Systems Biology Assignment

The deadline for the group assignment (Question I-II) is Sunday 19th May, 11:59pm. For your group work submit a single joint report (2-4 students per group) as a pdf file via the coursework website combining both questions I and II. Question I should be mostly straightforward. Question II is numerical simulation work and more time-intensive.

The deadline for Question III is Sunday 26th May, 11:59pm. Submit a separate pdf file for this part. Attempt Question III only if you want a top mark since its value is disproportionally small for how much effort it takes... Question III is an individual assignment for which you are not allowed to collaborate. Keep discussions about that question to an absolute minimum. If we believe you collaborated on the individual assignment we reserve the right to examine you orally. If you did discuss during any point, then state that on the first page.

Make sure to write your name(s) and email address(es) on the first page of your submitted files. Be concise! Submitting many disorganised figures will not earn any extra points. In fact, disorganised or indirect explanations will be interpreted as a lack of understanding and will be penalised.

If you email any questions to johan_paulsson@hms.harvard.edu or andreas.hilfinger@utoronto.ca we will send the response to the whole group to ensure fairness (provided there exists an answer to your question that doesn't give away the solutions to the problems).

Good Luck!

Question I: (30 marks, group work)

Consider the following mRNA dynamics of two genes that form an operon which is regulated by a transcription factor X_1.

$$\begin{aligned} x_1 &\xrightarrow{\alpha} x_1 + 1 \\ x_1 &\xrightarrow{x_1/\tau_1} x_1 - 1 \end{aligned} \tag{1}$$

$$\begin{aligned} x_2 &\xrightarrow{\lambda x_1} x_2 + 1 \\ x_2 &\xrightarrow{x_2/\tau_2} x_2 - 1 \end{aligned} \qquad \begin{aligned} x_3 &\xrightarrow{\lambda x_1} x_3 + 1 \\ x_3 &\xrightarrow{x_3/\tau_3} x_3 - 1 \end{aligned} \tag{2}$$

Here X_2 and X_3 are the mRNA species with lifetimes τ_2 τ_3, respectively.

A. For this three component model write down the drift and diffusion matrices M, D.

B. Use $M\eta + (M\eta)^T = D$ to solve for the components' normalised (co)variances in terms of average abundances and lifetimes. Is your answer an exact statement or an approximation?

C. Describe in words the scenario(s) in which fluctuations in mRNA levels will be dominated by extrinsic noise inherited from the transcription factor? Conversely, in which regime will the fluctuations in mRNA levels be dominated by intrinsic low-copy number noise in the transcriptional step?

Question II: (60 marks, group work)

Simulate the above model using the Gillespie algorithm:

A. Provide your code (structured and with comments) that simulates the above system for $\tau_2 = 2$, $\tau_3 = 4$ and pick one hundred different set of values for the remaining parameters such that your systems go from a regime in which intrinsic noise dominates the mRNA fluctuations to one in which extrinsic noise dominates the mRNA fluctuations. To illustrate that you have sampled systems ranging from intrinsic to extrinsic noise dominated regimes (and in between) make a diagram in which you plot the intrinsic versus the extrinsic contributions to the mRNA fluctuations as obtained from your simulation results (make two plots, one for X_2, and one for X_3).

B. For all your simulations show explicitly that the simulation ran long enough to describe the stationary state[1], i.e. that you have sampled enough reaction events that the flux balance $\langle R_i^+ \rangle = \langle R_i^- \rangle$ holds for all components. To illustrate that your simulations were accurate present three histograms characterising the relative errors in the flux balance relations for each of your simulations. Similarly, check whether your stochastic traces satisfy the normalised (co)variance equations derived in Question 1B. Illustrate the accuracy of your simulations by presenting six histograms corresponding to the relative deviations of η_{ij} from the analytical expressions.

C. Make a figure plotting the "observed" noise in mRNA levels from your simulations versus the theoretically expected mRNA noise (use the "observed" $\langle x_1 \rangle$, $\langle x_2 \rangle$ and $\langle x_3 \rangle$ from your simulations in those equations) from question 1B. Comment on the accuracy of your simulations.

D. Pick a specific set of parameter values and present one (short) time trace for the transcription factor and one of the mRNA at stationarity to illustrate how mRNA levels respond to changes in transcription factor levels. Hint: show just a handful of mRNA transitions and combine "two different y-axes" in the same plot, i.e. rescale the axes such that the long-term mRNA and transcription factor averages would coincide.

[1]Note, that the stationary state averages are to be derived from the probability distribution obtained from looking how much time X_1 and X_2 have spent in the respective states over the complete time trace.

E. Make a figure in which you plot the "observed" correlations between the mRNAs versus the ratio of their CVs, i.e. plot ρ_{23} versus CV_3/CV_2. Do your results exhibit a pattern? If not, explain how you arrived at your conclusion. If yes, describe the pattern.

Question III: (10 marks, individual work)

Consider the symmetric version of the model from Question 1 in which we allow for negative feedback of X_2 and X_3 onto X_1, i.e. consider the following system

$$\begin{aligned} x_1 &\xrightarrow{f(x_2,x_3)} x_1+1 \\ x_1 &\xrightarrow{x_1/\tau_1} x_1-1 \end{aligned} \tag{3}$$

$$\begin{aligned} x_2 &\xrightarrow{\lambda x_1} x_2+1 \\ x_2 &\xrightarrow{x_2/\tau_{mRNA}} x_2-1 \end{aligned} \qquad \begin{aligned} x_3 &\xrightarrow{\lambda x_1} x_3+1 \\ x_3 &\xrightarrow{x_3/\tau_{mRNA}} x_3-1 \end{aligned} \tag{4}$$

where $f(x_2, x_3)$ is a decreasing function that is symmetric with respect to exchanging x_2 and x_3.

A. For this model write down the drift and diffusion matrices M, D in terms of the unknown log-log sensitivity $h = -\frac{\partial \ln f(x_2,x_3)}{\partial \ln x_2} = \frac{\partial \ln f(x_2,x_3)}{\partial \ln x_3}$ (by symmetry).

B. Use $M\eta + (M\eta)^T = D$ to solve for the components' normalised (co)variances η_{22} and η_{33} in terms of average abundances, lifetimes and h. Is your answer an exact statement or an approximation?

C. Can you construct a feedback loop that suppresses noise levels in X_2 arbitrarily below Poisson noise levels $1/\langle x_2\rangle$? If not, why not? If yes, what are the conditions to achieve such effective noise suppression given your answer to the previous sub-question?

D. By considering $\frac{\eta_{22}}{1/\langle x_2\rangle} + \frac{\eta_{33}}{1/\langle x_3\rangle}$ determine whether you can construct a feedback loop that simultaneously suppresses noise levels in X_2 and X_3 arbitrarily below Poisson noise levels. If not, why not? If yes, what are the conditions to achieve such effective noise suppression given your answer to the previous sub-question?