Informing the Neural Network Activation Function with Graph Centrality Measures: the Case Study of Oscillating Chemical Reaction Simulation

Paola Lecca
Faculty of Engineering
Free University of Bozen-Bolzano, Italy
Paola.Lecca@unibz.it

25th International Conference on Computational Science
July 7 - 9, 2025 - Singapore

Goals

Oscillating systems are a challenge for neural networks

- Oscillating systems are abundant in nature. Examples include pendulum clocks, ocean waves, the human heartbeat, the sleep-wake cycle, but also network of interacting chemical species in a cell can exhibit a periodic dynamics.
- This kind of dynamics is a particularly challenging case for a neural network, especially when the oscillations have a sawtooth shape, i.e. with stretches along which the derivative is almost infinite.
- In this study, we show how a neural network with an appropriate sinusoidal activation function whose parameters are derived from the node vibrational centrality of the graph describing the set of chemical reactions is able to well approximate this type of oscillatory trends typical of various biological systems.

Case study: network of endogenous oscillatory enzyme reactions

Glycolytic oscillations

Glycolytic oscillations, identified almost 50 years ago, remain the prototypical example of periodic behaviour in a metabolic circuit.

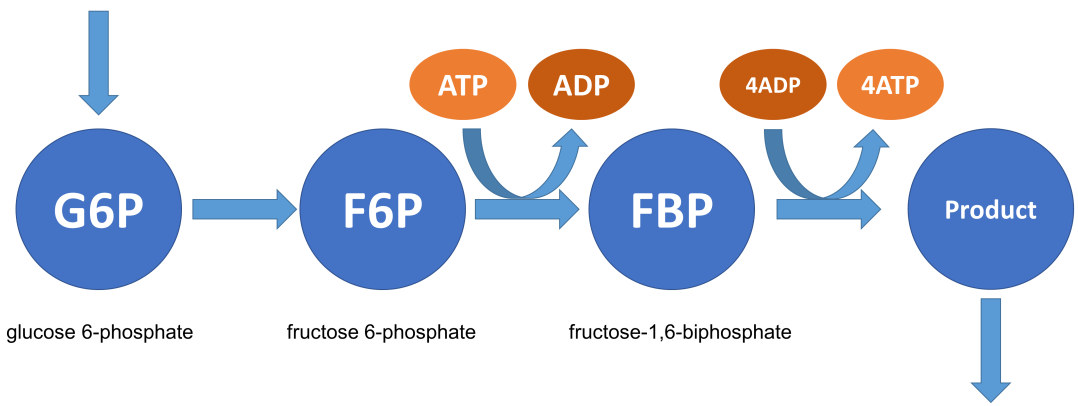
They last approximately 5-10 minutes in yeast when glucose is administered at a constant pace. The metabolic process periodically transforms the glycolytic substrate, which is given at a constant rate.

Glycolytic oscillations: a model

Goldbeter and Lefever, 1972

- The mechanism of glycolytic oscillations largely relies on the reaction catalysed by phosphofructokinase (PFK). The production of oscillations by PFK can be related to its activation by one of its reaction products, adenosine diphosphate (ADP), via adenosine monophosphate (AMP).
- PFK is an allosteric enzyme phosphofructokinase, which uses ATP as a phosphate donor to phosphorylate fructose-6-phosphate (F6P) in order to produce fructose-1,6-diphosphate and ADP.

This model is based on the concerted transition model for allosteric enzymes, to which is added the positive feedback exerted by the product. To exhibit oscillations such a system must be open and in non-equilibrium conditions. Therefore, in addition to PFK, the model includes the substrate input and the consumption of product in a second enzyme reaction, which may be of Michaelis-Menten type.

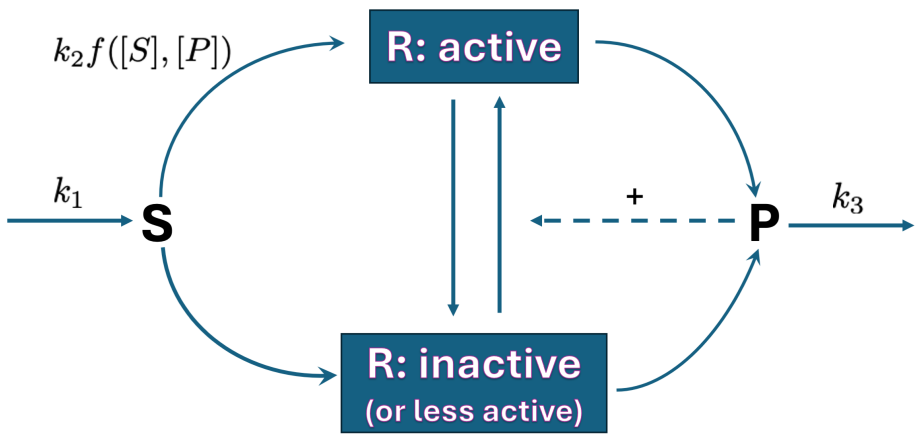


Adapted from: Jun-Hai Yang et al. 2008, THE JOURNAL OF BIOLOGICAL CHEMISTRY

Glycolytic oscillations: a scheme

1. Substrate S is supplied at a constant rate and binds to the active and inactive conformations (respectively R and T) of an allosteric enzyme, resulting in product P.
2. P is eliminated through a sink reaction facilitated by an enzyme with linear or Michaelis-Menten kinetics.
3. The allosteric enzyme is made up of subunits that move between two conformational states.
4. The product P, a positive effector, only binds to the R state, causing the allosteric enzyme to change from less active to more active.

The Figure below shows a schematic representation of the mechanism producing the oscillations.



Glycolytic oscillations: a mathematical model

Goldbeter, 2013

The dynamic of this system is governed by the following equations

$$\frac{d[S]}{dt} = k_1 - k_2 f([S],[P]) \tag{1}$$

$$\frac{d[P]}{dt} = r k_2 f([S],[P]) - k_3[P] \tag{2}$$

where $[S]$ and $[P]$ denote the normalized, dimensionless substrate and product concentrations, $f([S],[P])$ is the enzyme rate function which in the simple case where the substrate S binds exclusively to the most active conformation of the enzyme is given by:

$$f([S],[P]) = \frac{[S](1+[S])^{h-1}(1+[P])^h}{L + (1+[S])^h(1+[P])^h}, \tag{3}$$

Parameters k_1 and k_2 are the normalized substrate injection rate and maximum rate of the enzyme reaction, respectively; r is a normalization parameter, and $L \gg 1$ is the allosteric constant of the enzyme measuring the ratio of inactive (T) to active (R) conformation in the absence of ligand. $k_3[P]$ is the product sink function, or decay term.

Numerical solution via Runge-Kutta

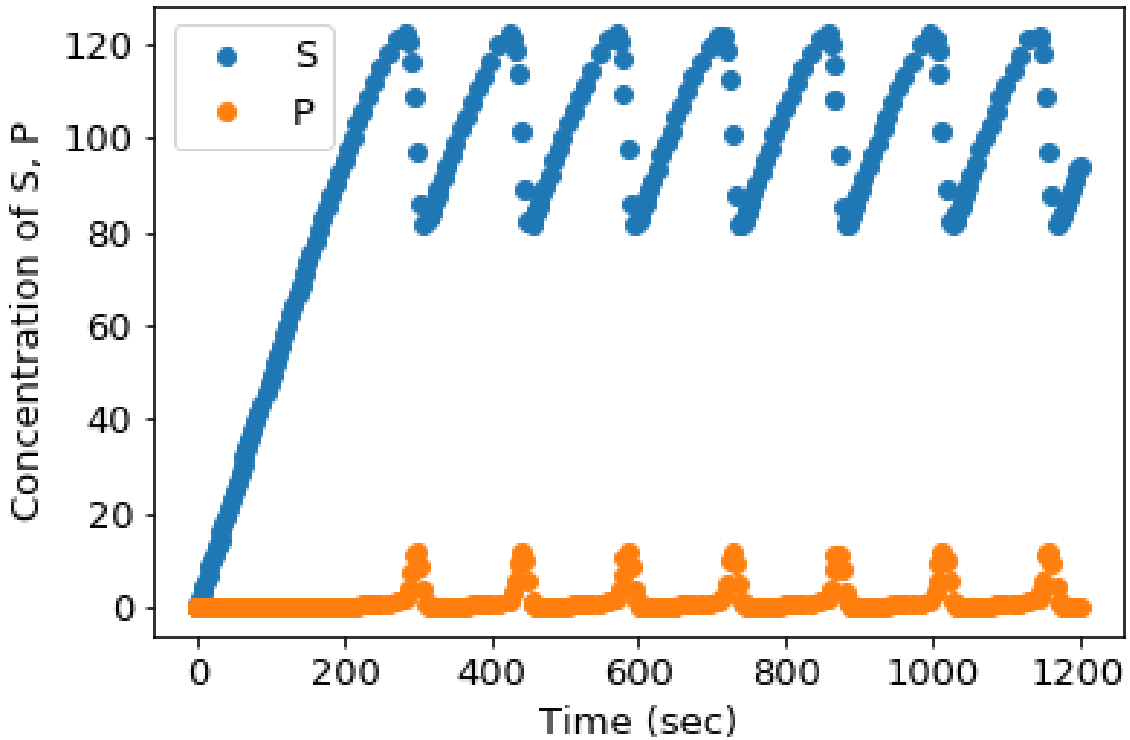


Figure: The curves obtained by numerical integration of Eqs. (1) and (2)with explicit Runge-Kutta method of order 5(4) for the following parameter values: $L = 10^6$, $k_1 = 0.5\ \mathrm{s}^{-1}$, $k_2 = 5.075\ \mathrm{s}^{-1}$, $r = 3$, $k_3 = 0.81\ \mathrm{s}^{-1}$, $h = 2$, and initial conditions $[S](t = 0) = [P](t = 0) = 0$.

Neural network

Formalisation - part I

A system of n ordinary differential equations (hereafter "ODEs") has the following form,

$$\begin{aligned} \frac{dx_1}{dt} &= F_1\left(t, x_1, x_2, \ldots, x_n\right) \\ \frac{dx_2}{dt} &= F_2\left(t, x_1, x_2, \ldots, x_n\right) \\ &\vdots \qquad\qquad \vdots \\ \frac{dx_n}{dt} &= F_n\left(t, x_1, x_2, \ldots, x_n\right) \end{aligned} \tag{4}$$

defined on $t_0 < t < T$ with given initial values, $x_1(0) = x_1^{(0)}, x_2(0) = x_2^{(0)}, \ldots, x_n(0) = x_n^{(0)}$.

Neural network

Formalisation - part II

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0, \tag{5}$$

where $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$ is the $n \times 1$ matrix of unknowns (for example, the concentration of the m chemical species in time), and

$$\mathbf{f}(t, \mathbf{x}) = \begin{bmatrix} F_1(t, x_1, x_2, \ldots, x_n) \\ F_2(t, x_1, x_2, \ldots, x_n) \\ \vdots \\ F_n(t, x_1, x_2, \ldots, x_n) \end{bmatrix}$$

is the $n \times 1$ matrix of functions. The solutions of the system are the functions describing the behaviour of $x_1, x_2, \ldots, x_n$ with respect to the variable t (that, when the systems of ODEs describe the dynamics of a system of m variables usually denotes the time).

Neural network

Formalisation - part III

The solution $\mathbf{x}$ calculated using a neural network can be expressed as:

$$\mathbf{x}(t, \mathbf{W}) = \mathbf{x}\left(t, \mathbf{W}_1, \ldots, \mathbf{W}_N\right) = \sigma\left(\mathbf{W}_N \ldots \sigma\left(\mathbf{W}_2 \sigma\left(\mathbf{W}_1 t\right)\right)\right), \quad (6)$$

and the neural network is a system of non-linear equations like

$$x(t, \mathbf{W}) = \sigma(\mathbf{W}t + \mathbf{b}), \quad (7)$$

where σ is the activation function.

Neural network

Scheme

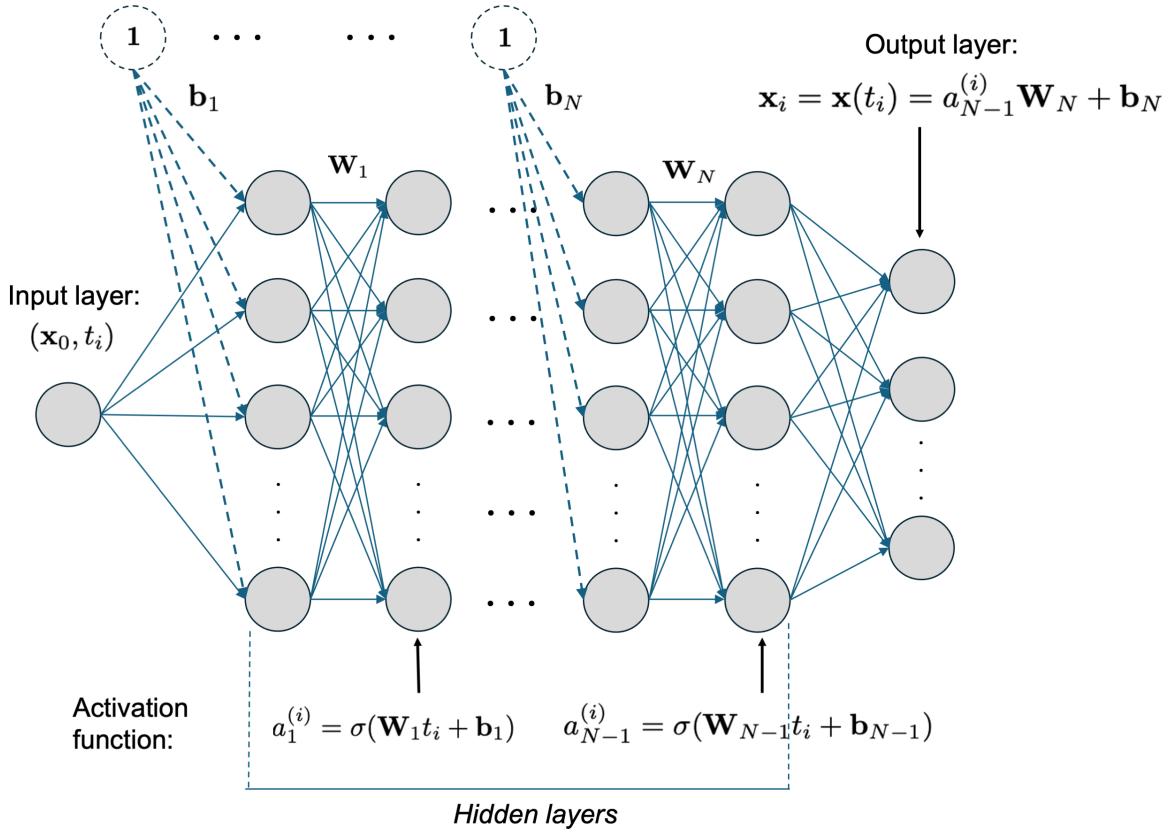


Figure: The network is fully connected and has one neuron in the input layer, and N hidden layers of p neurons each. The output layer has as many neurons as there are equations. The weight matrices $\mathbf{W}$ have dimensions as follows: $\mathbf{W}_1$ is a $p \times 1$ matrix, $\mathbf{W}_i$ (with $1 < i < N$) is a $p \times p$ matrix, and finally $\mathbf{W}_N$ is a $q \times p$ matrix, where q is the number of differential equations of the system. $\mathbf{b}_1, \ldots \mathbf{b}_N$ are the biases, $\mathbf{b}_1$ and and $\mathbf{b}_N$ have dimensions $p \times 1$, and $\mathbf{b}_N$ and $q \times 1$, respectively.

Toward the graph of the chemical reaction network

The node vibrational centrality - part I

Estrada et al. (2010) introduced a centrality measure, named bf node vibrational centrality.

In the vibrational centrality measure, the external stresses to which a system may be exposed are modelled through the concept of temperature. Herein temperature is meant to be a metaphor of all the different types of stress that the network can be submitted to. In line with this metaphor, nodes are rigid spheres and edges are elastic springs, submerged in a thermal bath at a given temperature T. Vibrational centrality quantifies the amplitude of the "oscillation" of a node in response to a stress.

Toward the graph of the chemical reaction network

The node vibrational centrality - part II

The n nodes of a network can be conceived as point in a n-dimensional Euclidean space, represented by the Moore-Penrose pseudo-inverse of graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D}$ is the diagonal matrix of degrees and A is the graph adjacency matrix of the network modelled as a graph. Henceforth we denote by $\mathbf{L}^+$ the pseudo-inverse of $\mathbf{L}$. Each diagonal entry of L^+, denoted as l_{ii}^+ for the i-th node, represents the squared distance of node i from the origin of the n-dimensional space and hence measures the nodes topological centrality, which is defined by

$$C(i) = \frac{1}{l_{ii}^+}. \tag{8}$$

Lower the value of l_{ii}^+, closer the node is to the origin more topologically central the node is.

Toward the graph of the chemical reaction network

The node vibrational centrality - part III

Two nodes connected by an arc are then represented as masses connected by springs (with elastic constant k). Furthermore, staying within the thermodynamics metaphor, a vibrational potential energy defined as

$$V(\mathbf{x}) = \frac{k}{2}\mathbf{x}^{\top}\mathbf{L}\mathbf{x} \tag{9}$$

is introduced, where x is the vector of node displacements. The probability distribution of node displacement is defined by the Boltzmann distribution

$$P(\mathbf{x}) = \frac{e^{-\frac{1}{T}V(\mathbf{x})}}{Z} = \frac{1}{Z}\exp\left(-\frac{k}{2T}\mathbf{x}^{T}\mathbf{L}\mathbf{x}\right) \tag{10}$$

where the partition function Z of the network is

$$Z \equiv \int d\mathbf{x}\exp\left(-\frac{k}{2T}\mathbf{x}^{T}\mathbf{L}\mathbf{x}\right).$$

Toward the graph of the chemical reaction network

The node vibrational centrality - part IV

Given $P(x)$, the mean displacement of the i-th node is, by definition,

$$\langle \Delta x_i \rangle \equiv \sqrt{\int x_i^2 P(\mathbf{x}) d\mathbf{x}} \tag{11}$$

It can be shown that the **vibrational centrality** is

$$\langle \Delta x_i \rangle = \sqrt{\frac{T}{k} l_{ii}^{+}}. \tag{12}$$

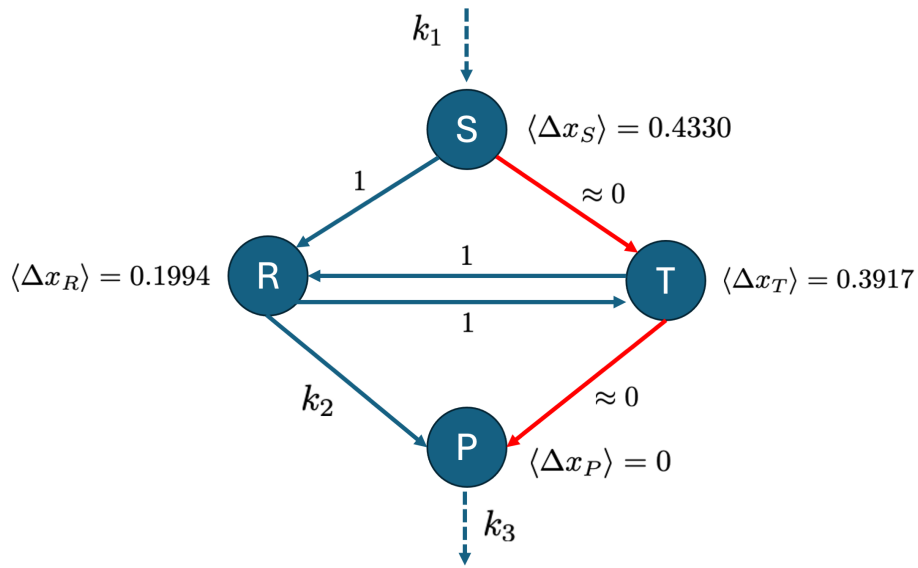
The graph of the chemical reaction network

The adjacency matrix **A** and the graph Laplacian **L** are then as follows.

$$\mathbf{A} = \begin{array}{c} \begin{matrix} S & T & R & P \end{matrix} \\ \begin{pmatrix} 0 & 1 & 1 & 0.000 \\ 0 & 0 & 1 & 1.000 \\ 0 & 1 & 0 & 5.075 \\ 0 & 0 & 0 & 0.000 \end{pmatrix} \end{array} \begin{matrix} S \\ T \\ R \\ P \end{matrix}$$

$$\mathbf{L} = \begin{array}{c} \begin{matrix} S & T & R & P \end{matrix} \\ \begin{pmatrix} 2 & -1 & -1 & 0 \\ 0 & 2 & -1 & -1 \\ 0 & -1 & 6 & -5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{matrix} S \\ T \\ R \\ P \end{matrix}$$

Therefore, we obtain the **vibrational centrality** shown in the Figure below. The substrate S is the node with the highest vibrational centrality, i.e. with the widest amplitude of oscillation (as also confirmed by the numerical solution of Eqs. (1).



Neural network activation function

Considerations in order to build it

- The numerical solution shows that product fluctuations occur after the linear increase of the substrate has reached a certain value of concentration.
- The substrate is the node with the greatest vibrational centrality, i.e. the node which, when perturbed is subject to oscillations of the greatest amplitude.

To take these two points into account, we write the activation function $a(z)$ as follows

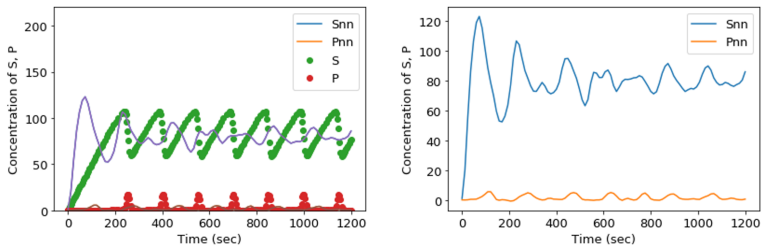
$$a(z) = a_1 \sin^2(\nu z) - \frac{a_2}{z + a_3} + a_4, \quad z \geq 0 \tag{13}$$

where z is the output of the node, $\nu = 1/\langle \Delta x_S \rangle$ (i.e. the reciprocal of the **vibrational centrality** of the substrate), $a_1 = a_2 = 100$, $a_3 = 0$, and $a_4 = 10$.

Results

Neural network solutions vs Runge-Kutta solutions in case of simple sinusoidal activation function

A.



B

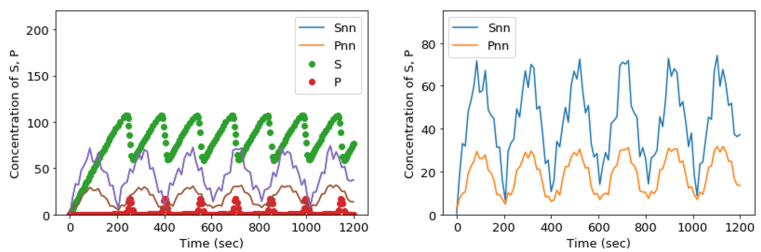


Figure: Numerical solution of the system of the differential equations (1) and (2) - obtained with Explicit Runge-Kutta method of order 5(4) - compared to the output of the neural network (Snn and Pnn) with the following parameters: (A.) learning rate: 0.01, size of input layer: 1 neuron, size of layer 1: 19 neurons, size of output layer; 2 neurons, activation function $1000\sin^2(0.8x)$, and 2000 epochs; (B.) learning rate: 0.01, size of layer 1: 19, activation function $1000\sin^2(0.8x)$, and 6500 epochs. The value of the objective function at the last iteration is 42.40 (sub-figures A.), and 62.20 (sub-figures B). The agreement between the numerical solution and the neural network output is suboptimal in both cases A. and B.

Results

Neural network solutions vs Runge-Kutta solutions in case of vibrational centrality infromed activation function

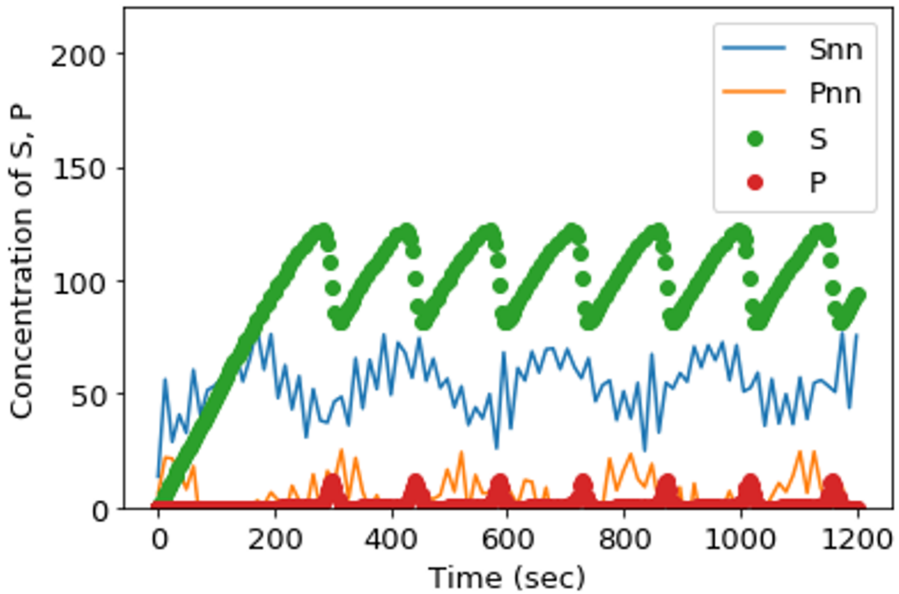


Figure: Numerical solution of the system of the differential equations (1) and (2) - obtained with Explicit Runge-Kutta method of order 5(4) - compared to the output of the neural network (Snn and Pnn) with the following parameters: learning rate: 0.025, size of layer 1: 15, activation function as in Eq. (13), and 2000 epochs. The agreement between the numerical solution and the neural network output is still non-optimal, but the oscillatory behaviour is maintained over time and the correct phase shift between product and substrate is obtained.

Discussion

- The oscillatory behaviour that best approximates the one given by the numerical solution was found only in the case in which ν is equal to the reciprocal of the vibrational centrality of the substrate, i.e. $1/0.43$.
- Experiments performed using values even slightly deviating from this one show a significant disagreement with the numerical solution and an incorrect phase relationship between the substrate and product curves. The correct phase relationship predicts that the product maximum immediately follows a substrate maximum.

Conclusions

- Although the parameters found in our experiments are the only ones that reproduce the oscillatory behaviour closest to that of the numerical solution, the agreement between the numerical solution and the approximation calculated by the neural network cannot be said to be optimal. We have performed an extensive exploration of the parameter space of the activation function and the hyper-parameters of the neural network. These experiments have shown us that the reason for this disagreement is not due to improper values chosen for these parameters. The reason should rather be sought in the temporal trend of the enzyme and substrate.

- The analyses of Goldbeter et al. (2013) show that when the enzyme responsible for product degradation approaches saturation, the oscillations take on a distinct triangle shape. When the product sink is linear, the product peak resembles a pulse. The decreasing sections of the impulse curve with the parameters used for this simulation have a very high slope (i.e. a derivative close to infinity) which the neural network cannot resolve appropriately.

THANKS FOR YOUR INTEREST!

Paola.Lecca@unibz.it

Please refer to the paper in conference proceedings for the details and the bibliographical references.