University of West Bohemia
Faculty of Applied Science
Department of Cybernetics

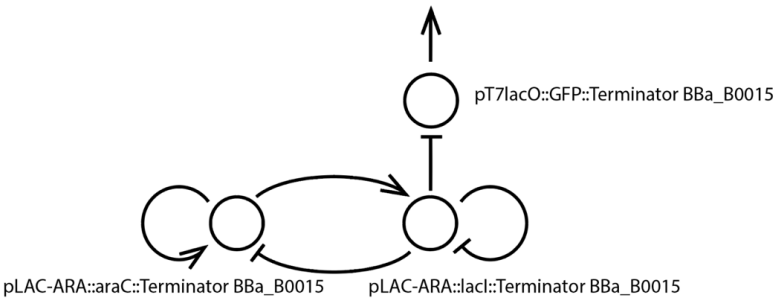
KKY/ZMB
FINAL PROJECT

Yauheni Petrachenka
August 4, 2024

1 Assigment

INTRODUCTION TO CELLULAR SYSTEM MODELING
FINAL PROJECT 2023 SUMMER
Due date: 30.6.2024

In this project you will virtually perform a bioengineering project. The objective of the project is to tune a system published in Nature by Stricker et al. in 2008. Note that in practice, tuning means changing the DNA sequence. In this project, the accompanying steps of DNA design, DNA assembly, and cell transformation are assumed to be automatic. In practice, however, each sequential genetic modification amounts to at least 1 week of laboratory work.



Mathematical modeling
The system described in the figure is a gene regulatory network comprising two interconnected single node motifs, a PAR connected to an NAR. If correctly tuned, the system generates oscillations in GFP. You will derive the differential equation model, construct a Matlab simulation, and find parameter values that result in oscillatory behavior. Note, the mathematical model introduced by Stricker et al. in 2008 is very detailed. In this section it is ok to use the simplified models introduced in lectures.

Break down modeling into the following steps and submit results from each part.

1) Describe the underlying chemical reaction network. Note, this chemical reaction network will indicate how activation and repression in the first gene jointly affect gene 1 expression (AND/OR gate).
2) Write down the system of differential equations using the mass action law.
3) Select physically relevant values for all reaction rates.
4) Simulate your system using MATLAB.
5) Identify reaction rates that yield oscillations in GFP. Oscillation period should be approximately 40min and the magnitudes of lacI and araC oscillation should be within an order of magnitude of each other.
6) Rank your parameter search by laboratory time required to arrive at this design. Assume 1 week for each sequential step. (Hint: tuning steps can also be performed in parallel)
7) (BONUS, not mandatory) Simulate the same system using SSA and compare the trajectories. Do you see the same oscillatory behavior?

2 Creating a model

First, we need to establish a model. The system under consideration consists of protein P1, which activates the transcription of Gene 1, leading to the production of protein P1 itself. Additionally, P1 activates the transcription of Gene 2, which produces protein P2. Protein P2 acts as a repressor for the transcription of both Gene 1 and Gene 2. To model this system, we will define a chemical reaction network. Parameters listed under the reaction arrow indicate repression, while parameters listed above the arrow indicate activation. The resulting chemical network can be expressed as follows:

$$\mathrm{G}_1 + \mathrm{P}_1 \xrightarrow{kD_{11},n_{11}} \mathrm{P}_1 \tag{1}$$

$$\mathrm{P}_1 + \mathrm{G}_2 \xrightarrow{kD_{12},n_{12}} \mathrm{P}_2 \tag{2}$$

$$\mathrm{G}_2 + \mathrm{P}_2 \xrightarrow[kD_{22},n_{22}]{} \mathrm{P}_2 \tag{3}$$

$$\mathrm{P}_2 + \mathrm{G}_1 \xrightarrow[kD_{21},n_{21}]{} \mathrm{P}_1 \tag{4}$$

We should combine the first and fourth reactions as well as the second and third reactions to derive the differential equations:

$$\frac{dp_1}{dt} = \beta_1 \cdot f_1(p_1) \cdot g_1(p_2) \tag{5}$$

$$\frac{dp_2}{dt} = \beta_2 \cdot f_2(p_1) \cdot g_2(p_2) \tag{6}$$

For modeling this system, we will use Hill functions to describe activation and repression, and we must include a decay term that accounts for dilution and degradation. The resulting equations for MATLAB simulation are as follows:

$$\frac{dp_1}{dt} = \beta_1 \cdot \frac{p_1^{n_{11}}}{p_1^{n_{11}} + kD_{11}^{n_{11}}} \cdot \frac{1}{1 + (\frac{p_2}{kD_{21}})^{n21}} - (\alpha_{delution} + \alpha_{degration}) \cdot p_1 \tag{7}$$

$$\frac{dp_2}{dt} = \beta_2 \cdot \frac{p_1^{n_{12}}}{p_1^{n_{12}} + kD_{12}^{n_{12}}} \cdot \frac{1}{1 + (\frac{p_2}{kD_{22}})^{n22}} - (\alpha_{delution} + \alpha_{degration}) \cdot p_2 \tag{8}$$

We will generate 100 simulations with random parameters α, β, and kD_{ij}, keeping n_{ij} fixed at 5. The concentration of p_2 will be plotted in each graph. The results are illustrated in Figure 1.

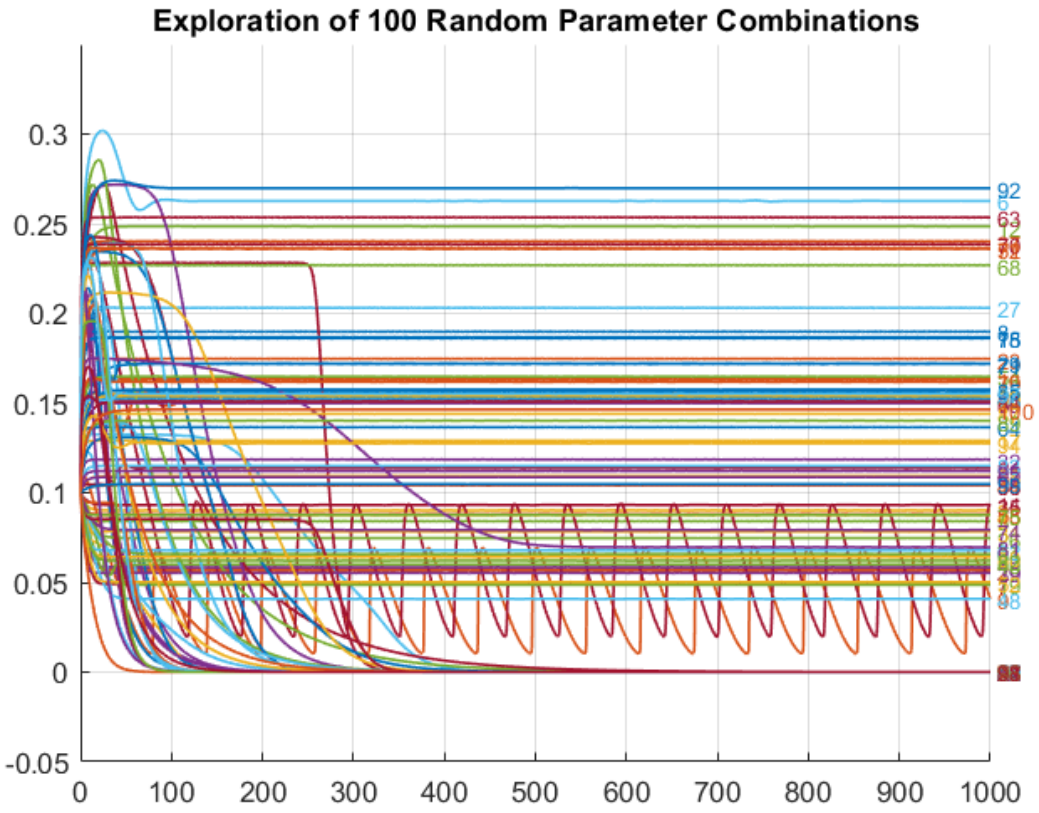


Figure 1: Exploration

As a result we have 100 curves we can see that only two of them oscillating. These are curves numbered 9 and 14, we will consider each curve separately. We will start with curve number 9, and plot concentration of p_2 and p_1.

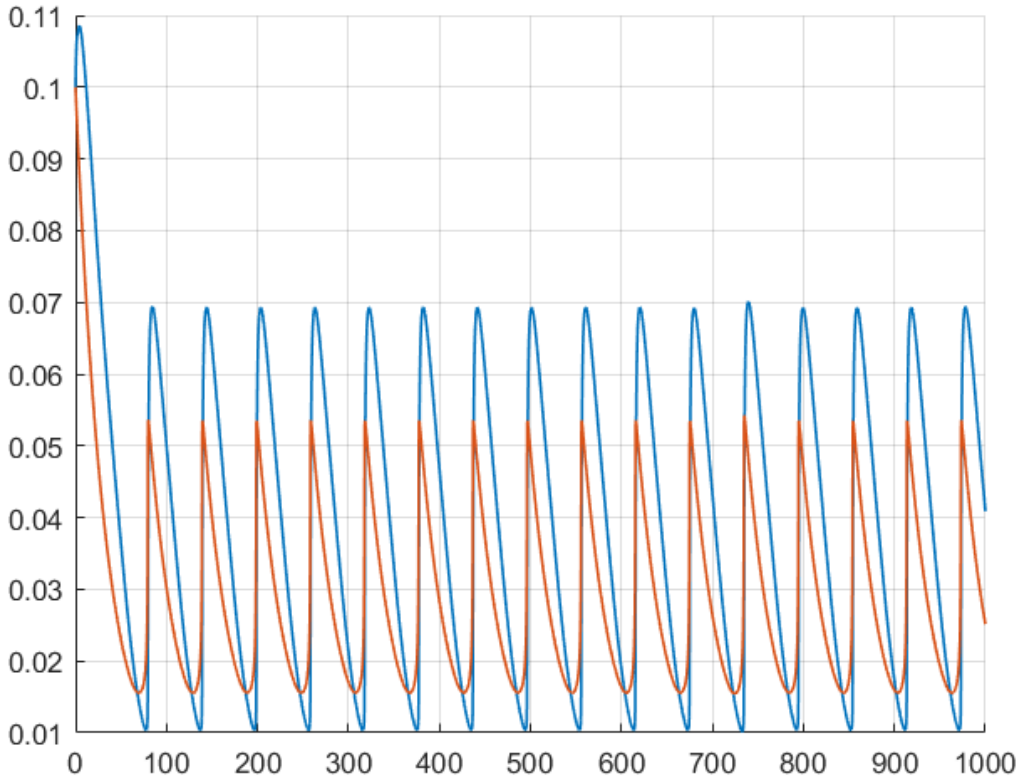


Figure 2: Curve number 9

We can see that there magnitude oscillating with the same phase and there period is approximately 70, that doesn't satisfies the condition of required period 40. So, curve number 9 fulfills only the oscillation condition and does not fulfill other conditions. That curve has such parameters:

- $\mathrm{kD}_{11} = 0.0765$
- $\mathrm{kD}_{12} = 0.0950$
- $\mathrm{kD}_{21} = 0.0140$
- $\mathrm{kD}_{22} = 0.0466$
- $\alpha_1 = 0.0307$
- $\alpha_2 = 0.0626$
- $\beta_1 = 2.9495$
- $\beta_2 = 1.4367$

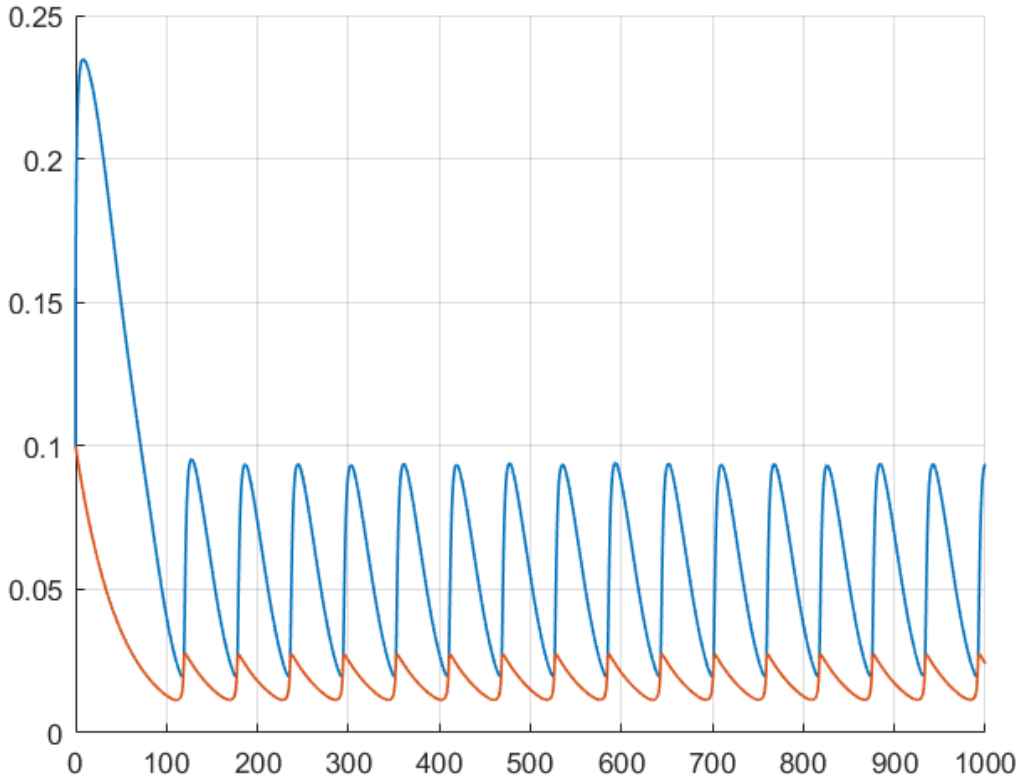


Figure 3: Curve number 14

As we can see fluctuations in the magnitude of the first protein do not perfectly correspond to the fluctuations of the second protein, however, this criterion is better met than in curve 9. The approximate oscillation period is 50, which almost corresponds to the required value. That curve has such parameters:

- $\mathrm{kD}_{11} = 0.0317$
- $\mathrm{kD}_{12} = 0.0676$
- $\mathrm{kD}_{21} = 0.0111$
- $\mathrm{kD}_{22} = 0.0902$
- $\alpha_1 = 0.0207$
- $\alpha_2 = 0.0570$
- $\beta_1 = 2.1950$
- $\beta_2 = 2.1686$

3 Labaratory experement

For laboratory experiments, we will also use GFP (Green Fluorescent Protein) as a reporter to indicate the transcription activity of the first and second genes. Ideally, we could use modified versions of GFP to distinguish which gene is being transcribed. Otherwise, we need to monitor the system and, in cases where GFP fluorescence remains constant, infer that the system might be fluctuating. In such cases, the experiment should be repeated twice: once with GFP attached to the first gene and once with GFP attached to the second.

In scenarios where modified GFP is available, we can conduct a single 40-minute experiment, observing colour changes, and then decompose the resulting fluorescence into green and the modified GFP's color. This will reveal the system dynamics. If no modified GFP is available, we will introduce GFP into both genes and monitor the fluorescence. If no change in fluorescence is observed within 40 minutes, we might suspect a potential system fluctuation. However, if fluorescence does change, this may indicate that the system either lacks oscillations or exhibits damped oscillations. Once we determine that the GFP fluorescence remains constant for 40 minutes, two more parallel experiments should be performed, with the GFP tag tagging only one gene at a time. This will allow us to deduce the dynamics of each gene's transcription and subsequently determine the system's physical parameters.

Thus, the time required for one experiment with modified GFP is 40 minutes plus the time needed for fluorescence decomposition. In the absence of modified GFP, the time needed for one experiment is a maximum of 40 minutes plus 40 minutes for two parallel experiments.

4 Conclusion

This study successfully developed a mathematical model to simulate the dynamics of a gene regulatory network involving positive and negative autoregulation. By randomly sampling 100 parameter sets, we identified conditions under which the system exhibits oscillatory behavior, essential for understanding gene expression patterns. However, achieving the precise period and amplitude of oscillations required for practical applications remains challenging, as demonstrated by the varied results from different parameter combinations.

In parallel, we proposed the use of GFP as a reporter to experimentally validate the model. The use of modified GFP could significantly streamline the process by allowing simultaneous tracking of multiple genes, while, in its absence, additional experiments would be needed to accurately capture the system's dynamics. Overall, while the model provides a solid foundation, further optimization and experimental refinement are necessary to fully realize its potential in synthetic biology.