CPIB Summer School
Computer Practicals
Stochastic Modelling

Dov Stekel / Markus Owen

7th September 2011

In this practical session, you will see how deterministic and stochastic methods can be used to model gene regulatory networks. In particular, we will compare results of the two simulation paradigms and see how they can give quite different results. We will use COPASI in this practical, because it is easy to switch between equivalent ordinary differential equation and stochastic models.

Exercise 1: Constitutively Expressed Gene

You will start by modelling a gene that is not regulated at all. The protein product P of this gene is produced at rate k and removed from the cell at rate m. We can represent the production and removal of P by two chemical reactions:

0 -> P (rate constant k) P -> 0 (rate constant m)

This reaction scheme is highly simplified because it assumes that protein is produced from DNA in one step and removed from the cell in one step. Can you think of any other assumptions we have made?

For the reaction scheme above, we can write down an equation for the rate of change of the number n_p of molecules of P in the cell:

$$\frac{dn_p}{dt} = k - mn_p \qquad \text{(Equation 1)}$$

This equation is easy to solve if you know how. We can work out the "steady state solution" – i.e. what happens to the system if we wait a long time until it has settled down and is not changing any more. In this case, we can put $dn_p/dt=0$, so that $n_p=k/m$.

We can also write a simulation scheme to model this gene regulatory network. Although this is not a very interesting network, it's useful to see how we do the modelling– we will make it more interesting later on.

We will start by setting up a stochastic simulation, using COPASI. The above model can be found in "constitutive.cps".

1.1 Open COPASI and load "constitutive.cps". Go through the model (Species, Reactions, etc) and make sure you understand how it is defined. Also look at the Plots that are defined in the "Output" sub-menu (click "Plots" to see a table of all defined plots). Two are for use with stochastic simulations, and one for the deterministic case. When you load the model, you should see that the stochastic ones are "active". The parameters in the model are "P-

production.v" (corresponding to *k* in (Equation 1), and initially set to 1 #/min), and "P-degradation.k1" (corresponding to *m* in (Equation 1), and initially set to 0.01/min).

1.2 For these parameters, what would be the steady state of the equivalent deterministic system? (HINT: the formula for this can be found above)

1.3 Carry out a stochastic simulation for the constitutively expressed gene by going to "Time Course" in the "Tasks" sub-menu, and ensuring that the "Method" is "Stochastic (Gibson + Bruck)". Then click "Run".

Two windows should appear with the results of your simulation (particle numbers over time) and a histogram of the particle numbers (which, for sufficiently large particle numbers or over a sufficiently long period, should peak around the mean particle number during the simulation).

1.4 Repeat the stochastic simulation run, using the same parameters.

Do you get the same time history? If not, why not?

Do a long simulation run. Does the mean protein number agree with the predicted deterministic value for the steady state?

1.5 Try changing the production rate. Click on the "Show sliders" icon (5^{th} from the left along the top of the main window), and a small window should appear with a slider for "(P-production).v" on a log scale between 0.001 and 100. Each time you adjust the slider, the simulation is re-run and displayed. What happens as you increase the production rate? As you decrease it?

Next we will simulate the same system deterministically, using an ordinary differential equation (ode) solver.

1.6 In the "Outputs", "Plots" table, make only the deterministic plot active (using the check boxes). Then go to the "Time Course" panel, change the "Method" to "Deterministic (LSODA)", and click "Run". A new window should appear showing the time course of the deterministic simulation: this is the number of protein molecules as a function of time as predicted by the deterministic model.

1.7 Reflect on the similarities and differences between the results of the stochastic and deterministic models. Do you think that if you averaged over many runs of the stochastic model you would get the results of the deterministic model?

Exercise 2: An Autoactivating Gene

Now we will investigate a slightly more complicated gene regulatory network: a gene which is activated by its own protein product. We are going to get interesting differences between the deterministic and stochastic models.

A reaction scheme for an autoactivating gene is shown below:

D -> D + P	(rate kbasal)	C -> C + P	(rate kact)
D + PDIM -> C	(rate dnaass)	C -> D + PDIM	(rate dnadis)
P -> 0	(rate m)	PDIM -> 0	(rate m)
P + P -> PDIM	(rate dimass)	PDIM -> P + P	(rate dimdis)

In this scheme, two molecules of protein P can associate to form a dimer PDIM, which can dissociate back to two monomers. Can you see which reactions correspond to this?

The dimer PDIM can bind to the DNA (D) to form a protein-DNA complex (C), which can also dissociate. Can you see which reactions correspond to this?

Protein P is produced when the DNA is not bound by PDIM at a low rate kbasal. When PDIM is bound to the DNA, protein P is produced at a higher rate kact. Can you see where this happens in the reaction scheme?

Finally, protein monomers and dimers are removed from the system at rate m.

We will start by running a stochastic simulation of this reaction scheme. For this we will use the COPASI file "autoactivating.cps".

2.1 Go through the model (Species, Reactions, etc) and make sure you understand how it is defined. Also look at the Plots that are defined in the "Output" sub-menu (click "Plots" to see a table of all defined plots). As before, two are for use with stochastic simulations, and one for the deterministic case. When you load the model, you should see that the stochastic ones are "active". The initial parameter values are: kbasal=0.03, kact=1, m=0.01, dimass=0.00005, dimdis=0, dnaass=0.01, dnadis=0.06
The initial particle numbers are set to: D_0=1, C_0=0, P_0=0, $PDIM_0$=0

2.2 Run the simulation by going to "Tasks", "Time Course" and clicking "Run" (you should check that the "Method" is "Stochastic").

You should get a window with a plot of the number of each kind of particle (molecule), plus the total protein (P+2*PDIM), as a function of time. The second window should show a histogram of particle numbers.

What happens in your simulation?

What should have happened is that your simulation showed flipping between two states: one state with only a few protein molecules and another state with many protein molecules. In the low protein state, the system is stable because the gene is not activated and little protein is made. In the high protein state, the system is also stable because the gene remains activated and makes lots of protein. How does the histogram relate to these observations about low and high protein states?

Try experimenting with changing the parameter values and see what happens to these two states and to the rate of flipping between them.

1.8 Try changing the dimer association rate. Click on the "Show sliders" icon, and a small window should appear with a slider for "(Dimer association).k1" on a linear scale between 0.00001 and 0.002. Each time you adjust the slider, the simulation is re-run and displayed. What happens as you increase the association rate? As you decrease it?

The autoactivator is an example of a bistable gene network – one with two alternative stable states. There are lots of other examples and several bistable gene networks have been constructed experimentally in *E. coli*, or found in naturally occurring gene circuits. What might be the function of such a network in a cell?

Now we will simulate the same reaction network with a deterministic ode scheme.

2.3 If you have modified parameters, initial conditions, etc, rerun the stochastic simulation with the parameter values
kbasal=0.03, kact=1, m=0.01,dimass=0.00005, dimdis=0, dnaass=0.01, dnadis=0.06;
and initial conditions: D_0=1, C_0=0, P_0=0, $PDIM_0$=0; and a duration of at least 200000. You will recover the plot with the flipping transitions.

In the "Outputs", "Plots" table, make only the deterministic plot active (using the check boxes). Then go to the "Time Course" panel, change the "Method" to "Deterministic (LSODA)", and click "Run". A new window should appear showing the time course of the deterministic simulation.
What happens?

2.4 Run the deterministic simulation again with the same parameters but this time with P_0=100 (i.e. change the initial number of P molecules, via the "Model", "Biochemical", "Parameter Overview").
What happens this time?

2.5 You should see that there are big differences in the behaviour of the stochastic and deterministic schemes for this system. The stochastic scheme flips between two long-lived states of gene expression, whereas the deterministic scheme only finds one state, but which state it finds depends on the initial parameter values.