

1. Stochastic protein dynamics. Consider a stochastic system in which particular protein in a cell has copy number n . Starting in state n , the transition rate for $n \rightarrow n + 1$ is β and the transition rate for $n \rightarrow n - 1$ is $-\alpha n$. The steady-state probability that the system is in state n is P_n .

- (a) Use detailed balance to obtain the steady-state relationship between P_n and P_{n+1} , and then derive a closed form expression for the probability distribution P_n .

$$P_n \beta = P_{n+1} (n+1) \alpha$$

$$P_{n+1} = \frac{\beta}{(n+1) \alpha} P_n$$

$$1 = \sum_n P_n$$

$$= P_0 + P_1 + P_2 + P_3 + \dots$$

$$= P_0 + \frac{\beta}{\alpha} P_0 + \left(\frac{\beta}{\alpha}\right)^2 \frac{1}{2} P_0 + \left(\frac{\beta}{\alpha}\right)^3 \frac{1}{3!} P_0 + \dots$$

$$= P_0 \sum_{n=0}^{\infty} \left(\frac{\beta}{\alpha}\right)^n \frac{1}{n!}$$

$$1 = P_0 e^{\beta/\alpha}$$

Thus,

$$P_0 = e^{-\beta/\alpha}$$

$$P_n = \left(\frac{\beta}{\alpha}\right)^n \frac{1}{n!} e^{-\beta/\alpha}$$

- (b) Provide a closed form expression for the generating function $\tilde{P}(\phi) = \sum_{n=0}^{\infty} e^{-\phi n} P_n$; express the mean $\langle n \rangle$ and the variance $\langle n^2 \rangle - \langle n \rangle^2$ in terms of derivatives of $\tilde{P}(\phi)$; and

evaluate the mean and variance in terms of model parameters.

$$\begin{aligned}
 \tilde{P}(\phi) &= \sum_{n=0}^{\infty} e^{-\phi n} P_n \\
 &= \sum_{n=0}^{\infty} e^{-\phi n} \left(\frac{\beta}{\alpha}\right)^n \frac{1}{n!} \\
 &= e^{-\beta/\alpha} \sum_{n=0}^{\infty} \left(e^{-\phi} \frac{\beta}{\alpha}\right)^n \frac{1}{n!} \\
 &= e^{-\beta/\alpha} e^{e^{-\phi} \frac{\beta}{\alpha}} \\
 &= \exp\left(\frac{\beta}{\alpha} (e^{-\phi} - 1)\right)
 \end{aligned}$$

$$\langle n \rangle = - \frac{d \ln(\tilde{P}(\phi))}{d\phi} \Big|_{\phi=0}$$

$$\langle n^2 \rangle - \langle n \rangle^2 = \frac{d^2 \ln(\tilde{P}(\phi))}{d\phi^2} \Big|_{\phi=0}$$

$$\langle n \rangle = - \frac{d \ln(\tilde{P}(\phi))}{d\phi} \Big|_{\phi=0}$$

$$= \frac{\beta}{\alpha} e^{-\phi} \Big|_{\phi=0}$$

$$= \frac{\beta}{\alpha}$$

$$\langle n^2 \rangle - \langle n \rangle^2 = \frac{d^2 \ln(\tilde{P}(\phi))}{d\phi^2} \Big|_{\phi=0}$$

$$= \frac{\beta}{\alpha}$$

- (c) Suppose you want to run a stochastic simulation. Starting in state n , what states can be reached? What is the mean time τ for this first transition? What is the probability distribution $p(\tau)$ for the first transition?

The total rate of transition out of state n is $\beta + n\alpha$, and the only possible future states, given the current state is n are $n - 1$, and $n + 1$.

$$Pr(n \rightarrow m) = \begin{cases} \frac{\alpha n}{\alpha n + \beta}, & m = n - 1 \\ \frac{\beta}{\alpha n + \beta}, & m = n + 1 \\ 0, & \text{otherwise} \end{cases}$$

The transition time τ has an exponential distribution with parameter $\beta + n\alpha$.

$$p(\tau) = (\alpha n + \beta) e^{-(\alpha n + \beta)\tau}$$

$$\langle \tau \rangle = \frac{1}{\alpha n + \beta}$$

- (d) Provide an ODE for a continuous variable $X(t)$ where $X(t)$ and $n(t)$ have the same dynamics in the limit of large n , and provide the mean and variance of $X(t)$ from the ODE starting with $X(0)$ equal to the mean.

$$\frac{dX(t)}{dt} = \beta - \alpha X(t)$$

According to the fluctuation dissipation theorem, the behavior of the ODE solution starting from a non-equilibrium initial condition approximates the stochastic behavior of the system.

Here, let's assume we start from $X(0) = \frac{\beta}{\alpha} + \delta$.

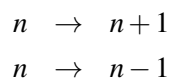
$$X(t) = \frac{\beta}{\alpha} + \delta e^{-\alpha t}$$

So, at steady state:

$$\langle X \rangle = \frac{\beta}{\alpha}$$

$$\begin{aligned} \langle X(t) - \langle X \rangle \rangle \langle X(t) - \langle X \rangle \rangle &= (\delta e^{-\alpha t}) (\delta e^{-\alpha t}) \\ &= \delta^2 e^{-2\alpha t} \\ &\rightarrow 0 \end{aligned}$$

2. Fluctuation-dissipation. Use your Gillespie simulation code for dynamics of a single protein species with copy number n and transitions



and corresponding rates β and αn . Use parameter values $\beta = 1/\text{min}$ and $\alpha = 0.1/\text{min}$. The auto-correlation function $S(\tau)$ is defined as $\langle n(\tau)n(0) \rangle - \langle n(t) \rangle \langle n(0) \rangle$.

- (a) Note that $S(0)$ is just the variance of n . Provide $S(0)$. According to the fluctuation-dissipation theorem, the shape of $S(t)$ should be identical to the shape of the relaxation of an equivalent ODE system for continuous variable $X(t)$, $S(t)/S(0) = [X(t) - \mu]/[X(0) - \mu]$, with μ equal to the steady-state value and $X(0)$ an arbitrary initial state (other than μ). Use the theorem to provide an analytic expression for $S(t)$.

$$\begin{aligned} S(0) &= \langle n^2 \rangle - \langle n \rangle^2 \\ &= \frac{\beta}{\alpha} \\ &= 10 \end{aligned}$$

In question 1(d), we have solved the $X(t)$ from the ODE starting with $X(0)$,

$$X(t) = X(0)e^{-\alpha t} + \frac{\beta}{\alpha}(1 - e^{-\alpha t})$$

So with $\mu = \frac{\beta}{\alpha}$,

$$\begin{aligned} \frac{S(t)}{S(0)} &= \frac{X(t) - \mu}{X(0) - \mu} \\ &= \frac{X(0)e^{-\alpha t} + \frac{\beta}{\alpha}(1 - e^{-\alpha t}) - \frac{\beta}{\alpha}}{X(0) - \frac{\beta}{\alpha}} \\ &= e^{-\alpha t} \\ S(t) &= S(0)e^{-\alpha t} \\ &= \frac{\beta}{\alpha}e^{-\alpha t} \end{aligned}$$

- (b) Now estimate the auto-correlation function from a trajectory. Start a simulation at time 0 with n equal to the equilibrium mean, and run a trajectory for $T = 10,000$ min saving the state every 0.1 min. For $0 \leq \tau \leq 100$ min, calculate $\langle n \rangle$ and $S(\tau)$ from the trajectory as the discrete-time equivalents of the integrals

$$\begin{aligned} \langle n \rangle &= T^{-1} \int_0^T dt n(t) \\ S(\tau) &= (T - \tau)^{-1} \int_0^{T-\tau} dt n(t)n(t + \tau) - \langle n \rangle^2. \end{aligned}$$

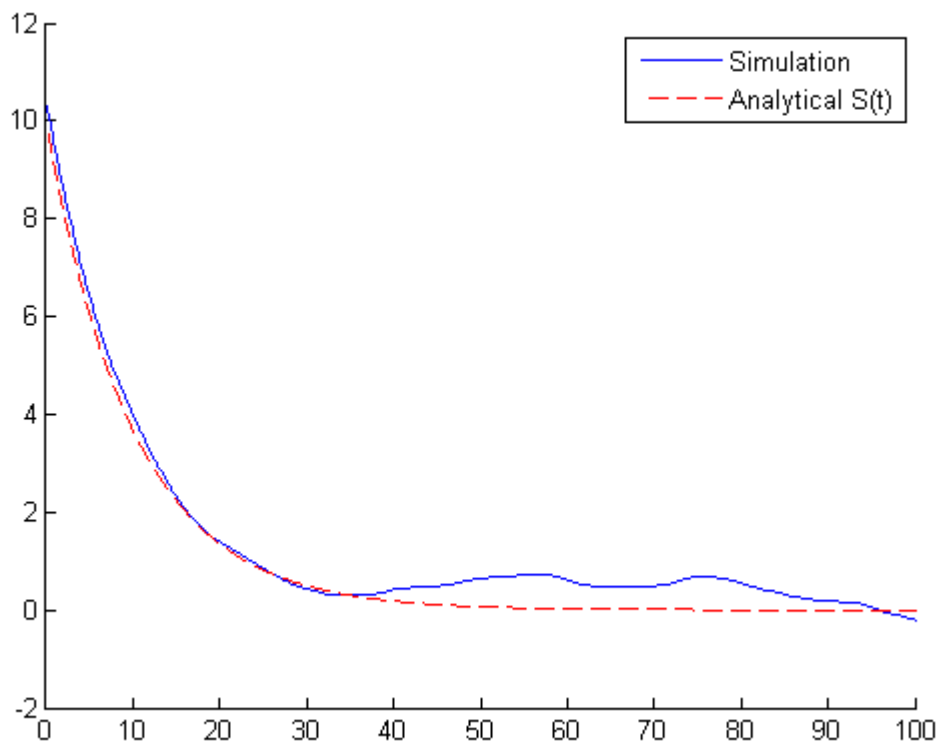
In other words, define $\Delta t = 0.1$ min, $t_i = i\Delta t$, and $M = T/\Delta t$. Since we are including both endpoints, $i = 0$ and $i = M$, the total number of time points is $M + 1$. The mean

and auto-correlation are

$$\langle n \rangle = (M+1)^{-1} \sum_{i=0}^M n(t_i)$$

$$S(j\Delta t) = (M-j+1)^{-1} \sum_{i=0}^{M-j} n(t_i)n(t_{i+j}) - \langle n \rangle^2.$$

In the discrete form, $\tau = j\Delta t$, and remember to only calculate results for $\tau \leq 100$ min, or $0 \leq j \leq 100/\Delta t$. Provide a plot comparing $S(t)$ from the simulation with the analytical result from the fluctuation-dissipation theorem.



The code is attached to the end of the key.

- (c) In the previous question you calculated an analytical result for P_n , the probability that the system has n copies of the protein. From the same trajectory you used for $S(t)$, estimate P_n as the fraction of time steps with particle count n , defined formally as

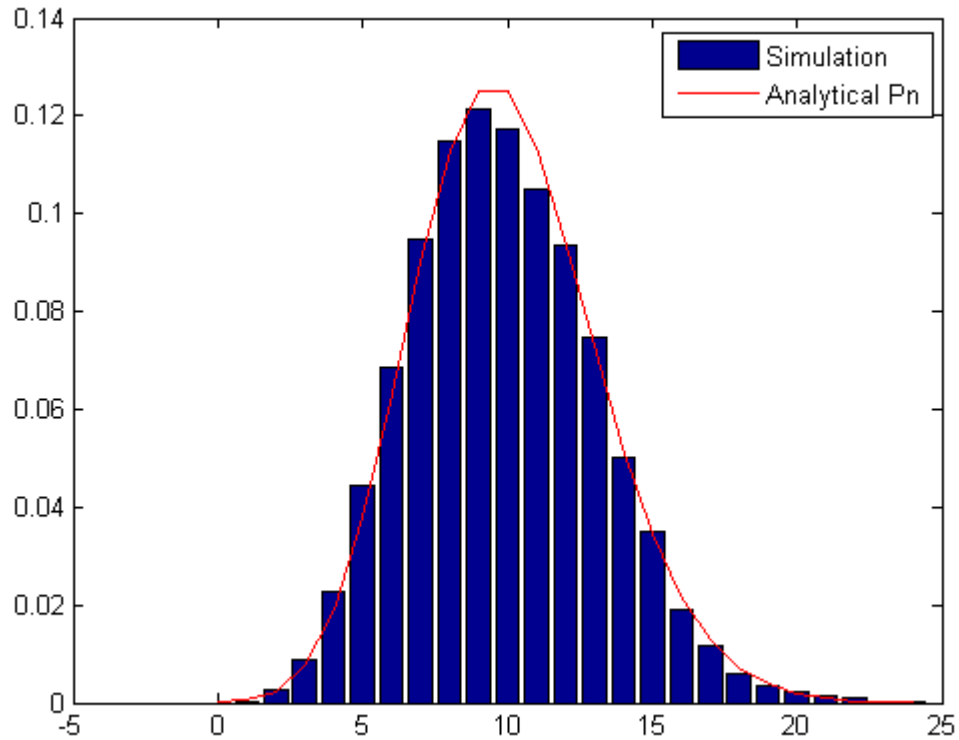
$$P_n = (M+1)^{-1} \sum_{i=0}^M \Theta[n(i\Delta t) = n],$$

where as usual Θ is 1 if its argument is True and 0 if False. Provide a table and a plot with the analytical and simulation results for P_n extending to the largest value for n observed in your simulation.

According to the simulation result in question 2(b), the largest value for n is 24. Then we can calculate the analytical results for P_n , comparing with simulation results, as shown in the table and plotting below.

Table 1: Analytical P_n vs Simulation P_n

n	Analytical P_n	Simulation P_n
0	0.00000	0.00005
1	0.00004	0.00045
2	0.00267	0.00227
3	0.00889	0.00757
4	0.02262	0.01892
5	0.04424	0.03783
6	0.06874	0.06306
7	0.09495	0.09008
8	0.11491	0.11260
9	0.12126	0.12511
10	0.11712	0.12511
11	0.10500	0.11374
12	0.09374	0.09478
13	0.07489	0.07291
14	0.05021	0.05208
15	0.03514	0.03472
16	0.01906	0.02170
17	0.01186	0.01276
18	0.00612	0.00709
19	0.00352	0.00373
20	0.00218	0.00187
21	0.00162	0.00089
22	0.00086	0.00040
23	0.00031	0.00018
24	0.00006	0.00007

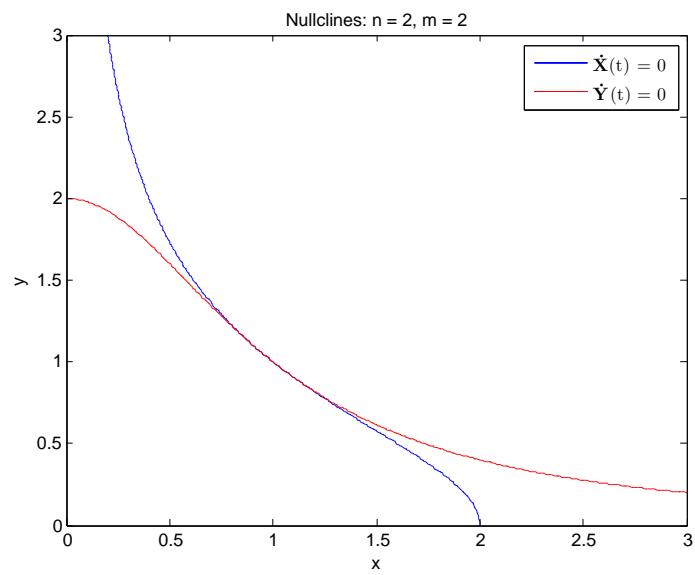
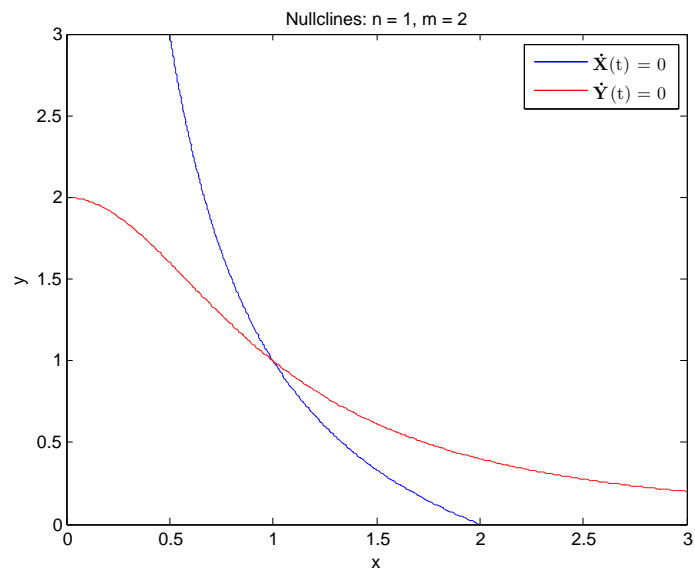


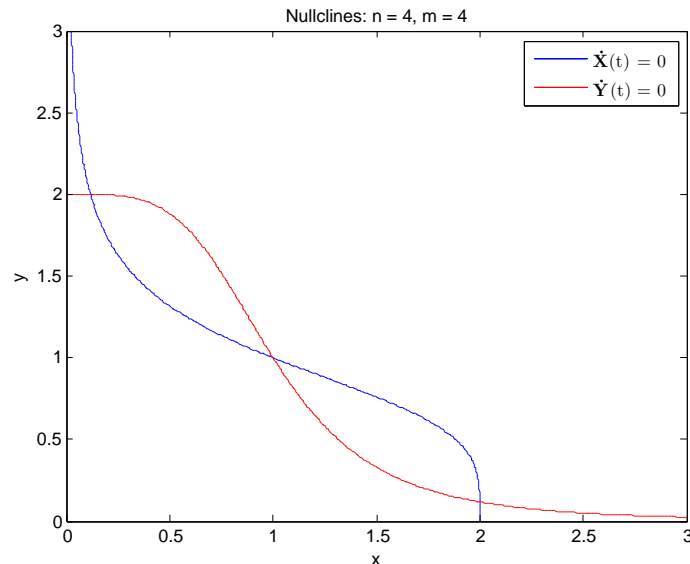
3. Stability analysis. Consider the Delta-Notch system with different exponents for proteins X and Y :

$$\begin{aligned}\dot{X}(t) &= \frac{2}{1+Y(t)^n} - X(t) \\ \dot{Y}(t) &= \frac{2}{1+X(t)^m} - Y(t).\end{aligned}$$

For convenience, we are working with units scaled so that the decay rate is 1/time, the threshold is 1 molecule. To simplify the problem, we have set the production rate to 2 molecules/time.

- (a) Provide plots of the nullclines for $(n = 1, m = 2)$, $(n = 2, m = 2)$, and $(n = 4, m = 2)$.





- (b) When n and m are sufficiently large, there are three fixed points: two asymmetric ($X \neq Y$) and one symmetric ($X = Y$). Provide approximate (X, Y) values for the asymmetric fixed points and provide the exact location of the symmetric fixed point.

It can be seen by plotting (and analytically) that if $(n, m) \rightarrow \infty$ then the two asymmetric solutions are given by: $(X, Y) = (0, 2)$ and $(2, 0)$.

- (c) For general n and m , provide the Jacobian at the symmetric fixed point and evaluate its eigenvalues in terms of model parameters. Provide the condition on the eigenvalues that makes the symmetric fixed point unstable (which permits patterning) and describe what conditions on the Hill coefficients permit patterning.

The Jacobian is the matrix of first partial derivatives with respect to X and Y . If we evaluate this matrix at the symmetric fixed point $(X, Y) = (1, 1)$ we get:

$$J = \begin{bmatrix} -1 & \frac{-n}{2} \\ \frac{-m}{2} & -1 \end{bmatrix}$$

To compute the eigenvalues we solve for the roots of the characteristic equation:

$$\begin{aligned} |J - \lambda I| &= 0 \\ (-1 - \lambda)^2 - \frac{mn}{4} &= 0 \\ \lambda &= -1 \pm \frac{1}{2}\sqrt{mn} \end{aligned}$$

For the symmetric fixed point to be unstable one of the eigenvalues must be greater

than 0. Therefore it is required that:

$$\begin{aligned}\lambda &> 0 \\ \sqrt{mn} &> 2\end{aligned}$$

Matlab Code

Problem 2

```

function Prob2 ()
global a b
a = .1; % alpha
b = 1; % beta

n_0 = b/a; % Initial count
nt = 1; % Number of trajectories
t_max = 10000; %Minutes for simulation (Approx).
G = cell(nt,2);
Var_int=zeros(1,t_max+1);
dt_int=0.1; % size of recording time interval
M = t_max / dt_int;

for i = 1:nt
    [T, N] = get_trajectory(n_0, t_max);
    G{i,1} = T;
    G{i,2} = N;
    [T_int, N_int] = const_intervals(T, N, t_max, dt_int);
end
Mean_n=sum(N_int)/(M+1);

S = zeros(1, (100/dt_int+1));
for j=0:1:100/dt_int
    S(j+1)=sum(N_int(1:M-j+1).*N_int(j+1:M+1)-Mean_n^2)/(M-j+1);
end
t=0:dt_int:100;
S_analytical=b/a.*exp(-a*t);

n_max = max(N_int);
P = zeros(1,n_max+1);
for i=0:n_max
    P(i+1) = sum (N_int==i)/(M+1);
end
n=0:n_max;
P_analytical = ((b/a).^n)./factorial(n).*exp(-b/a);

```

```

figure(1)
hold on
plot(t,S,'b',t,S_analytical,'--r')
legend('Simulation','Analytical S(t)')

```

```

figure(2)
bar(n,P)
hold on
plot(n,P_analytical,'r')
legend('Simulation','Analytical Pn')
end

```

```

function [T, N] = get_trajectory(n_0, t_max)
%T - transition times
%N - counts at each transition time
global a b
t = 0;
j = 1;
T(1) = 0;
N(1) = n_0;
while t < t_max
    j = j+1;
    dt = (-1/(a*N(j-1) + b)) * log(rand(1));
    t = t + dt;
    T(j) = t;
    if rand(1) < b/(b+a*N(j-1))
        N(j) = N(j-1) + 1;
    else
        N(j) = N(j-1) - 1;
    end
end
end
end

```

```

function [T_int, N_int] = const_intervals(T, N, t_max, dt_int)

% T_int - Times at even intervals.
% N_int - Count at times corresponding to T_int.
% T - Vector of transition times.
% N - Count of mRNA corresponding to times given in T.
% t_max - maximum of amount time
% dt_int - size of time interval

```

```
T_int = 0:dt_int:t_max;
N_int = size(T_int);
N_int(1) = N(1);
j = 1;
for i = 2:length(T_int)
    while T(j) < T_int(i)
        j = j+1;
    end
    N_int(i) = N(j-1);
end
end
```

Problem 3

```

%% First
n = 1;
m = 2;

fun1 = @(x,y) 2/(1+y^n) - x;
h1 = ezplot(fun1,[0,3,0,3]);
set(h1,'LineColor','b');
hold on

fun1 = @(x,y) 2/(1+x^m) - y;
h2 = ezplot(fun1,[0,3,0,3]);
set(h2,'LineColor','r');
title('Nullclines: n = 1, m = 2');
h1 = legend(' $\bf\dot{X}$ (t) = 0 ', ' $\bf\dot{Y}$ (t) = 0 ');
set(h1,'Interpreter','latex');
hold off;

%% Second
n = 2;
m = 2;

fun1 = @(x,y) 2/(1+y^n) - x;
h1 = ezplot(fun1,[0,3,0,3]);
set(h1,'LineColor','b');
hold on

fun1 = @(x,y) 2/(1+x^m) - y;
h2 = ezplot(fun1,[0,3,0,3]);
set(h2,'LineColor','r');
title('Nullclines: n = 2, m = 2');
h1 = legend(' $\bf\dot{X}$ (t) = 0 ', ' $\bf\dot{Y}$ (t) = 0 ');
set(h1,'Interpreter','latex');
hold off;

%% Third
n = 4;
m = 4;

fun1 = @(x,y) 2/(1+y^n) - x;

```

```
h1 = ezplot(fun1,[0,3,0,3]);
set(h1,'LineColor','b');
hold on

fun1 = @(x,y) 2/(1+x^m) - y;
h2 = ezplot(fun1,[0,3,0,3]);
set(h2,'LineColor','r');
title('Nullclines: n = 4, m = 4');
h1 = legend('$\bf\dot{X}(t) = 0','$\bf\dot{Y}(t) = 0');
set(h1,'Interpreter','latex');
hold off;
```