SBf12, Assignment 5 Solutions

Kevin Oishi

October 29, 2012

1 Problem 1

1.1 Part (a)

Keeping in mind the addendum about cell division in gro, and adding degradation reactions for gfp and rfp, here are is the chemical reaction network for Part 1a,

$$\emptyset \xrightarrow{1} X \tag{1}$$

$$X \xrightarrow{k} rfp + gfp \tag{2}$$

$$rfp \xrightarrow{0.05} \emptyset \tag{3}$$

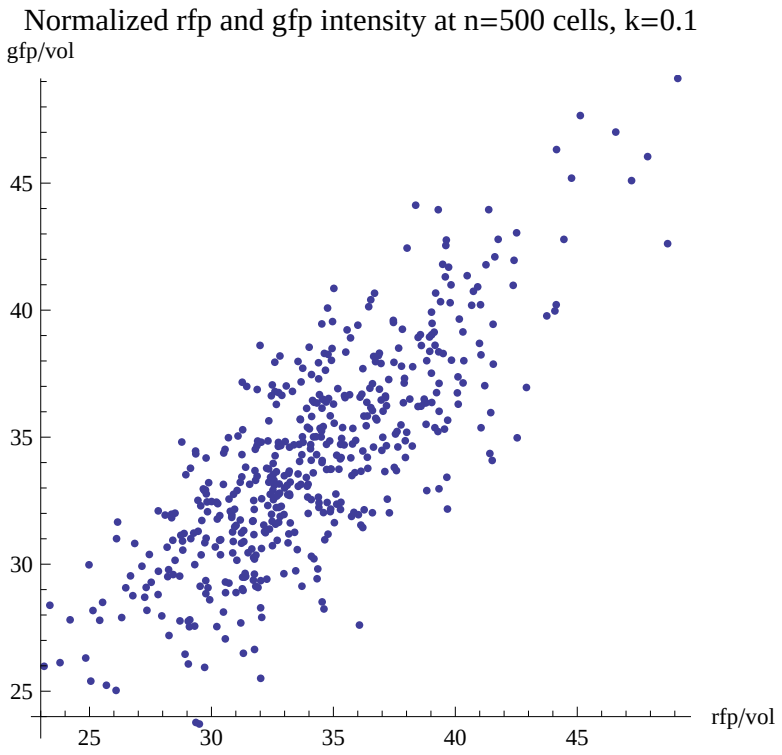
$$gfp \xrightarrow{0.05} \emptyset. \tag{4}$$

The following gro code implements the chemical reaction network:

```
include gro

outfile := fopen("/tmp/a5_1a.csv","w");
population_size := 1;

program p(k) := {
  gfp := 0;
  rfp := 0;
  X := 0;

  rate(volume) : { X := X + 1 }
  rate(k*X) :      { rfp := rfp + 1, gfp := gfp + 1 }
  rate(0.05*rfp) :   { rfp := rfp - 1 }
  rate(0.05*gfp) :   { gfp := gfp - 1 }

  just_divided & daughter : { population_size := population_size + 1 }
};
```

```
program main() := {
  population_size >= 500 : {
    maptocells fprint(outfile,X,",",rfp,",",gfp,",",volume,"\n") end,
    stop()
  }
};

ecoli([], program p(0.1));
```

The scatter plot of the state of cells at $population\_size = 500$ is shown below,



In this example, solving for intrinsic noise,

$$\eta_{int}^2 = \frac{\langle (gfp - rfp)^2 \rangle}{2 \langle gfp \rangle \langle rfp \rangle} \quad (5)$$

$$\eta_{int} = 0.0560378 \quad (6)$$

Solving for extrinsic noise,

$$\eta_{ext}^2 = \frac{\langle gfp \times rfp \rangle - \langle gfp \rangle \langle rfp \rangle}{\langle gfp \rangle \langle rfp \rangle} \quad (7)$$

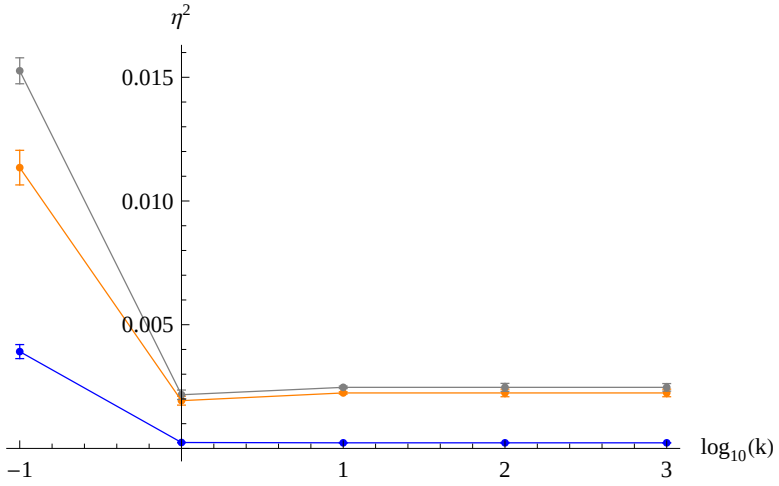
$$\eta_{ext} = 0.108718 \quad (8)$$

Then for total noise,

$$\eta_{tot}^2 = \eta_{int}^2 + \eta_{ext}^2 \tag{9}$$

$$\eta_{tot} = 0.122311 \tag{10}$$

And here is a plot of intrinsic, extrinsic, and total noise as a function of $\log_{10} k$. The blue line is η_{int}^2, the orange line is η_{ext}^2, and the gray line is η_{tot}^2. Error bars were found using the Jackknife method with subsample size $M = 100$.



Note that as k increases both intrinsic and extrinsic noise decrease, and appear to reach some nonzero asymptote. One way to think about this is that as k increases to infinity the conversion of X into rfp and gfp tends towards deterministic and instantaneous. The resulting noise comes from the production of X and degradation of rfp and gfp.

1.2 Part (b)

The CRN for Part (b) is as follows,

$$\emptyset \underset{0.1}{\overset{1}{\rightleftharpoons}} RNA_1 \tag{11}$$

$$\emptyset \underset{0.1}{\overset{1}{\rightleftharpoons}} RNA_2 \tag{12}$$

$$RNA_1 + RNA_2 \xrightarrow{k} \emptyset \tag{13}$$

$$RNA_1 \xrightarrow{1} RNA_1 + gfp \tag{14}$$

$$RNA_2 \xrightarrow{1} RNA_2 + rfp \tag{15}$$

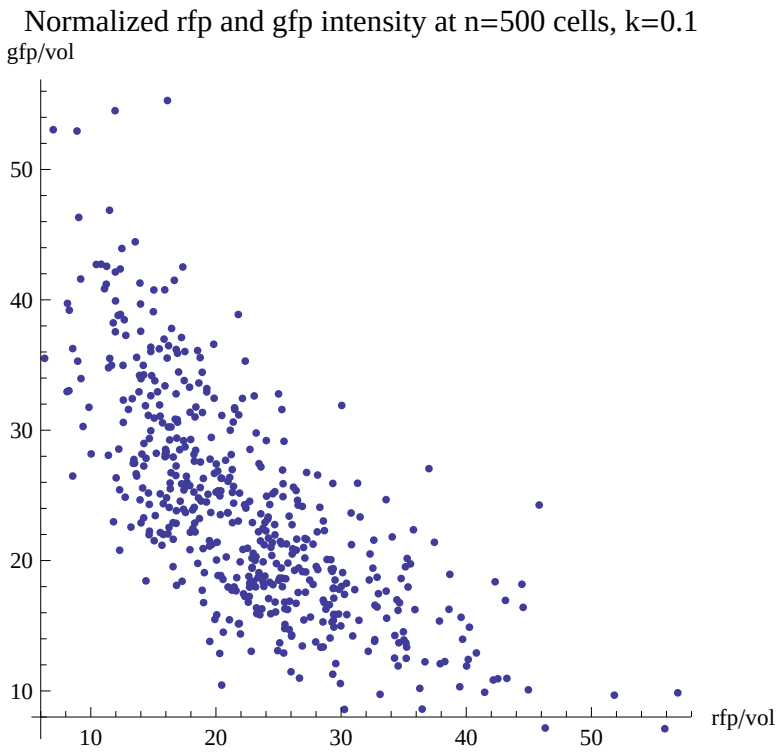
$$gfp \xrightarrow{0.05} \emptyset \tag{16}$$

$$rfp \xrightarrow{0.05} \emptyset. \tag{17}$$

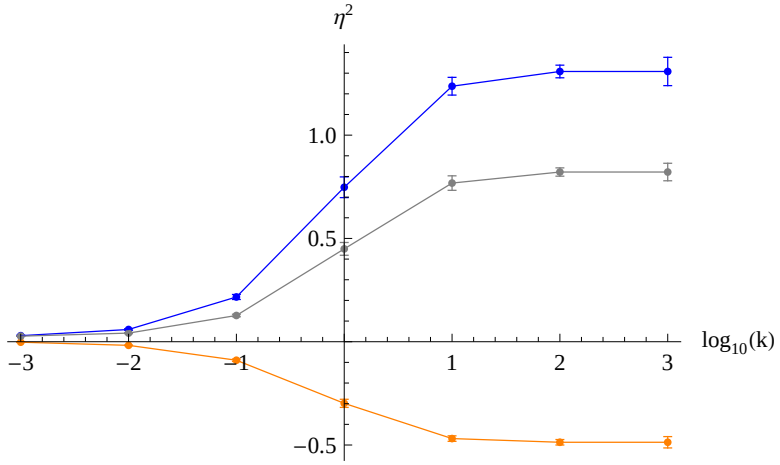
gro code that implements this CRN is show below,

```
include gro

outfile := fopen("/tmp/a5_1b_3.csv","w");
population_size := 1;

program p(k) := {
  gfp := 0;
  rfp := 0;
  rna1 := 0;
  rna2 := 0;

  rate(volume) :      { rna1 := rna1 + 1 }
  rate(volume) :      { rna2 := rna2 + 1 }
  rate(0.1*rna1) :    { rna1 := rna1 - 1 }
  rate(0.1*rna2) :    { rna2 := rna2 - 1 }
  rate(k*rna1*rna2) : { rna1 := rna1 - 1, rna2 := rna2 - 1 }
  rate(rna1) :        { gfp := gfp + 1}
  rate(rna2) :        { rfp := rfp + 1}
  rate(0.05*rfp) :    { rfp := rfp - 1 }
  rate(0.05*gfp) :    { gfp := gfp - 1 }

  just_divided & daughter : { population_size := population_size + 1 }
};

program main() := {
  population_size >= 500 : {
    maptocells fprint(outfile,rna1,",",rna2,",",rfp,",",gfp,",",volume,"\n") end,
    stop()
  }
};

ecoli([], program p(kr));
```

Here is a sample output where $k = 10^{-}1$,



And here is a plot of intrinsic, extrinsic, and total noise as a function of $\log_{10} k$. The blue line is η^2_{int}, the orange line is η^2_{ext}, and the gray line is η^2_{tot}. Error bars were found using the Jackknife method with subsample size $M = 100$.



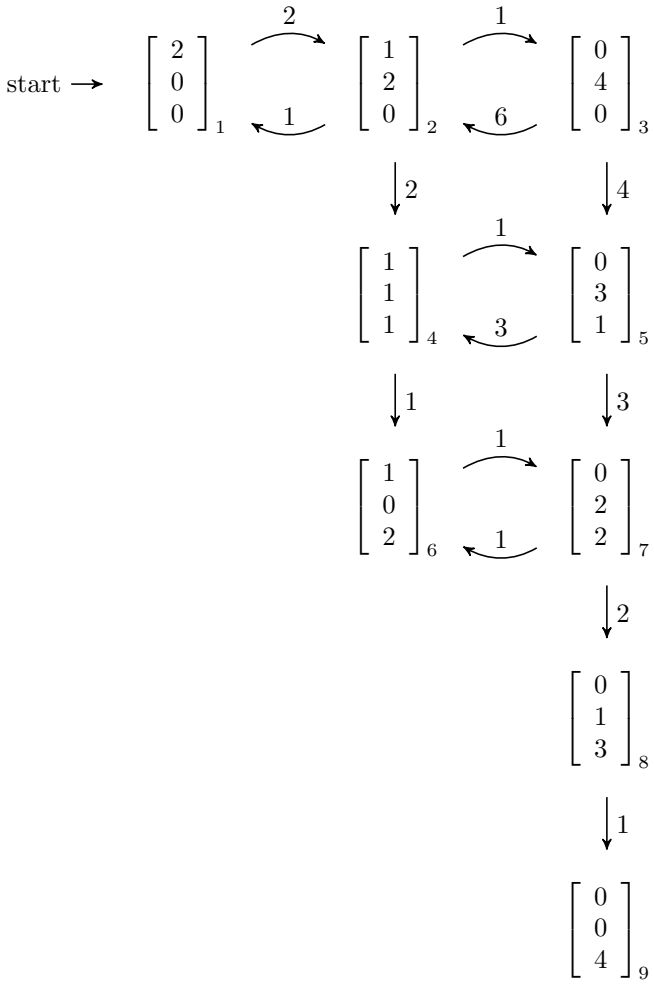
Note that here as k increases the distribution of $(RNA1, RNA2)$ becomes more and more bimodal, tending towards a distribution will some RNA_1 and very little RNA_2 or some RNA_2 and very little RNA_1. Intrinsic noise initially

increases, and approaches an asymptote, since as k increases the combined expression fo gfp and rfp are driven primarily by either the production and degradation of RNA_1 (reaction 11) or RNA_2 (reaction 12).

2 Problem 2

2.1 Part (a)

The state transition diagram for this system starting with $A = 2$ is shown below. In this diagram vector subscripts represent state indices used in constructing the rate matrix.



2.2 Part (b)

Using the state indices indicated in the transition diagram above, the rate matrix is as follows,

$$Q = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & -4 & 6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & -2 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 1 & -6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{18}$$

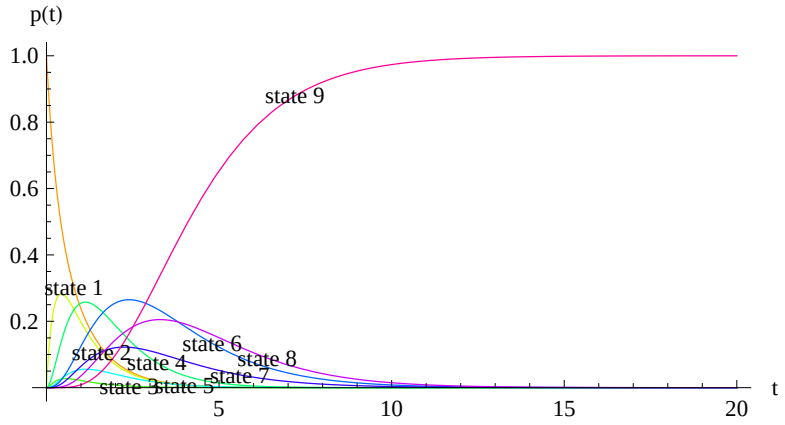
2.3 Part (c)

Using the rate matrix Q, the probability vector p is the solution to the following ODE,

$$\dot{p} = Qp \tag{19}$$

$$p(t) = e^{Qt}p(0). \tag{20}$$

Let $p(0) = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}^T$, then the probability of being in any particular state as a function of time, where state index is indicated in the state transition diagram, is shown in the plot below.



Note that an exact solution can be easily solved analytically by finding a similarity transformation that diagonalizes Q.

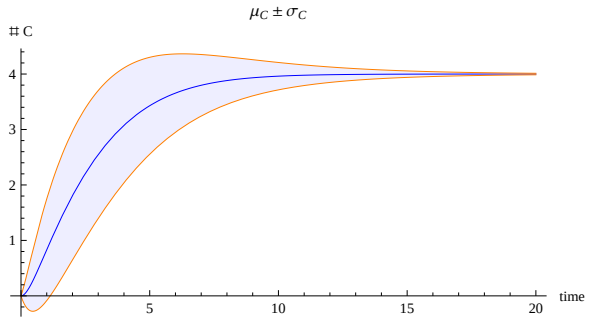
2.4 Part (d)

The mean and variance of C can easily be found knowing $p(t)$ and the value of C at each state in the Markov process,

$$\mu_C(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}^T p(t) \tag{21}$$

$$Var_C(t) = \begin{bmatrix} (0-\mu(t))^2 \\ (0-\mu(t))^2 \\ (0-\mu(t))^2 \\ (1-\mu(t))^2 \\ (1-\mu(t))^2 \\ (2-\mu(t))^2 \\ (2-\mu(t))^2 \\ (3-\mu(t))^2 \\ (4-\mu(t))^2 \end{bmatrix}^T p(t). \tag{22}$$

I used Mathematica to solve these functions numerically. The mean and one standard deviation window are plotted below.



3 Problem 3

Suppose a system of reactions over a finite number of species admits a mass vector with no zero entries. Let v be the vector representing the concentration of all species. Let the stoichiometric matrix of this system be A with nontrivial rate vector $K(x)$ and mass vector m. Then to say the system is conservative

means,

$$\begin{aligned} m^T \dot{x} &= m^T A K(x) && (23) \\ &= 0, \forall v \text{ with nonnegative entries.} && (24) \end{aligned}$$

Note that this condition holds only if $m^T A = 0$, meaning that conservation is independent of rates, and depends only on the topology of the reaction network.

In writing down the stochastic Markov process, note that any discrete state change can be written,

$$x_{t+1} = x_t + A\tilde{k} \quad (25)$$

where x_t is the current state, x_{t+1} is the new state, and $\tilde{k}$ is a rate vector where only one entry is nonzero, and that entry is 1. Looking at the mass of the discrete states,

$$\begin{aligned} m^T x_{t+1} &= m^T (x_t + A\tilde{k}) && (26) \\ &= m^T x_t + m^T A\tilde{k} && (27) \\ &= m^T x_t. && (28) \end{aligned}$$

This means that if the reaction network is conservative, then any discrete state change in the Markov process is also conservative.

Finally, note that the mass of a system at any discrete state is a linear combination of the molecular counts where every term is non-negative. This means that for some initial conditions x_0 with total mass $m_{tot} = m^T x_0$, there are a finite maximum number of species $x_{max,i} = \lfloor m_{tot}/m_i \rfloor$ that can ever occur. This means that the set of unique states that a trajectory x_t may pass through is a subset of the power set of states $\mathcal{P}$,

$$\begin{aligned} x_t &\in \mathcal{P} && (29) \\ \mathcal{P} &= \left\{ \begin{bmatrix} s_1 \\ \vdots \\ s_k \end{bmatrix} \right\}_{s_1=0\ldots x_{max,1},\ldots,s_k=0\ldots x_{max,k}} && (30) \\ |\mathcal{P}| &= \prod_i x_{max,i}. && (31) \end{aligned}$$

Since the order of $\mathcal{P}$ is finite, the mass conserving system must produce a finite number of states.

However, it is not true that a system that produces a finite number of states must be mass conserving! The simplest example of this is the network consisting of the degradation reaction,

$$A \rightharpoonup \emptyset. \quad (32)$$

4 [Extra Credit] Problem 4

$$\emptyset \underset{k_2=1}{\overset{k_1=5}{\rightleftharpoons}} X \tag{33}$$

Below are the plots for $n = 5, 10, 20, 30, 40$, as well as a figure of all of these plots overlayed on each other. Note that the steady state mean of the true infinite state process is at $X = 5$. As you increase n you should see the moments of the resulting finite process approach the moments of the infinite state process. This is the intuition behind the Finite State Projection (FSP) algorithm for approximating the solution of the chemical master equation for infinite or extremely large processes (Munsky and Khammash, 2006). For more information, check out this paper:

> Brian Munsky, and Mustafa Khammash. "The Finite State Projection Algorithm for the Solution of the Chemical Master Equation." The Journal of chemical physics 124, no. 4 (2006): doi:10.1063/1.2145882.