Miniproject 1: Genetic oscillator

Scientific Computing for Data Analysis, 2023

Reproducibility is an important principle underpinning the scientific method. The results obtained in e.g. an experiment should be achieved again when the study is replicated. In this mini-project you will reproduce parts of the results in a research article[1]. The article has been important in the field of systems biology, and a deterministic and a stochastic model are studied, both describing a "circadian clock". You won't (probably) be able to understand everything in it, as it requires special knowledge in biology, but you will be able to understand enough for this miniproject.

Background

Our environment contains many cyclic processes, such as the diurnal rhythm of light and darkness. Many organisms use molecular mechanisms known as "circadian clocks", to adapt to such changes. It simultaneously promotes the expression of a negative element, which in turn represses the positive element. The cycle completes itself upon degradation of the negative element and re-expression of the positive element. In the article[1], a circadian clock model containing two proteins is used: an *activator* protein A and a *repressor* protein B. The corresponding activator and repressor genes are denoted by D_A and D_R, respectively. (D'_A and D'_R denote the same genes with A and R, respectively, bound to them.) See the article[1] for a more detailed explanation.

Preparations

You will use Python to solve the problems in the assignments below. You have two options to choose from:

Option 1: You can use the course ODE and SSA codes, accessible from last part of Lab 1, or form Studium via `Modules => Lectures => Lecture 4`. You would need to adapt the ODE function, the propensity functions, parameters, and the state-change matrix according to the new model.

Option 2: Alternatively, you can use the Python library `Gillespy2`. This library is the same one used in the last part of Lab 1, and it is likely most convenient to work within the `stochSS` environment.

---

[1] *José M. G. Vilar, Hao Yuan Kueh, Naama Barkai, Stanislas Leibler : Mechanisms of noise-resistance in genetic oscillators*, PNAS April 30, 2002 vol. 99 no. 9 page 5988-5992. The article is available in Studium.
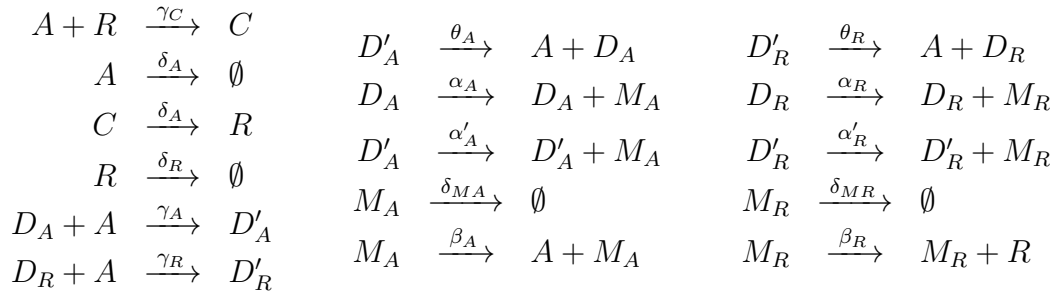
Assignment

a) (Deterministic) Reproduce the result in Fig 2a, and Fig 2b in the article by solving the ODE model given in page 1 (Eq. [1]) to see how both proteins A (activator) and R (repressor) vary over 400 hours. The model parameters and initial values to be used are given in the caption of Fig. 1.

For this task, if you opt for option 1, you will utilize the built-in function `solve_ivp` from the `scipy.integrate` library. In the provided code, you only need to introduce new parameter values, initial conditions, and adjust the timespan and the `ODEfun` function according to the right-hand sides in Eq. [1] of the article. Subsequently, plot only two variables A and R.

For option 2, use the library `Gillespy2` to simulate the model, here with a deterministic method (ODE). You do not need to set up the ODE-model described in Eq. [1] of the article. Instead, set up the stochastic model below (part b), and use method `ODE` in `Gillespy2`. You must choose a timespan (timesteps) that is not too coarse. If the curves get too choppy, you must probably increase the number of steps. However, if you use the ODE-switch in `Gillespy2`, it is equivalent to the ODE-model presented in the article. *Note:* If you simply adjust the code from the last part of lab 4 and use the command `results.plot`, you get 1 plot for all species and will not see anything. Instead plot the quantities that we are interested in (here A and R) separately by using `matplotlib.pyplot`.

b) (Stochastic) The stochastic model used in the article (it is not explicitly in the text) is the following: (16 reactions)

$$A + R \xrightarrow{\gamma_C} C \qquad D'_A \xrightarrow{\theta_A} A + D_A \qquad D'_R \xrightarrow{\theta_R} A + D_R$$
$$A \xrightarrow{\delta_A} \emptyset \qquad D_A \xrightarrow{\alpha_A} D_A + M_A \qquad D_R \xrightarrow{\alpha_R} D_R + M_R$$
$$C \xrightarrow{\delta_A} R \qquad D'_A \xrightarrow{\alpha'_A} D'_A + M_A \qquad D'_R \xrightarrow{\alpha'_R} D'_R + M_R$$
$$R \xrightarrow{\delta_R} \emptyset \qquad M_A \xrightarrow{\delta_{MA}} \emptyset \qquad M_R \xrightarrow{\delta_{MR}} \emptyset$$
$$D_A + A \xrightarrow{\gamma_A} D'_A \qquad M_A \xrightarrow{\beta_A} A + M_A \qquad M_R \xrightarrow{\beta_R} M_R + R$$
$$D_R + A \xrightarrow{\gamma_R} D'_R$$

This is a discrete Markov process with a continuous time parameter. Reproduce the results in Fig 2c and Fig 2d, i.e. simulate the circadian rhythm using the stochastic numerical method SSA and study how the proteins A and R vary during 400 hours. Use the same parameter values as in part (a).

What happens when you run the simulation several times? Does the solution agree with the ODE-solution in part (a), and does it exactly agree with the solution in Fig 2c and Fig 2d? If not, why?

c) The authors of the article point out that the presence of random white noise in the

stochastic model sometimes can lead to qualitative differences in comparison with the deterministic model. One illustration of this is shown in Fig. 5, and your task here is to reproduce the result in Fig 5. What is the difference between the deterministic and the stochastic solution in this particular case?

d) Contemplate and discuss the choice of mathematical models, i.e. a deterministic vs. stochastic model. When do you think the deterministic model is a better description of reality, and when is a stochastic model to prefer? Use the results above and also (part of) the article. For example, you can read on page 1: "The preceding equations [i.e. the ODE equations] would be strictly valid in a well stirred macroscopic reactor. At the cellular level, a more realistic approach has to consider the intrinsic stochasticity of chemical reactions". Also, read *Conclusions* at the end of the article.

Report

Your mini-project report should be submitted as a PDF file, and it should include your code. You can write the report using stochSS, Google Colab, or your preferred environment. For instructions on how to write a report using stochSS and Google Colab, visit the page "About Miniprojects" on the Studium course page.

If you do not want to work with StochSS and if you opt for option 2, you will need to install `Gillespy2` in your environment, which may require some effort. For option 1 just copy and paste the code and modify it for the new model.