Modeling Gene Expression with Differential Equations

Madison Kuduk

Department of Computer Science and Mathematics

Arcadia University

February 2020

Abstract

Gene expression is the process by which the information stored in DNA is converted into a functional gene product, such as protein. The two main functions that make up the process of gene expression are transcription and translation. Transcription and translation are controlled by the number of mRNA and protein in the cell. Gene expression can be represented as a system of first order differential equations for the rate of change of mRNA and proteins. These equations involve transcription, translation, degradation and feedback loops. In this paper, I investigate a system of first order differential equations to model gene expression proposed by Hunt, Laplace, Miller and Pham in their technical report, "A Continuous Model of Gene Expression", as well as past models that inspired theirs. I solve the model by Hunt et al. for various equilibrium points and analyze those points through eigenvalues and bifurcations to understand the biological relevance.

1 Introduction

Gene expression is the process by which the information stored in DNA is converted into a functional product, such as protein. This is how the gene takes its effect on the cell. Gene expression can be thought of as an on/off switch for the production of proteins. It also controls the volume in which proteins are made. Two basic processes make up gene expression: transcription and translation. In short, transcription obtains the genetic information within DNA and copies it to make mRNA. Translation decodes the mRNA to create protein.

Gene expression is controlled by both direct and indirect interaction with other genes and their products as a result of changing environmental stimuli. With that, the process of gene expression occurs within the cell simultaneously amongst many genes that are all affecting one another. This causes the process to become extremely complex, especially with higher-order species. Due to this majority of research has been focused on transcription in bacteria (prokaryotes) and simple eukaryotes. Eukaryotic cells have a nucleus, multiple organelles and a large amount of DNA arranged in the form of chromosomes. In prokaryotic

cells, there is no nucleus or organelles and the DNA is a significantly lesser amount that is compartmentalized in the form of a single circular chromosome. Time delay is not a factor in gene expression in prokaryotic cells because all the DNA is in one place. This is fact also simplifies the process in prokaryotes.

Hunt, Laplace, Miller and Pham propose a system of first order differential equations to model gene expression in their technical report, "A Continuous Model of Gene Expression."[4] The goals of the report are to create a mathematical model to describe gene expression in prokaryotes and analyze the behavior of the model through equilibrium solutions and bifurcations. With the analysis they would like to gain insight on the on the interrelation of transcription and translation in prokaryotic cells as well as possibly gain some insight on a way to model gene expression in higher order eukaryotic cells. My goals are to understand their model and its shortcomings, and possibly expand on their research in hopes of making strides towards a more accurate model.

1.1 The Biological Processes

The model proposed by Hunt et al. mirrors protein synthesis in prokaryotic bacteria.[4] The two processes in protein synthesis are transcription and translation. To begin transcription, DNA makes contact with the polymerase enzyme, whose main fuction is to synthesize DNA molecules. The polymerase enzyme moves along the DNA, breaks its bonds to unwind it and uses one side as a template to make messenger RNA(mRNA).

The role of mRNA is to carry genetic information from DNA to the ribosome, where the creation of proteins takes place. The structure of RNA is similar to that of DNA. However, the difference is that RNA is single stranded and the base pair thymine, found in DNA, is replaced with uracil. Once the mRNA reaches the ribosome, translation starts.

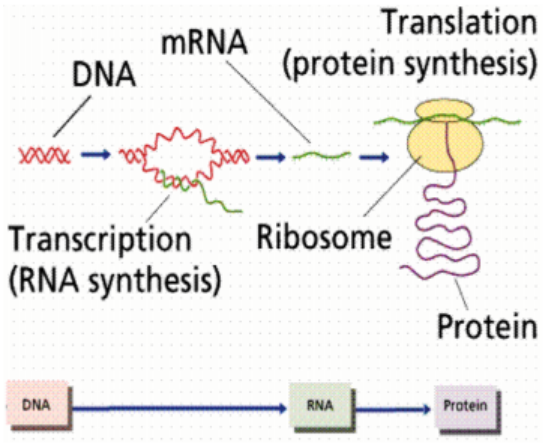

Figure 1: General Bacteria Protein Synthesis [4]

Ribosomes are essentially protein making factories. The mRNA is then decoded by transfer RNA(tRNA). The mRNA is read three base pairs at a time, this is called a codon. Each codon codes for an amino acid. Each amino acid is attached to its own tRNA molecule and all the tRNA deliver their amino acids to the ribosome. Here, the amino acids are

combined into a polypeptide chain that is then released by the ribosome and folded into a protein.[3]

1.2 General Model

Transcription, translation and protein degradation are controlled by the number of mRNA and protein present in the cell. In simple terms, more mRNA causes more translation and more protein causes more degradation. Protein degradation, also known as proteolysis, is a set of processes that result in the destruction of proteins. The rate and method of proteolysis is specific to the protein and the current state of the organism. The main reason for protein degradation is to rid abnormal or damaged proteins as well as proteins that are no longer useful. Gene expression can be represented as a system of first order differential equations where the rate of change of mRNA or protein is equal to some mathematical function representing the cause of the change. The system would look like

$$\frac{d\vec{r}}{dt} = g(\vec{r}, \vec{p}, t)$$

$$\frac{d\vec{p}}{dt} = h(\vec{r}, \vec{p}, t)$$

where $\vec{r}$ is a vector of all the mRNA, $\vec{p}$ is a vector of all the protein, and t is time. [4]

2 Past Models of Gene Expression

2.1 Chen's Linear Transcription Model

Chen, He, and Church proposed a model for gene expression in 1999[8].The model is a system of linear differential equations that incorporates a protein feedback loop to transcription. A protein feedback loop to transcription means that based on the concentration of the type of protein that has been created in the cell and the needs of the cell based on environmental stimuli, feedback is sent back to communicate to the cell to continue making the mRNA that creates that protein at a faster or slower rate. In this model a feedback loop to translation is ignored because it is assumed the transcription to translation process is relatively stable, meaning there is no need to tell the the cell to create more/less of the protein when it was already told to make more/less of the mRNA.

The system of equations is as follows:

$$\begin{aligned} \frac{d\vec{r}}{dt} &= f(\vec{p}) - V\vec{r} \\ \frac{d\vec{p}}{dt} &= L\vec{r} - U\vec{p} \end{aligned}$$

All variables are functions of time (t) and are defined as follows:

- n number of genes in the genome
- $\vec{r}$ mRNA concentrations, n-dimensional vector-valued functions of t

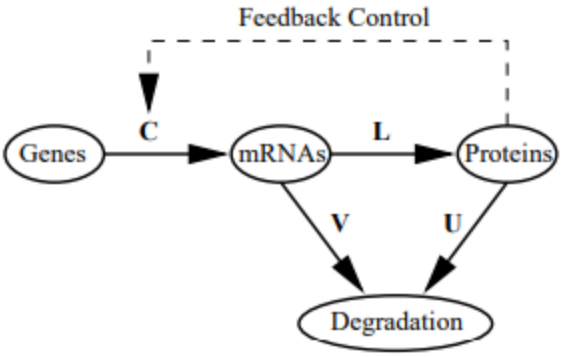

Figure 2: Chen's Linear Transcription Model for Gene Expression [8]

- $\vec{p}$ protein concentrations, n-dimensional vector-valued functions of t
- $f(\vec{p})$ transcription functions, n-dimensional vector polynomials on $\vec{p}$
- L translational constants, $n \times n$ non-degenerate diagonal matrix
- V degradation rates of mRNA, $n \times n$ non-degenerate diagonal matrix
- U degradation rates of proteins, $n \times n$ non-degenerate diagonal matrix

$$\begin{bmatrix} \frac{d\vec{r_1}}{dt} \\ \frac{d\vec{r_2}}{dt} \\ \vdots \\ \frac{d\vec{r_n}}{dt} \end{bmatrix} = \begin{bmatrix} f_1(\vec{p}) \\ f_2(\vec{p}) \\ \vdots \\ f_n(\vec{p}) \end{bmatrix} - \begin{pmatrix} V_{1,1} & 0 & 0 & 0 \\ 0 & V_{2,2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & V_{n,n} \end{pmatrix} \times \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_n(t) \end{bmatrix}$$

$$\begin{bmatrix} \frac{d\vec{p_1}}{dt} \\ \frac{d\vec{p_2}}{dt} \\ \vdots \\ \frac{d\vec{p_n}}{dt} \end{bmatrix} = \begin{pmatrix} L_{1,1} & 0 & 0 & 0 \\ 0 & L_{2,2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & L_{n,n} \end{pmatrix} \times \begin{bmatrix} r_1(t) \\ r_2(t) \\ \vdots \\ r_n(t) \end{bmatrix} - \begin{pmatrix} U_{1,1} & 0 & 0 & 0 \\ 0 & U_{2,2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & U_{n,n} \end{pmatrix} \times \begin{bmatrix} p_1(t) \\ p_2(t) \\ \vdots \\ p_n(t) \end{bmatrix}$$

Both $\vec{r}$ and $\vec{p}$ are vectors of dimension n because each gene has different mRNA and protein concentrations at every time t. The function $f(\vec{p})$ is n-dimensional because each gene has different transcription functions and the polynomials represent the combination of proteins present at time t, which all affect translation. Different rates of transcription allows for exponents. L, V and U are diagonal matrices because Chen assumes in his model that translation and degradation rates are constant. Each protein has different translation and degradation rates. Chen also chooses to ignore feedback from mRNA to genes and proteins to mRNA because it is all wrapped up in the protein to genes feedback loop. The system as it is is a nonlinear dynamic system.

2.1.1 Transformation to Linear Model

We can transform the system into a linear one by assuming $f(\vec{p})$ to be linear functions of $\vec{p}$ rather than polynomials. That gives us

$$f(\vec{p}) = C\vec{p}$$

We can justify this equation from a biological perspective by defining $f(\vec{p})$ as the combined effect of activators and inhibitors in transcription. This can be described as a linear function that is the contribution of all the inhibitors subtracted from the contribution of all the activators.

If the biological explanation is not convincing enough, we can justify the equation mathematically as well. Let $\vec{p_0}$ be the value of $\vec{p}$ at $t = 0$ and take the first order Taylor approximation. The definition of the first order Taylor approximation is $f(x) \approx f(a) + f'(a)(x - a)$. Apply this to our equation and we have

$$f(\vec{p}) = f(\vec{p_0}) + \frac{df(\vec{p})}{d\vec{p}}|_{\vec{p_0}}(\vec{p} - \vec{p_0})$$

This can be written as $f(\vec{p}) = C\vec{p} + s$ where $C = \frac{df(\vec{p})}{d\vec{p}}|_{\vec{p_0}}$ and $s = f(\vec{p_0}) - \frac{df(\vec{p})}{d\vec{p}}|_{\vec{p_0}}\vec{p_0}$. Now we can plug this back into our original equations to get a linear system.

$$\begin{aligned}\frac{d\vec{r}}{dt} &= C\vec{p} - V\vec{r} + s \\ \frac{d\vec{p}}{dt} &= \dot{L}\vec{r} - U\vec{p}\end{aligned}$$

We want to be able to eliminate s. In order to do this we set $\vec{r} = \vec{r} + \vec{r_s}$ and $\vec{p} = \vec{p} + \vec{p_s}$ to find $\vec{r_s}$ and $\vec{p_s}$ that will eliminate s. When we plug these back into the equations we get

$$\begin{aligned}\frac{d\vec{r}}{dt} &= C\vec{p} - V\vec{r} + (C\vec{p_s} - V\vec{r_s}) + s \\ \frac{d\vec{p}}{dt} &= L\vec{r} - U\vec{p} + (L\vec{r_s} - U\vec{p_s})\end{aligned}$$

where $\vec{r_s}$ and $\vec{p_s}$ can be solved by

$$\begin{pmatrix} -V & C \\ L & -U \end{pmatrix} \times \begin{pmatrix} \vec{r_s} \\ \vec{p_s} \end{pmatrix} = \begin{pmatrix} -s \\ 0 \end{pmatrix}$$

Since both V and U are both nonsingular diagonal matrices, we can assume the equation has a unique solution, which would allow us to eliminate s. Therefore we can consider the following system of equations.

$$\begin{aligned}\frac{d\vec{r}}{dt} &= C\vec{p} - V\vec{r} \\ \frac{d\vec{p}}{dt} &= L\vec{r} - U\vec{p}\end{aligned}$$

From this, the Linear Transcription Model can be defined as follows:

Let $x = (\vec{r}, \vec{p})^T$ be variables for mRNAs and proteins and M be a $2n \times 2n$ transition matrix.

$$\frac{dx}{dt} = Mx \qquad where \qquad M = \begin{pmatrix} -V & C \\ L & -U \end{pmatrix}$$

This equation looks like:

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \vdots \\ \frac{dx_n}{dt} \\ \vdots \\ \frac{dx_{2n}}{dt} \end{bmatrix} = \begin{pmatrix} -V_{1,1} & 0 & 0 & C_{1,1} & \dots & C_{1,n} \\ 0 & \ddots & 0 & \vdots & \ddots & \vdots \\ 0 & 0 & -V_{n,n} & C_{n,1} & \dots & C_{n,n} \\ L_{1,1} & 0 & 0 & -U_{1,1} & 0 & 0 \\ 0 & \ddots & 0 & 0 & \ddots & 0 \\ 0 & 0 & L_{n,n} & 0 & 0 & -U_{n,n} \end{pmatrix} \times \begin{bmatrix} r_1 \\ \vdots \\ r_n \\ p_1 \\ \vdots \\ p_n \end{bmatrix}$$

$\frac{dx}{dt}$ is the rate of change for the entire system. This includes the creation, degradation, and feedback of all mRNAs and proteins.

2.1.2 Solution to Linear Model

Assume M has $2n$ eigenvalues $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_{2n})^T$. The solution of the model is of the form

$$x(t) = Q(t)e^{\lambda t}$$

where $Q(t) = \{q_{ij}(t)\}$ satisfies

$$\sum_{j=1}^{2n} = deg(q_{ij}(t)) + 1 \leq 2n \quad for\ i = 1, 2, \dots, 2n$$

$Q(t)$ is a $2n \times 2n$ matrix whose elements are polynomial functions of t and $deg()$ returns the degree of the polynomial function. So the equation looks like:

$$x(t) = \begin{pmatrix} Q_{1,1}(t) & \dots & Q_{1,2n}(t) \\ \vdots & \ddots & \vdots \\ Q_{2n,1}(t) & \dots & Q_{2n,2n}(t) \end{pmatrix} e^{\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{2n} \end{bmatrix} t}$$

2.1.3 Repeated Eigenvalues

$Q(t)$ is a polynomial function of t rather than a constant to allow for the possibility of repeated eigenvalues. If m is the algebraic multiplicity of a given eigenvalue, it is possible there could be fewer than m independent eigenvectors associated with the eigenvalue. This means there will be fewer than m solutions of the form $ce^{\lambda t}$, where c is a constant. Therefore it is necessary to find solutions of a different form [1].

Suppose that A has an eigenvalue λ of multiplicity m. We have to find vector such that

$$(A-\lambda I)^k\vec{v}=\vec{0} \qquad but \qquad (A-\lambda I)^{k-1}\vec{v}\neq\vec{0}$$

These are called generalized eigenvectors. For the eigenvector $\vec{v_1}$ there is a chain of generalized eigenvectors $\vec{v_2}$ through $\vec{v_k}$ such that

$$\begin{aligned}(A-\lambda I)\vec{v_1} &= \vec{0}\\ (A-\lambda I)\vec{v_2} &= \vec{v_1}\\ &\vdots\\ (A-\lambda I)\vec{v_k} &= \vec{v_{k-1}}\end{aligned}$$

From here we can form linearly independent solutions such that

$$\begin{aligned}\vec{x_1} &= \vec{v_1}e^{\lambda t}\\ \vec{x_2} &= (\vec{v_2}+\vec{v_1}t)e^{\lambda t}\\ &\vdots\\ \vec{x_k} &= (\vec{v_k}+\vec{v_{k-1}}t+\vec{v_{k-2}}\frac{t^2}{2}+\cdots+\vec{v_2}\frac{t^{k-2}}{(k-2)!}+\vec{v_1}\frac{t^{k-1}}{(k-1)!})e^{\lambda t}\end{aligned}$$

[7]

Apply this to our model and we can see why $Q(t)$ could be a polynomial function of t.

2.1.4 System Analysis

The model should obey the rules of biology. The system can either be unstable, semistable, or stable. A positive value of λ would make the term $q_{ij}e^{\lambda_j t}$ an exponential function, therefore making the system unstable. The system is semistable if all the eigenvalues of λ are non-positive. The semistable system has a polynomial growth rate because of the polynomial term $q_{ij}(t)$. The system will be stable if all the eigenvalues of λ are non-positive and all the polynomials $q_{ij}(t)$ are constant. In order for the system to model real-life biology, the system must be a stable one because an exponential or polynomial growth rate of a gene or protein is unlikely to happen.

This means rather than having $q_{ij}(t)$, we can have q_{ij}. We are left with the solution $x(t)=Qe^{\lambda t}$ where Q is a $2n\times 2n$ constant matrix and $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_{2n})^T$. [8]

There are limitations to the Linear Transcription Model. With the biological system being so complex, there are many different interactions involved in the transcription process and it is very difficult to account for them all. Thus, some are ignored in Chen's model. Chen's model also does not account for time delay, however in prokaryotes we can ignore this issue in order to simplify the problem. Chen's model is a good start in attempting to model gene expression.

2.2 Kim and Tidor Limited NonLinear System

In 2003, Kim and Tidor proposed a nonlinear model for gene expression. They begin their model with the assumptions that there is no spatial dependence of concentrations or rate constants, no cross-talk between promoters, control of gene expression is at the transcription level only, and dependence of translation and transcription rates on protein and mRNA concentration is strictly monotonic. They created a general model that is

$$\begin{aligned} r_i &= deg_{r_i}(r_i) + tr_{y_i}(p_{y_i}) \\ p_i &= deg_{p_i}(p_i) + tl_{p_i}(r_i) \end{aligned}$$

The variables are as follows:

- r_i RNA concentration
- p_i protein concentration
- $deg_{r_i}(r_i)$ RNA degradation rate
- $tr_{y_i}(p_{y_i})$ transcription rate of RNA as a function of the repressor concentration p_{y_i} that controls RNA r_i expression
- $deg_{p_i}(p_i)$ protein degradation rate
- $tl_{p_i}(r_i)$ translation rate of RNA into protein

The degradation rates, $deg_{r_i}(r_i)$ and $deg_{p_i}(p_i)$, will always be negative because they correspond to the reduction in RNA and protein concentrations through the process of degradation. They also assume that $tr(p)$ is strictly monotonically decreasing for every repressor p and that $tr(r)$ is strictly monotonically increasing. The degradation rates are also strictly monotonically decreasing. y_i determines which transcription factor represses which gene. These assumptions make it easier to solve for steady states [5].

2.2.1 Simplification

We assume a steady state. This means that the rates of creation are equal to the rates of degradation.

$$\begin{aligned} 0 &= deg_{r_i}(r_i) + tr_{y_i}(p_{y_i}) \\ 0 &= deg_{p_i}(p_i) + tl_{p_i}(r_i) \end{aligned}$$

Due to the assumption that degradation rates are strictly monotonic with respect to concentration, we can invert the first equation and then eliminate r_i from the second equation.

$$\begin{aligned} r_i &= deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i})) \\ 0 &= deg_{p_i}(p_i) + tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i}))) \\ p_i &= deg_{p_i}^{-1}(-tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i})))) \end{aligned}$$

$deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i}))$ is monotonically decreasing because as protein concentration rises, transcription rate will increase, therefore making the inverse smaller. $-tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i})))$ is monotonically increasing because if $deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i}))$ is decreasing then that will cause $tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i})))$ to also be decreasing, so $-tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i})))$ will be increasing. This means that $deg_{p_i}^{-1}(-tl_{p_i}(deg_{r_i}^{-1}(-tr_{y_i}(p_{y_i}))))$ is monotonically decreasing. Therefore we can replace the right side in the last equation with some function that is strictly monotonically decreasing, which gives us

$$p_i = f_{iy_i}(p_{y_i})$$

This equation means that the steady-state level of any given protein has monotonically decreasing dependence on the concentration of the repressor controlling its expression [5].

Kim and Tidor's work is useful because they created a general model that can be used as a starting point for future work on models for gene expression. However they do not provide explicit functions.

3 Hunt et al. Proposed Model

In establishing the groundwork for their model, Hunt et al. pulled from the both the Chen and the Kim model. They took their model further by incorporating multiple feedback loops throughout the protein synthesis process, thus creating a nonlinear system. They continue to ignore time delay in their model because of the bacteria cell's compartmentalized DNA, as previously described. The model represents some time interval between formation and cytokinesis, which is the division of the cell into two daughter cells with identical DNA. The volume of the cell is considered to be constant during this time.

The model regards tRNA and rRNA to be excess in the cell, meaning that protein synthesis is only dependent on the concentration of mRNA, similar to the other models. The mRNA and proteins are grouped into three different types: type 1, type 2, and type 3. Type 1 mRNA produce type 1 proteins, who initiate and preform the transcription of all RNA. Type 2 mRNA create type 2 proteins which help to stabilize all RNA concentrations against degradation. Type 3 mRNA construct type 3 proteins which are not directly involved in protein synthesis. They are sent out to the cell for other purposes. They assume that each type of protein regulates the production of it's own protein type during transcription and translation [4].

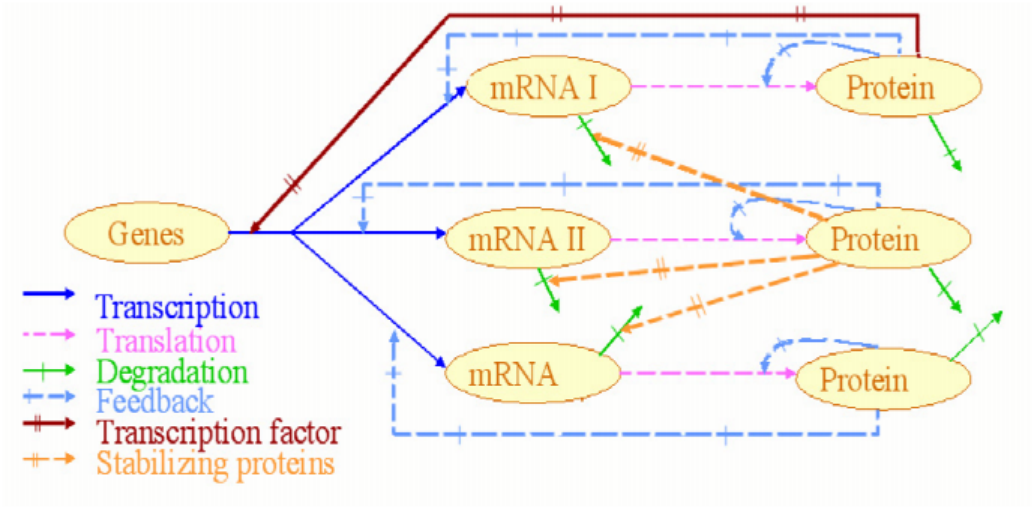

Figure 3: The figure shows the gene expression model. Group mRNA I, mRNA II and mRNA III produce type Protein I, Protein II and Protein III, respectively [4].

The transcription of any Type i mRNA can be expressed as a function of Type 1 proteins, since these affect the transcription of all mRNA, and Type i proteins, because of the feedback loop to transcription. The transcription function is a strictly increasing function of Type 1 proteins and a strictly decreasing function of Type i proteins. The reason for this is the more Type 1 proteins present, the more Type i mRNA produced. However, the more Type i protein present, the less transcription of Type i mRNA in order to avoid the creation of excess of one type of protein in the cell. The dependence of Type 1 proteins is assumed to be linear and somewhat logistic on Type i proteins.

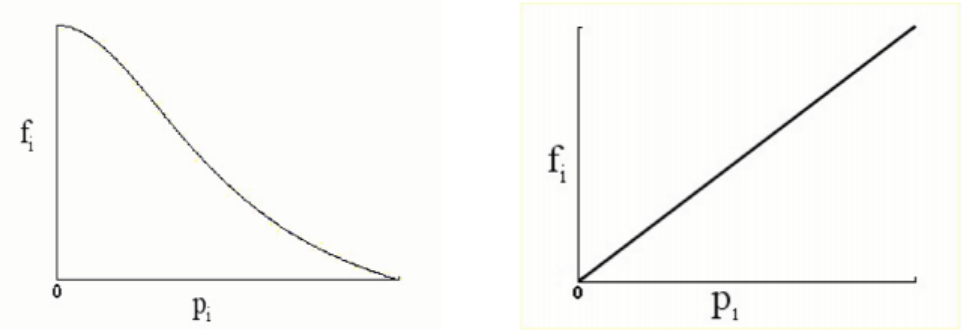

Figure 4: Left: The transcription function m strictly monotonically decreases as p_i increases. Right: The transcription function increases linearly with respect to p_1 [4].

As for translation, it is an increasing function of Type i mRNA and a decreasing function of Type i proteins. The presence of more Type i mRNA will increase the production of Type i proteins, however the presence of more Type i proteins will decrease the production of more Type i proteins, again in order to avoid an excess. Assume a linear dependence upon Type i mRNA.

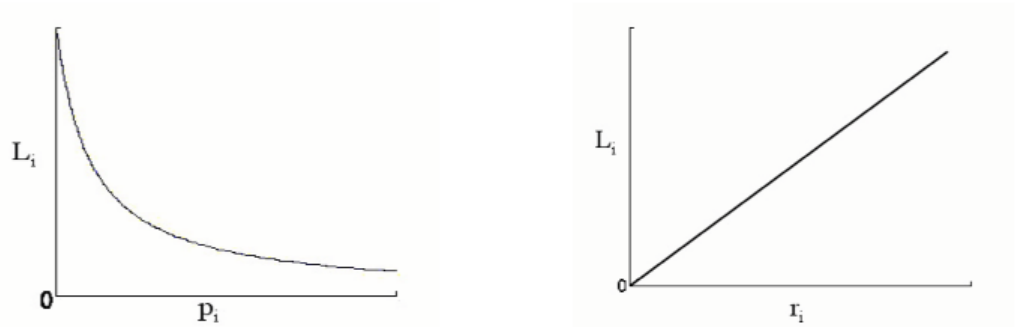

Figure 5: Left: The translation function strictly monotonically decreases as p_i increases. Right: The translation function of mRNA increases linearly with respect to r_i [4].

Degradation of Type i mRNA is an increasing function of the concentration of Type i mRNA present, but a decreasing function of Type 2 proteins, the stabilizers. The more mRNA there is, the faster it will degrade. However, with more stabilizing proteins present, the degradation will be slower. Again assume the degradation rate has a linear dependence on Type i mRNA.

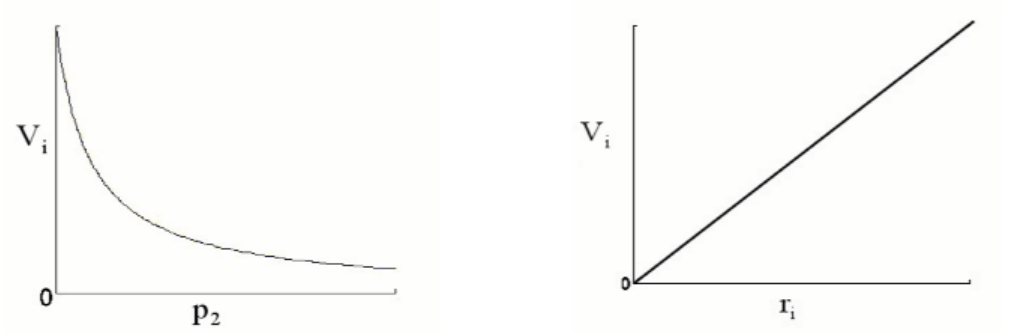

Figure 6: Left: The degradation function of mRNA strictly monotonically decreases as p_2 increases. Right: The degradation function of mRNA increases linearly with respect to p_i [4].

From these conclusions, the rates of change of mRNA and protein concentrations can be modeled by these general equations:

$$\begin{aligned}
\frac{dr_1}{dt} &= \frac{1}{1+\frac{p_1^2}{a_1^2}}C_1p_1 - \frac{1}{1+\frac{p_2}{b_1}}V_1r_1 & \frac{dp_1}{dt} &= \frac{1}{1+\frac{p_1}{d_1}}L_1r_1 - U_1p_1 \\
\frac{dr_2}{dt} &= \frac{1}{1+\frac{p_2}{a_2}}C_2p_1 - \frac{1}{1+\frac{p_2}{b_2}}V_2r_2 & \frac{dp_2}{dt} &= \frac{1}{1+\frac{p_2}{d_2}}L_2r_2 - U_2p_2 \\
\frac{dr_3}{dt} &= \frac{1}{1+\frac{p_3}{a_3}}C_3p_1 - \frac{1}{1+\frac{p_2}{b_3}}V_3r_3 & \frac{dp_3}{dt} &= \frac{1}{1+\frac{p_3}{d_3}}L_3r_3 - U_3p_3
\end{aligned} \tag{1}$$

r_i and p_i are the concentration of Type i mRNA and protein in unit of nMolar. C_i and L_i are the relative transcription and translation rates in the absence of feedback loops. U_i

and V_i are the relative natural degradation rates of mRNA and protein. a_i, b_i, and d_i are the effectiveness factors of the respective feedback loops with the same unit as p_i. The bigger the value of a_i, the smaller the effect of the feedback into the transcription term. b_i and d_i control the effectiveness of the feedback into the degradation of mRNA and translation of protein, respectively [4].

4 Analysis of the System

4.1 Stability at the Origin

It can be shown that the system has an equilibrium point at the origin. $\frac{dp_i}{dt} = 0$ and $\frac{dr_i}{dt} = 0$ for $i = 1, 2, 3$ at $(r_1, p_1, r_2, p_2, r_3, p_3) = (0, 0, 0, 0, 0, 0)$. The Jacobian matrix of the system is as follows:

$$\begin{bmatrix}
\frac{\partial f_1}{\partial r_1} = \frac{-1}{1+\frac{p_2}{b_1}} V_1 & \frac{\partial f_1}{\partial p_1} = \frac{-2p_1}{(1+\frac{p_1^2}{a_1^2})^2}(C_1 p_1) + \frac{C_1}{1+\frac{p_1^2}{a_1^2}} & \frac{\partial f_1}{\partial r_2} = 0 & \frac{\partial f_1}{\partial p_2} = \frac{\frac{-1}{b_1}}{(1+\frac{p_2}{b_2})^2}(V_1 r_1) & \frac{\partial f_1}{\partial r_3} = 0 & \frac{\partial f_1}{\partial p_3} = 0 \\
\frac{\partial f_2}{\partial r_1} = \frac{1}{1+\frac{p_1}{d_1}} L_1 & \frac{\partial f_2}{\partial p_1} = \frac{\frac{-1}{d_1}}{(1+\frac{p_1}{d_1})^2}(L_1 r_1) - U_1 & \frac{\partial f_2}{\partial r_2} = 0 & \frac{\partial f_2}{\partial p_2} = 0 & \frac{\partial f_2}{\partial r_3} = 0 & \frac{\partial f_2}{\partial p_3} = 0 \\
\frac{\partial f_3}{\partial r_1} = 0 & \frac{\partial f_3}{\partial p_1} = \frac{1}{1+\frac{p_2}{a_2}} C_2 & \frac{\partial f_3}{\partial r_2} = \frac{-1}{1+\frac{p_2}{b_2}} V_2 & \frac{\partial f_3}{\partial p_2} = \frac{\frac{-1}{a_2}}{(1+\frac{p_2}{a_2})^2} C_2 p_1 + \frac{\frac{1}{b_2}}{(1+\frac{p_2}{b_2})^2} V_2 r_2 & \frac{\partial f_3}{\partial r_3} = 0 & \frac{\partial f_3}{\partial p_3} = 0 \\
\frac{\partial f_4}{\partial r_1} = 0 & \frac{\partial f_4}{\partial p_1} = 0 & \frac{\partial f_4}{\partial r_2} = \frac{1}{1+\frac{p_2}{d_2}} L_2 & \frac{\partial f_4}{\partial p_2} = \frac{\frac{-1}{d_2}}{(1+\frac{p_2}{d_2})^2} L_2 r_2 - U_2 & \frac{\partial f_4}{\partial r_3} = 0 & \frac{\partial f_4}{\partial p_3} = 0 \\
\frac{\partial f_5}{\partial r_1} = 0 & \frac{\partial f_5}{\partial p_1} = \frac{1}{1+\frac{p_3}{a_3}} C_3 & \frac{\partial f_5}{\partial r_2} = 0 & \frac{\partial f_5}{\partial p_2} = \frac{\frac{1}{b_3}}{(1+\frac{p_2}{b_3})^2} V_3 r_3 & \frac{\partial f_5}{\partial r_3} = \frac{-1}{1+\frac{p_2}{b_3}} V_3 & \frac{\partial f_5}{\partial p_3} = \frac{\frac{-1}{a_3}}{(1+\frac{p_3}{a_3})^2} C_3 p_1 \\
\frac{\partial f_6}{\partial r_1} = 0 & \frac{\partial f_6}{\partial p_1} = 0 & \frac{\partial f_6}{\partial r_2} = 0 & \frac{\partial f_6}{\partial p_2} = 0 & \frac{\partial f_6}{\partial r_3} = \frac{1}{1+\frac{p_3}{d_3}} L_3 & \frac{\partial f_6}{\partial p_3} = \frac{\frac{-1}{d_3}}{(1+\frac{p_3}{d_3})^2} L_3 r_3 - U_3
\end{bmatrix}$$

To get the Jacobian matrix at the origin, the equilibrium point, we plug in 0 for r_1, p_1, r_2, p_2, r_3 and p_3. This gives us

$$J(0,0,0,0,0,0) = \begin{bmatrix}
-V_1 & C_1 & 0 & 0 & 0 & 0 \\
L_1 & -U_1 & 0 & 0 & 0 & 0 \\
0 & C_2 & -V_2 & 0 & 0 & 0 \\
0 & 0 & L_2 & U_2 & 0 & 0 \\
0 & C_3 & 0 & 0 & -V_3 & 0 \\
0 & 0 & 0 & 0 & L_3 & -U_3
\end{bmatrix}$$

To find the eigenvalues we take the determinate of the matrix that is the difference of λ times the identity matrix and $J(0,0,0,0,0,0)$. We set that determinate equal to 0.

$$det \begin{bmatrix}
\lambda + V_1 & -C_1 & 0 & 0 & 0 & 0 \\
-L_1 & \lambda + U_1 & 0 & 0 & 0 & 0 \\
0 & -C_2 & \lambda + V_2 & 0 & 0 & 0 \\
0 & 0 & -L_2 & \lambda + U_2 & 0 & 0 \\
0 & -C_3 & 0 & 0 & \lambda + V_3 & 0 \\
0 & 0 & 0 & 0 & -L_3 & \lambda + U_3
\end{bmatrix} = 0$$

The characteristic equation is

$$(U_2 + \lambda)(U_3 + \lambda)(V_2 + \lambda)(V_3 + \lambda)(-C_1 L_1 + U_1(V_1 + \lambda) + V_1\lambda + \lambda^2) = 0$$

. It is clear that $-U_2$, $-U_3$, $-V_2$ and $-V_3$ are eigenvalues. To find the last two eigenvalues we solve $-C_1L_1 + U_1(V_1 + \lambda) + V_1\lambda + \lambda^2 = 0$.

$$\lambda^2 + (U_1 + V_1)\lambda + (-C_1L_1 + U_1V_1) = 0$$

Apply the quadratic formula

$$\begin{aligned}
\frac{-(U_1 + V_1) \pm \sqrt{(U_1 + V_1)^2 - 4(-C_1L_1 + U_1V_1)}}{2} &= \lambda \\
-\frac{1}{2}(U_1 + V_1 \pm \sqrt{4C_1L_1 - 4U_1V_1 + (U_1 + V_1)^2}) &= \lambda \\
-\frac{1}{2}(U_1 + V_1 \pm \sqrt{4C_1L_1 - 4U_1V_1 + U_1^2 + 2U_1V_1 + V_1^2}) &= \lambda \\
-\frac{1}{2}(U_1 + V_1 \pm \sqrt{4C_1L_1 + (U_1 - V_1)^2}) &= \lambda
\end{aligned}$$

So we have the six following eigenvalues:

$$\lambda_1 = -U_2, \qquad \lambda_2 = -U_3, \qquad \lambda_3 = -\frac{1}{2}(U_1 + V_1 + \sqrt{(U_1 - V_1)^2 + 4C_1L_1}),$$
$$\lambda_4 = -V_2, \qquad \lambda_5 = -V_3, \qquad \lambda_6 = -\frac{1}{2}(U_1 + V_1 - \sqrt{(U_1 - V_1)^2 + 4C_1L_1})$$

Eigenvalues $\lambda_1, \ldots, \lambda_5$ will always be negative because C_1, V_1, L_1, and U_1 are positive. As for λ_6, it will be negative when $\alpha = C_1L_1 - U_1V_1$ is negative. This is the product of the production terms minus the product of the degradation terms of Type 1 mRNA and protein. The stability of the system is dependent on only Type 1 products because the transcription of all types of mRNA depends on Type 1 proteins. α will be negative when the rate of degradation is faster than the rate of production. If the products are decaying faster than they are being produced, everything will eventually die out. Thus, the concentrations of all mRNA and protein will approach 0 as time goes on. This agrees with our mathematical analysis because when all eigenvalues are negative, the origin will be a stable equilibrium node.

When $\alpha > 0$, λ_6 will be positive, resulting in a saddle at the origin. A saddle node bifurcation means that almost all initial concentrations of mRNA and proteins near the origin will move away from it. The system will only approach the origin in certain conditions. When $r_1 = 0$ and $p_1 = 0$, no mRNA will ever be produced and the products will die out, meaning the system will approach 0. A saddle occurs when there is one eigenvalue that is 0 [6]. In our case, $\lambda_6 = 0$ when $\alpha = 0$. We can conclude that no limit cycles exist around the origin at any time because none of the eigenvalues can ever be purely imaginary.

4.2 Simplification of the System

We can see that the first four equations of the system are independent of the variables r_3 and p_3, meaning they can be solved independently of the other two. This means the system can be simplified to those four equations. As we recall, r_3 and p_3 are the concentrations of type 3 mRNA and protein. Type 3 mRNA and proteins are not directly involved in the process

of protein synthesis. This is further justification for disregarding the last two equations in order to simplify our system. Any equilibrium point of the six-dimensional system will also be one of the four-dimensional system. With that, we can analyze the first four equations and make inferences about the behavior of the full system from there.

$$\frac{dr_1}{dt} = \frac{1}{1+\frac{p_1^2}{a_1^2}}C_1p_1 - \frac{1}{1+\frac{p_2}{b_1}}V_1r_1 \qquad \frac{dp_1}{dt} = \frac{1}{1+\frac{p_1}{d_1}}L_1r_1 - U_1p_1 \tag{2}$$

$$\frac{dr_2}{dt} = \frac{1}{1+\frac{p_2}{a_2}}C_2p_1 - \frac{1}{1+\frac{p_2}{b_2}}V_2r_2 \qquad \frac{dp_2}{dt} = \frac{1}{1+\frac{p_2}{d_2}}L_2r_2 - U_2p_2$$

Another beneficial simplification of the system is to obtain values for the constants in the system in order to limit the number of parameters. It is difficult to get exact values for the rates of production and decay of mRNA and protein, however estimates can be made. Hunt et al. suggest these values based on biological literature: $C_i \approx 0.03$ mRNA/(protein min), $L_i \approx 2$ protein/(mRNA min), $U_1 \approx 0.15$ min, $U_i \approx 0.015$ min, for $i = 2, 3$ and $V_i \approx 0.03$ min [4]. Note that these values are for the case when $\alpha > 0$.

If we plug in the constants the system looks like,

$$\frac{dr_1}{dt} = \frac{1}{1+\frac{p_1^2}{a_1^2}}0.03p_1 - \frac{1}{1+\frac{p_2}{b_1}}0.03r_1 \qquad \frac{dp_1}{dt} = \frac{1}{1+\frac{p_1}{d_1}}2r_1 - 0.15p_1$$

$$\frac{dr_2}{dt} = \frac{1}{1+\frac{p_2}{a_2}}0.03p_1 - \frac{1}{1+\frac{p_2}{b_2}}0.03r_2 \qquad \frac{dp_2}{dt} = \frac{1}{1+\frac{p_2}{d_2}}2r_2 - 0.015p_2$$

4.3 Extreme Cases

Even the simplified four-dimensional system is difficult to analyze due to it's fourteen parameters and several nonlinearities. In order to simplify the system so we can analyze other potential equilibrium solutions, we can look at several special cases in which certain parameters and terms can be ignored.

Case 1. $b_i \gg p_2$ and $d_i \gg p_i$

This case assumes that the stabilization of mRNA and the feedback from proteins to translation are negligible in comparison to the other interactions of the system. The system simplifies to

$$(1.1)\quad \frac{dr_1}{dt} = \frac{1}{1+\frac{p_1^2}{a_1^2}}C_1p_1 - V_1r_1 \qquad (2.1)\quad \frac{dp_1}{dt} = L_1r_1 - U_1p_1 \tag{3}$$

$$(3.1)\quad \frac{dr_2}{dt} = \frac{1}{1+\frac{p_2}{a_2}}C_2p_1 - V_2r_2 \qquad (4.1)\quad \frac{dp_2}{dt} = L_2r_2 - U_2p_2$$

The system has a trivial equilibrium point at the origin. In order to find the other equilibrium points, we set the equations equal to 0 and solve for the concentration variables.

We can first solve for r_1 and p_1 just by looking at equations (1.1) and (2.1). First use equation (2.1) to get r_1 in terms of p_1.

$$\begin{aligned} 0 &= L_1 r_1 - U_1 p_1 \\ r_1 &= \frac{U_1}{L_1} p_1 \end{aligned}$$

Then plug our solution for r_1 into equation (1.1) to solve for p_1.

$$\begin{aligned} 0 &= \frac{C_1}{1 + \frac{p_1^2}{a_1^2}} p_1 - \frac{V_1 U_1}{L_1} p_1 \\ \frac{C_1}{1 + \frac{p_1^2}{a_1^2}} &= \frac{V_1 U_1}{L_1} \\ 1 + \frac{p_1^2}{a_1^2} &= \frac{C_1 L_1}{V_1 U_1} \\ p_1 &= \pm \sqrt{\frac{C_1 L_1 a_1^2}{V_1 U_1} - a_1^2} \end{aligned}$$

Plug the solution for p_1 into the previous equation for r_1 in order to get the solution for r_1.

$$r_1 = \pm \frac{U_1}{L_1} \sqrt{\frac{C_1 L_1 a_1^2}{V_1 U_1} - a_1^2}$$

We can simplify these equations to make our analysis easier, which gives

$$p_1 = \pm \frac{a_1}{\sqrt{V_1 U_1}} \sqrt{\alpha} \qquad , \qquad r_1 = \pm \frac{a_1 U_1}{L_1 \sqrt{V_1 U_1}} \sqrt{\alpha}$$

where $\alpha = C_1 L_1 - V_1 U_1$. If $\alpha > 0$, then $p_1, r_1 \in \mathbb{R}$. If $\alpha < 0$, then p_1, r_1 are purely imaginary. If $\alpha = 0$, then $p_1, r_1 = 0$ which is the trivial solution. The only case that is of interest is when $\alpha > 0$ and p_1, r_1 are in the real number space.

To solve for p_2 and r_2, first use equation (4.1) to get r_2 in terms of p_2.

$$\begin{aligned} 0 &= L_2 r_2 - U_2 p_2 \\ r_2 &= \frac{U_2}{L_2} p_2 \end{aligned}$$

Now plug solutions for p_1, r_2 into equation (3.1).

$$0 = \frac{1}{1+\frac{p_2}{a_2}} C_2 \left(\pm \sqrt{\frac{C_1 L_1 a_1^2}{V_1 U_1} - a_1^2} \right) - \frac{V_2 U_2}{L_2} p_2$$
$$0 = \frac{V_2 U_2}{L_2} p_2 \left(1 + \frac{p_2}{a_2} \right) \pm C_2 \sqrt{\frac{C_1 L_1 a_1^2}{V_1 U_1} - a_1^2}$$
$$0 = \frac{V_2 U_2}{L_2 a_2} p_2^2 + \frac{V_2 U_2}{L_2} p_2 \pm C_2 a_1 \sqrt{\frac{C_1 L_1}{V_1 U_1} - 1}$$

This equation can simplify to

$$0 = p_2^2 + a_2 p_2 \pm \frac{a_1 a_2 C_2 L_2}{V_2 U_2 \sqrt{V_1 U_1}} \sqrt{\alpha}$$
$$p_2 = \frac{-a_2}{2} \pm \frac{1}{2} \sqrt{a_2^2 \pm \frac{4 a_1 a_2 C_2 L_2}{V_2 U_2 \sqrt{V_1 U_1}} \sqrt{\alpha}}$$

Let $\beta = \frac{4 a_1 C_2 L_2}{V_2 U_2 \sqrt{V_1 U_1}}$ so, $p_2 = -\frac{a_2}{2} \pm \frac{1}{2} \sqrt{a_2 (a_2 \pm \beta \sqrt{\alpha})}$. Return to the equation for r_2 and substitute p_2 to get $r_2 = -\frac{U_2 a_2}{2 L_2} \pm \frac{U_2}{2 L_2} \sqrt{a_2 (a_2 \pm \beta \sqrt{\alpha})}$.

To summarize, the possible equilibrium solutions of the system are

$$r_1 = \pm \frac{a_1 U_1}{L_1 \sqrt{V_1 U_1}} \sqrt{\alpha} \quad , \quad p_1 = \pm \frac{a_1}{\sqrt{V_1 U_1}} \sqrt{\alpha}$$
$$r_2 = -\frac{U_2 a_2}{2 L_2} \pm \frac{U_2}{2 L_2} \sqrt{a_2 (a_2 \pm \beta \sqrt{\alpha})} \quad , \quad p_2 = -\frac{a_2}{2} \pm \frac{1}{2} \sqrt{a_2 (a_2 \pm \beta \sqrt{\alpha})}$$

where $\alpha = C_1 L_1 - V_1 U_1$ and $\beta = \frac{4 a_1 C_2 L_2}{V_2 U_2 \sqrt{V_1 U_1}}$. When $\alpha < 0$, r_1, r_2, p_1, p_2 are purely imaginary. When $\alpha = 0$, $r_1, r_2, p_1, p_2 = 0$, bringing us back to the trivial solution. Therefore, the only cases we care about are when $\alpha > 0$. There are 5 different equilibrium points to consider with different cases. All the cases have the condition $\alpha > 0$.

- 1) If $p_1 = \frac{a_1}{\sqrt{V_1 U_1}} \sqrt{\alpha}$, then $r_1 = \frac{a_1 U_1}{L_1 \sqrt{V_1 U_1}} \sqrt{\alpha}$, $p_2 = \frac{-a_1}{2} \pm \frac{1}{2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})}$, $r_2 = \frac{-a_1 U_2}{2 L_2} \pm \frac{U_2}{2 L_2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})}$ and $a_2 + \beta \sqrt{\alpha} > 0$. This yields these two equilibrium points.

  $$\left(\frac{a_1 U_1}{L_1 \sqrt{V_1 U_1}} \sqrt{\alpha}, \frac{a_1}{\sqrt{V_1 U_1}} \sqrt{\alpha}, \frac{-a_1 U_2}{2 L_2} + \frac{U_2}{2 L_2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})}, \frac{-a_1}{2} + \frac{1}{2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})} \right),$$

  $$\left(\frac{a_1 U_1}{L_1 \sqrt{V_1 U_1}} \sqrt{\alpha}, \frac{a_1}{\sqrt{V_1 U_1}} \sqrt{\alpha}, \frac{-a_1 U_2}{2 L_2} - \frac{U_2}{2 L_2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})}, \frac{-a_1}{2} - \frac{1}{2} \sqrt{a_2 (a_2 + \beta \sqrt{\alpha})} \right)$$

- 2) If $p_1 = \frac{-a_1}{\sqrt{V_1U_1}}\sqrt{\alpha}$, then $r_1 = \frac{-a_1U_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}$, $p_2 = \frac{-a_1}{2} \pm \frac{1}{2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}$, $r_2 = \frac{-a_1U_2}{2L_2} \pm \frac{U_2}{2L_2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}$ and $a_2 - \beta\sqrt{\alpha} > 0$. This yields these two equilibrium points.

$$\left(\frac{-a_1U_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_1}{\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_1U_2}{2L_2} + \frac{U_2}{2L_2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}, \frac{-a_1}{2} + \frac{1}{2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}\right),$$

$$\left(\frac{-a_1U_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_1}{\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_1U_2}{2L_2} - \frac{U_2}{2L_2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}, \frac{-a_1}{2} - \frac{1}{2}\sqrt{a_2(a_2-\beta\sqrt{\alpha})}\right)$$

- 3) If $a_2 - \beta\sqrt{\alpha} = 0$, then $p_1 = \frac{-a}{\sqrt{V_1U_1}}\sqrt{\alpha}$, $r_1 = \frac{-aU_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}$, $p_2 = \frac{-a_2}{2}$ and $r_2 = \frac{-a_2U_2}{2L_2}$. This yields this equilibrium point.

$$(\frac{-aU_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a}{\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_2U_2}{2L_2}, \frac{-a_2}{2})$$

In order to analysis the stability of these equilibrium points, we get it's eigenvalues. The Jacobian matrix for the system is as follows

$$\begin{bmatrix} \frac{\partial f_1}{\partial r_1} = -V_1 & \frac{\partial f_1}{\partial p_1} = \frac{-2p_1}{(1+\frac{p_1^2}{a_1^2})^2}(C_1p_1) + \frac{C_1}{1+\frac{p_1^2}{a_1^2}} & \frac{\partial f_1}{\partial r_2} = 0 & \frac{\partial f_1}{\partial p_2} = 0 \\ \frac{\partial f_2}{\partial r_1} = L_1 & \frac{\partial f_2}{\partial p_1} = -U_1 & \frac{\partial f_2}{\partial r_2} = 0 & \frac{\partial f_2}{\partial p_2} = 0 \\ \frac{\partial f_3}{\partial r_1} = 0 & \frac{\partial f_3}{\partial p_1} = \frac{C_2}{1+\frac{p_2}{a_2}} & \frac{\partial f_3}{\partial r_2} = -V_2 & \frac{\partial f_3}{\partial p_2} = \frac{\frac{-1}{a_2}}{(1+\frac{p_2}{a_2})^2}(C_2p_1) \\ \frac{\partial f_4}{\partial r_1} = 0 & \frac{\partial f_4}{\partial p_1} = 0 & \frac{\partial f_4}{\partial r_2} = L_2 & \frac{\partial f_4}{\partial p_2} = -U_2 \end{bmatrix}$$

$$det\begin{bmatrix} -V_1-\lambda & \frac{\partial f_1}{\partial p_1} & 0 & 0 \\ L_1 & -U_1-\lambda & 0 & 0 \\ 0 & \frac{\partial f_3}{\partial p_1} & -V_2-\lambda & \frac{\partial f_3}{\partial p_2} \\ 0 & 0 & L_2 & -U_2-\lambda \end{bmatrix} = 0$$

$$0 = (\text{-V}_1-\lambda)\begin{vmatrix} -U_1 & 0 & 0 \\ \frac{\partial f_3}{\partial p_1} & -V_2-\lambda & \frac{\partial f_3}{\partial p_2} \\ 0 & L_2 & -U_2-\lambda \end{vmatrix} - (\frac{\partial f_1}{\partial p_1})\begin{vmatrix} L_1 & 0 & 0 \\ 0 & -V_2-\lambda & \frac{\partial f_3}{\partial p_2} \\ 0 & L_2 & -U_2-\lambda \end{vmatrix}$$

$$= (\text{-V}_1-\lambda)(-U_1-\lambda)\begin{vmatrix} -V_2-\lambda & \frac{\partial f_3}{\partial p_2} \\ L_2 & -U_2-\lambda \end{vmatrix} - (\frac{\partial f_1}{\partial p_1})(L_1)\begin{vmatrix} -V_2-\lambda & (\frac{\partial f_3}{\partial p_2}) \\ L_2 & -U_2-\lambda \end{vmatrix}$$

$$= (\text{-V}_1-\lambda)(-U_1-\lambda)((-V_2-\lambda)(-U_2-\lambda) - (L_2)(\frac{\partial f_3}{\partial p_2})) - (\frac{\partial f_1}{\partial p_1})(L_1)((-V_2-\lambda)(-U_2-\lambda) - (L_2)(\frac{\partial f_3}{\partial p_2}))$$

$$= (\text{-V}_1-\lambda)(-U_1-\lambda) - (\frac{\partial f_1}{\partial p_1})(L_1)$$

$$\lambda^2 + (V_1 + U_1)\lambda + (V_1U_1 - L_1(\frac{\partial f_1}{\partial p_1})) = 0$$

This gives two eigenvalues, $\lambda = \frac{-V_1 - U_1 \pm \sqrt{(V_1 - U_1)^2 + 4L_1(\frac{\partial f_1}{\partial p_1})}}{2}$. We can now plug in solutions for $\frac{\partial f_1}{\partial p_1}$ for each case to get specific eigenvalues in order to analysis the stability of each equilibrium point.

For the first case, we have $p_1 = \frac{a_1\sqrt{\alpha}}{\sqrt{V_1U_1}}$. After plugging this into the equation for $\frac{\partial f_1}{\partial p_1}$, we get $\frac{\partial f_1}{\partial p_1} = \frac{V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha}{(V_1U_1+\alpha)^2}$. Now to plug this into our eigenvalues.

$$\lambda_1 = \frac{-V_1 - U_1 + \sqrt{(V_1 - U_1)^2 + \frac{4L_1(V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha)}{(V_1U_1+\alpha)^2}}}{2}$$

$$\lambda_2 = \frac{-V_1 - U_1 - \sqrt{(V_1 - U_1)^2 + \frac{4L_1(V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha)}{(V_1U_1+\alpha)^2}}}{2}$$

Both eigenvalues yield negative values for the condition $\alpha > 0$. This means that both points are stable equilibrium points. The first fixed point is the one of interest to us because it has all positive components. Negative components are not relevant in terms of biology.

For the other two cases, $p_1 = \frac{-a_1\sqrt{\alpha}}{\sqrt{V_1U_1}}$, so they will have the same eigenvalues. With this value for p_1, we get $\frac{\partial f_1}{\partial p_1} = \frac{V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha}{(V_1U_1-\alpha)^2}$. This gives the eigenvalues

$$\lambda_1 = \frac{-V_1 - U_1 + \sqrt{(V_1 - U_1)^2 + \frac{4L_1(V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha)}{(V_1U_1-\alpha)^2}}}{2}$$

$$\lambda_2 = \frac{-V_1 - U_1 - \sqrt{(V_1 - U_1)^2 + \frac{4L_1(V_1^2U_1^2C_1^2 - 2V_1U_1C_1a_1^2\alpha)}{(V_1U_1-\alpha)^2}}}{2}$$

In this case, again when $\alpha > 0$, both eigenvalues are negative, meaning the equilibrium points are stable. However, all these points have negative components, meaning they are biologically irrelevant. This means that in the case where stabilization of mRNA and feedback to translation are negligible, we find one solution that may be of interest to biologist. Transcription rates can be adjusted (values of a_i), in order to study mRNA and protein concentration and how long it will take them to reach saturation. Since feedback to transcription is the only one considered in this case, the discovery of a stable relevant equilibrium point proves that feedback to transcription is a essential element in gene expression. Again, the equilibrium point of interest is

$$\left(\frac{a_1U_1}{L_1\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{a_1}{\sqrt{V_1U_1}}\sqrt{\alpha}, \frac{-a_1U_2}{2L_2} + \frac{U_2}{2L_2}\sqrt{a_2(a_2 + \beta\sqrt{\alpha})}, \frac{-a_1}{2} + \frac{1}{2}\sqrt{a_2(a_2 + \beta\sqrt{\alpha})}\right).$$

Hunt et al. also found one biologically relevant equilibrium point and expressed it in the following way:

$$\left(X_{-eq}\frac{V_2U_1}{C_2a_2U_2L_1}(a_2U_2 + X_{-eq}L_2), X_{-eq}\frac{V_2}{C_2a_2U_2}(a_2U_2 + X_{-eq}L_2), X_{-eq}, X_{-eq}\frac{L_2}{U_2}\right)$$

where,

$$X_{-eq} = -\frac{B_f - \sqrt{B_f^2 + 4A_f C_f \sqrt{\alpha}}}{2A_f}$$

and

$$\alpha = C_1 L_1 - V_1 U_1, \qquad A_f = L_2 V_2 V_1 U_1,$$
$$B_f = a_2 U_2 V_2 V_1 U_1, \qquad C_f = \sqrt{V_1 U_1} U_2 C_2 a_2 a_1.$$

When plugging in the approximate biologically relevant values in the Table 1 below, both my solution and their solution equates to $(15.8035, 210.713, 9.51519, 1268.96)$. Figure 7 shows a plot of the simplified system (System 3) for Case 1 with the initial conditions and parameter values as seen in the table. Figure 8 uses the same initial conditions and parameter values, but is a plot of the general four-dimensional system(System 2). As you can see, in both plots each of the concentration variables stabilize near our equated equilibrium values. This proves that this case was successful in finding a true equilibrium point for the four-dimensional system. Further analysis can be done to see if this is a stable equilibrium point of the six-dimensional system.

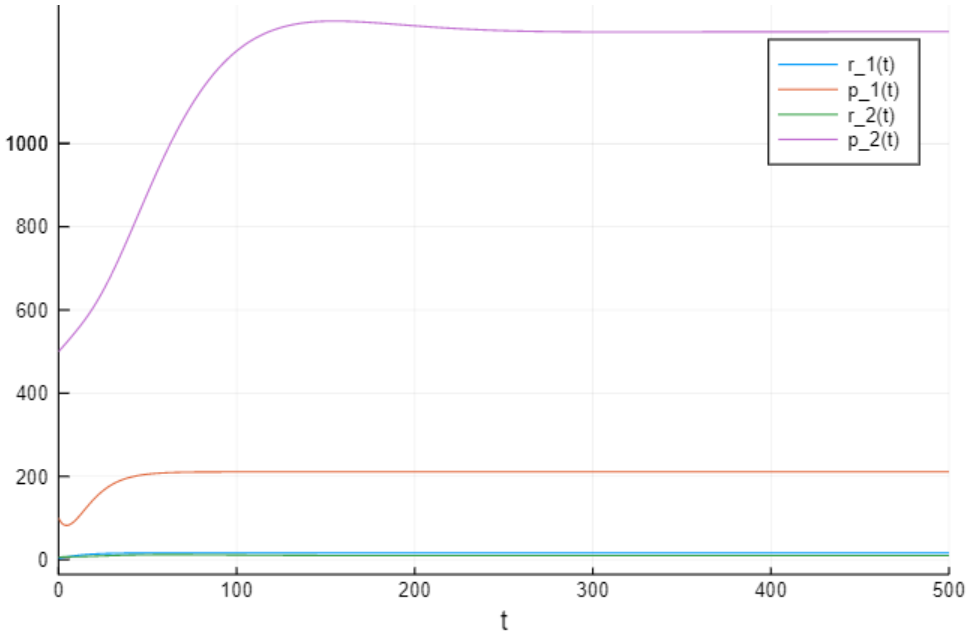

Figure 7: Case 1 specific system(System 3) with I.C. in Table 1.

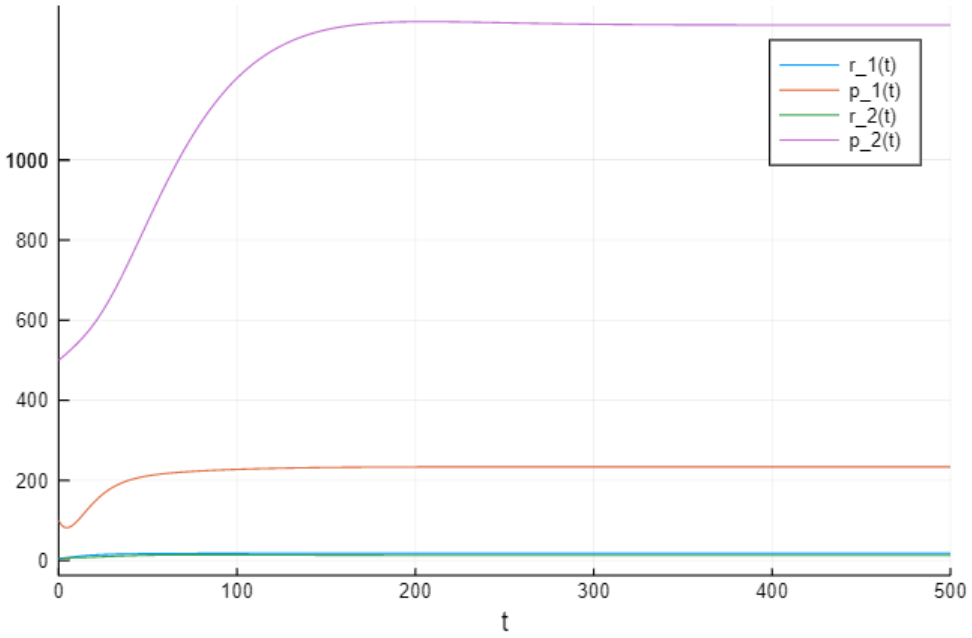

Figure 8: General four-dimensional(System 2) system with I.C. in Table 1.

i	r_i	p_i	C_i	L_i	V_i	U_i	a_i	b_i	d_i
1	3	100	0.03	2	0.03	0.15	60	5000	5000
2	6	500	0.03	2	0.03	0.15	60	5000	5000

Table 1: Initial values and parameters for plots in Figures 7 and 8.

Case 2. $a_i, d_i \gg p_i$

In this case, both the feedback from proteins to transcription and the feedback from proteins to translation are negligible. In order to find other equilibrium solutions other than $(r_1, r_2, p_1, p_1) = (0, 0, 0, 0)$, again set all the equations equal to 0 and solve for the concentration variables. The system looks like:

$$(1.2)\quad \frac{dr_1}{dt} = C_1 p_1 - \frac{1}{1+\frac{p_2}{b_1}} V_1 r_1 \qquad (2.2)\quad \frac{dp_1}{dt} = L_1 r_1 - U_1 p_1 \tag{4}$$

$$(3.2)\quad \frac{dr_2}{dt} = C_2 p_1 - \frac{1}{1+\frac{p_2}{b_2}} V_2 r_2 \qquad (4.2)\quad \frac{dp_1}{dt} = L_2 r_2 - U_2 p_2$$

First get $r_2 = \frac{U_2}{L_2} p_2$ and plug into equation (3.2). Get p_1 in terms of p_2.

$$0 = C_2 p_1 - \frac{1}{1+\frac{p_2}{b_2}} \times \frac{V_2 U_1}{L_2} p_2$$
$$C_2 p_1 = \frac{V_2 U_2}{L_2 + \frac{L_2 p_2}{b_2}} p_2$$
$$p_1 = \frac{V_2 U_2 b_2}{C_2 L_2 b_2 + C_2 L_2 p_2}$$

Get $r_1 = \frac{U_1}{L_1} p_1$ plug in solution for p_1 in terms of p_2. This gives $r_1 = \frac{V_2 U_2 U_1 b_2}{C_2 L_2 L_1 b_2 + C_2 L_2 L_1 p_2} p_2$. Now plug p_1, r_1 into equation(2.2).

$$0 = \frac{C_1 V_2 U_2 b_2}{C_2 L_2 b_2 + C_2 L_2 p_2} p_2 - \frac{V_1}{1+\frac{p_2}{b_1}} \times \frac{V_2 U_2 U_1 b_2}{C_2 L_2 L_1 b_2 + C_2 L_2 L_1 p_2} p_2$$
$$\frac{C_1 V_2 U_2 b_2}{C_2 L_2 b_2 + C_2 L_2 p_2} = \frac{V_1 V_2 U_1 U_2 b_2}{(1+\frac{p_2}{b_1})(C_2 L_2 L_1 b_2 + C_2 L_2 L_1 p_2)}$$
$$C_1 C_2 L_1 L_2 V_2 U_2 b_2 (1+\frac{p_2}{b_1})(b_2 + p + 2) = C_2 L_2 V_1 V_2 U_1 U_2 b_2 (b_2 + p_2)$$
$$C_1 L_1 (1+\frac{p_2}{b_1}) = V_1 U_1$$
$$p_1 = \frac{b_1 V_1 U_1 - C_1 L_1}{C_1 L_1}$$

To simplify, we can let $\alpha = C_1 L_1 - V_1 U_1$, which yields $p_2 = -\frac{b_1 \alpha}{C_1 L_1}$. Plug this back into our equations for the other concentration variable and we get the equilibrium solution:

$$\left(\frac{U_1 U_2 V_2 b_1 b_2 \alpha}{C_2 L_2 L_1 (-b_2 C_1 L_1 + b_1 \alpha)}, \frac{U_2 V_2 b_1 b_2 \alpha}{C_2 L_2 (-b_2 C_1 L_1 + b_1 \alpha)}, -\frac{U_2 b_1 \alpha}{C_1 L_1 L_2}, -\frac{b_1 \alpha}{C_1 L_1} \right)$$

Note that when $\alpha > 0, r_1, p_1, r_2, p_2 < 0$, $\alpha < 0, r_1, p_1, r_2, p_2 > 0$ and when $\alpha = 0, r_1, p_1, r_2, p_2 = 0$, the trivial solution. The eigenvalues for this point can be found from the Jacobian matrix below, however they become extremely complex and difficult to solve for by hand.

$$\begin{bmatrix} \frac{\partial f_1}{\partial r_1} = -\frac{V_1}{1+\frac{p_2}{b_1}}, & \frac{\partial f_1}{\partial p_1} = C_1, & \frac{\partial f_1}{\partial r_2} = 0, & \frac{\partial f_1}{\partial p_2} = \frac{-\frac{V_1 r_1}{b_1}}{(1+\frac{p_2}{b_1})^2} \\ \frac{\partial f_2}{\partial r_1} = L_1, & \frac{\partial f_2}{\partial p_1} = -U_1, & \frac{\partial f_2}{\partial r_2} = 0, & \frac{\partial f_2}{\partial p_2} = 0 \\ \frac{\partial f_3}{\partial r_1} = 0, & \frac{\partial f_3}{\partial p_1} = C_2, & \frac{\partial f_3}{\partial r_2} = -\frac{V_2}{1+\frac{p_2}{b_2}}, & \frac{\partial f_3}{\partial p_2} = \frac{-\frac{V_2 r_2}{b_2}}{(1+\frac{p_2}{b_2})^2} \\ \frac{\partial f_4}{\partial r_1} = 0, & \frac{\partial f_4}{\partial p_1} = 0, & \frac{\partial f_4}{\partial r_2} = L_2, & \frac{\partial f_4}{\partial p_2} = -U_2 \end{bmatrix}$$

According to Hunt et al. when $\alpha < 0$, the origin is stable and the nontrivial point is unstable, residing in the positive hyper-octant. Remember that α is the degradation terms subtracted from the production terms. This means that for $a < 0$, the product of the degradation terms is greater than than the product of the production terms. If the mRNA and proteins are degrading faster than they are being produced, the system will eventually

die out. Therefore it makes biological sense for the origin to be a stable equilibrium point under this condition. When $\alpha = 0$, the origin is the only equilibrium point and it's stability is unknown. There are again two solutions when $\alpha > 0$, the origin and the nontrivial point. The stability swaps between the two points. Although the nontrivial point is stable, under this condition all it's components are negative, making it biologically irrelevant. Since there is no positive stable equilibrium point and the origin is unstable, the concentrations will grow without bounds for realistic initial values, which can be seen in Figure 9 below. It would be impossible for a cell to survive under these conditions. Therefore, in biology, we cannot not neglect feedback loops to transcription and translation. This case proves that feedback loops to transcription and translation are critical components in gene expression.

Note that when using the same initial conditions and parameters in Table 2 to plot the general four-dimensional system (System 2), it looks as if there could be a possibility of stabilization. While this case is not biologically relevant and feedback loops to transcription and translation cannot be ignored, it is possible that this could be a significant point in the six-dimensional system. Further analysis is needed to confirm this.

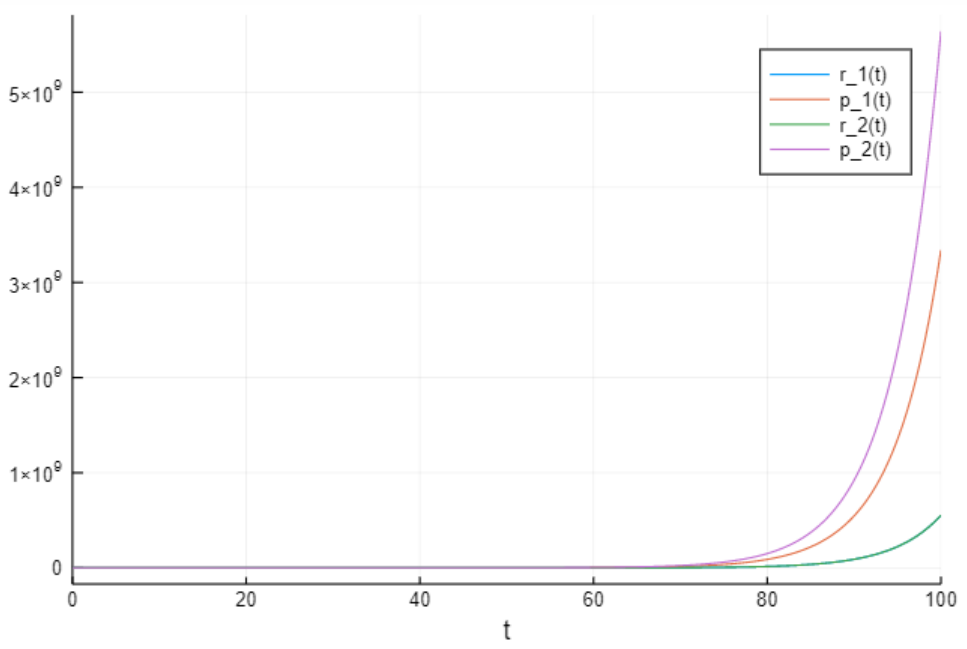

Figure 9: Case 2 specific system (System 4) with I.C. in Table 2

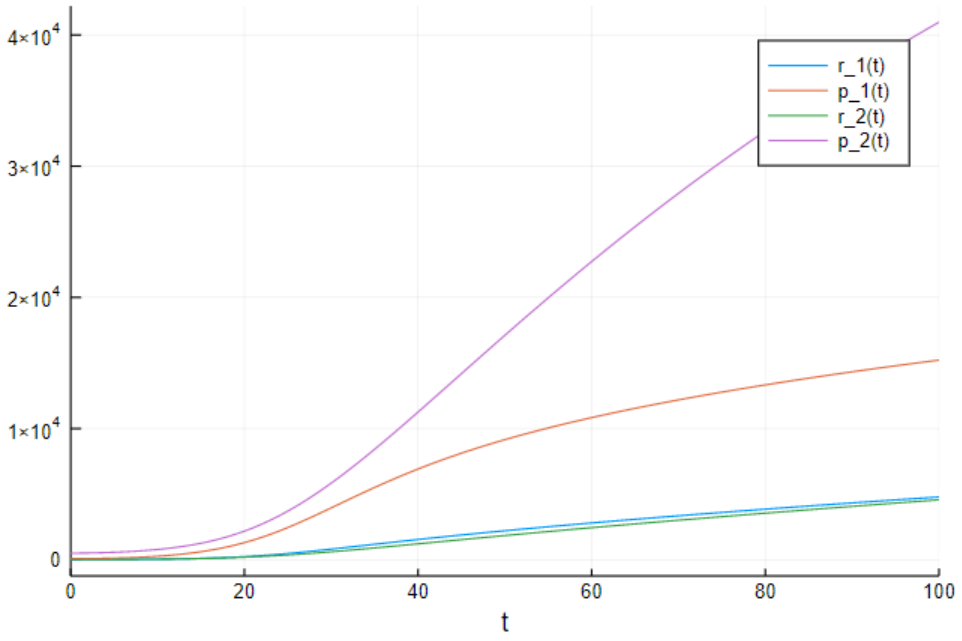

Figure 10: General four-dimensional system (System 2) with I.C. in Table 2

i	r_i	p_i	C_i	L_i	V_i	U_i	a_i	b_i	d_i
1	3	100	0.03	2	0.03	0.15	5000	100	5000
2	6	500	0.03	2	0.03	0.15	5000	100	5000

Table 2: Initial conditions and parameters for plot in Figures 9 and 10.

5 Conclusion

Case 1 was able to provide a stable equilibrium point that is of biological relevance, not only for the case specific system, but for the larger four-dimensional system as well. This equilibria can most likely be interpreted as cellular homeostasis. This means maintaining a steady state in order to keep the cell healthy. A cell's health is affected by many factors, including environmental stimuli. The concentration values for the stable state can be further analyzed by adjusting the transcription and translation rates. The difficulty that comes with this is factoring the effects of environmental stimuli into the model because it is so unpredictable. That is a flaw in the model by Hunt et al. as well as many other models. In the case that $\alpha > 0$, where we found our stable equilibrium point, the origin is also an equilibrium point. This equilibria yields negative real-values eigenvalues as well as complex eigenvalues. This is good because rather than the system dying out as it approaches the origin, it spirals around it. This is ideal in terms of biology. In the case that $\alpha = 0$, meaning the production rates are equal to the degradation rates ($C_1L_1 = V_1U_1$), the origin is the only equilibrium point. This point yields a zero eigenvalue, meaning it is neither stable nor unstable. We do not want the origin to be stable in any biologically relevant case.

For case 2, feedback from proteins to both transcription and translation are neglected. This case produced no biologically relevant equilibrium points, therefore all concentrations grew without bounds. We know that this situation could not occur in biology. Although this case did not produce any relevant equilibria, it proved that feedback to transcription and translation must be accounted for in the model. In previous models, such as Chen's, feedback to translation had been ignored. Therefore Hunt et al. has taken a step in the right direction by including that feedback loop in their model as well as proved that it is an essential component. However, the case 2 results were not completely congruent when applied out of the case-specific system. This demonstrates the limitations that can occur with case based analysis.

Overall, the model provided explicit functions for gene expression and produced one promising stable solution. I was able to obtain the real complicated differential equations solutions using numerical approximations and compare with the solutions given by the approximate system that the paper uses. My numerical approximations did in fact agree with the solutions derived both in the paper and by myself. I also was able to present the solutions in a simpler manner than that of Hunt et al. If I were to further by research, I would apply my solutions to the six-dimensional system and again test using numerical approximations.

This model does disregard many other factors that affect gene expression, as many models do, because the cell environment is constantly undergoing change. The consideration of overall cell health lacks in the model, which is the essential factor in maintaining cellular homeostasis. The model is also only applicable to prokaryotic cells and simple eukaryotic cells, such as yeast. Newer models of gene expression are emphasizing the stochasticity of gene expression and attempting to statistically analyze rather than predict. Research published in 2020 by Cao and Grima presents a two state model for gene expression in eukaryotic cells that describes promoter switching, transcription, translation and mRNA and protein decay. The model includes mRNA maturation, cell division, gene replication, dosage compensation, growth-dependent transcription and auto-regulatory feedback. Data from yeast, mouse and human cells are used to confirm their model. However, as expected they were not able to get an exact solution for their model[2]. If more of the factors considered in the model by Cao and Grima were able to be considered in an ODE system like the one proposed by Hunt et al. that could be revolutionary in the field of biology.

References

[1] William E. Boyce and Richard C. DiPrima. *Elementary Differential Equations*. John Wiley Sons, Inc., 111 River Street, Hoboken, NJ, 2012.

[2] Zhinxing Cao and Ramon Grima. Analysticla distributions for detailed models of stochastic gene expression in eukaryotic cells. *Proceedings of the National Academy of Sciences*, 2020.

[3] Your Genome. What is gene expression?

[4] Elizabeth Miller Jason Pham Joseph Hunt, Lissette Laplace. A continuous model of gene expression. *California State Polytechnic University, Pomona*, 2005.

[5] Philip M. Kim and Bruce Tidor. Limitations of quantitive gene regulation models: A case study. *Genome Research*, 2003.

[6] Y. A. Kuznetsov. Saddle-node bifurcation. *Scholarpedia*, 1(10):1859, 2006. revision #151865.

[7] Jiří Lebl. Notes on diffy qs: Differential equations for engineers.

[8] George Church Ting Chen, Hongyu He. Modeling gene expression with differential equations. 1999.