***Modelling Gene Expression 2***

**INVESTIGATION USING A STOCHASTIC MODEL**

In the computer room practical sessions we build up the steps necessary to carry of a stochastic analysis of a simple system (red and green balls drawn from a bag).

We can use the same steps to model other systems that we have previously encountered. These are provided as files in the stochastic examples folder.

**Coding the stochastic system**

*Read through the examples and check you understand the structure of the stochastic versions and how they relate to the deterministic models.*

The template file gene\_exp\_stoch1.py is provided to run a simulation of stochastic gene expression (without including negative feedback via repression of transcription).

*1a) Complete the code to simulate the system stochastically:*

*- Fill the blanks left in the function* gillespie\_gene\_expression

*Steps 0 and 1 are already filled out.*

*Steps 2, 3 and 4 need completing.*

*- Input the correct initial conditions and parameter values.*

*When complete the code should plot the species concentrations over the time period 0 < t < 30000.*

*b) Contrast the observed behaviour of the stochastic simulation to the output of the deterministic model.*

*c) Add code that plots the average for the protein number over 100 simulation runs. How does this averaged behaviour compare to the deterministic model?*

Note. To avoid long waits set the number of runs to a lower value (e.g. 20) until your code is working. Once you have got it working increase it to 100.

**Measuring the noise expressed by the stochastic system**

*2a) Work with the results of 100 runs. Analyse the final state attained by the system at t=30000. Create a histogram showing the copy numbers of:*

*i) mRNA*

*ii) protein distribution*

*Find the mean and standard deviation of these distributions and comment briefly on their shape.*

Note. Remember to use only a low number of runs while getting your code working!

*b) The coefficient of variation CV is a measure of the relative noise level exhibited by a system.*

*Calculate this value for the mRNA and protein distributions at time t=30000.*

*c) Adjust the model to include negative feedback via repression of transcription, using the same parameters as used in the deterministic model.*

*Repeat the calculation of CV and compare your finding to the previous result.*

**Simulating transcriptional bursting**

In the model parameters used so far we have assumed that the average switching time for the gene to transition between the active and inactive state occurs on a timescale of seconds.

Recent experimental developments mean that we can observe mRNA transcription occurring it situ and able to quantify mRNA copy numbers on an absolute basis (i.e. count the number of mRNA transcripts).

It has been observed that often mRNA is produced in bursts (on a time scale of minutes or longer) that are separated by periods of inactivity.

In our model we can simulate this by adjusting the rates that the gene switches between its active and inactive state. The original values we used are:

|  |  |
| --- | --- |
| Model parameter | Value (seconds-1) |
|  |  |
|  |  |

|  |  |
| --- | --- |
| Model parameter | Value (seconds-1) |
|  |  |
|  |  |

Let’s adjust these in to the following values:

Note: this value was originally set at 5x60

(a typo – if you use this level you still get similar behaviour but its harder to compare)

*3a) Repeat the analysis that estimated average time spent in the active/inactive states and the average proportion of time that the gene will spend in the active state (exercises 5a. b. and c.) Comment on your findings.*

*b) How does the parameter change affect the behaviour of the model using a deterministic simulation? Comment this result.*

*c) How does the parameter change affect the behaviour of the stochastic model? Repeat the analysis of you completed in exercises 1 and 2 for the updated system.*

*(Consider the model without negative feedback only).*

*Comment on your findings.*