Solution to Assignment 1

1 Q1

(*Refer to Figures 1 and 2*).

1. The reversible reaction between A and B will yield reachability graph shown in Figure 1. The two species have conservation of mass where the total number of species may not exceed the initial amount of 2.
2. Similar with previous reaction networks, A and B may switch to one another reversibly. However, B's transformation to C is a terminal reaction. Thus, once all of B (rising from A $\leftrightarrow$ B reaction) becomes C, no more reactions occur (*sink state*).
3. Only three possible states exist when the initial condition is "A = 2, B = 0", because of the 2-to-1 stoichiometric relationship.
4. Figure 2, shows an infinite reachability graph (shown only from 0 to 3 copy numbers for each species) that arise from the given chemical reaction network. There are three different reactions occurring in this network : 1) synthesis of A (left-to-right horizontal arrows), 2) A transforms to B (right-to-left diagonal arrows), and 3) B degradation (downward arrows).

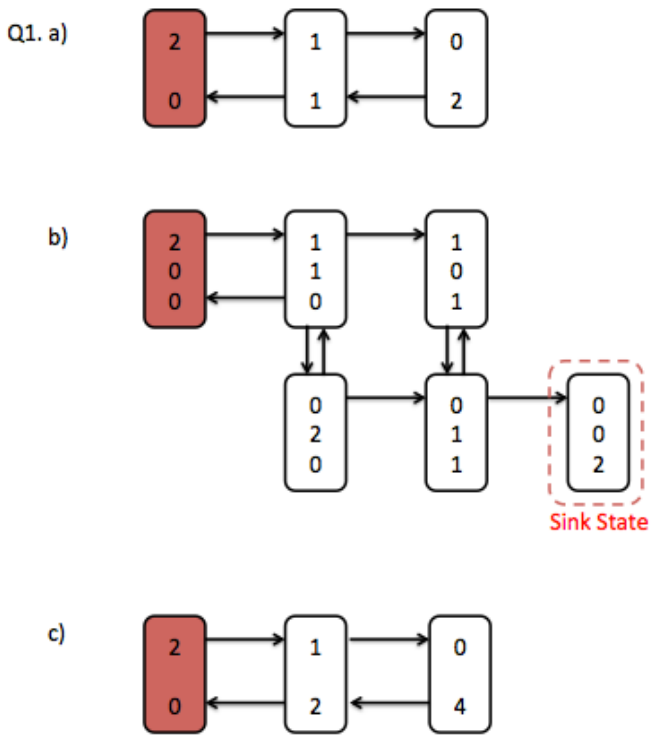

Fig. 1: Reachability graphs of the chemical reaction networks given in Q1. a, b, and c. The initial state is highlighted red.

2 Q2

Using gro, normalized concentrations of GFP and RFP are outputted at the time when the cell population reaches 500 (refer to attached gro code). The scatter plot of $[GFP]$ vs. $[RFP]$ is shown in Figure 3. Notice that when the synthesis of GFP and RFP are independent (transcribed from separate mRNAs), the internal noise is larger. The external noise distribution is unaffected by the change in transcription correlation.

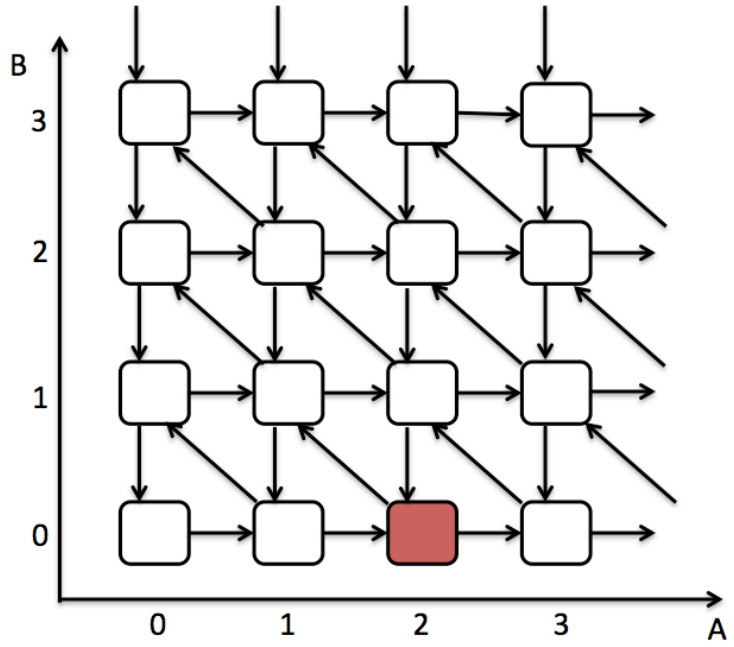

Fig. 2: Reachability graphs of the chemical reaction networks given in Q1. d. The initial state is highlighted red.

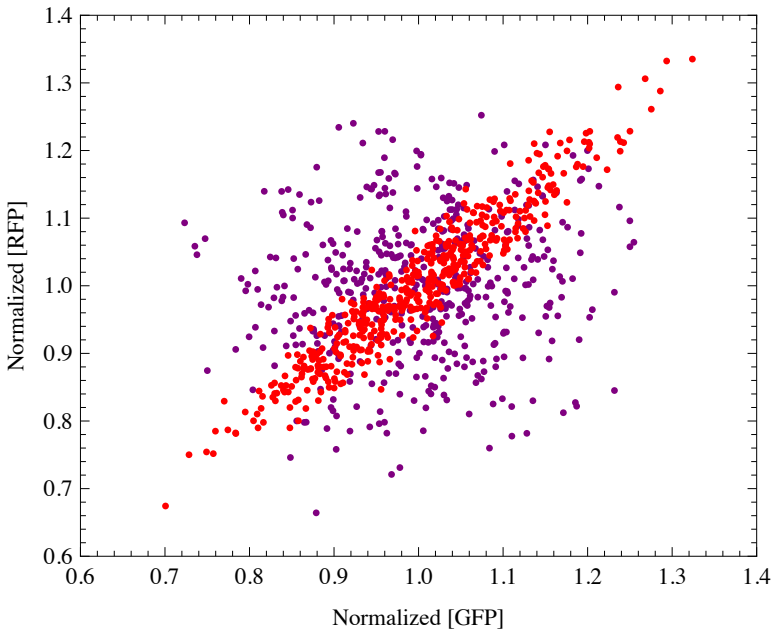

Fig. 3: Quantification of noise. Each point represents the mean concentration of a single cell. (a) Purple dots : GFP and RFP on a separate operon. (b) Red dots : GFP and RFP are on a same operon (shares a promoter).

3 Q3

Figure 4 shows the normalized GFP vs normalized RFP plots obtained using three different sets of rate constants. The purple dots correspond to Q2a solution (synthesis rate = 1, degradation rate = 0.1), the green dots correspond to Q3a solution (synthesis rate = 2, degradation rate = 0.1), and the orange dots correspond to Q3b solution (synthesis rate = 1, degradation rate = 0.05). (Refer to gro code). Table 1 shows the internal, external and total noise computed from each simulation. (The equations are obtained from Supplementary Material [1]) It shows that when the synthesis rate of the chemical species increase, the internal noise to external noise ratio reduces. The external noise (spread of dots along x = y line) are more or less similar to each other, however.

Tab. 1: Internal, external and total noise of GFP, RFP simulation in gro.

	η_{int}	η_{ext}	η_{tot}
Example 2a	9.4×10^{-2}	3.6×10^{-2}	1.0×10^{-1}
Example 3a	1.0×10^{-2}	4.8×10^{-2}	4.9×10^{-2}
Example 3b	8.4×10^{-2}	3.9×10^{-2}	9.3×10^{-2}

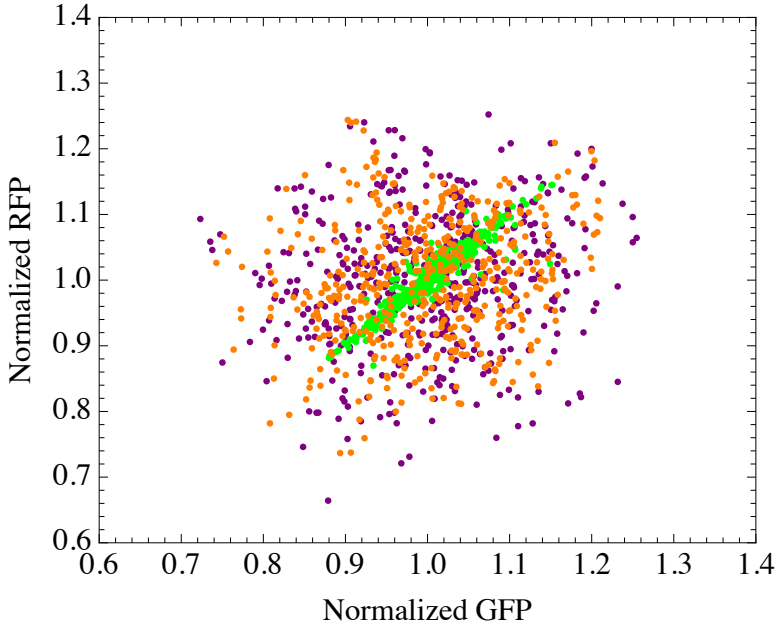

Fig. 4: Normalized GFP vs Normalized RFP plot. (a) Purple : $\alpha = 1$, $\beta = 0.1$, (b) Green : $\alpha = 2$, $\beta = 0.1$, (c) Orange : $\alpha = 1$, $\beta = 0.05$.

4 Q4

The output GFP copy numbers in all cells are normalized to plot the distribution shown in Figure 5 (Purple). Poisson distribution is fitted to the output and the result is shown in yellow ($\alpha \approx 322.6$). (Used Mathematica FindFit function, where the optimization cost is measured in 2-norm). When plotting the function, you will notice that a 'true' probability distribution will have much lower distribution overall, resulting in experimental data looking much more taller. This can be fixed by 'approximating' the Poisson distribution by binning the discrete random variables together. Even with all the tricks, the fit will not look great. This is because using the Poisson distribution to predict the GFP population does not take into account the fact that cells are dividing. At each division, the molecules will get divided up asymmetrically (and this noise can be approximate with binomial distribution), thus introducing a secondary source of noise for GFP copy number.

The comparison of fano factors for the two different simulations (protein only, mRNA + protein) shows that the two step reaction has higher value (three-fold). This indicates that the more sequential reactions are incorporated, the noisier the downstream process may get.

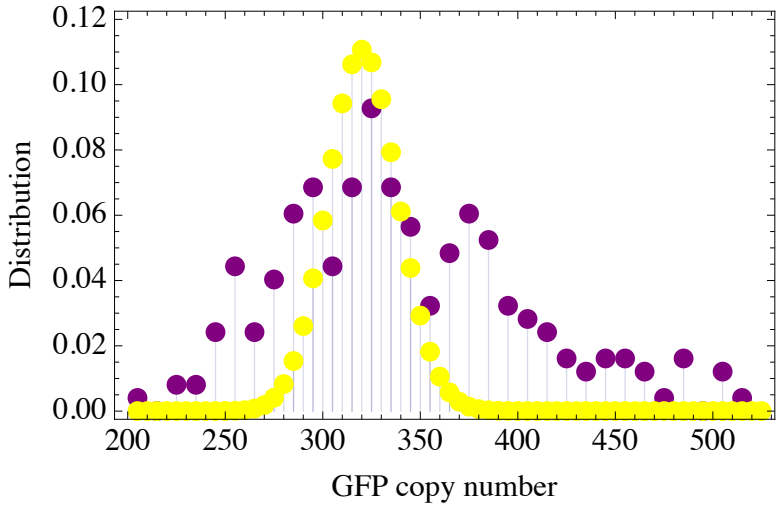

Fig. 5: GFP copy number distribution (Purple) and estimated Poisson distribution (Yellow)

```
include gro
numCells := 0;
synth := 1;
deg := 0.1;

program p() := {
  mRNA1 := 0;
  gfp := 0;
  mRNA2 := 0;
  rfp := 0;
  rate ( synth * volume ) : { mRNA1 := mRNA1 + 1 };
  rate ( deg * mRNA1 ) :    { mRNA1 := mRNA1 - 1 };
  rate ( synth * mRNA1 ) :   { gfp := gfp + 1 };
  rate ( synth * volume ) : { mRNA2 := mRNA2 + 1 };
  rate ( deg * mRNA2 ) :    { mRNA2 := mRNA2 - 1 };
  rate ( synth * mRNA2 ) :   { rfp := rfp + 1 };
  true : {
    numCells := numCells + 1,
  };
};

program report() := {
  needs mRNA1, gfp, mRNA2, rfp;
  selected : { message ( 1,
          "cell "  <> tostring(id)
     <> ": mRNA1 ="  <> tostring(mRNA1)
     <> ", GFP="   <> tostring(gfp)
     <> ", [GFP]=" <> tostring(gfp/volume)
     <> ": mRNA2="  <> tostring(mRNA2)
     <> ", RFP="   <> tostring(rfp)
     <> ", [RFP]=" <> tostring(rfp/volume)
     <> ", Num = " <> tostring(numCells) ) }
};

program output(delta) := {
  needs mRNA1, gfp, mRNA2, rfp;
  p := [ t := 0, s := 0 ];
  true : {
    p.t := p.t + dt;
    p.s := p.s + dt;
  };
```

```
  p.s > delta : {
    print( id, ", ", p.t, ", ", volume, ", ", mRNA1, ", ", gfp, ", ", mRNA2, ", ", rfp, "\n" );
    p.s := 0;
  };
};

program main() := {
  numCells > 500 : {
    exit();
  };
  true : {
    numCells := 0;
  };
};

set( "gfp_saturation_min", 0);
set( "gfp_saturation_max", 100 );
set( "rfp_saturation_min", 0 );
set( "rfp_saturation_max", 100 );

ecoli ( [ x := 0, y := 0 ], program p()
+ report() sharing mRNA1, gfp, mRNA2, rfp
+ output(dt*10) sharing mRNA1, gfp, mRNA2, rfp );
```

```
include gro
synth := 5;
deg  := 0.5;
numCells := 0;
program p() := {
  // version 1.
  //gfp := 0;
  //rate ( synth * volume ) : { gfp := gfp + 1 };
  // version 2
  mRNA := 0;
  gfp := 0;
  rate ( synth * volume ) : { mRNA := mRNA + 1 }
  rate ( deg * mRNA ) : { mRNA := mRNA - 1 }
  rate ( synth * mRNA ) : { gfp := gfp + 1 };
  true : {
    numCells := numCells + 1,
  };
};

program report() := {
  needs gfp;
  selected : { message ( 1,
          "cell "  <> tostring(id)
     <> ": GFP ="  <> tostring(gfp)
     <> ": [GFP] =" <> tostring(gfp/volume) ) }
};

program output(delta) := {
  needs gfp;
  p := [ t := 0, s := 0 ];
  true : {
    p.t := p.t + dt;
    p.s := p.s + dt;
  };
  p.s > delta : {
    print( id, ", ", p.t, ", ", volume, ", ", gfp, "\n" );
    p.s := 0;
  };
};

program main() := {
  numCells > 250 : {
    exit();
  };
  true : {
    numCells := 0;
  };
};

set( "gfp_saturation_min", 0);
set( "gfp_saturation_max", 800 );
ecoli ( [ x := 0, y := 0 ], program p() + report() sharing gfp + output(dt*10) sharing gfp );
```

References

[1] M Elowitz, A Levine, E Siggia, and P Swain (2002) Science

Intro to Synthetic Biology

Assignment 2

Grade metric

1. (2pts)
    a. 0.5pts for each correct reachability graph
2. (3pts)
    a. (2pts) Correct scatter plot of Normalized GFP vs RFP for a) and b)
    b. (1pt) Noise analysis
3. (3pts)
    a. (2pts) Correct scatter plot
    b. (1pt) Noise analysis
4. (2pts)
    a. (1pt) experimental Distribution and estimated Poisson distribution
    b. (1pt) reasons for discrepancy