Prof. Dr. Joachim Krug
Alexander Klug
Benjamin Schmiegelt
Suman G. Das

Summer term 2021

Biological Physics II

Problem Set 6

Please hand in your solutions before 12:00 noon on Wednesday, July 7, 2021.

1. Gene expression noise

2+7+5+6+4+6 = 30 pts

Consider the "standard model" of gene expression discussed in the lectures[1]. We use the same notation as in the lectures, so that n_1, n_2 and n_3 are the numbers of a gene, its mRNA transcripts and the corresponding protein respectively. The maximum value for gene copy number is $n_1^{\max}$. Gene activation and deactivation happen with rate constants λ_1^+ and λ_1^- respectively. The mRNA synthesis rate and lifetime have parameters λ_2 and τ_2, and the corresponding parameters for protein are λ_3 and τ_3. We will use $n_1^{\max} = 1$. (Refer to the lectures for any further details.)

a) Consider the special case where the gene is always on. Write down the two-variable master equation for this case.

b) From the master equation, derive closed equations for the evolution of the first and second moments of $n_2(t)$. Solve these equations to obtain the moments as a function of time, with the initial condition $n_2(0) = 0$. Use your solution to calculate the Fano factor of mRNA as a function of time. Now, using the parameters $\lambda_2 = 1, \tau_2 = 2$, plot your analytical expressions of $\langle n_2(t) \rangle$ and the Fano factor starting from $t = 0$ up to sufficient time so that the convergence to steady state is observed. Also, compute the variance of mRNA numbers in steady state.

c) Use the Gillespie algorithm (or any other method) to simulate the above system and numerically compute and plot the same quantities on the same graph. Also, compute the distribution of n_2 in steady state and plot the appropriate Poisson distribution that fits the data.

d) Now use the same parameters as before, but allow the gene to activate and de-activate with rates $\lambda_1^+ = 0.4$ and $\lambda_1^- = 0.2$. Analytically find $\langle n_1 \rangle$ and $\langle n_2 \rangle$ in steady state. Numerically compute the same quantities and compare them. (Hint: For the numerical computation, one can sample values from a single trajectory at evenly-spaced time points, with the value of the variable at a time point corresponding to the most recent event in the past.) Plot the steady state distribution of n_2 and show that it is not Poissonian.

[1] Detailed treatments of this model can be found in [1] *Paulsson, Johan. "Models of stochastic gene expression." Physics of life reviews 2.2 (2005): 157-175*, [2] *Peccoud, Jean, and Bernard Ycart. "Markovian modeling of gene-product synthesis." Theoretical population biology 48.2 (1995): 222-234*, and [3] *Raj, Arjun, et al. "Stochastic mRNA synthesis in mammalian cells." PLoS Biol 4.10 (2006): e309*, and other papers.

e) The steady-state variance of mRNA can be shown to be

$$V_2 = \frac{\lambda_1^+ \lambda_2 \tau_2}{\lambda_1^+ + \lambda_1^-}\Big[1 + \frac{\lambda_1^+}{(\lambda_1^+ + \lambda_1^-)}\frac{\lambda_2}{(\lambda_1^+ + \lambda_1^- + 1/\tau_2)}\Big]. \tag{1.1}$$

Numerically compute and plot V_2 as a function of λ_1^+ (while keeping the other parameters same as in the previous part) starting with $\lambda_1^+ = 0.1$ and increasing it till V_2 effectively converges to an asymptotic value. Show that this asymptotic limit of V_2 matches with your analytical results in part b).

f) So far we have ignored the translation process. Now consider the same problem with the protein synthesis and degradation parameters $\lambda_3 = 0.5$ and $\tau_3 = 5$, and the rest of the parameters same as in part d). Numerically compute and plot the steady state distribution of n_3. Plot one set of sample trajectories of $n_1(t)$, $n_2(t)$ and $n_3(t)$ in steady state over a time interval of 25 units. Both analytically and numerically obtain $\langle n_3 \rangle$ in steady state and compare them.