Homework 3

Hubert Gruniaux

Contents

I Introduction

In this report, we embark on a journey through diverse mathematical models to understand the intricate dynamics of neural systems. Our exploration begins with self-connected neurons, followed by models featuring mutual excitation, Hopfield neural networks, and concludes with a PCA study on a specialized "ring" network model.

The various studies focus mainly on the mathematical aspects of the underlying dynamic systems. Analytical and qualitative fixed-point studies are provided.

The different parts are mostly independent.

II Neuron with self-connection

II.1 The model

In the first part we start with a simple model for neurons: a neuron with an autapse – a chemical or electrical synapse from the neuron to itself. The firing rate r of the neuron is described by the following differential equation:

$$\frac{dr(t)}{dt} = -r(t) + f(wr(t) + I) \tag{1}$$

with

- ω, the synaptic strength,
- I, the external input and
- $f(s) = 60(1+\tanh(s))$, the activation function. The curve of f is given as a reference in Figure 1. Some notable points: $f(s) \underset{s\to-\infty}{\longrightarrow} 0$, $f(s) \underset{s\to+\infty}{\longrightarrow} 120$ and $f(0) = 60$.
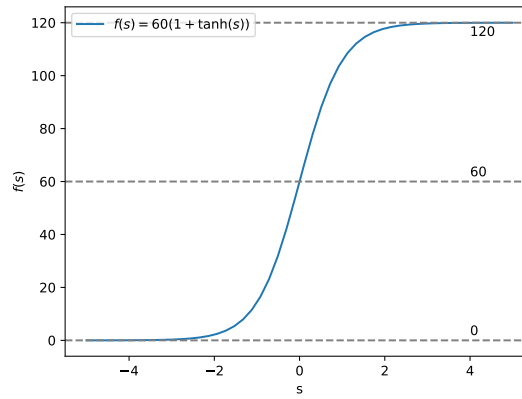

Figure 1: Plot of the activation function $f(s) = 60(1 + \tanh(s))$

Unless explicitly specified, the constants ω and I are taken as $0.05 \text{ nA} \cdot \text{Hz}^{-1}$ and -3 mV respectively the rest of the section.

II.2 Study of the dynamics

The Figure 2 show the solutions of the Equation 1 for some initial conditions (namely $r(0) \in \{50, 60, 61\}$). We see that depending on the initial condition, the solution converge towards different fixed points.

In particular, the special case $r(0) = 60 \overset{\text{def}}{=} r_{\text{crit}}$ gives a constant solution (that is equals to 60 obviously). However, if $r(0) < r_{\text{crit}}$, the solution converges towards $\min f = 0$, and if $r(0) > r_{\text{crit}}$, the solution converges towards $\max f = 120$. Therefore, three fixed points are conjectured.
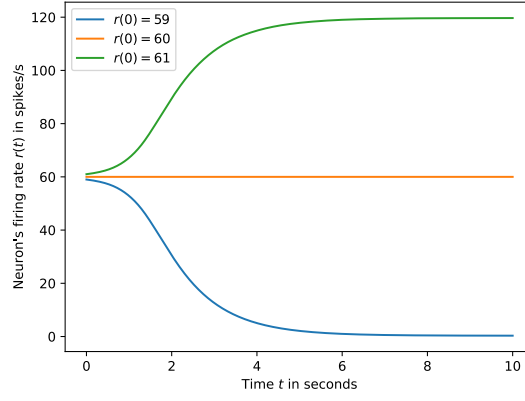
Figure 2: Plot of the dynamics for different initial conditions

A more interesting case (and more realistic) appears when we introduce some noise into the Equation 1:

$$\frac{dr(t)}{dt} = -r(t) + f(wr(t) + I) + \sigma\eta(t) \tag{2}$$

where η is a gaussian white noise and $\sigma \in \mathbb{R}_+$ the noise amplitude. Of course, we retrieve Equation 1 when $\sigma = 0$ (no noise).

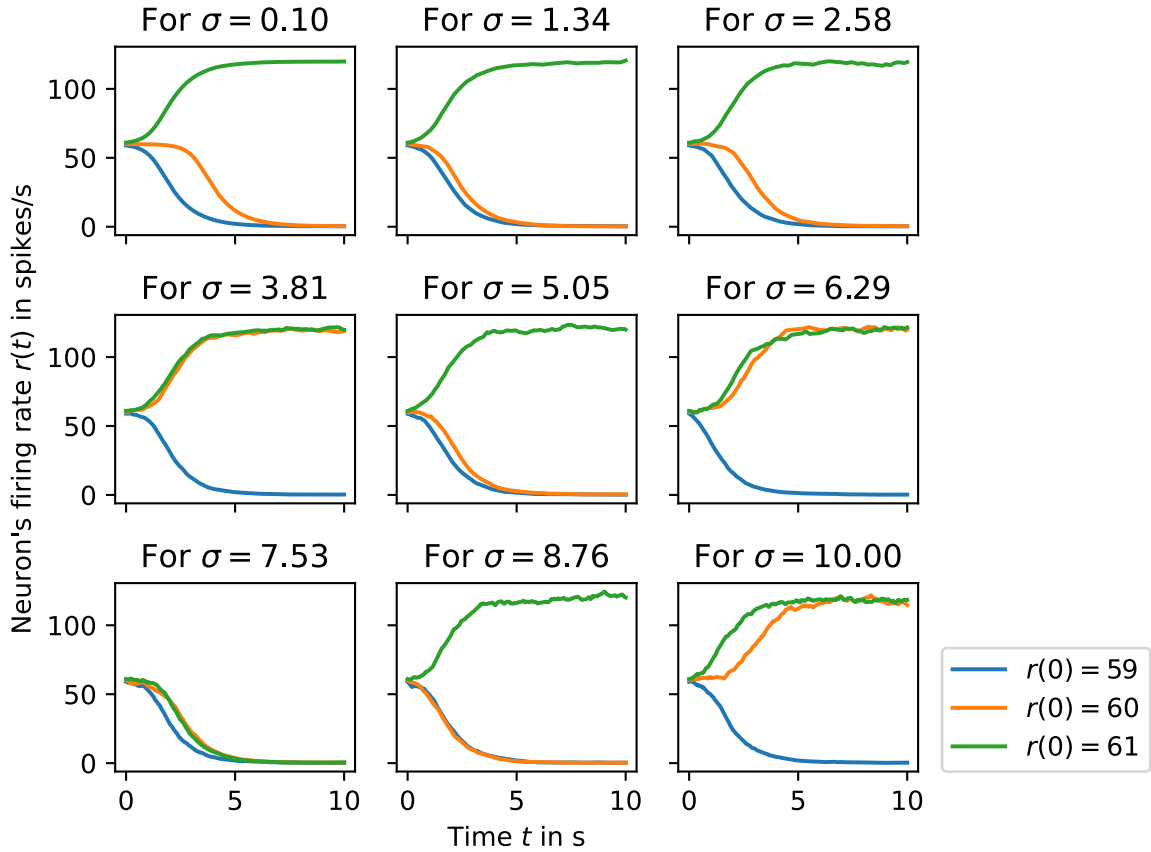The solutions of Equation 2 are given in Figure 3 for different initial conditions and σ.

Figure 3: Plot of the dynamics for different initial conditions and noise amplitude

We then observe that in neither case does the curve converge towards 60, even though it initially starts at 60. This is true even with very low noise: the fixed point 60 is unstable.

However, we still observe the same two other fixed points (0 and 120), even with a lot of noise. These are stable fixed points.

II.3 Study of the fixed points

This section provides a more formal and in-depth study of the behavior of the dynamic system. To do this, we study the model in phase space.

In particular, the curve of the (time) derivative of r as a function of r itself is provided in the Figure 4.
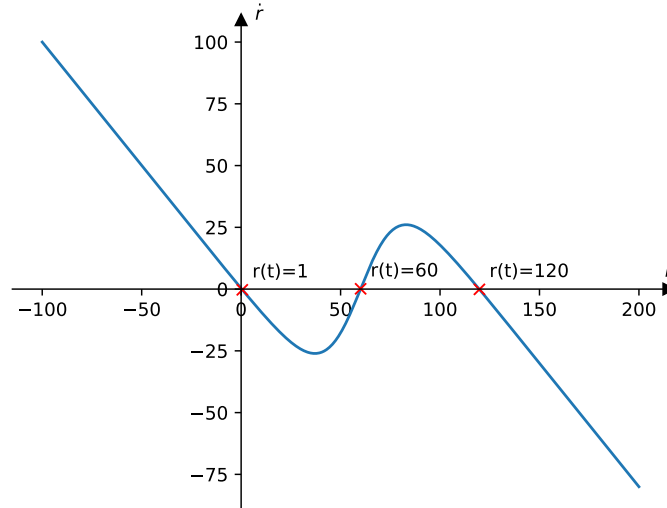
Figure 4: The derivative $\frac{dr}{dt}$ as a function of r for $\omega = 0.05$ nA $\cdot$ Hz^{-1} and $I = -3$ mV

The cancellation points on the Figure 4 correspond to fixed points in the system (points of convergence of the solutions). In this case, we find the three fixed points we've already encountered: 0 (due to numerical instability, we got 1 on the diagram), 60 and 120.

If we plot the count of zeros (fixed points) as a function of ω, the synaptic strength, and I, the external input, we get the bifurcation diagram shown in Figure 5.
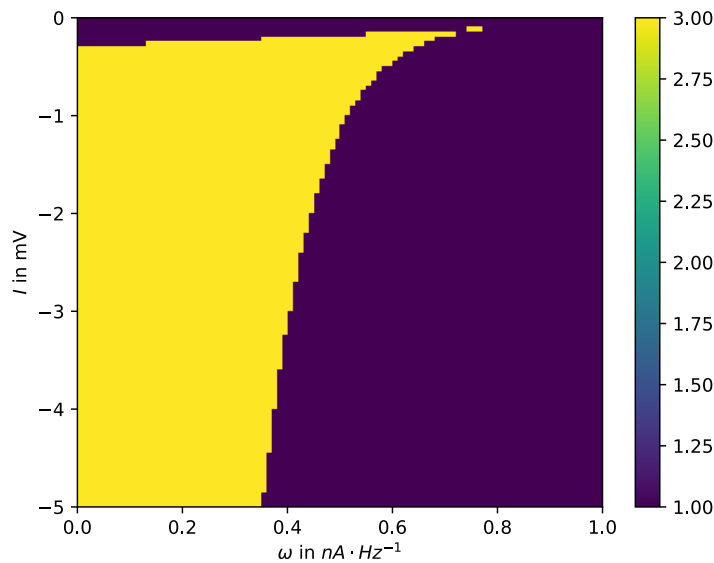
Figure 5: Bifurcation diagram (count of zeros of the curve $\dot{r}$ as a function of r)

The yellow part correspond to three concellation points, whereas the purple part correspond to one concellation point. As a remainder, the concellation points in the state space are linked to the fixed points of the dynamic system.

We observe that the weaker the external input and the lower the sypnatic force, the more unstable the system (more fixed points). We also note that it's mainly the sypnatic force that reduces the number of fixed points (when it's large). The external input has less impact.

III Mutual excitation

III.1 The model

In this second part, we will consider a little more complex system. Now, we will study a circuit of two neurons that excite each other.

Let consider the two following functions:

$$F_1 : x, y \mapsto -x + f(\omega_2 y + I)$$
$$F_2 : x, y \mapsto -y + f(\omega_1 x + I) \tag{3}$$

with $f(s) = 50\sigma(s)$ and $\sigma(s) = (1 + e^{-s})^{-1}$. The function f is the activation function and $f(\mathbb{R}) = [0, 50]$ with $f(s) \underset{s \to -\infty}{\longrightarrow} 0$ and $f(s) \underset{s \to +\infty}{\longrightarrow} 50$.

Then, the firing rate model is described by the below system of differential equations:

$$\dot{x}(t) = F_1(x(t), y(t))$$
$$\dot{y}(t) = F_2(x(t), y(t)) \tag{4}$$

with $\dot{x}(t) = \frac{dx}{dt}(t)$ and $\dot{y}(t) = \frac{dy}{dt}(t)$.

A more in depth mathematical study of the system will be provided in the third subsection (when studying analitically the fixed points).

III.2 Study of the dynamics

A simulation of the dynamics for some random initial conditions is provided in Figure 6. In that particular case, they all converge towards a single fixed point: 60. It will be showed later that this model may have more than one fixed point, including a potential unstable point.
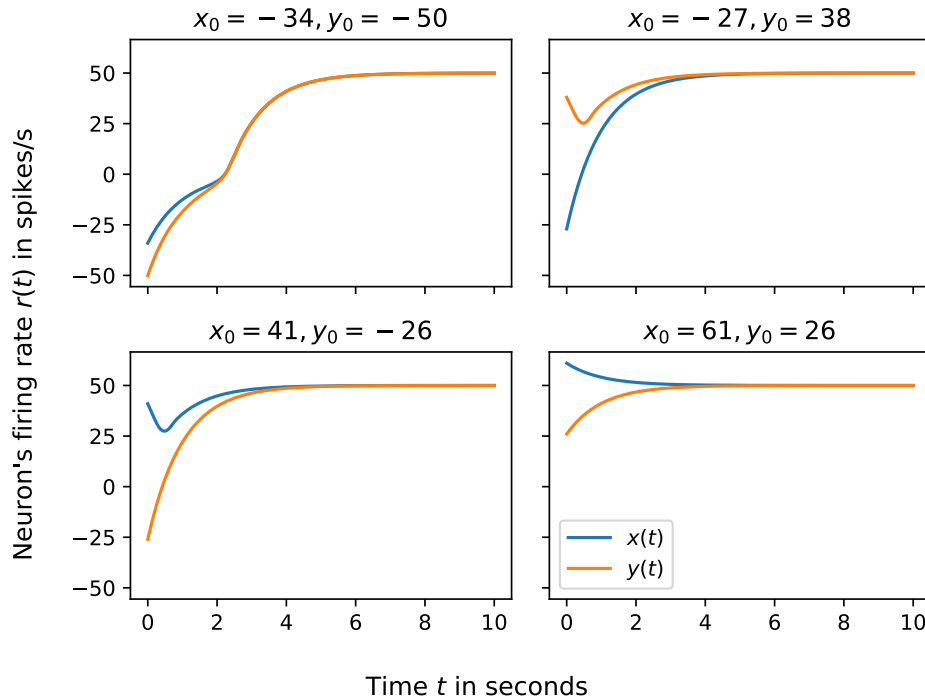

Figure 6: The dynamics

III.3 Study of the fixed points

Like with the first model, we study the model in the state space to better analyze the fixed points.
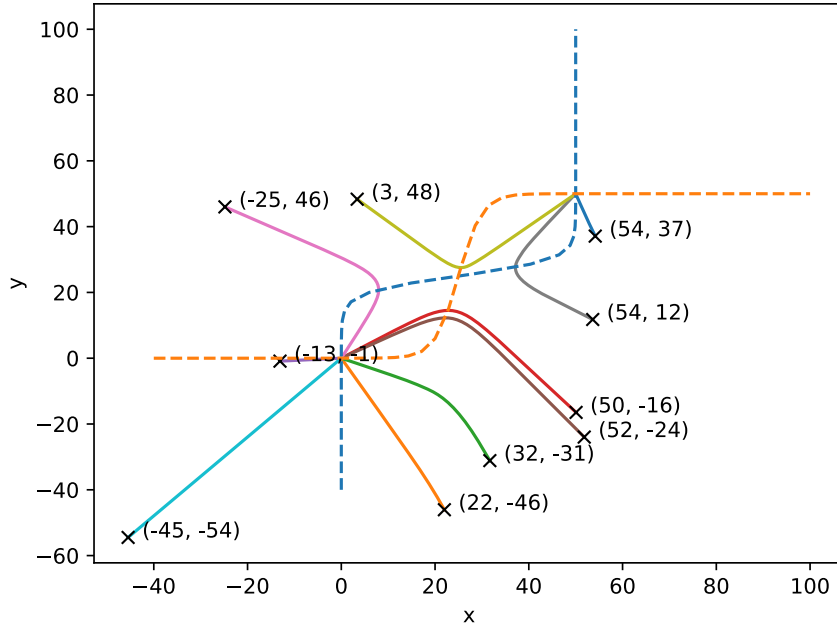
Figure 7: Trajectories in the state space for random initial conditions (and nullclines)

The two dashed lines are the nullclines of the system. Their crossing points indicates the fixed points of the system. Indeed, we can observe that the different trajectories converges towards one of those points.

We can see that the points at $(0, 0)$ and $(50, 50)$ appear to be stable. The point at $(25, 25)$, on the other hand, appears to be unstable (no random trajectory converges on it). These three points are the fixed points of our system.

We can analitically study the stability of the fixed points. For that, we first need to compute the Jacobian matrix of the system at each point:

$$\forall x, y, J(x, y) = \begin{pmatrix} \frac{\partial F_1}{\partial x}(x, y) & \frac{\partial F_1}{\partial y}(x, y) \\ \frac{\partial F_2}{\partial x}(x, y) & \frac{\partial F_2}{\partial y}(x, y) \end{pmatrix} \tag{5}$$

Now we have:

$$\frac{\partial F_1}{\partial x}(x, y) = \frac{\partial F_2}{\partial y}(x, y) = -1$$

$$\frac{\partial F_1}{\partial y}(x, y) = 50\omega_2\sigma(\omega_2 y + I)(1 - \sigma(\omega_2 y + I)) \tag{6}$$

$$\frac{\partial F_2}{\partial x}(x, y) = 50\omega_1\sigma(\omega_1 x + I)(1 - \sigma(\omega_1 x + I))$$

And with the assumption $\omega_1 = \omega_2 = \omega$, we get

$$\forall x, y, \mathrm{tr}(J(x, y)) = -2 < 0 \tag{7}$$

And

$$\forall x, y, \det(J(x, y)) = 1 - 2500\omega^2\sigma(\omega x + I)\sigma(\omega y + I)(1 - \sigma(\omega x + I))(1 - \sigma(\omega y + I)) \tag{8}$$

The signs of $\mathrm{tr}(J)$ and $\det(J)$ will tell us the stability of the fixed point.

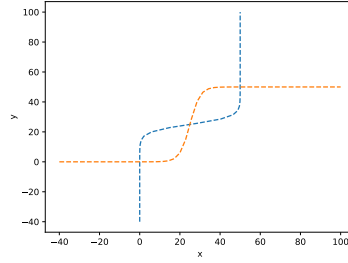In our case, the fixed points are $(0, 0)$, $(25, 25)$ and $(50, 50)$:

- $\det(J(0,0)) = \det(J(50,50)) \approx 1 > 0$
- $\det(J(25,25)) \approx -24 < 0$

Therefore, because $\text{tr}(J) < 0$, $(0,0)$ and $(50,50)$ are stable but $(25,25)$ is an unstable fixed point.
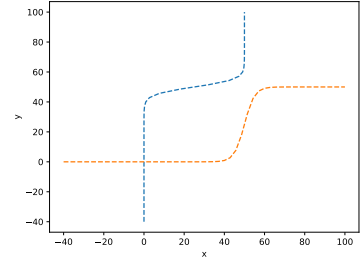
The count of fixed points (number of crossing points of the nullclines) change as a function of the input current. The exact function is given by the bifurcation diagram on Figure 12. Whereas, different nullclines are shown for different input currents on Figure 8.
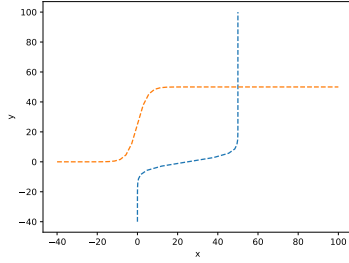
Subfigure 9: For $I = 0$ nA   Subfigure 10: For $I = -10$ nA   Subfigure 11: For $I = -20$ nA

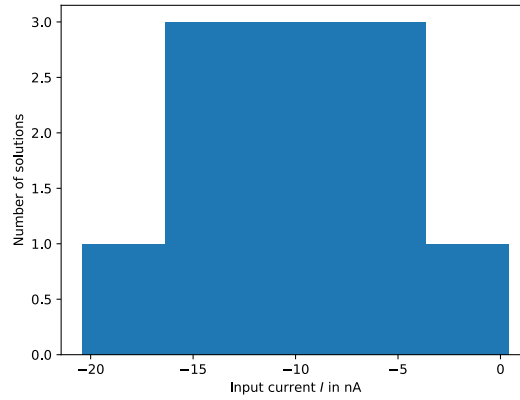Figure 8: Nullclines for different input current I

Figure 12: Bifurcation diagram

From Figure 12, we observe two regimes. For $I \in [-16, -6]$, we have three fixed points including one unstable. Otherwise, we only have a single fixed point that is stable.

IV Hopfield neural network

IV.1 The model

We consider two patterns p_1 (the cat) and p_2 (the man/astronaut) in $\mathbb{R}^d$ with $d = 16 \times 16 = 256$. Both patterns are shown in Figure 13.
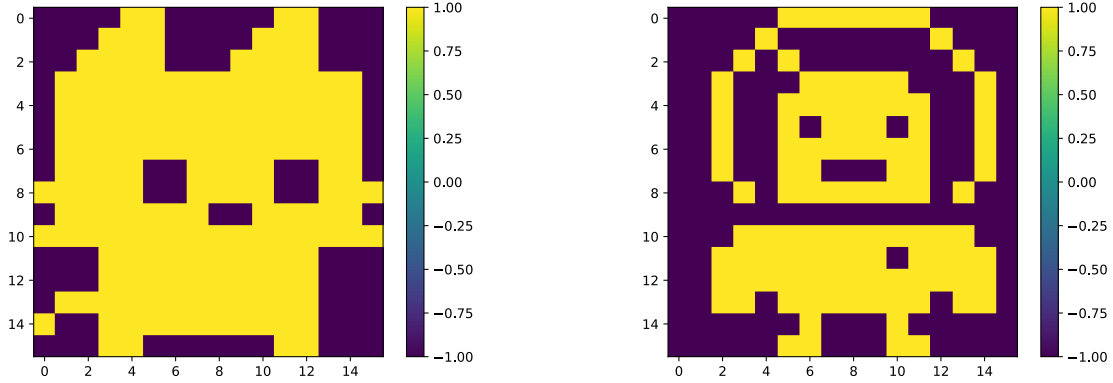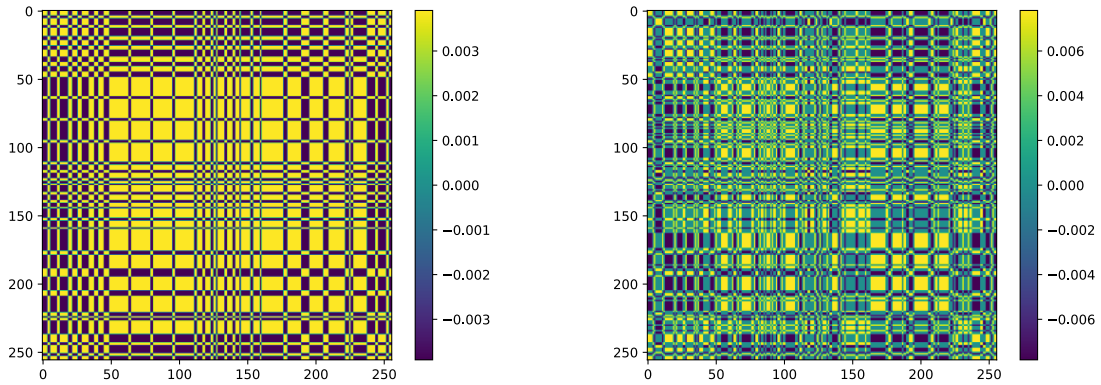

Figure 13: The cat and man patterns

Let $p_1, ..., p_n \in \mathbb{R}^d$ be the n patterns, then we define the weight matrix W to be:

$$W = \frac{1}{N} \sum_{i=1}^{n} p_i p_i^T \tag{9}$$

In particular, we note W_1 the weigth matrix for the pattern p_1 and W_2 the weight matrix for the patterns p_1 and p_2. Both the matrices are shown in Figure 14.

Subfigure 15: For the cat pattern, W_1 Subfigure 16: For the cat + man patterns, W_2

Figure 14: Different weight matrices W

The dynamics of the Hopfield neural network is given by the following system (with $x(t)$ the "predicted" pattern at time t):

$$\dot{x}(t) = -x(t) + \text{sign}(Wx(t)) + \sigma\eta(t) \tag{10}$$

with

$$\text{sign}(x) = \begin{cases} +1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases} \tag{11}$$

and where η is a gaussian white noise and $\sigma \in \mathbb{R}_+$ the noise amplitude.

IV.2 The dynamics and fixed points of the model

The figures Figure 17 and Figure 18 both show two examples of the evolution of the system for W_1 and W_2 respectively, and an initial uniform random pattern. We see that in the first case, the pattern converges towards the cat pattern (as W_1 only memorized the cat pattern). In the second case, the model converges towards the man pattern, but for a different initial pattern, the model can also converge towards the cat. Notably, because W_2 memorizes both the cat and man patterns.

In both cases, the convergence is quite fast. Only a few iterations are needed before observing one of the memorized patterns clearly.
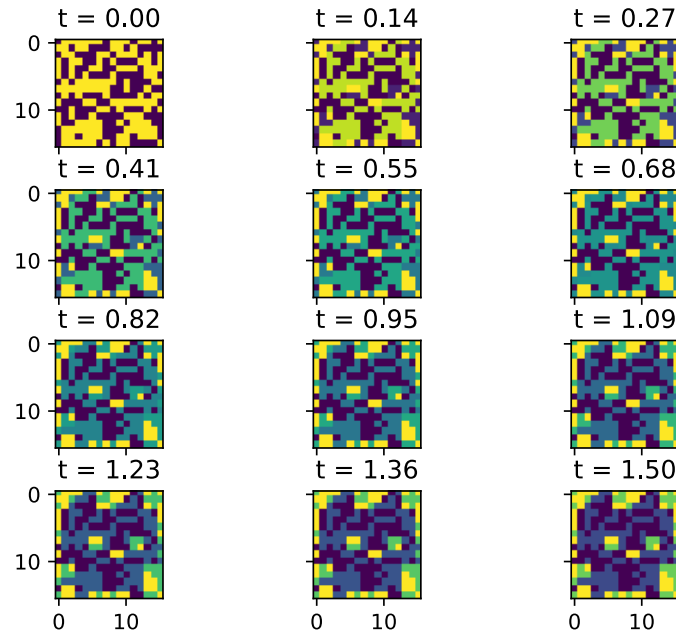
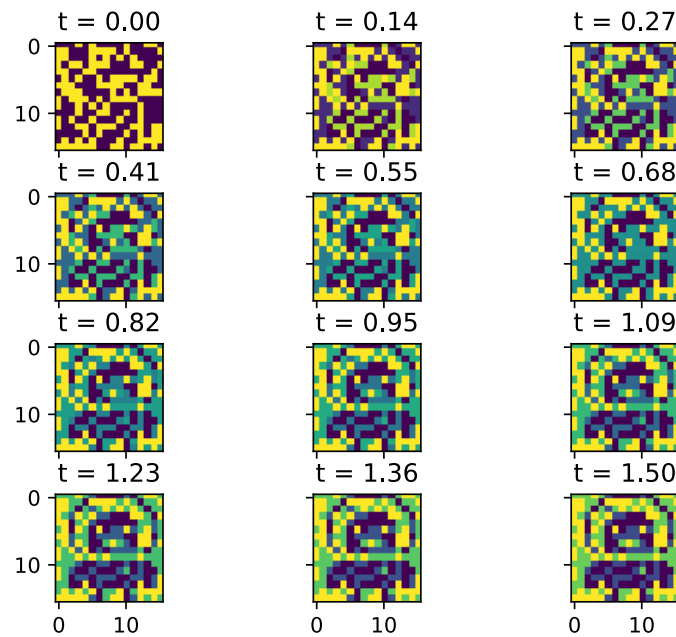Figure 17: Simulation of the network for W_1 and an initial random pattern

Figure 18: Simulation of the network for W_2 and an initial random pattern

Now, let consider the man pattern and add some white noise to it (to corrupt it). This way, we get Figure 19. The man is still recognizable with some effort, but it is not the same as the uncorrupted man pattern shown in Figure 13.
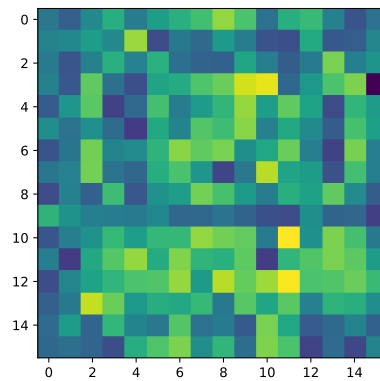

Figure 19: The man pattern corrupted with noise

Still, the network will converge towards the clean pattern (without noise) as shown in Figure 20. Moreover, the convergence is quite fast (the image is clearly recognizable for t small).

This example show the resilience of the network to noise. And the convergence towards the nearest "memorised" pattern.
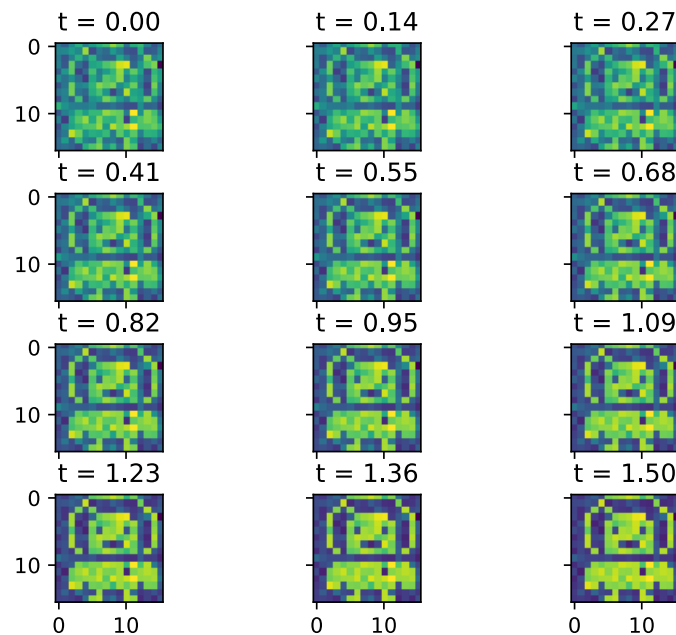

Figure 20: Simulation of the network for the corrupted man pattern

IV.3 Network capacity

An important property of our network is how much distinct patterns can be memorized by N neurons.

We note C the capacity of our network, which is how many patterns can be memorized until the network is not able to relax back to the pattern after adding small noise. We can prove[1] that $C \approx 0.138N$ where N is the count of neurons. So, with $N = 256$, we get $C_{256} \approx 35$ and with $N = 64$, we get $C_{64} \approx 9$.
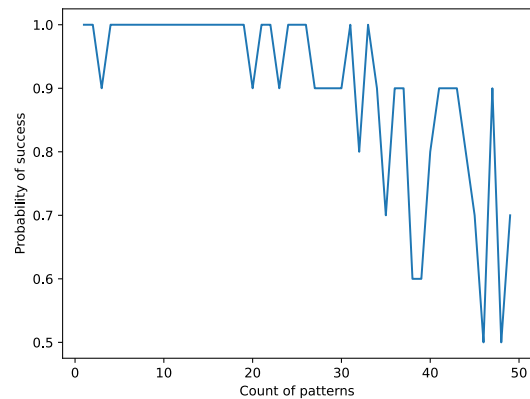
We can retrieve these results experimentally using the following protocol:

1. Create a weight matrix for M random patterns.
2. Select one of the random patterns and add noise to it.
3. Simulate the network with the corrupted pattern as the initial condition.
4. If the network converge towards the uncorrupted pattern, then OK. Otherwise, it is a failure.
5. Repeat from the step 2 as many times as you wish.
6. At the end, we get a probability of success for the network to relax back to the initial pattern for M memorized patterns.

The Figure 21 show the experience for $N = 64$ and $N = 256$ and for different initial pattern counts. We clearly see a decrease in probability after the computed theoretical capacity $C = 0.138N$ is reached.
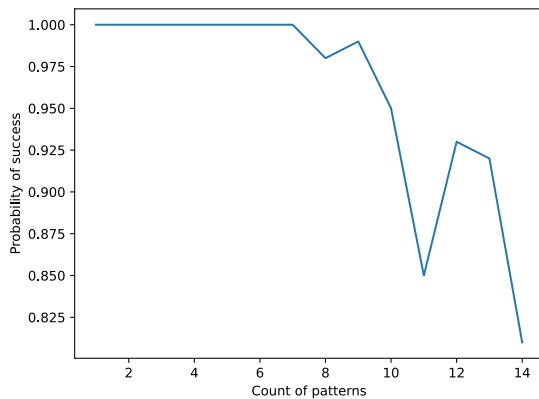
Subfigure 22: For $N = 64$ Subfigure 23: For $N = 256$

Figure 21: Simulations for 10 trials each time

---

[1]Hertz, J.A. (1991). Introduction To The Theory Of Neural Computation (1st ed.). CRC Press.

V PCA and the ring network model

V.1 The model

The dynamics of the network is given by the following system:

$$\dot{x}(t) = -x(t) + J\Phi(x(t)) \tag{12}$$

where the activation function is $\Phi(s) = \tanh(s)$.

We define the connectivity matrix of this network, J, such that the strength between neurons i and j is given by

$$J_{i,j} = 2\cos(\theta_i - \theta_j) \tag{13}$$

The Figure 24 gives the connectivity matrix that will be used for the rest of this section. It was generated using random $\theta_i \in [0, 2\pi)$.
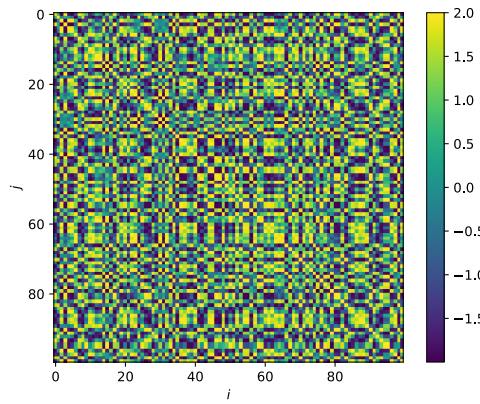

Figure 24: The connectivity matrix for randoms θ_i

We observe that the connectivity matrix is symmetric. Indeed, this fact can be proved mathematically:

$$\forall i, j, J_{i,j} = 2\cos(\theta_i - \theta_j) \overset{\cos \text{ is even}}{=} 2\cos(\theta_j - \theta_i) = J_{j,i} \tag{14}$$

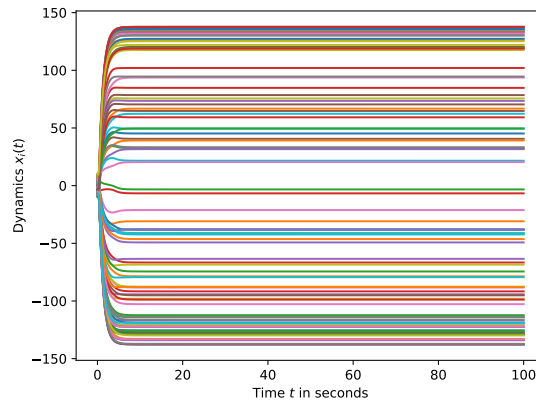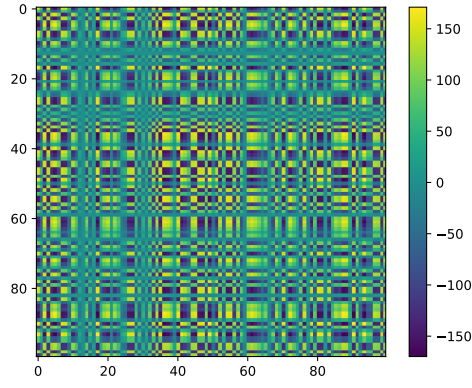The simulated dynamics of the system for $N = 100$ neurons and $t \in [0, 100]$ ms is provided on Figure 25.

Figure 25: Dynamics of $x_{i(t)}$ for $i \in [\![1, N]\!]$

One can do a principal component analysis (PCA) on the matrix X build from the previously simulated dynamics (matrix of shape $N \times T$). The covariance matrix and the principal components (eigen values) are provided on Figure 26.

Subfigure 27: The covariance matrix
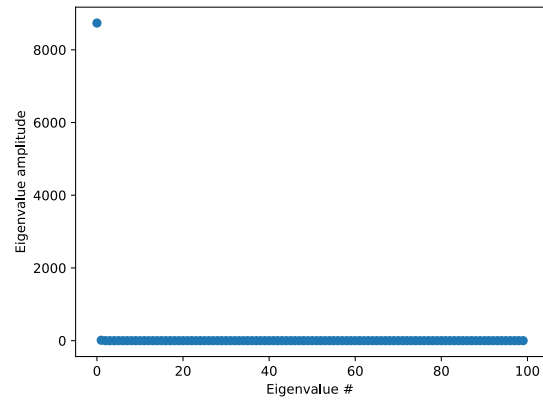
Subfigure 28: The eigenvalues

Figure 26: Results of PCA

The explained variance for the i-th principal component is defined as:

$$\text{EV}_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \tag{15}$$

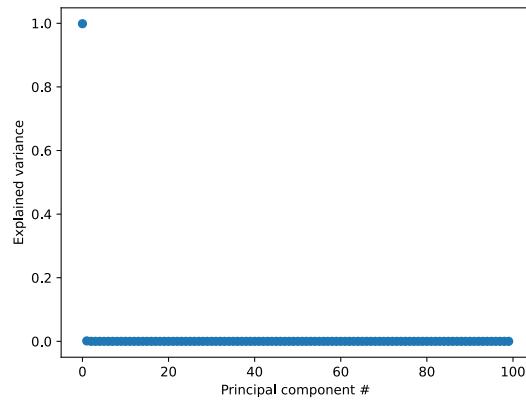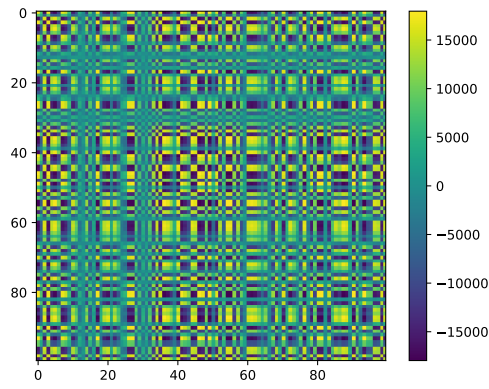Their values are provided on Figure 29 for each principal component.

Figure 29: The computed explained variances
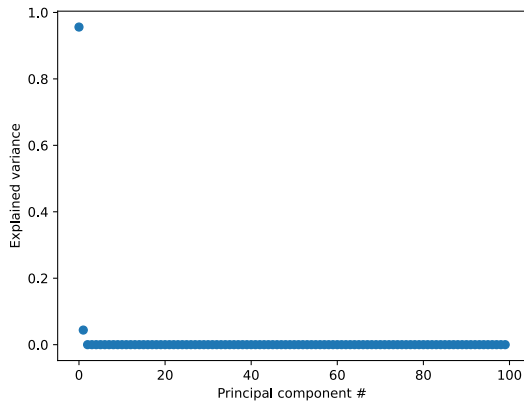
V.2 PCA on set of attractors

In this section, we built a 100×500 matrix that contains the collected 500 attractors (the fixed points) for different random initial conditions. The results of PCA on this matrix are provided in Figure 30.
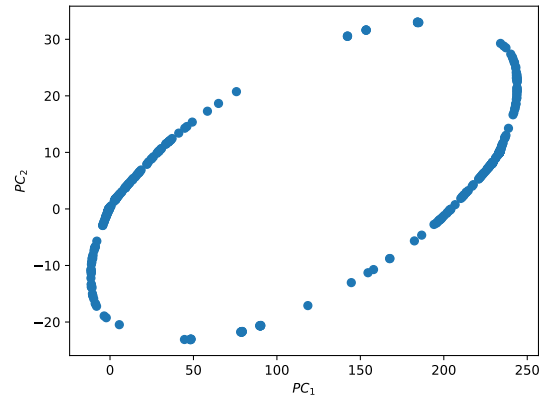
Subfigure 31: The covariance matrix
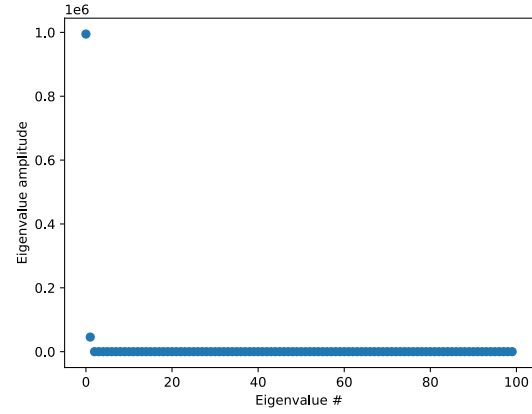
Subfigure 32: The eigenvalues

Subfigure 33: The computed explained variances

Subfigure 34: Principal components in the plane $\mathrm{PC}_1 - \mathrm{PC}_2$

Figure 30: Results of the PCA

We observe that there is two interesting components (of explained variance 0.956 and 0.0439 respectively). The thid component has an explained variance 10^6 times smaller than the second. Likewise, the second component has an explained variance 21 times smaller than the first one (which can be easily seen in the plot).

If we plot the principal components in the plane $\mathrm{PC}_1 - \mathrm{PC}_2$, we can see a ring. The name of the model comes from that fact.

VI Conclusion

Our exploration of diverse mathematical models has provided valuable insights into the dynamics of neural systems. From self-connected neurons to Hopfield networks and a specialized "ring" network model, each part contributed to a deeper understanding of neural models.

The different models studied have their own particularities: advantages and disadvantages. The main differences can be summed up in the activation function and the number of neurons considered. We note that networks with more neurons (like Hopfield's) can implement more advanced behaviors: pattern memorization. In particular, the number of patterns memorized is proportional to the number of neurons.

One final point: we showed in the last section that BCP analysis can be used to reduce the complexity of a problem by preserving only the relevant components (and thus reducing the size of the problem).