6

Stochastic genes and persistant decisions

6.1 Simulating stochastic production and decay

6.1.1 *Collect 10 000 numbers x, each selected by $x = -n_c \cdot \ln(ran_2)$ where ran_2 is another random number selected in $[0;1]$ and $n_c = 10$. Plot the histogram $h(x)$ in a graph where the x axis is linear and the y axis is logarithmic.*

6.1.2 *Repeat the Gillespie simulation of Eqs. (6.1) and (6.2), including that the mRNA lifetime is exponentially distributed with a mean lifetime that gives $n_c = 20$ proteins per mRNA. Compare with a simulation where we set $n_c = 10$, but maintain the mean production level of proteins.*

Answer Consider the production of mRNA as reaction number 1:

$$r_1 = \frac{\alpha}{n_c} \cdot \left(0.01 + \frac{(N/200)^2}{1 + (N/200)^2 + (N/200)^4} \right) \tag{6.1}$$

If this happens, the number of proteins is updated by $N \to N - n_c \ln(\text{ran})$ with $\text{ran} \in [0,1]$. In this way, a positive number of proteins is selected with the average expectation of n_c. The other event is a degradation event with rate:

$$r_2 = \frac{N}{\tau} \tag{6.2}$$

After this reaction is chosen the number of proteins is updated as $N \to N-1$.

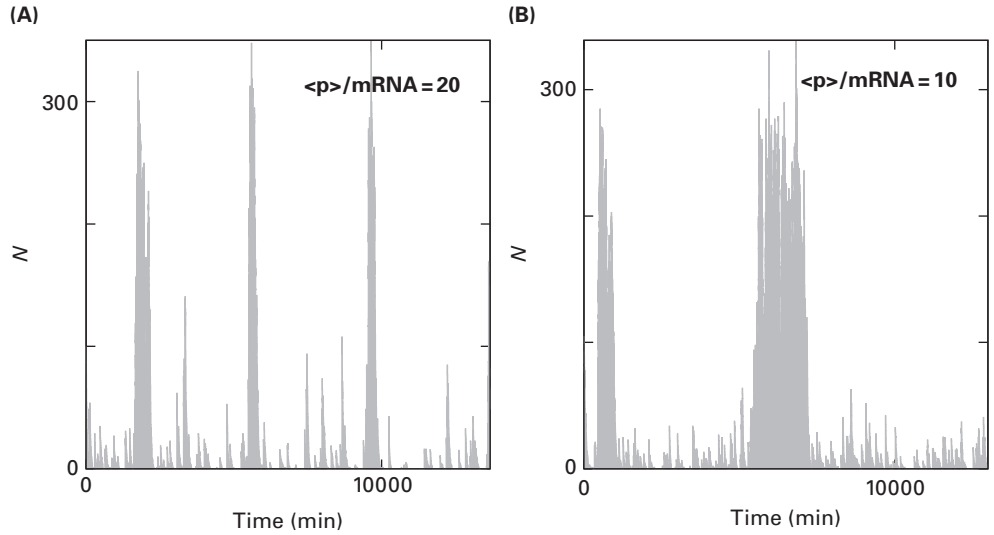

Figure 6.1 Simulation of bistable system with two different proteins per mRNA levels. (A) $n_c = 20$, (B) $n_c = 10$.

The Gillespie algorithm, in the limit where the mRNA lifetime is very short, reads: select the smallest of:

$$t_1 = -\ln(\text{ran}_1)/r_1 \tag{6.3}$$
$$t_2 = -\ln(\text{ran}_2)/r_2 \tag{6.4}$$

and update N accordingly, while increasing time $t \to t + \min(t_1, t_2)$. Setting $\tau = 30$ min, $\alpha = 20$ min^{-1} and $n_c = 20$, we obtain the behaviors shown in Fig. 6.1.

6.1.3 *Consider protein and associated mRNA production and decay, through rate Ω for $M \to M + 1$, rate ω for $M \to M + P$ for each M, rate γ_m for $M \to M - 1$ for each M and rate γ_P for $P \to 0$ for each protein. What is the average number of proteins per mRNA? Simulate the process using the Gillespie algorithm and the rates $\Omega = 0.02$, $\omega = 0.1$ (protein per second per mRNA), $\gamma = 0.01$ (mRNA lifetime of 100 s) and $\gamma_p = 0.002$ (protein lifetime of 500 s).*

Answer The average number of proteins is given by average lifetime $=1/\gamma_m =$ 100 s multiplied by protein production per time unit, $\omega \cdot \gamma_m = 20$ proteins per message.

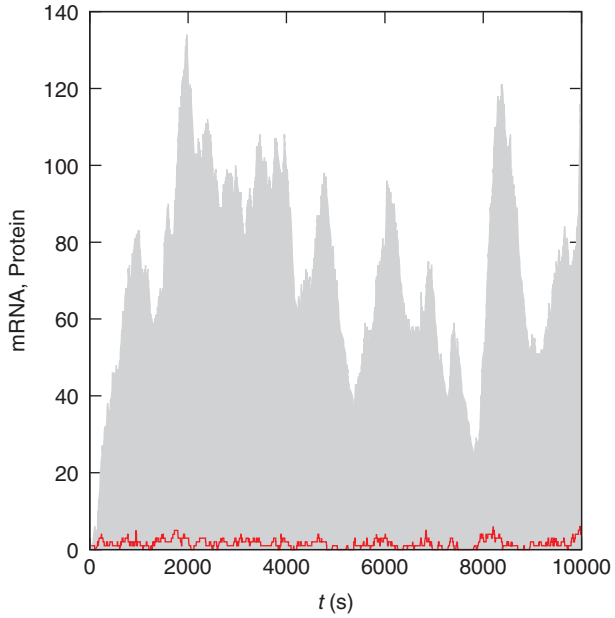

Figure 6.2 Simulation of production of mRNA and associated protein.

With m as the available number of mRNA and p the number of proteins, the simulation is set up as:

(1) $t_1 = -\ln(\text{ran}_1)/\Omega$ for $m \to m+1$
(2) $t_2 = -\ln(\text{ran}_2)/(m \cdot \omega)$ for $p \to p+1$
(3) $t_3 = -\ln(\text{ran}_3)/(m \cdot \gamma_\text{m})$ for $m \to m-1$
(4) $t_4 = -\ln(\text{ran}_4)/(p \cdot \gamma_\text{P})$ for $p \to p-1$

where ran_i is random numbers independently chosen uniformly in the interval $[0,1]$. Notice that in this coarse-grained simulation we put all mRNA together in one scalar variable m with a probability of reactions that is multiplied by m. As the mRNAs are independent, this simplification is allowed. The result of simulations is shown in Fig. 6.2.

6.1.4 *Repeat the investigation of the toggle switch [238] shown in Fig. 6.5:*

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \frac{a}{1+v^2} - u \quad \text{and} \quad \frac{\mathrm{d}v}{\mathrm{d}t} = \frac{b}{1+u^2} - v \tag{6.5}$$

for $a = b = 10$, by determining fixed points for both equations, as well as their crossing points. Hint: plot the right-hand side of the equations for u and v in $[0,10]$. Simulate the equations using the Gillespie algorithm by assuming that both u and v are produced and decay in discrete units of size 1.

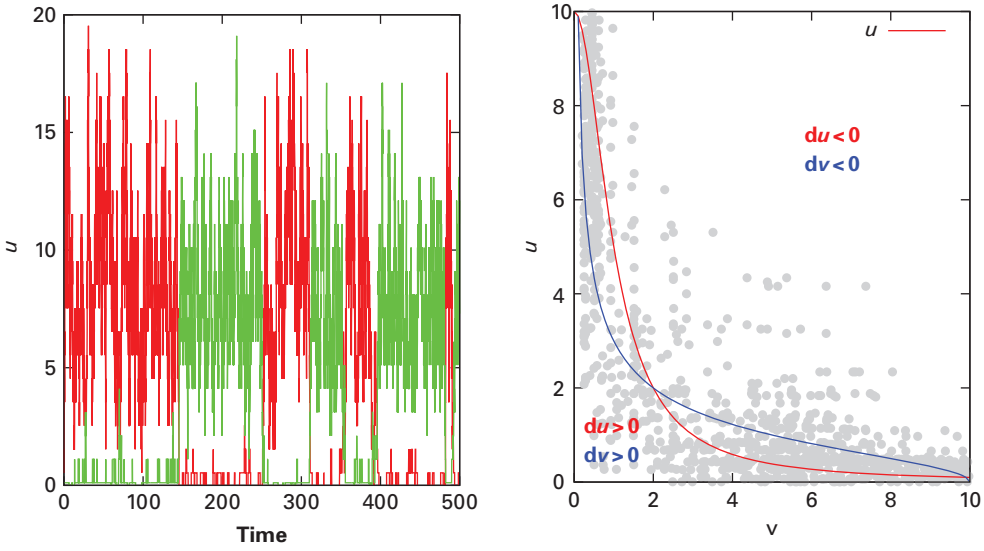

Figure 6.3 Simulation of production of repressilator. Right-hand panel shows null-clines where the grey dots reflect system positions during a small part of the simulation on the left.

Answer Define four different processes:
(1) $u \to u+1$ with rate $r_1 = \frac{10}{1+v^2}$
(2) $u \to u-1$ with rate $r_2 = u$
(3) $v \to v+1$ with rate $r_3 = \frac{10}{1+u^2}$
(4) $v \to v-1$ with rate $r_4 = u$.
The updates proceed as in standard Gillespie with selecting four times, $t_i = \ln(-\mathrm{ran}_i)/r_i$ and performing the associated update for the shortest selected time. For eventual better representation of production one may update $u \to u - \ln(\mathrm{ran})$, respectively $v \to v - \ln(\mathrm{ran})$ when increasing the u or v number. Results are shown in Fig. 6.3.

6.2 Stochastic simulation of the λ-switch

6.2.1 *Consider the following model for a lambdoid like phage:*

$$\frac{\mathrm{d}CI}{\mathrm{d}t} = \frac{0.1 + (CI/0.2)^2}{(1+(CI/0.2)^2)\cdot(1+(Cro/0.5)^2)} - CI$$
$$\frac{\mathrm{d}Cro}{\mathrm{d}t} = \frac{1}{(1+(CI/0.2)^2)} - Cro$$

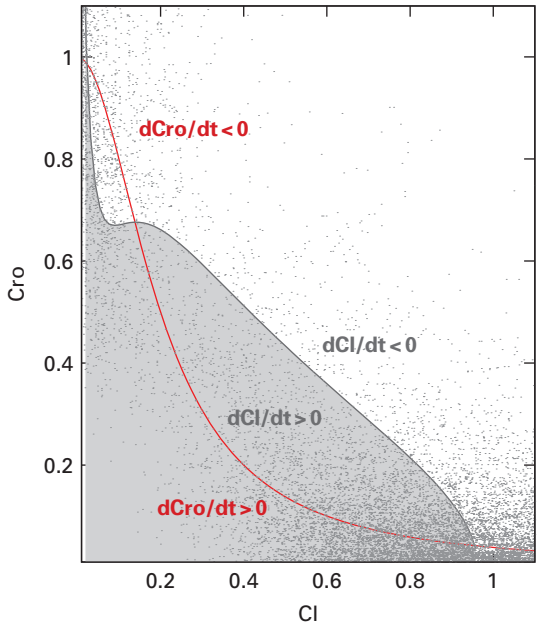

Figure 6.4 Null clines for changes in CI and Cro in a simplified model for phage lambda switch. Above the lines both concenrations will decline, below they will increase. Thus determinstc drift will tend to focuss dynamics along intersection of the lines.

resembling PRM activated by CI and repressed by Cro, and PR repressed by CI (see Fig. 6.11). What does the product in the denominator of the "PRM term" correspond to in terms of operator design? Plot regions of $(CI, Cro) \in [0,1] \times [0,1]$ *where* $dCI/dt > 0$*, and* $dCro/dt > 0$*. Identify stable and unstable fixed points. Hint: identify null-clines by solving steady-state equations in terms of Cro as a function of CI.*

Answer The product in the denominator corresponds to a design where one operator can be occupied by CI, whereas the other can be independently occupied by Cro. Both operators influence the promoter for CI, whereas only one represses the promoter for Cro.

In steady state the equations can be rephrased:

$$\frac{\mathrm{d}CI}{\mathrm{d}t} = \frac{0.1 + (CI/0.2)^2}{(1 + (CI/0.2)^2) \cdot (1 + (Cro/0.5)^2)} - CI = 0 \Rightarrow$$

$$Cro = 0.5 \cdot \frac{0.1 + (CI/0.2)^2}{\sqrt{(1 + (CI/0.2)^2) \cdot CI}}$$

$$\frac{\mathrm{d}Cro}{\mathrm{d}t} = \frac{1}{(1 + (CI/0.2)^2)} - Cro = 0 \Rightarrow Cro = \frac{1}{1 + (CI/0.2)^2}$$

which is plotted in Fig. 6.4 with an indication of regions where CI increases, and Cro increases. One see that there are three intersections with lines where $\mathrm{d}CI = 0$ and $\mathrm{d}Cro/\mathrm{d}t = 0$, of which the middle is unstable.

6.2.2 *Simulate the model in Question 6.2.1 using Gillespie algorithm with discretized production into units of CI and Cro of characteristic size* $\delta = 0.1$*, but with dilution being considered to occur in very small units, say 0.01. Discuss the timescale in the simulation, and the step size* δ *in terms of promoter activity. Hint: define four processes, two production events and two decay events, each happening with rates given by the respective terms in Eq. 6.20 divided by the size of the change (*$\delta = 0.1$ *for production and 0.01 for decay). If any variable becomes less than* 0*, then set it equal to zero, since concentrations cannot be negative.*

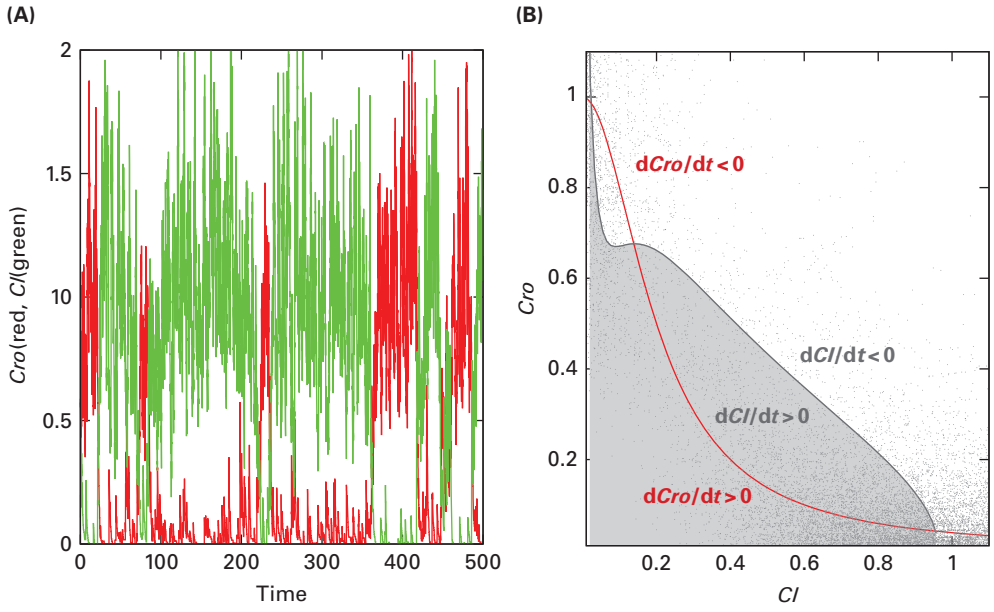

Figure 6.5 Gillespie simulation for CI and Cro in a simplified model for the phage λ switch. The dots in panel (B) show positions of the systems during the simulations shown in panel (A).

Answer The timescale 1 is the decay/dilution time of both the proteins in the equations. In the simulation we use $\delta = 0.1$ and define the processes:

(1) $CI \to CI + \delta$ with rate $r_1 = \frac{1}{\delta} \cdot \frac{0.1+(CI/0.2)^2}{(1+(CI/0.2)^2)\cdot(1+(Cro/0.5)^2)}$
(2) $CI \to CI - 0.01$ with rate $r_2 = CI/0.01$
(3) $Cro \to Cro + \delta$ with rate $r_3 = \frac{1}{\delta} \cdot \frac{1}{(1+(CI/0.2)^2)}$
(4) $Cro \to Cro - 0.01$ with rate $r_4 = Cro/0.01$

and assign respective times randomly, as in the standard Gillespie procedure ($t_i = -\ln(\text{ran})/r_i$, ran $\in [0, 1]$). In Fig. 6.5 we show a simulation where we have replaced the fixed δ protein production with exponentially distributed bursts, thus adding $-\delta \cdot \ln(\text{ran})$ at each production event.

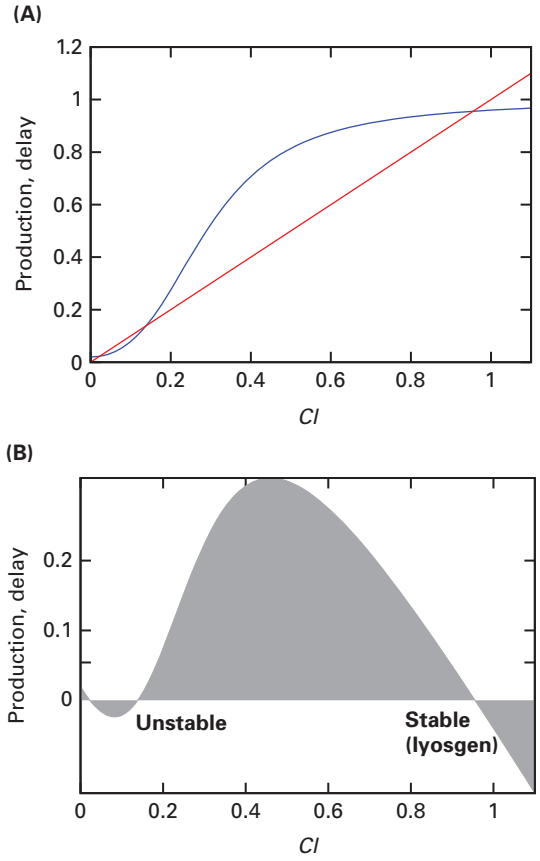

Figure 6.6 Production and decay of a simplified λ switch, assuming in addition that Cro is a quickly adjusting variable.

6.2.3 *Rewrite the equations in Question 6.2.1 in nM units, where we assume that CI binding has binding constant $K = 40\,nM$, and Cro has $100\,nM$, instead of 0.2 and 0.5, respectively. How many proteins per message does $\delta = 0.1$ then correspond to?*

Answer Basal production of CI of 1 should be repressed by binding $K = 0.2$, implying that a real $K_{\rm CI} = 40$ nM should be supported by a basal production of 200 nM per time unit. Similarly, fully expressed Cro production should be repressed by binding $K = 0.5$, implying that a real $K_{\rm Cro} = 100\,{\rm nM}$ should be supported by Cro production of 200 nM per time unit.

$$\begin{aligned}\frac{{\rm d}[CI]}{{\rm d}t} &= 200\,{\rm nM}\cdot\frac{0.1+([{\rm CI}]/40\,{\rm nM})^2}{(1+([{\rm CI}]/40\,{\rm nM})^2)\cdot(1+([{\rm Cro}]/100\,{\rm nM})^2)} - [CI]\\ \frac{{\rm d}[Cro]}{{\rm d}t} &= 200\,{\rm nM}\cdot\frac{1}{(1+([{\rm CI}]/40\,{\rm nM})^2)} - [{\rm Cro}]\end{aligned} \tag{6.6}$$

in units of the CI and Cro dilution time $\sim$ cell generation of *E. coli*. A $\delta = 0.1$ accordingly correspond to 20 nM of protein per mRNA (or 20 proteins per message).

6.2.4 *Rewrite the equations in Question 6.2.1 as a one-variable simulation, by assuming that Cro is a fast variable that instantly adapts to the steady-state value set the equation for Cro:*

$$\frac{dCI}{dt} = \frac{(0.1+(CI/0.2)^2)\cdot(1+(CI/0.2)^2)}{(1+(CI/0.2)^2)^2+4} - CI = Prod - CI$$

Plot the production and decay terms separately, as well as their sum.

Answer See Fig. 6.6.