# Solution to Assignment 3

## 1 Q1

The nominal values of biochemical reaction rates ( transcription, translation, and degradation ) used in the gro simulation are shown in Table 1. The various biochemical reaction rates were varied by multiplying $10^{[}1, 0.5, 0, -0.5, -1, -1.5, -2]$. Figure 1 shows the result

**Tab. 1:** Nominal rates of translation($\alpha_r$), transcription($\alpha_p$), and degradation($\beta_r, \beta_p$).

| $\alpha_r$ | $\alpha_p$ | $\beta_r$ | $\beta_p$ |
|---|---|---|---|
| 2.96 | 3.0 | 0.1878 | 0.1 |

of gro simulation with varied transcription rate. The histograms show the distribution of GFP concentration at Ts when the transcription rate is decreased from 1000% to 10% of nominal value. The bottom row plot show the mean and the variance of these distributions. The distributions show that as the transcription rate is decreased, the GFP concentration decreases as well. This relationship is easily derived from solving the ODE of [RNA] and [GFP] discussed in class. For an open-loop expression, the steady-state of [GFP] is $\alpha_r\alpha_p/\beta_r\beta_p$ (Note that the parameters do not exactly reflect the simulation mechanism of gro : this relationship is based on mass-action kinetics continuous model). Thus the steady-state [GFP] is directly proportional to $\alpha_r$. This relationship will be shown through the rest of Question 1, where the transcription rate, degradation rates of ran and protein are varied. The variance of steady-state [GFP], however, does not exhibit a well-characterizable trend as shown in the Figure. This is because the dynamics of second-order cumulants (variance) are slightly more complex than the first-order cumulants (mean). More detail of the relationship between first- and second-order cumulants will be discussed in Assignment 4.

Some of the notable features in the relationship between the biochemical reaction rates and the mean and the variance of GFP are : 1) Increase in the synthesis rates or the decrease in the degradation rates result in higher mean of GFP, 2) The higher degradation rate reduces the noise (variance) of average GFP values, whereas higher synthesis rates increase the noise. The tunability of synthetic gene circuits are thus demonstrated through the four primary biochemical reactions that contribute to the synthesis and degradation of GFP.

## 2 Q2

The reachability graph of the chemical reaction network is shown in Figure 4. There are six possible states given the initial condition of $X_0 = [2,0,0,2,0]^T$ (where $X$ is a vector of copy number for each species). Let $\lambda_1, \lambda_2$ and $\lambda_3$ be the reaction rates, and $p_i$ be the probability distribution of being in state $i = 1, \cdots, 5$. Then we can write down the Chemical Master Equation (CME) as follows.

$$\dot{p_0} = -2\lambda_1 p_0 \tag{1}$$
$$\dot{p_1} = 2\lambda_1 p_0 + \lambda_3 p_3 - (\lambda_1 + 2\lambda_2) p_1 \tag{2}$$
$$\dot{p_2} = \lambda_1 p_1 + \lambda_3 p_4 - 4\lambda_2 p_2 \tag{3}$$
$$\dot{p_3} = 2\lambda_2 p_1 - (\lambda_3 + \lambda_1) p_3 \tag{4}$$
$$\dot{p_4} = \lambda_1 p_3 + 4\lambda_2 p_2 + 2\lambda_3 p_5 - (\lambda_2 + \lambda_3) p_4 \tag{5}$$
$$\dot{p_5} = \lambda_2 p_4 - 2\lambda_3 p_5 \tag{6}$$

The matrix form of Eq (1)-(6) is

$$\frac{d}{dt}\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} = \begin{bmatrix} -2\lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 2\lambda_1 & -(\lambda_1+2\lambda_2) & 0 & \lambda_3 & 0 & 0 \\ 0 & \lambda_1 & -4\lambda_2 & 0 & \lambda_3 & 0 \\ 0 & 2\lambda_2 & 0 & -(\lambda_3+\lambda_1) & 0 & 0 \\ 0 & 0 & 4\lambda_2 & \lambda_1 & -(\lambda_2+\lambda_3) & 2\lambda_3 \\ 0 & 0 & 0 & 0 & \lambda_2 & -2\lambda_3 \end{bmatrix} \begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} \tag{7}$$

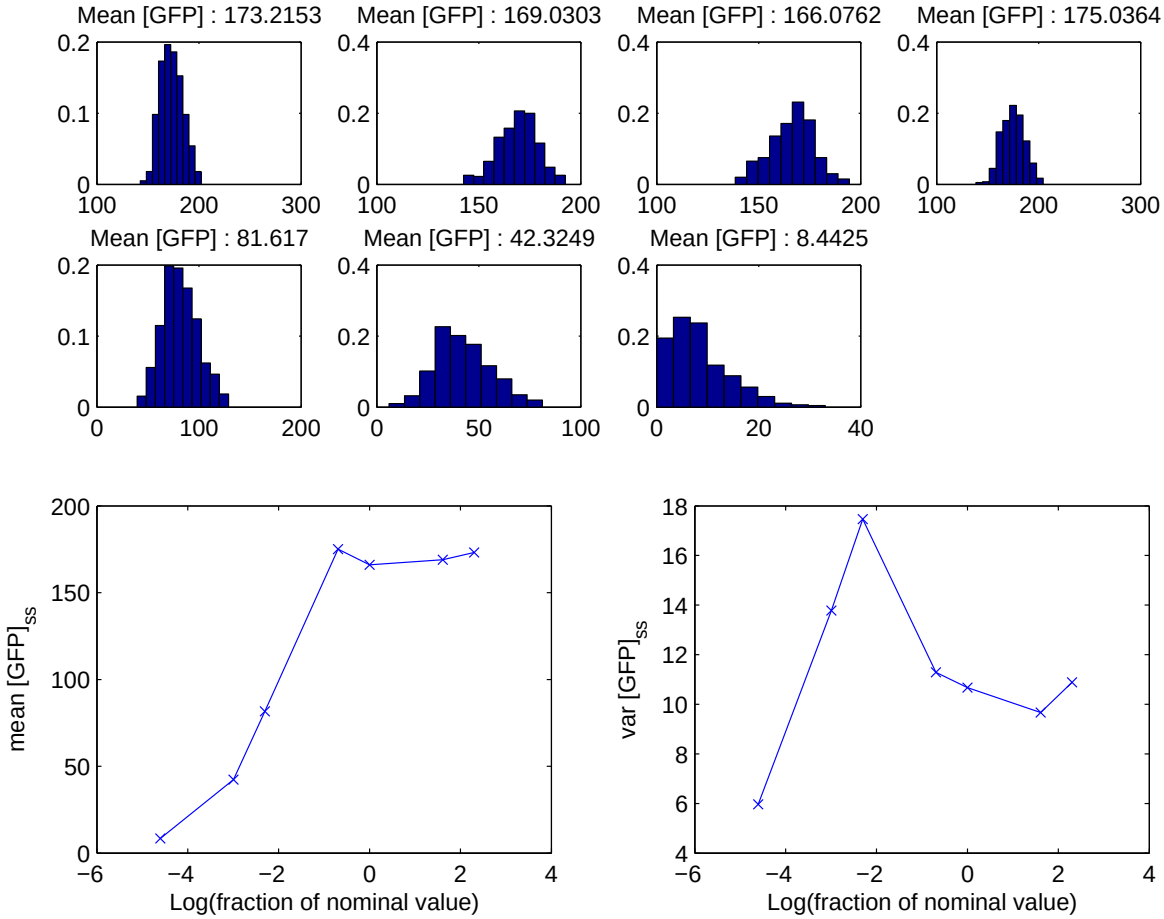


**Fig. 1:** Distributions of GFP concentration at settling time Ts and the mean and the variance of GFP concentration plotted against the transcription rate variation ratio to the nominal value.

(Notice that each column sum to zero). The square matrix in Eq (7) is the transition matrix ($A$) for the Markov process representation of our CRN. The equation has a closed form solution $\mathbf{p}(t) = e^{At}\mathbf{p}_0$ and can be solved analytically to determine the probability distribution of the states at a given time, $t$. Setting $\mathbf{p}_0 = [1,0,0,0,0,0]^T$, we get the following plot. To compute the mean and the variance of the copy number of species E ($N_E$) with respect to time, we can use the following definition.

$$\langle N_E(t)\rangle = \sum_{i=1} N_{E,i}p_i(t) \tag{8}$$

$$= [0,0,0,1,1,2]^T e^{At}\mathbf{p}_0 \tag{9}$$

$$\left\langle N_E^2(t)\right\rangle - \langle N_E(t)\rangle^2 = \sum_{i=1} N_{E,i}^2 p_i(t) - \left(\sum_{i=1} N_{E,i}p_i(t)\right)^2, \tag{10}$$

where the $N_{E,i}$ denotes the copy number of E in state $i$ (e.g. $N_{E,6} = 2$ ). The trajectory of mean and 1 standard deviation window is shown in Figure 6.

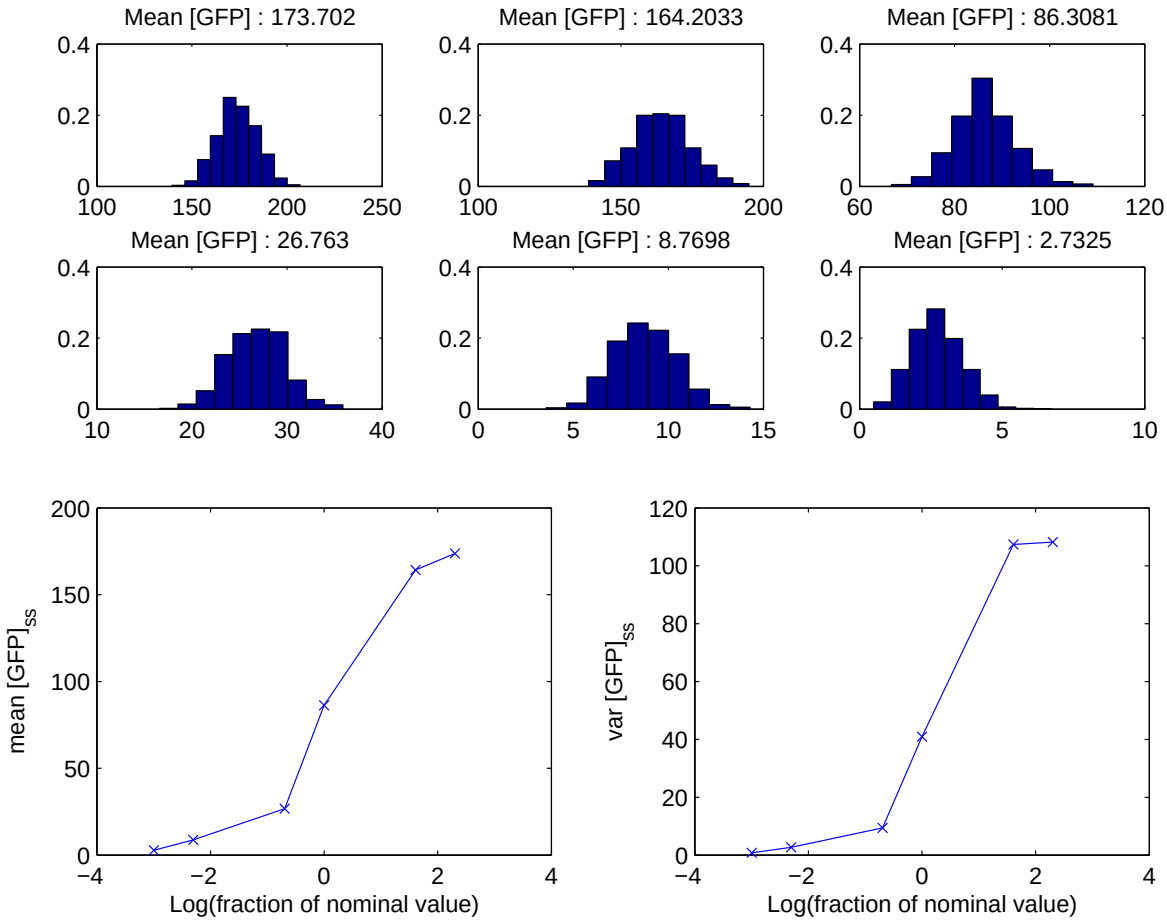


**Fig. 2:** Distributions of GFP concentration at settling time Ts and the mean and the variance of GFP concentration plotted against the translation rate variation ratio to the nominal value.

## 3 Q3

The RBS sequences used in the bistable switch circuit are B0031 ( TCACACAGGAAACC ) and B0034 ( AAAGAGGAGAAA )from the Bioparts Registry (`http://partsregistry.org/Part:BBa_B0031`). They are one of the most frequently used RBSs when constructing synthetic gene networks as their relative strengths are well-defined in certain strains of E. coli. Without a priori information, it can be tricky to pin-point the exact sequences of RBS. However, from knowing the general structure and mechanism of transcription/translation, there are certain clues to infer the location of an RBS. RBS sequences are recognized by ribosomes to begin translation, usually 8 bp upstream of AUG (prokaryotes), and sometimes referred to as the Shine-Dalgarno sequence.

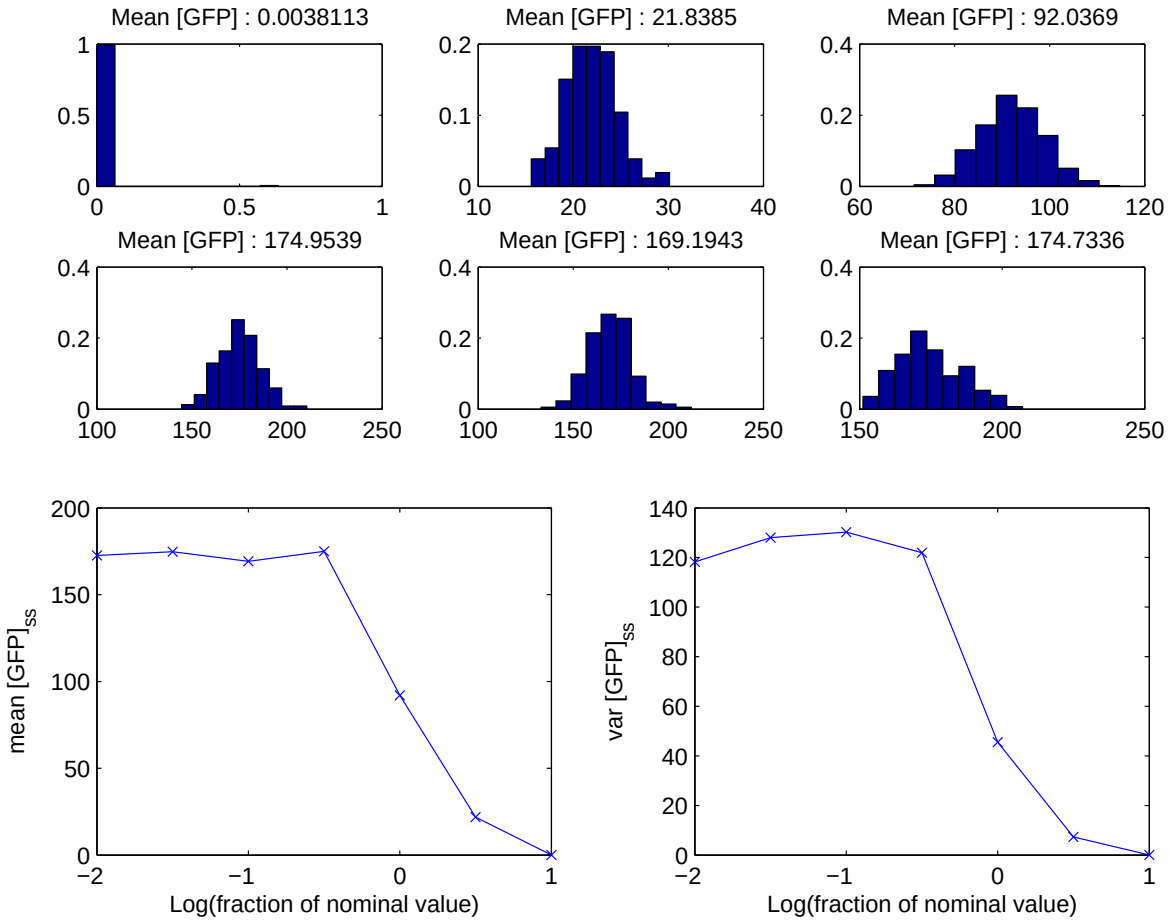


**Fig. 3:** Distributions of GFP concentration at settling time Ts and the mean and the variance of GFP concentration plotted against the mRNA degradation rate variation ratio to the nominal value.

## 4 Grade Metric

1. (3 pts)
   - 0.5 pts for correct demonstration of the relationship between various chemical reaction rates and mean [GFP]
   - 1 pt for summary of ”tunability”
2. (2.5 pts)
   - 0.5 pts for correct Markov process diagram
   - 1 pt for correct rate matrix
   - 0.5 pts for correct probability distribution trajectory
   - 0.5 pts for correctly showing the mean and the std of E copy number using the CME
3. (2.5 pts)
   - 0.5 pts for correctly identifying the RBS sites
   - 1 pt for correct result using RBS calculator

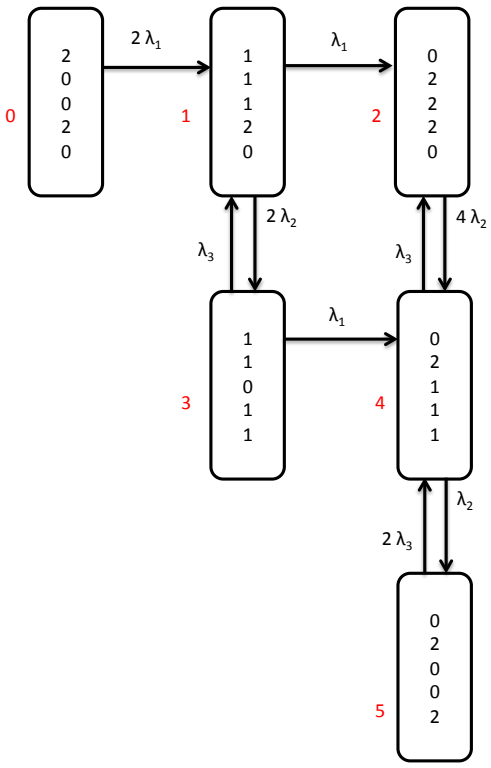


**Fig. 4:** Reachability graph of the chemical reaction networks given in Q2. There are six possible states (enumeration shown in red)

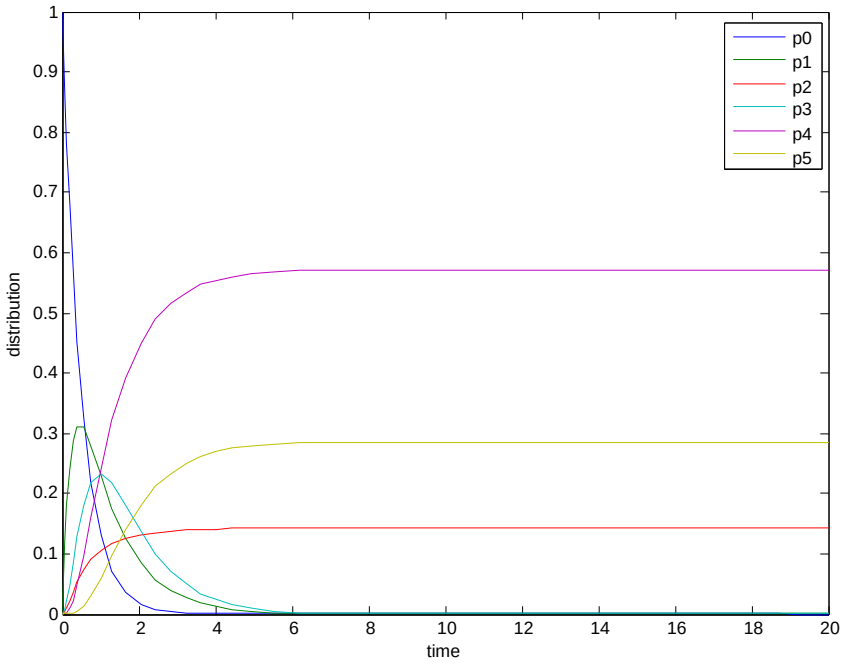


**Fig. 5:** Probability distributions of the states with respect to time.

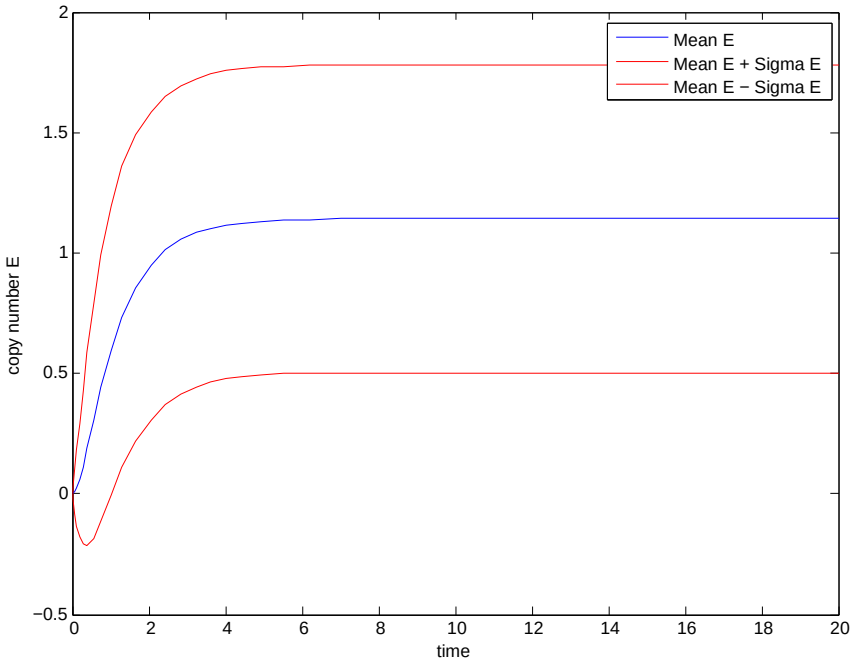


**Fig. 6:** Mean and 1 standard deviation window of copy number E with respect to time.

- 1 pt for the new designs

4. (2 pts)
   - Description and summary of each iGEM project
5. (2pts)
   - Extra credit