

Problem 1a)

The volume basis beta = $\langle m_c \rangle * N_c * V = 3.96E-10 \text{ gDW/cell} * 1E8 \text{ cells/ml} * 1 \text{ ml} = 3.96E-2 \text{ gDW}$ in each sample. To convert to the specific volume basis, we first change the number of copies of mRNA to nmoles, by dividing the number in the table by avogadro's constant ($6.023E23$), and converting to nmoles by multiplying by $1e9$. Next, we can convert to nmoles by multiplying by the concentration of cells per sample, $1E8 \text{ cells/ml}$, and the sample volume ($1/\text{ml}$). Finally, divide by the volume basis calculated above to obtain nmoles/gDW.

IPTG (mM)	mRNA/cell	moles/cell	nmoles/cell	nmoles/gDW
0	19	$3.15457E-23$	$3.15457E-14$	$7.97E-05$
$5.00E-04$	21	$3.48663E-23$	$3.48663E-14$	$8.80E-05$
0.005	41	$6.80724E-23$	$6.80724E-14$	$1.72E-04$
0.012	67	$1.1124E-22$	$1.1124E-13$	$2.81E-04$
0.053	86	$1.42786E-22$	$1.42786E-13$	$3.61E-04$
0.216	93	$1.54408E-22$	$1.54408E-13$	$3.90E-04$
1	93	$1.54408E-22$	$1.54408E-13$	$3.90E-04$

	Value	Bionumber or formula
mass per cell (g/cell) A	$1.00E-09$	111609
Water fraction in cell B	$6.04E-01$	103689
gDW/cell	$3.96E-10$	$A*(1-B)$

Problem 1b.)

At pseudo-steady state, we can set the mRNA rate equation to 0. We can then rearrange to solve for m.

$$0 = r_x \bar{u} - (\mu + \theta_m) m^* \Rightarrow (\mu + \theta_m) m^* = r_x \bar{u}$$

$$m^* = \frac{r_x \bar{u}}{\mu + \theta_m} = \frac{r_x}{\mu + \theta_m} \bar{u}$$

$$r_x = k_E^X R_{X,T} \left(\frac{G}{\tau_X K_X + (\tau_X + 1)G} \right) \text{ where } k_E^X R_{X,T} = Vmax \left(\frac{\text{nmol}}{\text{gDW} * \text{hr}} \right) \text{ and } k_E^X = \frac{e_x}{L}$$

$$\bar{u} = \frac{W_1 + W_2 f_I}{1 + W_1 + W_2 f_I} \text{ where } f_I = \frac{I^n}{K^n + I^n} \text{ and } I = \text{inducer}$$

$$\text{Next, set } K_X = \frac{r_x}{\mu + \theta_m}, \text{ so that } m^* = K_X(\theta, G) \bar{u}(I, \kappa)$$

With this, we have derived the gain function, K_X , and the mRNA pseudo steady state abundance. Since the degradation and gene abundance is constant, the gain function is actually a constant.

Matthew Tan
5-12-2020
Prelim
CHEME 7770
Problem 1c.)

Now we need to estimate parameters to fit the data in the table. First, let's work with what we're given. To convert gene copies per cell to a concentration, first convert to nmoles/cell using similar steps as in 1a, then multiply by number of cells/ml (10^8) to obtain, nmoles/mL, and convert to nmoles/L for nM:

Parameter	Value
μ specific growth rate (hr ⁻¹)	$1/(\tau_D \text{ in hr}) * \ln(2) = 1.039720771$
θ_m degradation rate (hr ⁻¹)	$1/(\text{half-life in hr}) * \ln(2) = 8.317766167$
G gene concentration (nM)	3.32×10^{-4} (2 gene/cell $> 3.32 \times 10^{-15}$ nmoles/cell $> 3.32 \times 10^{-7}$ nmoles/mL)
$\mu + \theta_m$ (hr ⁻¹)	9.357

Other parameters that we need can be estimated or derived from outside sources. RNA polymerase was converted to nmole/gDW using identical calculations as in 1a.

Parameter	Value
K, dissociation con. IPTG (uM)	49.6 (Bio#: 101976)
K_x (uM)	0.03328 (from McClure) ($K_x = \text{slope/intercept fig 2, avg for A2 and D}$)
$R_{x,T}$ RNA pol (nmole/gDW)	5000 RNAP/cell $> 2.10 \times 10^{-2}$ (Bio#: 101440, for 40 min)
L LacZ length (nt)	3075 (bio#: 102070)
e_x elongation rate (nt/sec)	42 (bio#: 108488)
K_i (sec ⁻¹)	0.02381 to 0.04 (McClure)

Now some calculated parameters

Parameters	Value
K_E^x (hr ⁻¹)	$42/3075 = 0.013658537 \text{ sec}^{-1} = 49.17$
Vmax (nmole/gDW-hr)	$2.10 \times 10^{-2} * 49.17 = 1.03$
τ_x (dimensionless)	$K_E^x/K_i = 0.468994$ (avg of $(49.17/0.0238 \text{ to } 49.17/0.04)$)

My original value for τ_x made the saturation term much too small to match the nmoles/gDW in the data table. We can deduce that the gain function should equal the max mRNA amount, as this will occur when the control function is 1. From the data, we can see that the max amount is 3.9×10^{-4} nmoles/gDW, while the τ_x I found gives a gain function of about 2.4×10^{-6} , which is smaller than the 0 inducer value. My other values seem correct, so to make this work, we're going to apply a correction factor to our gain function. If we know that the gain function should be about 3.90×10^{-4} , and we get a calculated gain function of 2.40×10^{-6} , then our correction factor, α should be:

$$\alpha = \frac{\text{Max mRNA}}{\text{Calculated gain function}} = 162.2993$$

Thus, I will multiply my mRNA formula by this correction factor α . My parameter values are derived from literature but not for this specific case, so I'll assume that this correction factor accounts for a degree of uncertainty in what the actual parameters should be for this particular gene system.

Matthew Tan

5-12-2020

Prelim

CHEME 7770

Now, we need to determine control functions. At 0 inducer, we know what the value of mRNA is, and what $f_I = 0$. Thus:

$$7.97 * 10^{-5} = 3.90 * 10^{-4} \frac{W_1}{1 + W_1}, W_1 = 0.257$$

Now, to solve for W_2 , we first need to find out what F_I looks like. By doing a least squares difference fit to a agonist vs response function with variable slope and 4 parameters function in Graph pad prism for the data in table 1, I obtained an $n = 1.5$. Next, to calculate the W_2 , I picked a point in the middle of the data set, $[0.012, 2.81E-4]$, and used Wolfram alpha to solve for W_2 , noting that:

$$2.81 * 10^{-4} = 3.90 * 10^{-4} \left(\frac{W_1 + W_2 f_I}{1 + W_1 + W_2} \right), \text{ where } f_I = \frac{0.012^{1.5}}{K^{1.5} + 0.012^{1.5}}$$

Subbing in the appropriate W_1 and the K (dissociation constant of IPTG from LacZ), I found that W_2 is about 21.8249. Now I have everything I need to compute the model! (Note: I tried using the inducer value of 1 and the max mRNA to find W_2 , but it kept giving me very large W_2 (trillion), so no).

Problem 1D.

I've my model data and the table data on a plot with the x-axis a log scale axis. That plot is shown below:

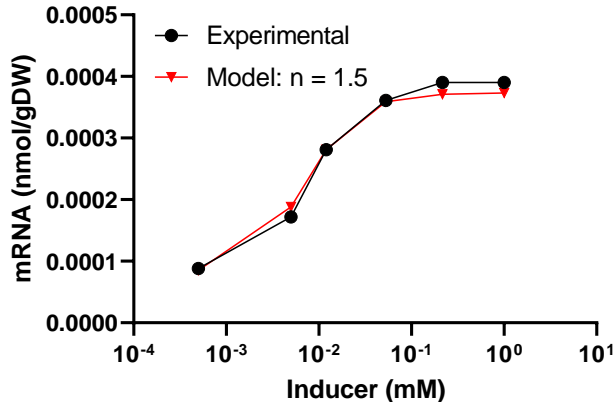


Figure 1: Semilog plot of mRNA amount from the experimental data in the first table along with my model predictions.

That is a really good fit! The values themselves were computed using Excel, and plotted using GraphPad Prism. Prism reports that the likelihood that the two data-sets share the same curve is >99.99%, based on Akaike's Information Criterion, which indicates an extremely good fit. One way I've found to change the shape of the curve is to change the hill coefficient, n , for f_I . This seems to have a large influence on how well the shape of the model matches the data. You can also change the dissociation constant for IPTG. You can also change the fit itself (relative magnitudes) by changing the weights of the microstates, and the gain function (which shouldn't be changed based on the data we know). All of the data, calculations, and parameters can be found in the excel file `prelim_v2.xlsx`, as well as the prism file `Prelim problem 1.pzfx`.

Problem 2a and Problem 2b

See the attached sheet for the derivation of the systems of equations and subsequent non-dimensionalization. Note, there was an error in equation 6 of the non-dimensionalization steps in the paper, where they had $X(\sim)$ multiplied by δx . Since $X(\sim)$ has units of concentration, and δx has units of time (or per time), it wouldn't become unitless when divided by $\alpha_Z(\sim)$. This was also seen in their non dimensional forms of Y and Z . I have corrected this in my steps.

Problem 2c

This was tricky. The brute force way would be to find steady state values of X and different S values (where some S values can give multiple steady state values), and then evaluate each steady state using eigenvalues to determine stability. While this could've been implemented using loops, I decided to use Bifurcations.jl to make my life easier.

To make this work, I need to give an initial guess of a steady state. The easiest steady state to find would be when $S = 0$. I set $S = 0$ in the function and phase portrait code (Bistable.jl) that was given for the last homework, and instructed it to give me the steady state solution at long t (60). For $S = 0$, this appears to be $[X, Z] = [0.1168485008013625, 0.9989844651644064]$, which is pretty close to $[0, 1]$. The solution plotted on the phase portrait confirms this. So, my initial guess for the steady state for a Bifurcation problem will be $[0, 1]$.

I then ran bistablebifur.jl, which evaluates the bifurcation problem using the Bifurcations.jl package. It then generated a plot of all steady state values, which is shown below in Figure 2.

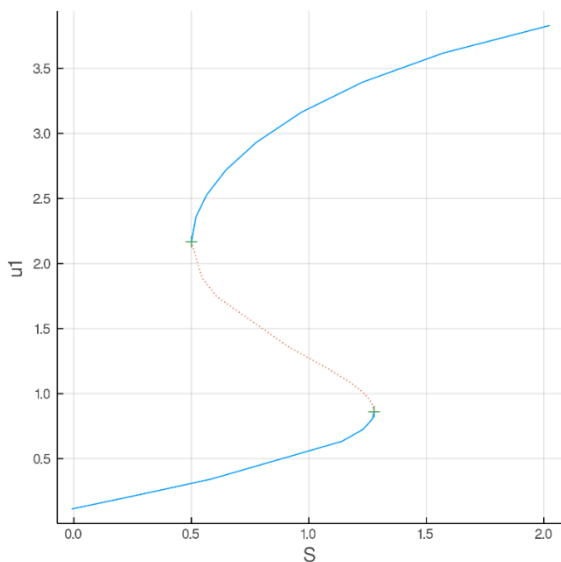


Figure 2: Full bifurcation plot for the bistable toggle switch. The blue lines represent stable steady states, and the dotted red represent unstable steady states. The green crosses show saddle node points. Note that $u1 = X$ in this plot.

Additionally, I found the location of the saddle node points, which appear to be:

$$[X, S] = [0.85, 1.28] \text{ and } [2.05, 0.51]$$

From the plot, it appears that the plot shown in Figure 1B of the paper is indeed reproducible.

For 2D, I used Julia's ODE solver once again to solve the system of equations shown in Equation 1 of the paper. I've attached the plots (Figure 3) that show X vs time for various S values. (2d_prelim.jl)

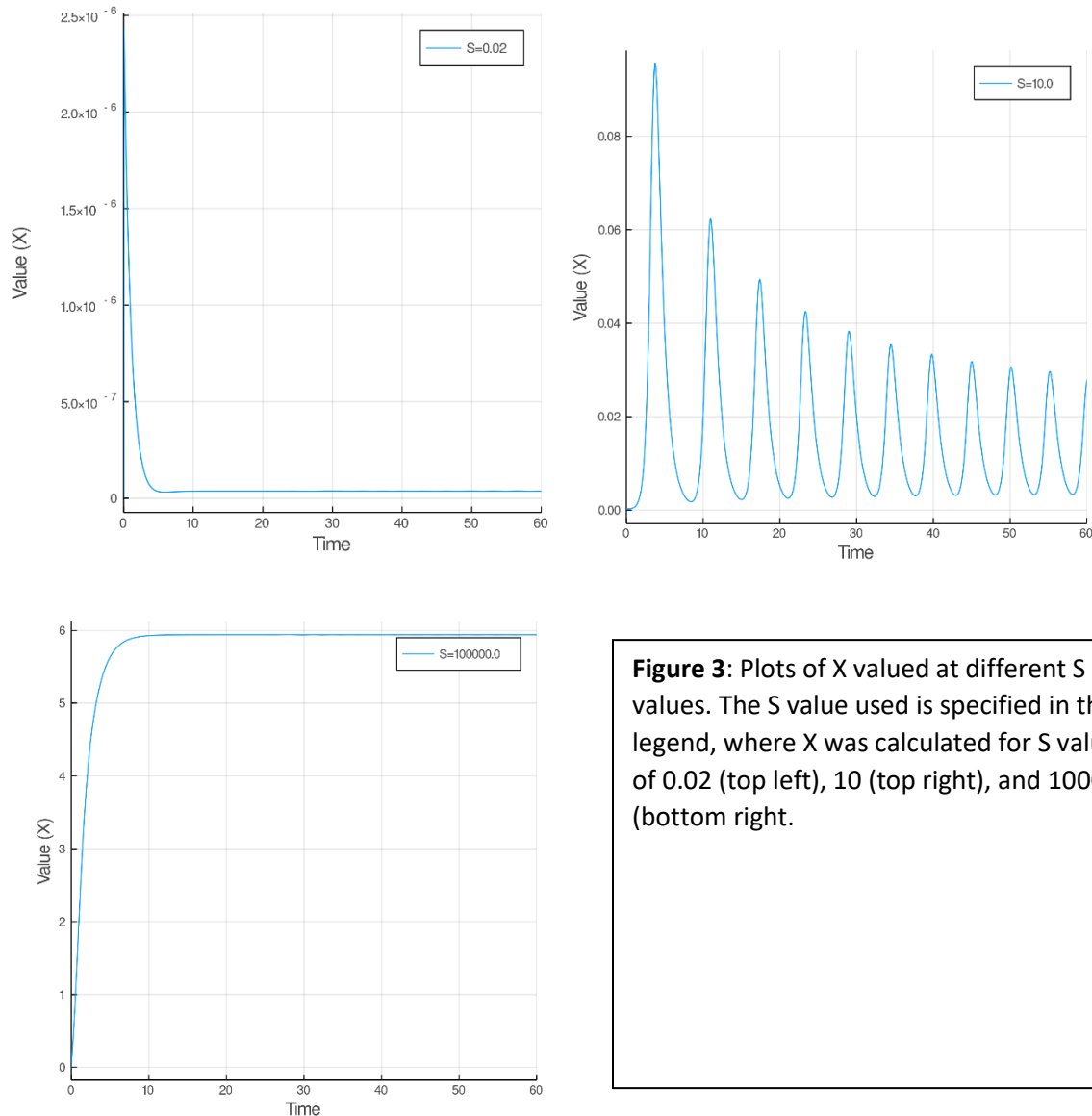


Figure 3: Plots of X valued at different S values. The S value used is specified in the legend, where X was calculated for S values of 0.02 (top left), 10 (top right), and 100000 (bottom right).

From the plots above, it looks like we can recreate Figure 2B of the paper. I've also included plots that show X, Y, and Z for each S values at the end of this document.

For this problem, I first tried to figure out roughly where the Hopf bifurcation and the saddle node occurs. I accomplished this by repeatedly plotting values of Z at increasing S values. Whenever Z began to oscillate, I noted that's where the Hopf bifurcation was. When Z returned to a single value, I noted that's where the saddle node was. From my testing, I discovered that the Hopf bifurcation occurs between 0.2 and 0.3 S , and that the saddle node occurs at about 35000 to 36000 S .

For the first case, I chose an S value of 0.075, which gave X , Y , Z values of :

3.5353541873779054e-5

0.3818357211822364

0.000797330990777635

For the second case, I chose an S value for 37500, which gave X , Y , and Z values of:

5.551000598740573

0.004120014320430175

0.0004491125975008558

Once I had these values, I set the initial X , Y , and Z values to these points for Cell 1, 25% larger for Cell 2, and 25% smaller for Cell 3. I then reran the ODE solver for S of 100. These plots are shown below (Figure 4). To generate these plots, I used 2e_prelim.jl.

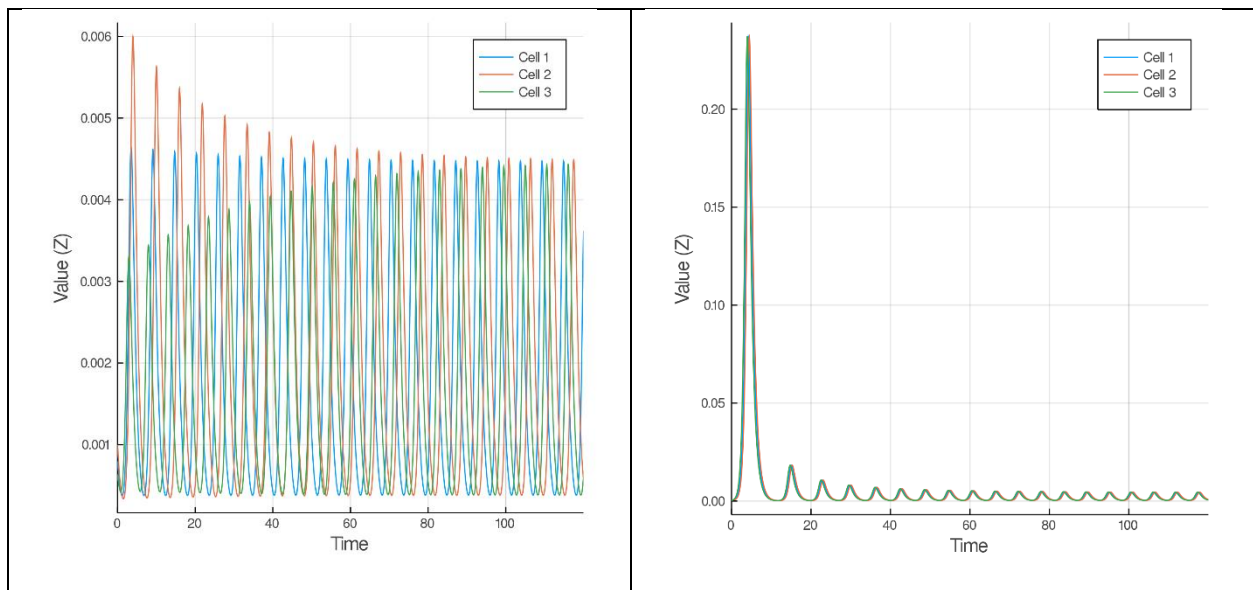


Figure 4: Plots of Z values for cells that started at an S of 0.075 (left), or cells that started at an S of 37500 (right), and then instantaneously changed to $S = 100$ at time = 0. The initial steady state values are given in the text, as well as in plot form in the appendix.

Matthew Tan

5-12-2020

Prelim

CHEME 7770

For the first case, where S started below the Hopf bifurcation, the oscillations are incoherent, based on the fact that they're out of phase with each other. For the second case, where S started above the saddle node, the oscillations are coherent, as we can see that they completely overlap each other. This is in line with the results in the paper, which state that when going through the Hopf bifurcation, small initial differences (the 25% in our case) are amplified, since the Hopf bifurcation cause oscillations to become a repulsing spiral. In contrast, cells that pass through the saddle node are far from expression levels associated with the oscillatory regime. Although in both cases, the initial steady state did not have noticeable oscillations (see appendix).

Problem 2F.

I don't think they were able to obtain coherent oscillations by decreasing from 105 to 100, at least not in the way that it was described. This is because at $S = 105$, you're still below the saddle point and above the Hopf bifurcation, meaning that Z was already oscillating instead of being at steady state. See figure S3 for the baseline "steady-state" at $S = 105$, which is not a steady state value at all. Therefore, I couldn't actually reproduce their plot in Figure 3E, since I don't know how to computationally change the S value during the ODE problem. Regardless, the saddle point appears to be well above $S = 105$, so I don't think that their description in Figure 3E is correct.

APPENDIX

Problem 2A and 2B derivations and non dimensionalization.

$$\frac{d\tilde{X}}{d\tilde{t}} = \frac{\tilde{\alpha}_x + \tilde{\beta}_x S}{1 + S + (\tilde{Z}/\tilde{x}_x)^{n_{zx}}} - \tilde{\delta}_x \tilde{X} \quad 1.) \quad \leftarrow \underline{\underline{\text{Part A.}}}$$

$$\frac{d\tilde{Z}}{d\tilde{t}} = \frac{\tilde{\alpha}_z}{1 + (\tilde{X}/\tilde{x}_z)^{n_{xz}}} - \tilde{\delta}_z \tilde{Z} \quad 2.)$$

Part B.) To non-dimensionalize, $\delta_z = \frac{\tilde{\delta}_z}{\tilde{\delta}_x}$, $t = \tilde{t} \delta_x \quad 3.)$

$$\alpha_x = \frac{\tilde{\alpha}_x}{\tilde{\alpha}_z}, \quad \beta_x = \frac{\tilde{\beta}_x}{\tilde{\alpha}_z} \quad 4.)$$

$$z_x = \frac{\tilde{z}_x \tilde{\delta}_x}{\tilde{\alpha}_z}, \quad x_z = \frac{\tilde{x}_z \tilde{\delta}_x}{\tilde{\alpha}_z} \quad 5.)$$

$$X = \frac{\tilde{X}}{\tilde{\alpha}_z}, \quad Z = \frac{\tilde{Z}}{\tilde{\alpha}_z} \quad 6.) \quad \star$$

So the non-dimensionalized form is

$$\boxed{\begin{aligned} \frac{dX}{dt} &= \frac{\alpha_x + \beta_x S}{1 + S + (Z/z_x)^{n_{zx}}} - X \\ \frac{dZ}{dt} &= \frac{1}{1 + (X/x_z)^{n_{xz}}} - \delta_z Z \end{aligned}}$$

Corrected!

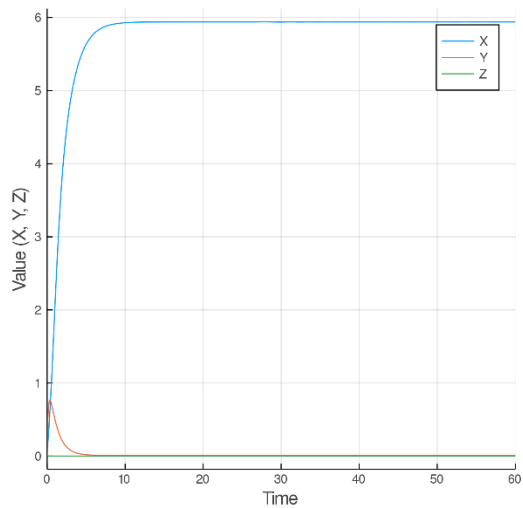
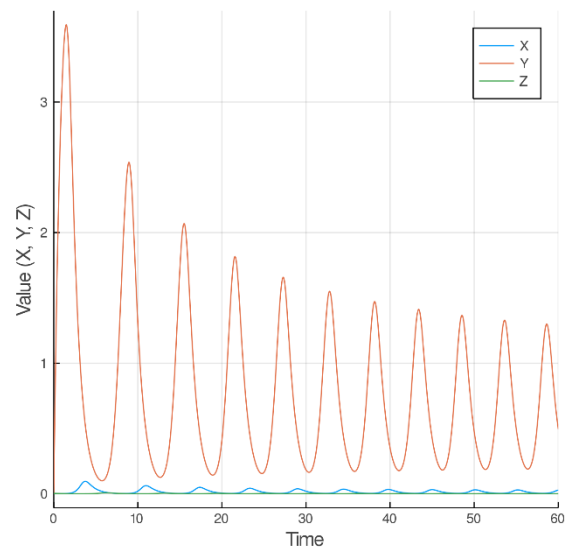
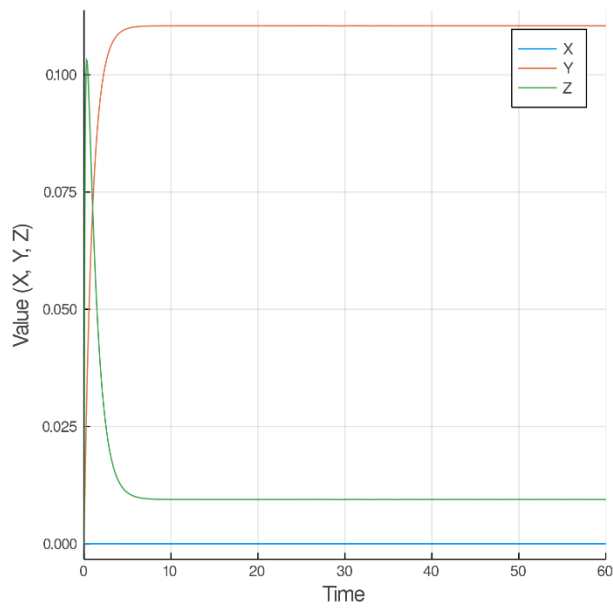


Figure S1: Plots of X, Y, and Z values at different S values. The S value used is specified in the legend, where X was calculated for S values of 0.02 (top left), 10 (top right), and 100000 (bottom right).

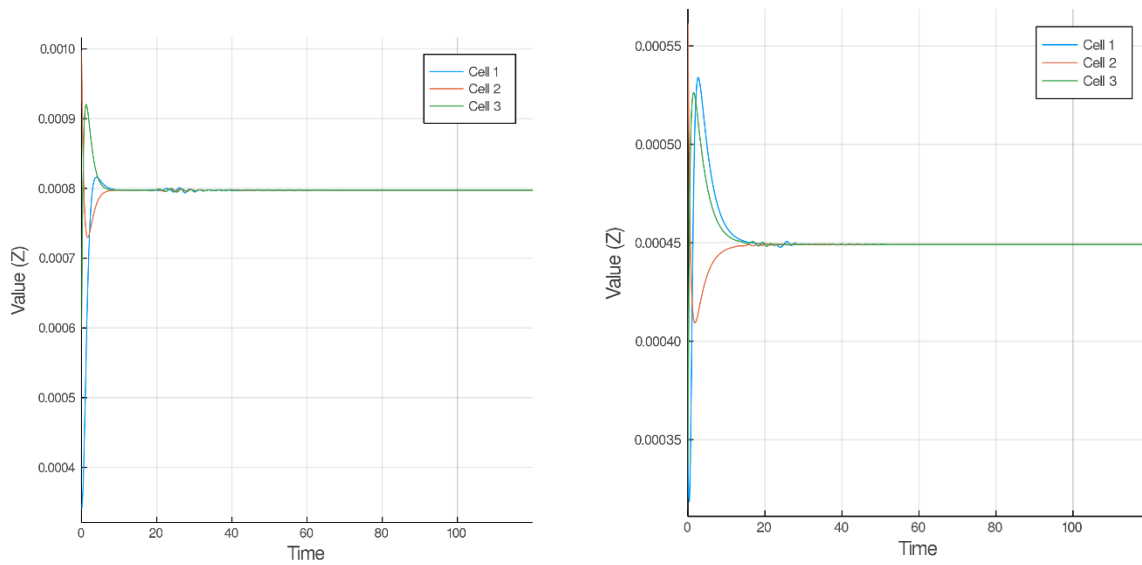


Figure 2S: Initial plots of cell 1, 2, and 3 before changing the S value to 100. Initial S values are $S=0.075$ (left) and $S=37500$ (right).

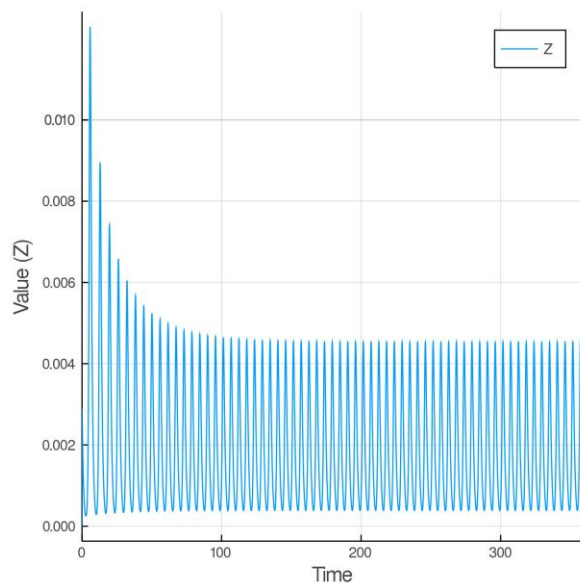


Figure S3: The “steady state” at $S = 105$, which is not a steady state at all like what’s shown in Figure 3E of the Perrez-Carrasco paper.