Stochastic dynamics of regulated expression

Javier Castillo Uviña
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.**

In this paper will be studied the determistic and stochastic dynamics behind the gene expression and the gene regulation processes for a single protein specie production from mRNA molecules. The deterministic case will be reproduced thank to a RungeKutta 4 algorithm and the stochastic case will be reproduced thank to a Gillespie algorithm. The resuls will be discussed from the point of view of the Master Equation for the constitutive mRNA expression.

I. INTRODUCTION

The protein synthesis easier process is the mRNA sequence translation into a polypeptide in the cell, what is well known as gene expression. In this paper, gene expression is summarized as the process by a gen is used to generate mRNA which is read to produce a protein. We will study a system were are involved the mRNA transcription and a unique specie of protein translation. This system is described with the following expressions:

$$\begin{aligned} \frac{dm}{dt} &= \alpha_m - \delta_m m \\ \frac{dp}{dt} &= \alpha_p m - \delta_p p \end{aligned} \tag{1}$$

The first equation describes the mRNA concentration variation thorugh: the mRNA generation regulated by the α_m rate and the mRNA destruction based on the δ_m rate and the mRNA concentration. The second equation describes the protein concentration variation through: the protein generation regulated by the α_p rate which is also related with the mRNA concentration and the protein destruction based on the δ_p rate and the protein concentration. All this four rates are positive real numbers. The stationary values can be computed analytically from 1, where the concentration rates are equal to 0:

$$\begin{aligned} m^{st} &= \frac{\alpha_m}{\delta_m} \\ p^{st} &= \frac{\alpha_m \alpha_p}{\delta_m \delta_p} \end{aligned} \tag{2}$$

A more complex system which is also studied in this paper is the gene regulation. This process describes a system where the protein affect to the mRNA creation due to the fact that the protein created affect to the transcription. This process, where exists a regulation, is known as a motif. We will study the case that the generated proteins repress the mRNA production, but there exists cases where protein can activate the mRNA creation. For the repression case the desciptive reaction is the following:

$$\begin{aligned} D + p &\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [Dp] \\ D &\underset{k_t}{\rightarrow} D + m \end{aligned}$$

D is the DNAPromoter the different k are the constants associated to each reaction. As we can see in [1] the resulting model is:

$$\begin{aligned} \frac{dm}{dt} &= \alpha_m f_-(p) - \delta_m m \\ \frac{dp}{dt} &= \alpha_p m - \delta_p p \end{aligned} \tag{3}$$

where:

$$f_-(p) = \frac{K}{K + p} \tag{4}$$

The stationary values can be computed analytically from 3, where the concentration rates are equal to 0:

$$\begin{aligned} m^{st} &= \frac{\alpha_m}{\delta_m} f_-(p^{st}) \\ p^{st} &= \frac{\alpha_m \alpha_p}{\delta_m \delta_p} f_-(p^{st}) \end{aligned} \tag{5}$$

II. METHODOLOGY

In order to make this study a code is wrote to describe the dynamics behind 1 and 3. It is solved numerically with a Runge-Kutta 4 in order to show the deterministic behaviour. Besides, the Gillespie algorithm is used to simulate stochastic dynamics in a volume $V = \mu m^3$. The code has been written in fortran 77 and the graphics has been plotted thank to gnuplot.
What we will se in all three following cases is a time evolution from initial condition m=0 and p=0. Each case has different equation rates parameters and final time, all cases show the deterministic behaviour and two stochastic simulations. The mRNA and protein concentration probability densities have been computed with $N = 1000$ simulations from the deterministic stationary values until a final time of 10 minutes for each simulation in order

*Electronic address: jcastiuv7@alumnes.ub.edu

to have a profile of stochastic steady values.
In order to know how the system will reach the stationary values for the stochastic dynamics we can study the Master Equation for the constitutive mRNA expression in [1] we obtain these expressions for the distribution shape and parameters of probability density for the mRNA and the protein concentration. What we get for Gillespie algorithm is a gaussian for the mRNA concentration distribution with parameters:

$$\begin{aligned} \mu &= \frac{\alpha}{\delta_m} \\ \sigma &= \sqrt{\frac{\alpha}{\delta_m V}} \end{aligned} \tag{6}$$

For the protein concentration we obtain a gaussian distribution too but the parameters are:

$$\begin{aligned} \mu &= \frac{P^{st}}{V} \\ \sigma &= \frac{\sqrt{P^{st}}}{V} \end{aligned} \tag{7}$$

Being P the number of molecules in the Gillespie algorithm not the concentration.

III. RESULTS

A. Constitutive expression: default case

Parameter values for this first default case are:

- $\alpha_m = 100nMmin^{-1}$
- $\delta_m = 1min^{-1}$
- $\alpha_p = 10min^{-1}$
- $\delta_p = 0.1min^{-1}$
- Stationary point (100, 10000)

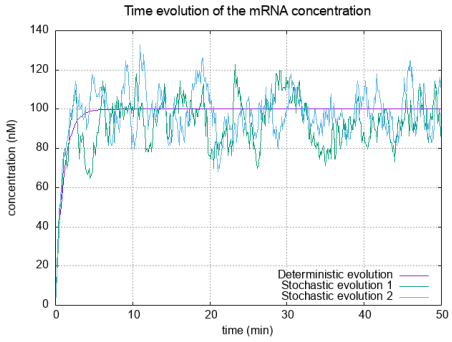

FIG. 1: It is shown the time evolution of the mRNA concentration through 50 min.

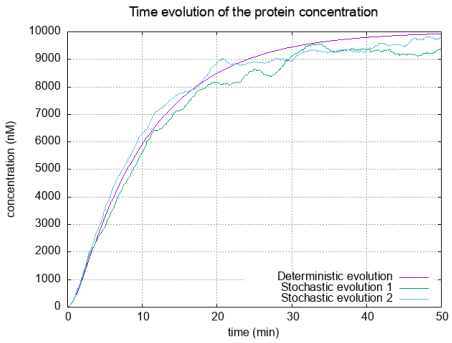

FIG. 2: It is shown the time evolution of the protein concentration through 50 min.

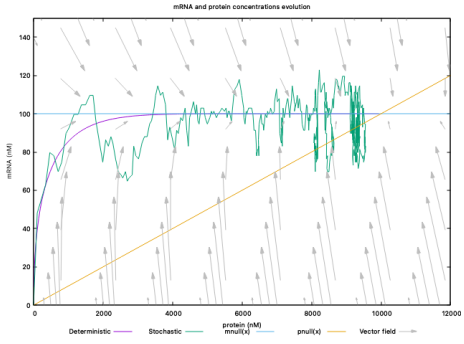

FIG. 3: It is shown the phase portrait for the mRNA and protein concentration evolution.

The nullclines shown in Fig 3 show the stationary value which is reached for the deterministic case. The stochastic dynamics only reaches the stationary value with a certain probability.

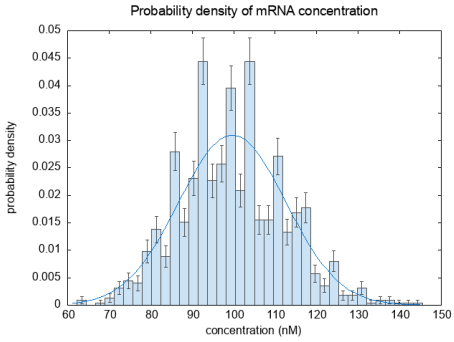

FIG. 4: mRNA concentration probability density profile for the $N = 1000$ simulation at the steady state shown in boxes and a gaussian distribution with the simulation parameters plotted in a blue line.

The value parameter for the gaussian fitted in Fig 4 are G(100,12.87) where G is the gaussian with (μ,σ) parameters. For the stochastic dynamic the mRNA concentration approaches a gaussian because we have variations due to a white noise. The parameter for the gaussian

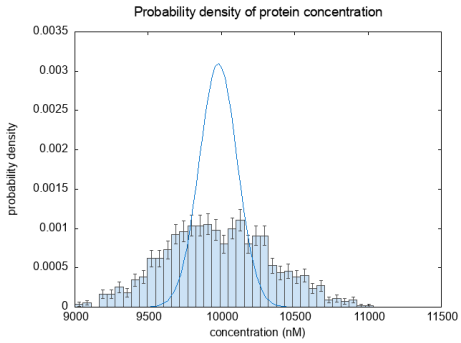

FIG. 5: Potein concentration probability density profile for the $N = 1000$ simulation at the steady state shown in boxes and a gaussian distribution with the simulation parameters plotted in a blue line.

in Fig 5 are G(10000,128.86), the probability density for the protein concentration does not fit the gaussian distribution because is related with the mRNA concentration and it accumulates the variations due to the noise in the mRNA concentration and the variations due to the noise in the protein concentration.

B. Constitutive expression: effect of transcription and translation rates

Parameter values for the second case are:

- $\alpha_m = 1000nMmin^{-1}$
- $\delta_m = 1min^{-1}$
- $\alpha_p = 1min^{-1}$
- $\delta_p = 0.1min^{-1}$
- Stationary point (1000, 100000)

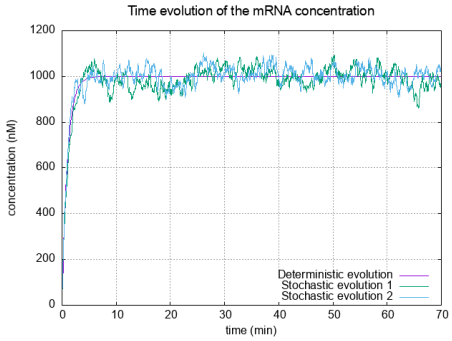

FIG. 6: It is shown the time evolution of the mRNA concentration through 70 min.

For this case the time evolution shape should be similar to the previous but the steady values change according to 2. Besides, the mRNA concentration should reach the

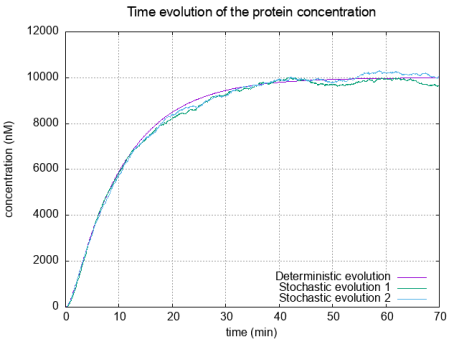

FIG. 7: It is shown the time evolution of the protein concentration through 70 min.

stationary value earlier than the case before due to the bigger α_m and this is what is shown in Fig 6, on the other hand the protein concentration take more time to reach the stationary value due to the diminution in the α_p, as we can see in Fig 7.

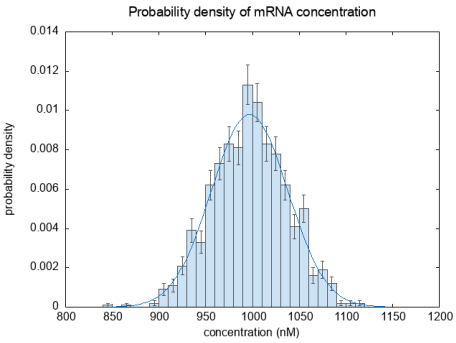

FIG. 8: mRNA concentration probability density profile for the N = 1000 simulation at the steady state shown in boxes and a gaussian distribution with the new simulation parameters plotted in a blue line.

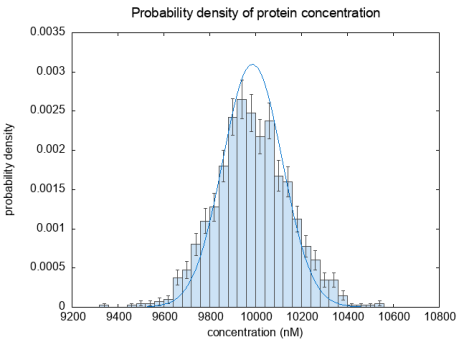

FIG. 9: Potein concentration probability density profile for the N = 1000 simulation at the steady state shown in boxes and a gaussian distribution with the new simulation parameters plotted in a blue line.

In this case the gaussian fitted is G(1000,40.75) in Fig

8 and G(10000,128.86) in Fig 9. As we can see mRNA concentration follos the gaussian as the case before and the protein concentrationts fit better the gaussian distribution this can be possible due to the fact that the lower the generation rate the lower the variations they can suffer.

C. Regulated expression: effect of negative feedback

- $\alpha_m = 10100 * f_-(p) nM min^{-1}$
- $\delta_m = 1 min^{-1}$
- $\alpha_p = 10 min^{-1}$
- $\delta_p = 0.1 min^{-1}$

For the last case the parameter have been chosen in order to obtain the same stationary values than the first case. So the stationary point is (100, 10000).

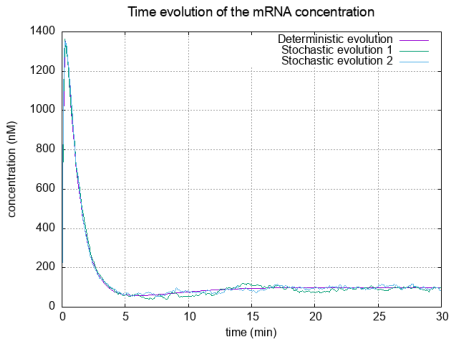

FIG. 10: It is shown the time evolution of the mRNA concentration through 30 min.

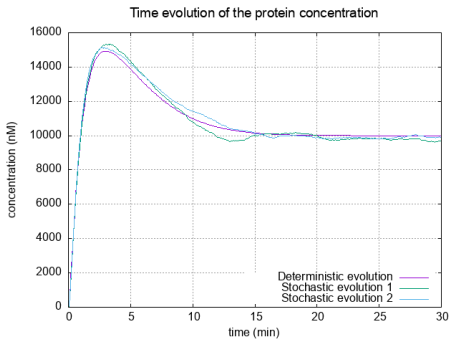

FIG. 11: It is shown the time evolution of the protein concentration through 30 min.

The stationary point is well-reached in Fig 10 and Fig 11 as it is expected. The shape can be explained because we have a big constant α_m so it is expected to grow rapidly, but as it is said in the Introduction a repression case is studied so they decay as proteins grow. For the proteins case They grow due to the fact they are related to the mRNA concentration. As they grow the m decays so the protein concentration has to decay to but both concentrations find the stationary value in a shorter time than the non-regulated case.

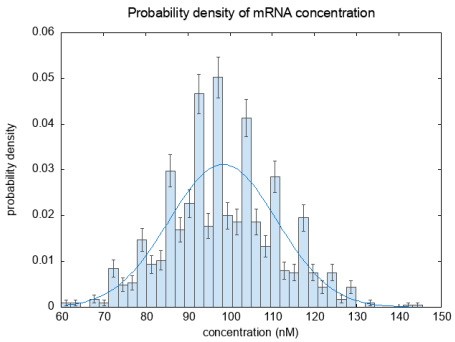

FIG. 12: mRNA concentration probability density profile for the N = 1000 simulation at the steady state shown in boxes and a gaussian distribution with the new simulation parameters plotted in a blue line.

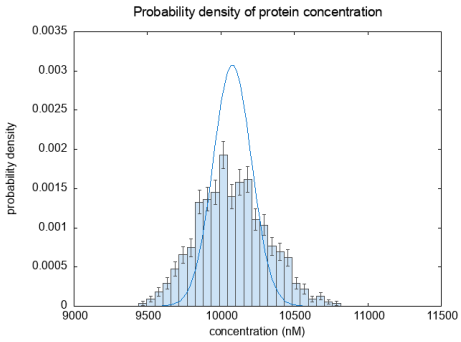

FIG. 13: Potein concentration probability density profile for the N = 1000 simulation at the steady state shown in boxes and a gaussian distribution with the new simulation parameters plotted in a blue line.

In Fig 12 the gaussian distribution parameters are G(100,12.87) and G(10000,128.86) for the protein concentration distribution in Fig 13. Again the mRNA fits better than the protein the proposed gaussian profile. Nevertheles, the protein concentration profile is more similar to the analytic master equation approach than the non-regulated case this is due to the fact that a negative feedback causes this effect on the stationary value convergence.

IV. CONCLUSIONS

The Gillespie algorithm let us accuratelly illustrate how the stochastic gene expression and gene regulation dynamics work. They previous hypotheses are correctly fulfilled. The stationary point 2 is reached with a certain

probability. This probability follows a gaussian profile with parameters 6 for the mRNA concentration and 7 for the protein concentration. It also reproduce the protein probability density does not fit the gaussian for the non-regulated case due to the mRNA concentration variations. Eventually it also reproduce correclty the negative feedback effects.

---

[1] *Computational Systems Biology* notes of Complex Systems and Biophysics master degree