearance of high-activity attracting periodic orbits and drives the system to the quasi-inactive state. Finally, the mechanisms involved in the *Recall* behaviour were studied in detail. The stationary analysis revealed the importance of the input amplitude in the switching from the resting state to the active state.

Going beyond the previous work [3], we proposed an application of the model in a setting of working memory (WM). This last section 8 was inspired by the results of the PhD thesis [4]. The WM task consists in presenting a sample with coloured circles in different locations to the subject and remove it after a few seconds. Once a delay period finishes, the subject is asked to report either the location of an asked colour or, alternatively, the colour of a cued location. To simplify the neuronal architecture described in [4], we considered only four interacting neuronal populations, encoding two colours (Blue and Green) and two positions (Left and Right). When analysed individually, each one of the neuronal population was bistable and its behaviour was described by the model studied in previous sections.

In this thesis report, we propose and study two complete models that try to reproduce the neuronal behaviour in the explained WM setting. In both, it was decided to model the neuronal link between the colour and the location presented in the sample by phase-synchronizing the corresponding populations. The theory of the higher dimensional version of the model presented in [3] (Section 8.1.2), as well as the previously obtained results, were essential to design a plausible WM model that enabled to decode the correct colour-location pairs. In particular, we decided to use high-frequency oscillations in the model because they proved to maintain the existent state.

The first model goes through all stages in a WM context and, under some parameter conditions, leads to an accurate decoding. However, the model presents several drawbacks and needs several changes to be robust. Specifically, the correct decoding is very sensitive on the oscillatory phase of the targeted population and the *Clearance* phase that follows has to be achieved manually. To overcome these issues, the second model improves the first one thanks to synaptic facilitation. The results given by a first rough version seem to indicate a gained robustness in the model (a correct decoding) and a more natural induction of

Clearance. Even though we did not refer to neurological observations in order to prove our implementation of the synaptic adaptation, this second model simulates remarkably well the behaviour of neuronal networks involved in the completion of the WM task. We think that this approach could be further investigated.

References

[1] G. Buzsáki (2006) Rythms of the brain. *Oxford University Press*, New York.

[2] E. Montbrió, D. Pazó, A. Roxin (2015) Macroscopic description for networks of spiking neurons. *Phys Rev X.; 5:021028*.

[3] H. Schmidt, D. Avitabile, E. Montbrió, A. Roxin (2018) Network mechanisms underlying the role of oscillations in cognitive tasks. *PLoS Comput Biol 14(9): e1006430*. https://doi.org/10.1371/journal.pcbi.1006430.

[4] J. Moura Barbosa (2019) Neural network mechanisms of working memory interference, doctoral thesis in Universitat de Barcelona. http://hdl.handle.net/10803/669202.

[5] Wulfram Gerstner, Werner M. Kistler, Richard Naud and Liam Paninski (2014) Neuronal Dynamics: From single neurons to networks and models of cognition. *Cambridge University Press*. Online book: https://neuronaldynamics.epfl.ch/online/Ch1.S3.html.

[6] G.B. Ermentrout, Terman, D. (2010) Mathematical Foundations of Neuroscience. *Springer*, New York. http://dx.doi.org/10.1007/978-0-387-87708-2

[7] Juri D. Kropotov (2009) Quantitative EEG, Event-Related Potentials and Neurotherapy. *Academic Press*. ISBN 9780123745125, https://doi.org/10.1016/B978-0-12-374512-5.00003-7. https://www.sciencedirect.com/science/article/pii/B9780123745125000037

[8] Z. Mei (2000). A Numerical Bifurcation Function for Homoclinic Orbits. In: Numerical Bifurcation Analysis for Reaction-Diffusion Equations. *Springer Series in Computational Mathematics, vol 28. Springer*, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-04177-2_8

[9] T. Breunung, G. Haller (2019). *Springer Nature*. When does a periodic response exist in a periodically forced multi-degree-of-freedom mechanical system?

[10] C. Stangor, J. Walinga (2014). Introduction to Psychology – 1st Canadian Edition. *BCcampus*. Print ISBN: 978-1-77420-004-9. https://opentextbc.ca/introductiontopsychology/

[11] SL. Jackman, WG. Regehr (2017). The Mechanisms and Functions of Synaptic Facilitation. *Neuron*. doi: 10.1016/j.neuron.2017.02.047. PMID: 28472650; PMCID: PMC5865607.

A. Numerical details

A.1 Computing invariant manifolds:

Consider f_t the vector field (12) describing the studied dynamical system. Let $s^* = (r_{SP}, v_{SP})$ be the saddle point of the system, that has suitable parameters so that the saddle point exists ((J, η) in the bistable range in Fig. 3). The first step to construct the invariant manifolds of s^* is to compute the eigenvalues of the Jacobian matrix of f_t evaluated at point s^* (as in 10). As s^* is a saddle point, one eigenvalue is positive (λ_+) and the other negative (λ_-), and their respectively associated eigenvectors can be called e_+ and e_-. If the system (equation 3) is integrated numerically backwards from $s^* \pm \epsilon \cdot e_-$, for ϵ sufficiently small, then a numerical approximation of the stable invariant manifolds (dotted curves in Figure 6) is obtained. Similarly, to obtain a numerical approximation of the unstable manifolds, a forward integration of the system starting at points $s^* \pm \epsilon \cdot e_+$ must be done.

A.2 Computing the bistability boundary:

The aim is to numerically compute a set of 4-dimensional points (J, η, r, v) such that, when the model has parameters (J, η), there is a saddle node bifurcation and the conjunction of the two implicated stable points has coordinates (r, v). This will allow us to compute the model's bistability boundary (Fig. 3).

The computations are based in the Newton-Raphson method. Without loss of generality, this subsection will only explain the computation for the left bound of the boundary, i.e. the points corresponding to the first saddle node bifurcation with parameter η_{c1} in Fig 15 (left branch of the curve in Fig. 24). The computation of the right bound follows the same steps.

To start the procedure a J_1 is chosen and fixed. Then, the system of equations (4), (5), (6) is numerically solved for $J = J_1$ and a small initial η. This computation provides a first point of the left boundary $x_1 = (J_1, \eta_1, r_1, v_1)$. The following points will be found by approaching the boundary curve with small steps.

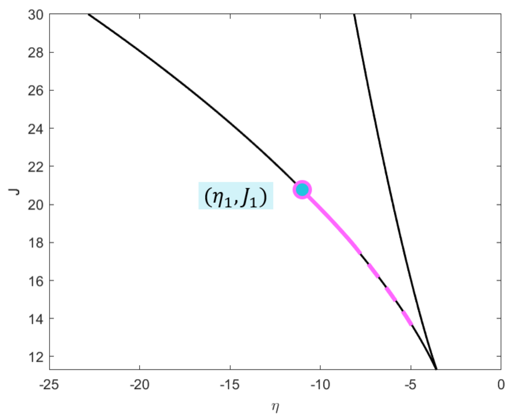


Figure 24: Coordinates J, η of the bistability boundary. The point (η_1, J_1) is an example of initial point for the computation.

The Newton-Raphson method allows to find a good approximation for the root of a real-valued function. In the case of the bistable boundary, the function of interest is the one given by equations (4), (5), (6), but considering J also as a variable. This vector-field function will be called F :

$$F(J, \eta, r, v) = \begin{pmatrix} \frac{\Delta}{\pi\tau^2} + \frac{2}{\tau} \cdot r \cdot v \\ \frac{v^2}{\tau} + J \cdot r + \frac{\eta}{\tau} - \pi^2\tau \cdot r^2 \\ \left(\frac{2v}{\tau}\right)^2 - \frac{2r}{\tau}(J - 2\pi^2\tau \cdot r) \end{pmatrix}. \tag{16}$$

We have coloured the variables to emphasize that, instead of considering the equations with J and η as parameters as we did in the main part of this thesis, here we take them as variables.

Notice that, if we want to find solutions for $F(J, \eta, r, v) = 0$, we have three equations and four variables. This defines a curve in the space and to compute it we use a predictor-corrector method. The corrector step is based in the Newton method and Langrange multipliers.

Consider the Jacobian matrix of F:

$$DF(J, \eta, r, v) = \begin{pmatrix} 0 & 0 & \frac{2}{\tau} \cdot v & \frac{2}{\tau} \cdot r \\ r & \frac{1}{\tau} & J - 2\pi^2 \tau r & \frac{2}{\tau} \cdot v \\ -\frac{2}{\tau} \cdot r & 0 & -\frac{2}{\tau}(J - 2\pi^2 \tau \cdot r) + 4\pi^4 \cdot r & \frac{8}{\tau^2} \cdot v \end{pmatrix}. \quad (17)$$

First, starting at the point $x_1 = (J_1, \eta_1, r_1, v_1)$, the lower part ($J$ decreases) of the boundary's left side (pink curve in Fig.24) will be progressively computed. Let s be a vector in the kernel of $DF(x_1)$, and define $w = x_1 - \epsilon \cdot s$, for a small constant ϵ (0.1 for the plot of Fig. 3). Then, if the norm of $F(w)$ is sufficiently small, up to some chosen tolerance (tol$= 10^{-5}$ in Fig. 3), the point w can be consider as a new point x_2 of the bistability boundary. Otherwise, as long as norm $(F(w)) > tol$, the code follows a sequence of steps to decrease this norm. It becomes a problem of finding the local minima subject to an equality constraint, which we will solve using the method of Lagrange multipliers. We compute $K := DF(w)$ and $p := -K^T(KK^T)^{-1} \cdot F(w)$. Then we define a new $w* \leftarrow w + p$ and checks the tolerance condition for $F(w*)$. If $w*$ does not verify it, $w \leftarrow w*$ and the steps sequence is repeated until a suitable point x_2 is found. (See Fig. 25 for an illustration when $F : \mathbb{R} \to \mathbb{R}$).

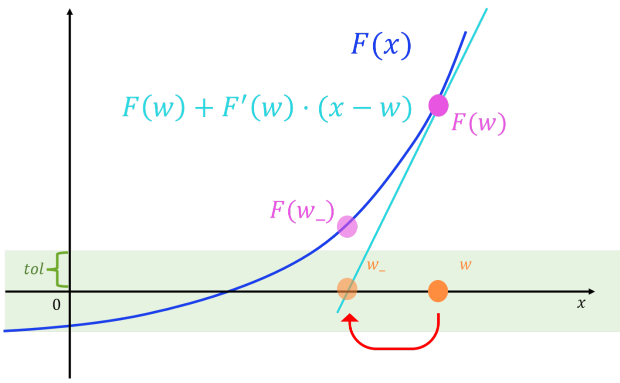


Figure 25: Newton-Raphson method for the approximation of a root of a function $F : \mathbb{R} \to \mathbb{R}$.

Then, from the new point x_2, the code runs again the procedure to find the following point x_3 and so on until all the lower part of the boundary's left side is computed. In fact, there must be a condition on the norm of the matrix KK^T's determinant. If the norm is too small, the matrix cannot be correctly inverted and the code should stop. This will happen when one reaches a point too near to the cusp bifurcation, which corresponds to the collision of the three fixed points (node, saddle and focus). So once such a problematic point is reached, the computations for this part of the boundary are done.

The procedure to compute the upper part of this branch is the same as the one previously explained with the only difference that the new points are chosen by adding $\epsilon \cdot s$, and not subtracting it.

B. Other applications in a WM model

B.1 First trial

Theory:
All populations start activated, forced by high frequency oscillations (80 Hz), where the Blue and the Left populations have the same phase and the Green and Right populations share a shifted phase. Rapidly, the two associated pairs get synchronized.
At 0.5 seconds the high frequency external input is removed. This symbolises the removal of the image that the subject was memorising. Then, at x seconds is added a sinusoidal current (of the studied form as in expression (7)), with frequency of 14 Hz to the Green population.
Observe that this frequency is in the range of "*Clearance*". As the Green population and the Right population were in phase and mutually excite each other, it is expected that, as the Green population goes to the resting state according the *Clearance* phenomenon, the Right population would follow too and get inactive. On the other hand, the Blue population would be less inhibited by the green population and thus remain active. And, with the previous reasoning, the Left population is expected to remain active too and thus the subject's answer can be decoded: the Blue circle was located in the Left.

Results:

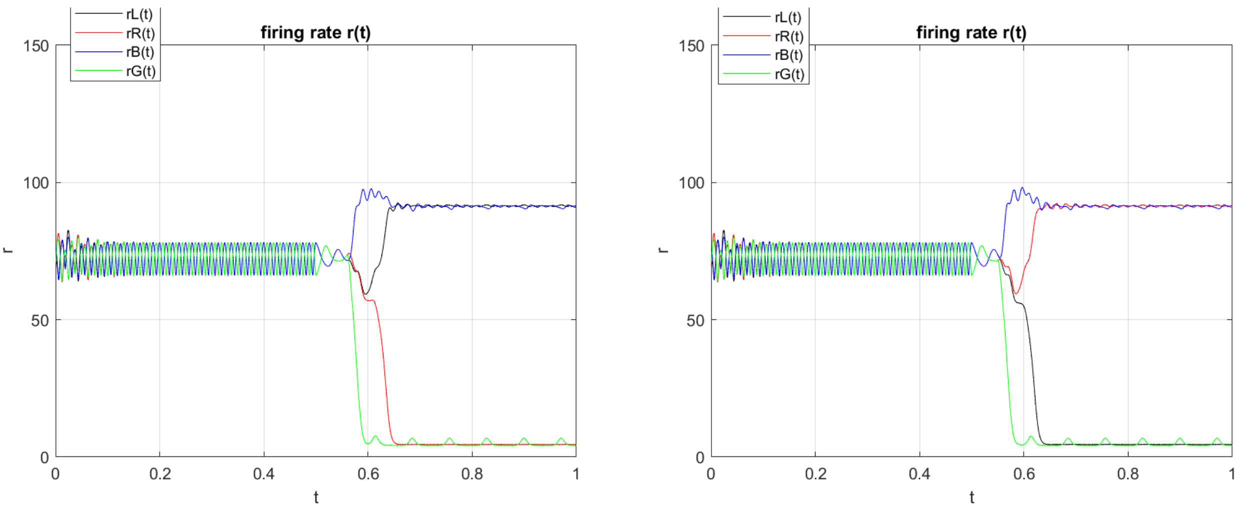


Figure 26: Results of the WM model first trial. The decoding input is added at 0.56 seconds (left) and at 0.55 seconds (right). For both plots the parameters used are: $\tau = 0.02, \eta = -10, \Delta = 2, A = 1, \gamma = 5.675463855030419$, $J_{SE} = 15\sqrt{\Delta} - 1$, $J_E = 3.4$, $J_I = 5.8$.

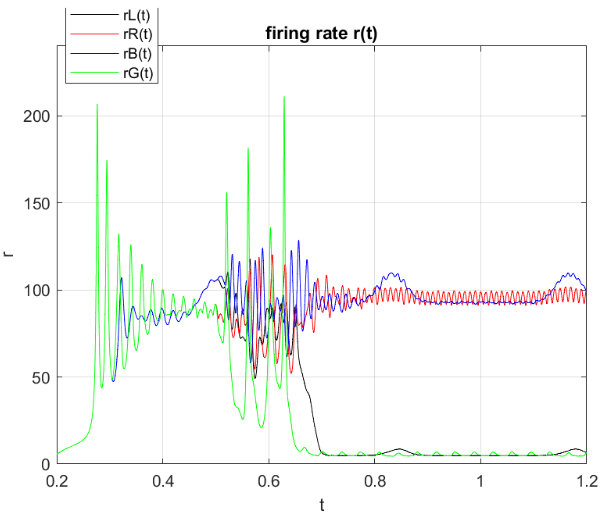
Observations:
The figures show that, indeed, the applied input with 14Hz frequency makes the Green population go to 0 and the Blue population remain active. However, depending on the time this current is applied, either the Left or the Right populations become inactive (Fig. 26). Thus, this model is not functional because it is not possible to correctly decode which location was memorized with each color.

B.2 Frequency synchronization

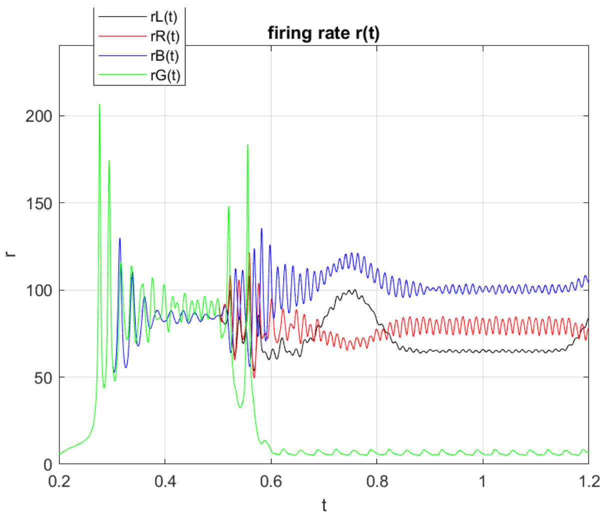
Theory:

All four populations start in a resting state without oscillations. Then they are all activated thanks to a strong transient input simulating the moment where the image is presented.
When the subject is memorizing, oscillations of the form $A(\gamma \cdot sin(\pi \cdot freq \cdot t)^{20} - 1)$ arise in each population, but, for each associated pair (Blue with Left and Green with right), the chosen firing frequency is different. Inspired by the previous observations about the different frequency ranges and their impact in changing or not the state, it is chosen to give a slow firing rate to one pair of populations (as in Fig. 6 (a) and (b)) and a fast firing frequency (as in Fig. 6 (d)) to the other pair. It is expected that with these chosen frequencies, all populations will remain active and synchronized by pairs.
For the decoding part an oscillatory input is injected to the color population that is not queried (Green for instance). This input has frequency in the range responsible for the *Clearance* phenomenon. The aim is that the Green population, followed by the population of its associated location, go to the resting state. This way, the population coding for the location where the Blue circle will remain active and can be decoded.

Results:



a) Pair Blue-Left: frequency 3Hz, pair Green-Right: frequency 100 Hz, *Clearance*: frequency 28Hz.



b)Pair Blue-Left: frequency 2Hz, Pair Green-Right: frequency 80 Hz, *Clearance*: frequency 30Hz.

Figure 27: **Results second model WM.** Parameters used: $\tau = 0.02, \eta = -10, \Delta = 2, A = 1, \gamma = 5.675463855030419$, $J_{SE} = 15 \cdot \sqrt{\Delta}$, $J_E = 3$, $J_I = 4$. First transient input (activation): $I_1 = 200$.

Observations:

The synchronization via the firing rate frequency does not seem to work. In the numerical results the difference between the two pairs cannot be appreciated. Unfortunately the expected result is not obtained. Although the Green population indeed goes to the inactive state and the Blue population remains active, what happens with the populations coding for color seems very random and sensible to small frequency changes. With different simulations it was not possible to find characteristics for which the desired result is obtained.