

A Comparison Between Stochastic and Deterministic Models of a Biological Oscillator

H. JI, J. SOLHEID, AND J. GU

Abstract - We look at the behavior of biological oscillators, specifically analyzing a genetic circuit that has oscillatory behavior. Implementing the system in deterministic and stochastic models, we compare the two models using various tests and analyze the effect of stochastic noise on these oscillations. We first investigate the effect of volume and find that at small system volumes, stochastic events cause the system to exhibit more sporadic oscillations and a longer period of oscillation. Next, we develop classification systems for discerning the boundary in the parameter space where the model begins to show oscillations. This is more challenging in the stochastic model, due to persistent fluctuations. We compare different methods of analysis of our deterministic and stochastic models and visualize the conditions which lead to oscillation in each model. We find that the parameter range of oscillation is larger in the stochastic model than in the deterministic model. Finally, we find an increased rate of mRNA production can create greater noise in a system and amplify the difference between a stochastic and deterministic system.

Keywords : repressilator; stochastic model; deterministic model; limit cycle; trajectory

Mathematics Subject Classification (2020) : 37N25

1 Introduction

Biological oscillators occur in many manners, such as controlling circadian rhythms, mitosis, and DNA synthesis [10]. We are interested in understanding the application of biological oscillators to a repressilator system, discussed in [1] and [5], which features multiple repressor proteins, each inhibiting the transcription of mRNA. The repressor binds to the mRNA promoter, inhibiting the RNA polymerase enzyme from attaching to the promoter and from beginning transcription.

Previously discussed by Novák and Tyson [10], a delayed negative feedback loop is required to create an oscillating system. The loop consists of multiple mRNA concentrations and its translated protein concentrations. Let us start with a species of mRNA called m_X , which is translated into a repressor protein called p_X . Repressor p_X inhibits the transcription of another mRNA m_Y , which limits the translation of repressor p_Y . p_Y inhibits the expression of m_Z , which limits the production of protein repressor p_Z . p_Z inhibits the expression of m_X , completing the self-regulating feedback loop. This feedback loop is negative since a high concentration of one type of mRNA will cause the concentration of its respective repressor protein to also increase, which results in its concentration decreasing.



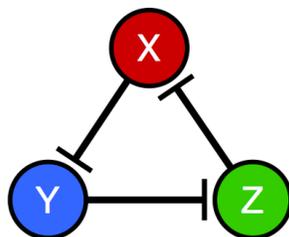


Figure 1: Model repressilator system with 3 different repressors; the bars represent which repressor is repressing which protein (p_X represses p_Y , p_Y represses p_Z , p_Z represses p_X)

We can model the behaviors of protein repressors by using the protein synthesis and degradation rates [8]. The synthesis rate is determined by a transcription factor and the degradation rate is determined by the rate at which the repressor protein is broken down by a protease enzyme. As discussed in [10], protein synthesis systems like a repressilator system have an inherent time delay, which introduces oscillations of the concentration of each repressor protein into the system. The delay is built into how the species interact, rather than the algorithm or model.

Two of the main approaches used to study repressor based oscillatory systems are deterministic and stochastic models. For a repressilator system, a deterministic model shows the change in the rate of a reaction over time, while a stochastic model accounts for the possibility of noise using dynamic probability distributions. The behavior of a deterministic model is described by a system of differential equations, while a stochastic model is described by probabilities of discrete events.

At larger volumes of molecules, stochastic biochemical models are known to agree with the behavior of the deterministic model as discussed by Gillespie [3]. However, at smaller volumes of molecules, the behavior of a stochastic model can disagree with a deterministic one. For example, qualitative differences have been found in steady state behavior, as presented by Lipshtat [4], Ma [7], and Schultz [11].

Lipshtat discusses how in a low volume system, a stochastic model can predict multiple steady states for one system, whereas a deterministic model will not predict multiple steady states [4]. Ma demonstrates how in small systems, there is additional fluctuation that cannot be accurately modeled just by using equations that do not account for internal noise [7].

At these low volumes, the stochastic model may better predict the behavior of a reaction system, notably in systems with bimodal outcomes, where a stochastic model predicts multiple varying outcomes which a deterministic model fails to accurately portray [11]. As discussed in a paper by Elowitz and Leibler, a deterministic system also neglects the discrete nature of molecules in these systems and the processes in which they interact [1]. They also mention how the real world repressilator system exhibited fluctuations not observed in the deterministic model, which further justifies a stochastic analysis of the repressilator system.

It was brought to our attention in review of this manuscript that a stochastic version of the repressilator model has been studied previously in Loinger [5], where the effect of binding sites and plasmid number on the presence and characteristics of oscillation was studied. We specifically focus on the model described in Elowitz [1], which Loinger models from, and investigate and classify the oscillations in the deterministic and stochastic models. We further expand by performing a parameter space study and a periodogram analysis to look at the effects of mRNA production rate on oscillation behavior. The paper in [5] provides some results on the expected differences between the deterministic and stochastic models, which we corroborated, as described below.

In summary, while stochastic repressilator models have been studied previously in [1] and [5], the major new contribution of our present work is to study in detail the boundary in parameter space where oscillations begin for the repressilator system, in both the stochastic and deterministic models. To our knowledge this has not been done previously.

We were curious as to whether stochastic noise would qualitatively change the nature of oscillations in our repressilator model. We apply various tests to investigate similarities and differences between deterministic and stochastic models. Tests included noting the effects of volume and examining the parameter space to identify different classes of behavior, notably looking at the reaction rates and size of the system, as well as using frequency domains to further examine differences.

2 Methods

2.1 Deterministic

In our deterministic and stochastic models, we state two conditions, taken from [1]: the repressors act identically to one another (i.e. the system is homogeneous), and reaction rates are linearly proportional to the concentration of a repressor. The deterministic model assumes that variables are continuous within time, with each instant in time having a given rate of change. Given a set of initial conditions for the biological characteristics of each mRNA and repressor protein, the deterministic model, as a set of equations, can predict concentrations of each species at each time point. Our variables, which represent the concentration of each mRNA and repressor protein species in units of moles per unit volume U , are continuous variables in the deterministic model. As continuous variables, we can use ordinary differential equations (ODE) to describe such a system. The time delay is inherent to the interaction between species, and is not explicitly implemented through a delay differential equation. We use the ODE model discussed in Elowitz and Leibler [1].

$$\frac{dm_i}{dt} = -m_i + \frac{\alpha}{(1 + p_j^n)} + \alpha_0 \quad (1)$$

$$\frac{dp_i}{dt} = -\beta(p_i - m_i) \quad (2)$$



Variable i , labeled X , Y , and Z , represents the mRNA (m_i) that translates the repressor protein p_i . Variable j , labeled Z , X , and Y , represents the mRNA (m_j) that translates the repressor protein (p_j). p_j is the repressor protein which inhibits the production of m_i . Note the correspondence between j & i of Z & X , X & Y , and Y & Z . These pairs follow the negative feedback loop shown in the introduction. The variables α_0 , α , β , and n are parameters which are determined based on the inherent biological qualities of the system and can be adjusted to simulate different systems.

m_i	Concentration of mRNA (number/U), and i is labeled (X, Y, Z)
p_i	Concentration of repressor protein (number/U), where i is labeled (X, Y, Z)
p_j	Concentration of repressor protein (number/U), where j is labeled (Z, X, Y)
α_0	Flat rate of mRNA production
α	Rate of mRNA production dependent on concentration of repressor present
β	Ratio of rate of mRNA production to protein production
n	Hill coefficient; estimates the extent of cooperativity between repressor and mRNA

Table 1: Description of variables in reaction rate equations. All concentrations are in arbitrary units and all parameters are dimensionless.

To derive Equation 1, the concentration of mRNA transcribed can be described as a function dependent on the concentration of repressor and mRNA present in the system. First, the rate of mRNA transcription decreases due to the completion of the negative feedback loop and degradation of mRNA, which is directly proportional to the concentration of m_i . When there is no repressor bound to the mRNA promoter, the mRNA is transcribed at a rate of $\alpha + \alpha_0$. Otherwise, when the repressor is bound on the mRNA promoter, the rate of transcription is α_0 , which accounts for possible “leakage” in the mRNA transcription system [1]. α_0 is related to the probability that a repressor protein can bind to the promoter, while $\alpha + \alpha_0$ is related to the probability that a repressor does not bind to the promoter. Equation 3 describes the conditions of such a system:

$$\frac{dm_i}{dt} = -m_i + \alpha_0(\text{probability } p_j \text{ is bound}) + (\alpha_0 + \alpha)(\text{probability } p_j \text{ is unbound}) \quad (3)$$

The probability that the repressor binds to the promoter is $1/(1 + p_j^n)$, discussed in [1]. Because the probability that the repressor does not bind to the promoter is complementary, the probability that the repressor is unbound can be described as $1 - 1/(1 + p_j^n)$. Substituting these values in and rescaling by the mRNA lifetime, we get Equation 1. Note that in this paper, we set $\alpha_0 = 0$ for all systems to specifically focus on the effects of α and β . The case when $\alpha_0 \neq 0$ is also discussed in [1].



To derive Equation 2, we assume a positive linear relation between the amount of m_i and the translation rate $\frac{dp_i}{dt}$. There is also a negative linear relationship between the amount of p_i and translation rate $\frac{dp_i}{dt}$ because an increased amount of p_i would lead to the completion of the negative feedback loop and lower the translation rate. We use nondimensionalization to rescale m_i by the mRNA translation efficiency, which is the rate of protein production per mRNA production $\frac{dp_i}{dm_i}$. This gets us Equation 2. Note that equations 1 and 2 represent a nondimensionalized form of the ODE system. The nondimensionalization process is done in more detail in [1].

2.2 Stochastic

A stochastic simulation algorithm (SSA) model incorporates many random biological processes and factors, such as cell-to-cell variation in volume and size, differences in expression capacity, and fluctuations in the pathways of upstream molecular components, which are not accounted for in a deterministic model [9]. The stochastic model is implemented using the Gillespie Algorithm [2], a method to efficiently simulate stochastic chemical systems. For a stochastic model, the Gillespie Algorithm uses the probability of a reaction occurring in some unit time step rather than reaction rates. This reaction probability is determined by a joint probability density function, where each reaction is given specific parameters to occur depending on the structure of the biological system. Time is increased by a random amount, then one reaction in the system is chosen to occur at the time-step, where each reaction's probability of occurring is calculated by taking its reaction rate and dividing it by the sum of all reaction rates of the system. This provides an accurate stochastic simulation for the behavior of the repressilator system.

Here, we assume that at each timestep, each reaction's probability of occurring is independent. Looking at the wait time for one specific reaction, we have an exponential distribution, where the probability that the wait time for a specific reaction to occur will decrease at an exponential rate as time increases. The change in a probability distribution cannot be modeled as a single output at a given time. We instead must average multiple instances (replicas) of a stochastic model. The final display will show the concentration of each species in units of moles of per unit volume as a function of time.

The stochastic model follows these reaction rules:

1. m_i is produced under a reaction probability rate of k_1 dependent on the protein p_j concentration. The reaction probability rate k_1 can be described by an expression, rather than a constant.

$$k_1 = \frac{\alpha}{(1 + p_j^n)} + \alpha_0 \quad (4)$$

2. m_i decays under a normalized reaction probability rate dependent on its own mRNA concentration as the presence of mRNA completes the negative feedback loop, inhibiting itself. The normalization of the reaction rate is due to rescaling by the translation efficiency of mRNA.



3. p_i is translated in a forward reaction at a nondimensionalized probability rate of β dependent on the mRNA concentration.
4. p_i decays at a nondimensionalized probability rate of β dependent on its own protein concentration as the presence of a repressor completes the loop and inhibits itself.

Note that the deterministic and stochastic model follow the same behavior. Equation 1 of the deterministic model corresponds with reaction rules 1 and 2 of the stochastic model. Thus, the derivation of k_1 can be attained from referring to derivation of Equation 1. Equation 2 of the deterministic model corresponds to reaction rules 3 and 4 of the stochastic model.

2.3 Implementation

The deterministic and stochastic approaches both model the concentration level of each mRNA and protein repressor species. At a given concentration, a change in volume represents a change in the number of biomolecules. If the volume decreases, then the quantity of particles decrease, increasing the significance of random stochastic fluctuations in the system. Hence, volume plays an important role in a stochastic model which measures such random fluctuations. However, a deterministic model is unable to notice such random fluctuations, and therefore is unaffected by changes in volume. Thus, we assume volume only acts in our stochastic model in the form of a variable called volume (measured in unit volume U), which acts as a multiplier against the concentration of the species present. Unit volume U is a scale factor and an arbitrarily chosen variable in the stochastic system. It represents an effective volume where the molecules can be considered well-mixed. This volume stays constant for a given system. The model simulates a stochastic system at a default of $1U$ unless otherwise specified.

The differential equations are solved using Solution IVP [12], a Python-based ordinary differential equation solver using a 4th order Runge-Kutta algorithm with 5th order error. The time step is adaptive and dependent on the gradient of the system. The stochastic model uses PySB [6], a Python package, and all runs (replicas) are averaged for our results.

We assume each species of mRNA has the same rate of production, α_0 , each species of repressor protein has the same rate of dynamic production, α , and the ratio of mRNA production to protein production, β , is the same for all species. To compare both models, we test different parameters in a loop to determine the sets of conditions that would cause noise to be significant. We disregard the importance of the initial conditions, as we only analyze the behavior of the model after sufficient time has passed. The stochastic model is run many times, and each individual simulation or replica, is used to find an average stochastic trajectory. The data received was then quantitatively analyzed using the time domain and frequency domain. The frequency domain was achieved by using a periodogram, a method based on Fourier transform, on standardized data. We also analyzed the data qualitatively by comparing different trends.



3 Results

3.1 Variance of Models

We first look at the consistency of the stochastic model. In each model, we use the average of many individual replicas to get a stochastic trajectory. In Figure 2, we compare individual stochastic replicas to the average.

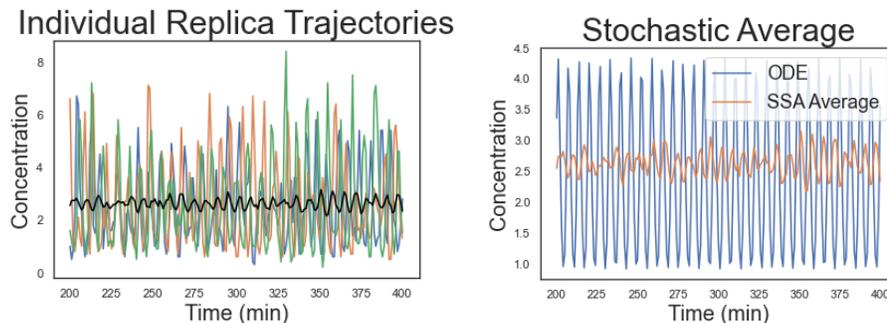


Figure 2: (Left) A plot of 3 different replicas of the stochastic model of one mRNA species, compared with the average of the stochastic replicas of the same species (in black). (Right) For analysis, we compare this average stochastic trajectory (now in orange) with the deterministic model. The black and orange curve represent the same average stochastic trajectory. $\alpha = 10$, $\beta = 4$, $a_0 = 0$, $n = 2$, $V = 1U$, 100 replicas.

We notice that compared to the individual replicas, the stochastic average has a much smaller and more consistent amplitude. More importantly, the oscillations present in each stochastic replica are also present in the average. When comparing this stochastic average, we see that it is much closer to the behavior of the deterministic model than any of the individual replicas. The average of each replica's standard deviation is approximately 0.1957, while the standard deviation of the stochastic model average is much lower at approximately 0.0267.

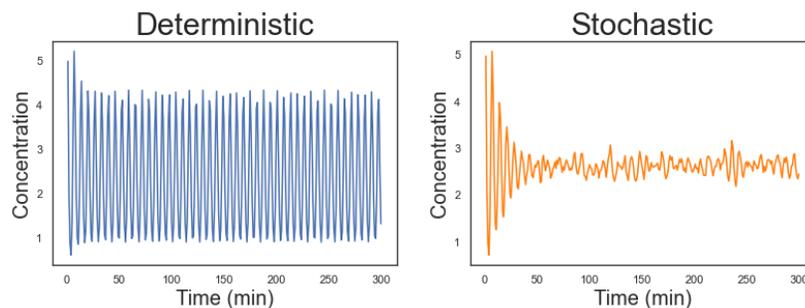


Figure 3: Average of 100 replicas for one mRNA species (m_X). The parameters are $\alpha = 10$, $\beta = 4$, $a_0 = 0$, $n = 2$, $V = 1U$, and 100 replicas.



To investigate the limit behavior and oscillations in the system, we ignore the initial conditions and only analyze the behavior after a “burn-in” period has passed. To do so correctly, we identified a burn-in period of 300 minutes. This means that when we analyze the trajectory of a model, we are looking at data points after 300 minutes. We choose this “burn-in” time because the parameter space does not exhibit changes in behavior after this time; an example can be seen in Figure 3.

3.2 Analysis of Low Volume Systems

We analyze the effect that volume of the system has on the stochastic noise. In the deterministic model, the dynamic variables are concentrations of molecular species (number per volume). Thus, changing the volume of the system has no effect on dynamics. In contrast, in the stochastic model the dynamic variables are the number of copies of each molecular species. Thus, changing the volume can affect system dynamics since the number of particles also changes (assuming constant concentration). The expectation is that as the volume of a system decreases, the noise would have a more significant effect on the system behavior, and therefore a stochastic model would differ more from a deterministic one. Figures 4 and 5 display the presence of oscillation for a stochastic and deterministic system with different volumes for one specific set of parameters.

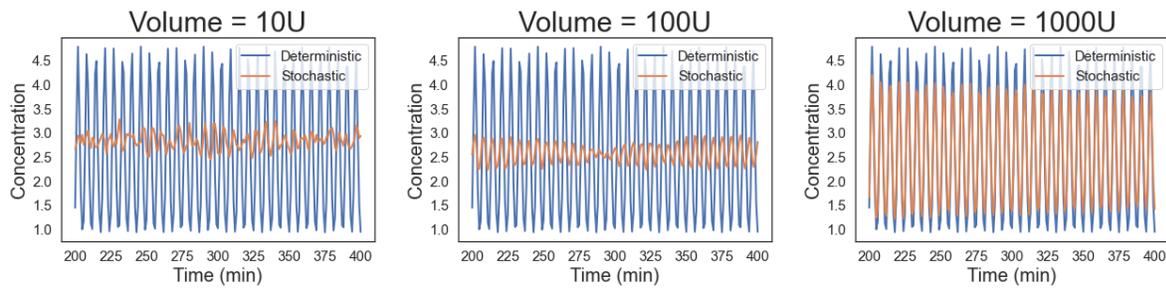


Figure 4: Concentration graphs for one species of mRNA in an oscillating state with varying volumes, with the deterministic model shown in blue and the stochastic in orange. The parameters are $\alpha = 11.5$, $\beta = 4.5$, $\alpha_0 = 0$, $n = 2$, 5 replicas.

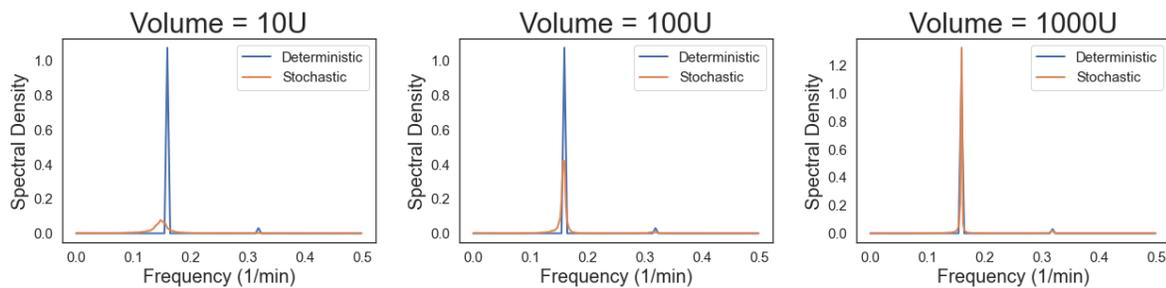


Figure 5: Frequency graphs of the models from Figure 4



After analysis, we notice that for oscillatory systems (those with large α and small β), a stochastic model with low volume will have a lower amplitude and a weaker prevalent oscillation frequency when compared to a deterministic model with the same parameters. As the volume increases, the stochastic model will increase in amplitude and exhibit a strong oscillation frequency. When the volume is very large, the stochastic model agrees with the deterministic model, which can be seen in the rightmost graphs in Figure 4 and Figure 5. We now examine parameters which do not fall into this category of high α and low β .

3.2.1 Divide Point

In volumes greater than $5U$ and less than $30U$, adjusting the parameters of α and β , whilst having $\alpha_0 = 0$ and $n = 2$, results in an interesting trend in the frequency domain graphs. Figure 6 shows how for a certain β , increasing values of α to a certain point leads the stochastic and deterministic graphs to more closely align. Past that point, increasing values of α will lead to the stochastic and deterministic graphs diverging from each other. For reference sake, we call this point the divide point.

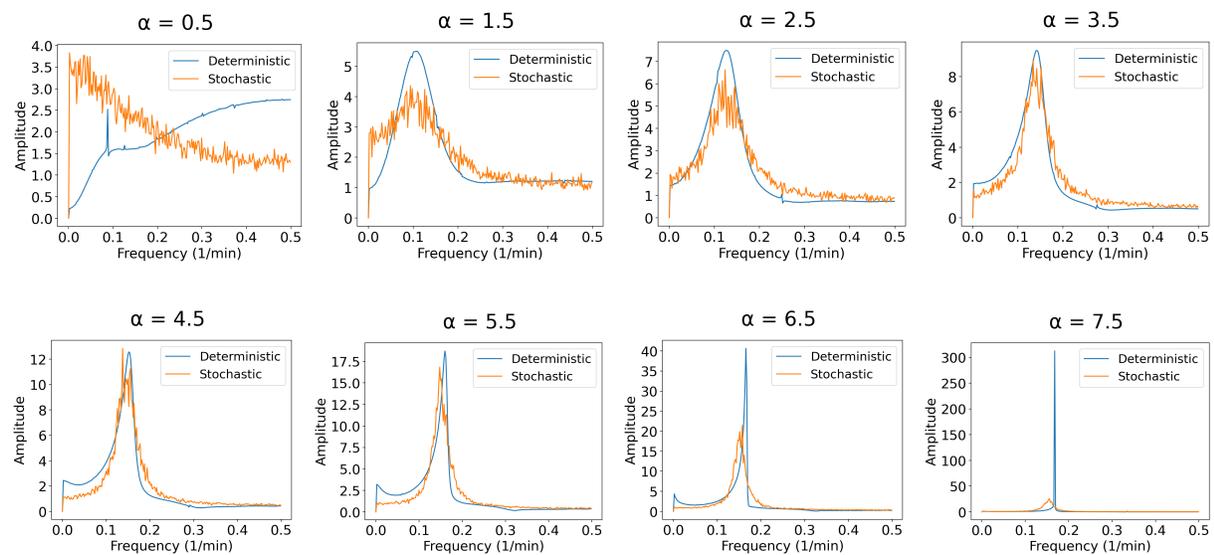


Figure 6: Frequency graphs, from left to right, showing increasing values of α from 0.5 to 7.5 under constant conditions of $\beta = 4.5$ and $V = 20U$

As depicted by Figure 6, the stochastic and deterministic graphs align in a particular way. To explain this thoroughly, let us look at the trend seen in the Figure 6. β is kept constant at 4.5, while α is incrementally increased from 0.5 to 7.5. In the beginning, the stochastic graph's amplitude increases relative to the deterministic graph's amplitude. The stochastic graph slowly rises to match the graph of the deterministic, leading to an increased alignment between the two graphs. The best alignment is reached at the divide point. Past the divide point, the stochastic graph's amplitude decreases relative to the



deterministic graph's amplitude. This leads to the graphs increasingly diverging from each other after the divide point.

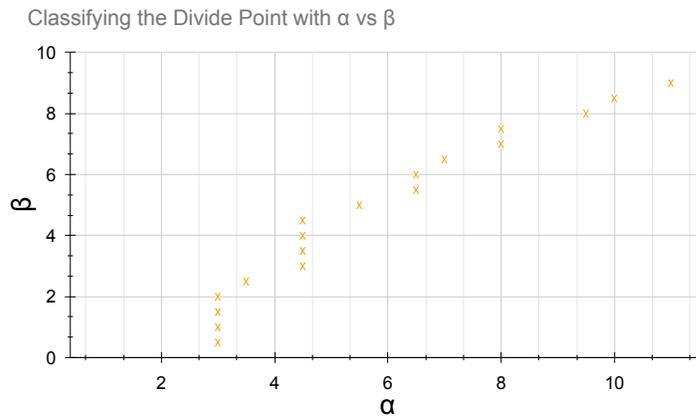


Figure 7: Plotting possible divide points on an α vs. β chart. α and β are from 0.5 to 11.5 and 0.5 to 9.5, respectively. Throughout all runs, these variables are kept constant: $V = 20U$, $\alpha_0 = 0$, $n = 2$, Replica = 100.

Noting this trend, we plot the divide point in a chart of α and β . We define the divide point as where the stochastic graph stops increasing/decreasing in amplitude relative to the deterministic graph's amplitude. Under a constant volume of $20U$, values of α and β are cycled through and possible divide points are plotted in Figure 7.

3.3 Classification of Oscillations

We first look at the deterministic and stochastic models under volumes of $1U$. When adjusting model parameters, we see two classes of behavior, which we call "constant oscillation" and "no oscillation". In the deterministic model, the "constant oscillation" behavior is just a classic limit cycle. However, in the stochastic model, it is less clear whether persistent oscillations are truly due to limit cycle behavior versus persistent fluctuations, so that is why we adopt the new term "constant oscillation". In the "no oscillation" case, we are looking at dynamics that either show damped oscillation or no oscillation whatsoever.

Notice how noise is present in the stochastic model in Figure 8. This noise makes it difficult to determine the threshold between predictions of "constant oscillation" and "no oscillation" classification. Elowitz [1] uses the same deterministic model as us, classifying their graphs as "steady state stable" ("no oscillation") and "steady state unstable" ("constant oscillation"). They specifically focus on the effect of changing α and β , keeping other parameters constant. Their results show that for a system where $\alpha_0 = 0$ and $n = 2$, changes in β have little to no effect on the limit cycle behavior of the system when α is low. As α increases past a certain threshold, β begins to have a larger influence on the behavior of the system, eventually reaching a positive linear relationship where changes in β are about as effective as changes in α .



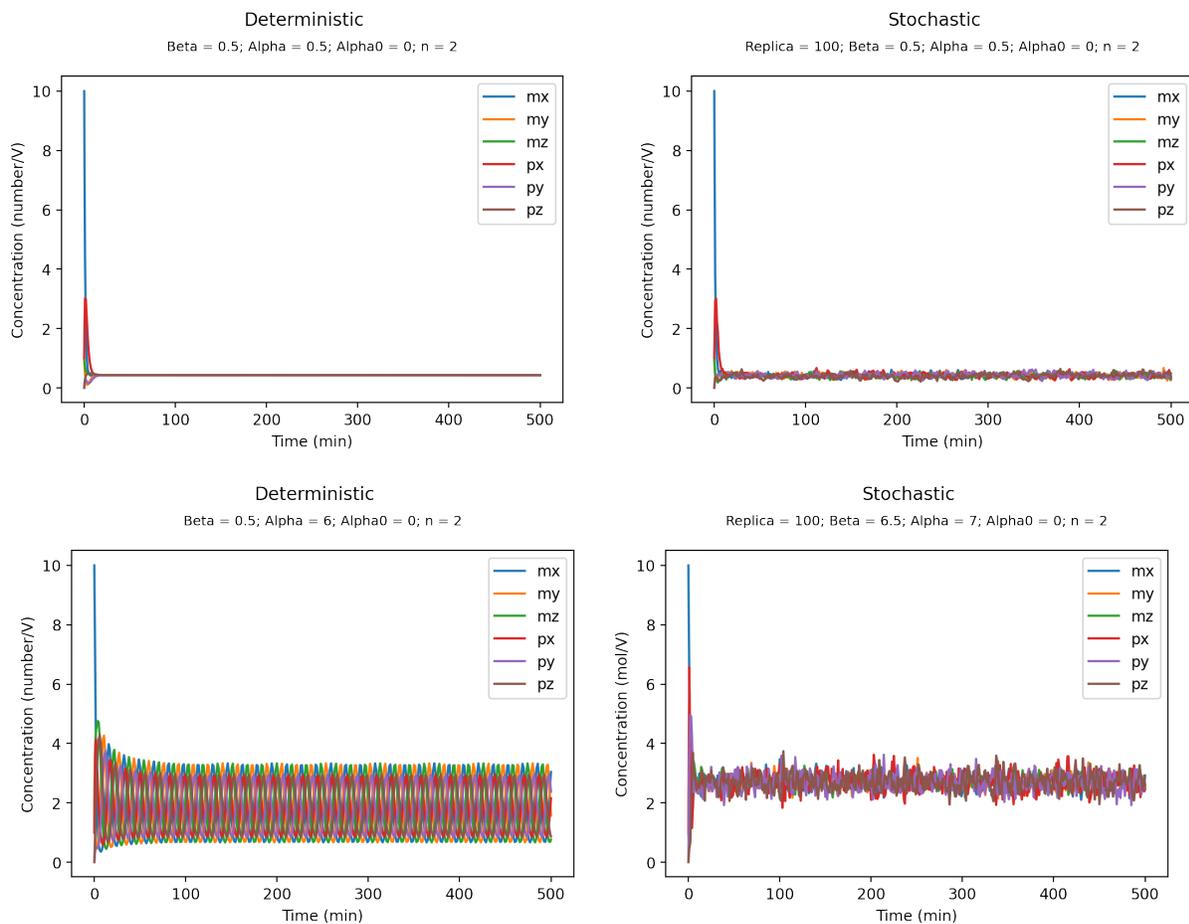


Figure 8: No oscillation is produced in the top deterministic and stochastic graph; constant oscillation is produced in the bottom deterministic and stochastic graph.

3.3.1 Periodogram

Because of the limitations of time-domain graphs, we use a periodogram to analyze the frequency domain of both deterministic and stochastic graphs. A periodogram is a modified form of Fourier transformation, and displays the prevalence of an observed frequency. The x-axis denotes possible frequencies of oscillation, and the y-axis denotes the prevalence of the oscillation for a specific frequency, which is referred to as the amplitude or spectral density. The data of a single species (p_z) is chosen and then standardized to enable us to make comparisons. A common trend is that the stochastic graphs had a lower amplitude and were more spread out compared to the deterministic graphs. Figure 9 is an example.



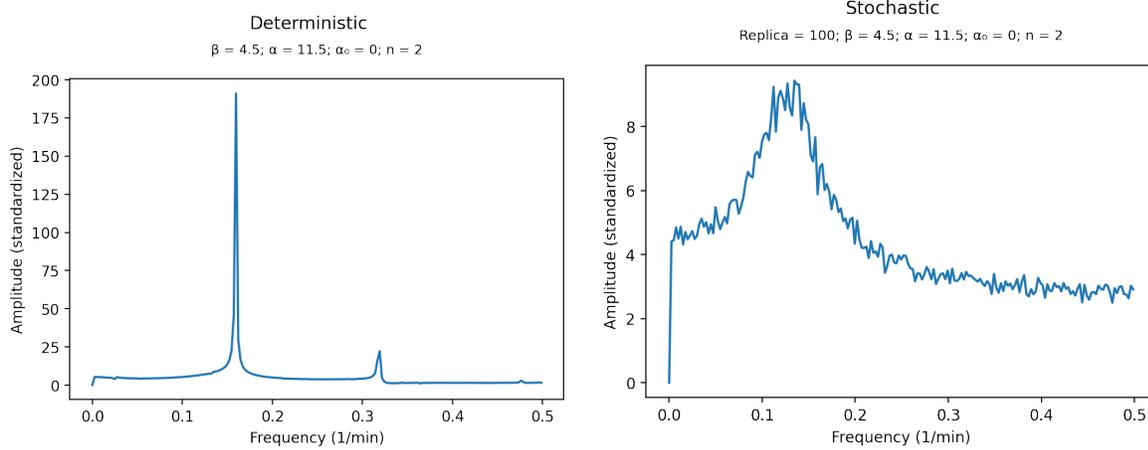


Figure 9: The deterministic model (left) and the stochastic model (right) with $a_0 = 0, \alpha = 11.5, \beta = 4.5, n = 2$. The stochastic model has a greater spread and lower amplitude compared to the deterministic model.

We first use the periodogram to classify the deterministic model into two groups of “no oscillation” and “constant oscillation.” The distinction between the two is clear, as seen by the following examples in Figure 10.

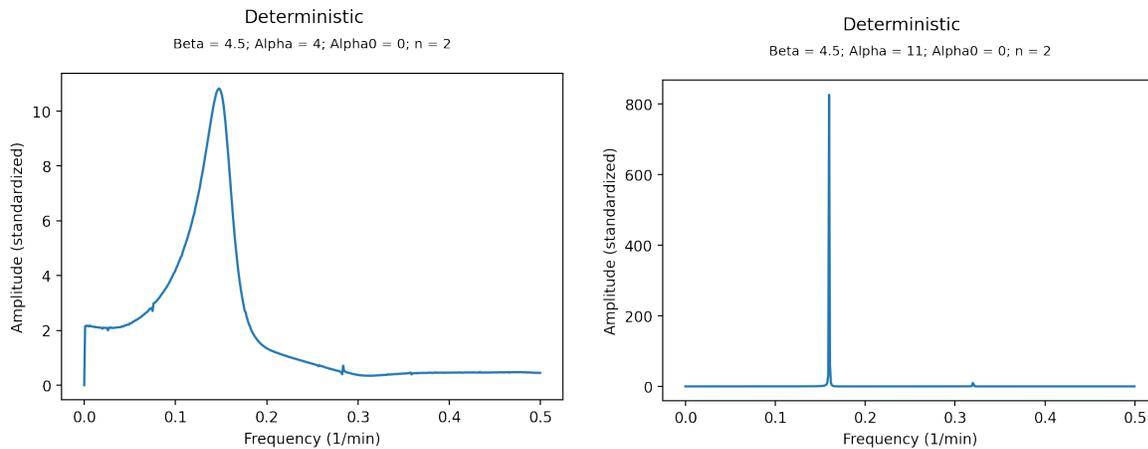


Figure 10: The graph on the left displays a wide spread, indicating how it lacks a clear oscillation and should be classified as “no oscillation.” The graph on the right shows a concentrated peak, indicating a clear, constant oscillation.



α vs β Deterministic Model (Frequency Domain Analysis)
 No Oscillation (1), Constant Oscillation (2)

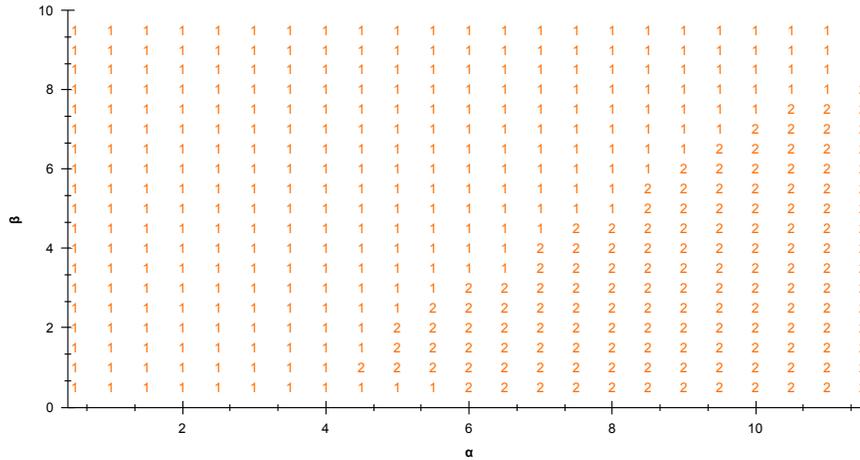


Figure 11: Using the differential equations, we tested α from 0.5 to 11.5 and β from 0.5 to 9.5. For all runs $\alpha_0 = 0$ and $n = 2$. This shows a classification of model behavior at different points in the parameter space. This was a qualitative classification of time-domain graphs, done by periodgram inspection.

We then proceed on classifying the stochastic model into the two groups of “no oscillation” and “constant oscillation.” Using the frequency domain graphs, we establish criteria to identify graphs as “no oscillation” and “constant oscillation.” This is difficult to explain without understanding the trend of α increasing under a certain β , which can be seen in Figure 12.

A spread out graph represents “no oscillation” as the wide variance in the frequency represents the effect of noise. A spiked graph represents a “constant oscillation” as it shows the prevalence of a single frequency. When analyzing the frequency-domain graphs, we cycle through values of β , and for each β , we cycle through increasing values of α . When looking at a certain β , the goal is to find a “cut-off” between “no oscillation” and “constant oscillation,” where we propose that graphs whose α is under this “cut-off” are classified as “no oscillation” and those whose α meet/exceed this cut-off are classified as “constant oscillation.” This “cut-off” is defined as when the first non-zero frequency has an amplitude of at most half of the peak. This is when a clear jump in amplitude can be seen as shown in Figure 12. Using this condition, we created a graph to show the distinction between “no oscillation” and “constant oscillation” using a frequency domain analysis in Figure 13. Notice that, in general, increasing α correlates with an increase in the strength of the oscillations, and in general, increasing β correlates with a decrease in the strength of oscillations.



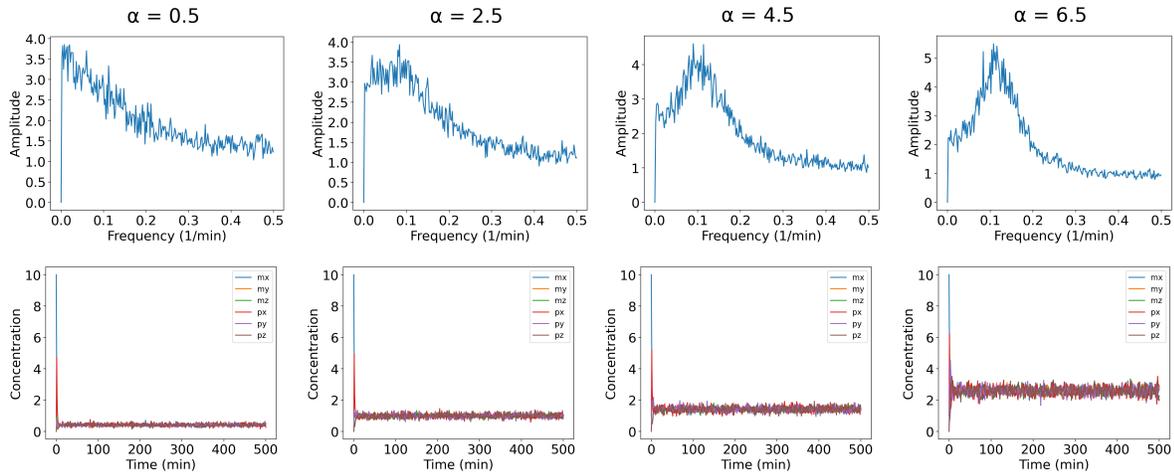


Figure 12: Top graphs show the stochastic frequency domain, while the bottom graph show the stochastic time domain. The frequency domain graphs show a trend from being spread out to being spiked, while the time domain graphs show a change from “no oscillation” to “constant oscillation.” The parameter α is increased from left to right. In general, increasing parameter α increases the strength of oscillations, both in the deterministic and stochastic models. All runs are preformed under 100 replicas, with $\alpha_0 = 0$, $\beta = 4$, and $n = 2$.

α vs β Stochastic Model (Frequency Domain Analysis)

No Oscillation (1), Constant Oscillation (2)

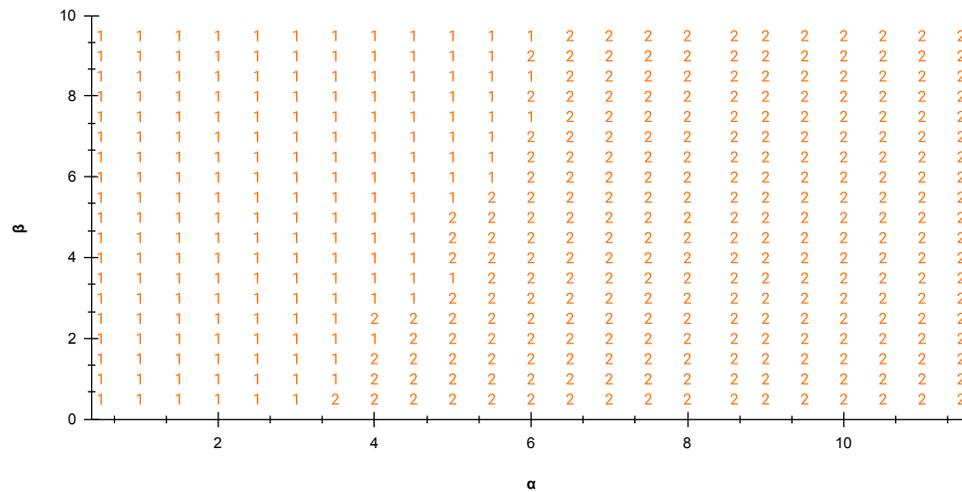


Figure 13: Using the stochastic model, we tested α (0.5 to 11.5) and β (0.5 to 9.5). For all runs (each 100 replicas), $\alpha_0 = 0$ and $n = 2$. This was a qualitative analysis of frequency-domain graphs.



3.3.2 Amplitude Analysis

To further improve our criteria for determining oscillation versus no oscillation, we look at more rigorous methods of classification. Beginning with a time domain analysis, we can examine the amplitude more accurately to see how an oscillating system's amplitude behaves. We implement a peak-finding algorithm to identify a set of peaks in each simulation trajectory, then compare this to the mean over the entire trajectory to create a set of amplitudes for a specific trajectory. The peaks are restricted to the limit cycle to account for "burn in" time due to the initial conditions. We then plot this amplitude on a heat map. In Figure 14, the bar on the right indicates that squares which are colored red exhibit the largest relative amplitude, while squares in blue exhibit small amplitudes. An unmarked white square indicates no detected oscillation, indicating that the system reaches a steady state.

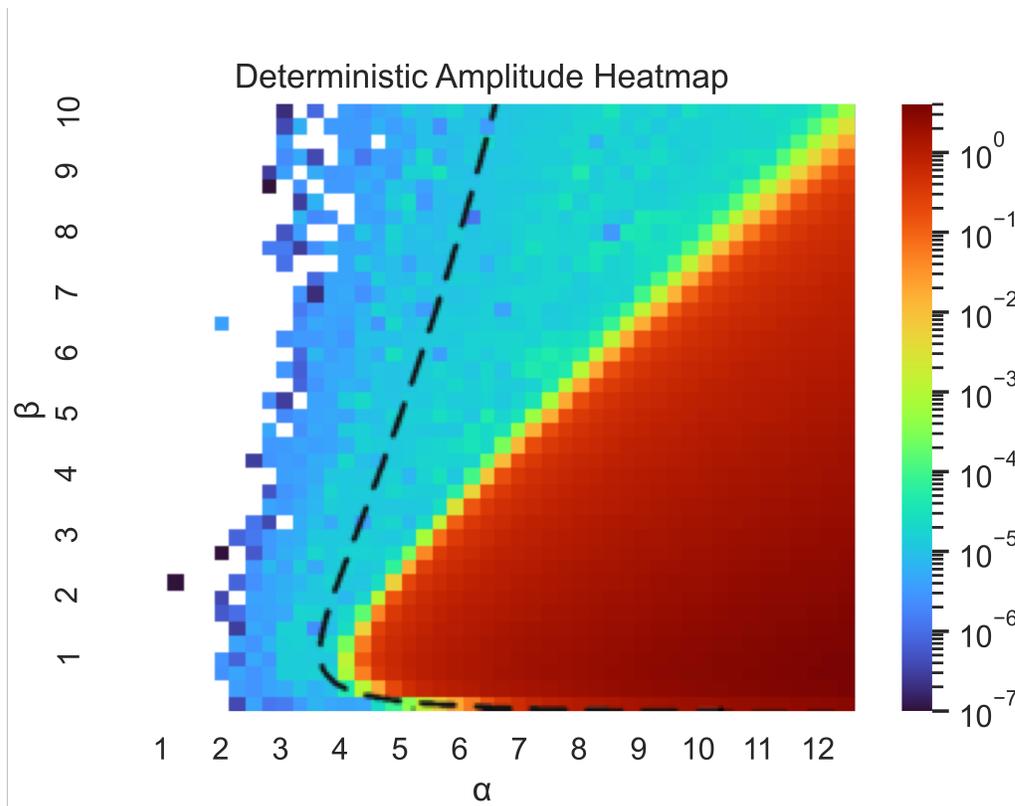


Figure 14: A log plot of average amplitude for the deterministic model, using the previously described peak-finding method of determining amplitude. The parameter space has α increasing from 0 to 12 along the x-axis and β increasing from 0 to 10 along the y-axis, both in increments of 0.25. The dashed black line marks the border where the analytical solution becomes unstable according to [1]. Throughout all runs, these variables are kept constant: $\alpha_0 = 0$, $n = 2$.

From this, we get a picture of different areas of oscillation depending on the amplitude measured. We can compare this to our previous classification of the deterministic model in Figure 11 and the solution from [1]. This amplitude heat map corroborates with the previous visual analysis done in Figure 11. Notice the discrepancy between the analytical solution and our visual analysis. We further examine this discrepancy and the systems for which our classification disagrees with the analytical solution.

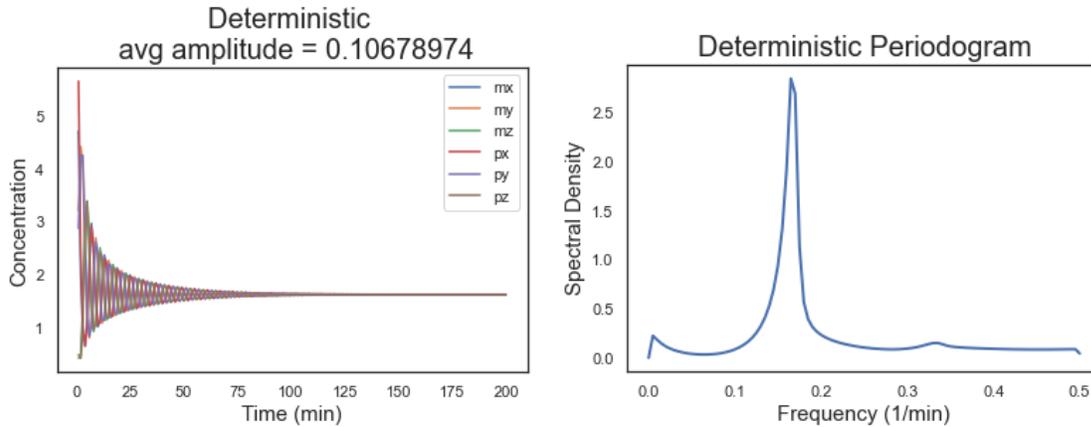


Figure 15: A time-domain trajectory plot (left) and corresponding periodogram plot (right) for the deterministic model of one example parameter set ($\alpha = 6$ and $\beta = 5$); predicted to oscillate from the analytic solution since it is to the right of the line in Figure 14.

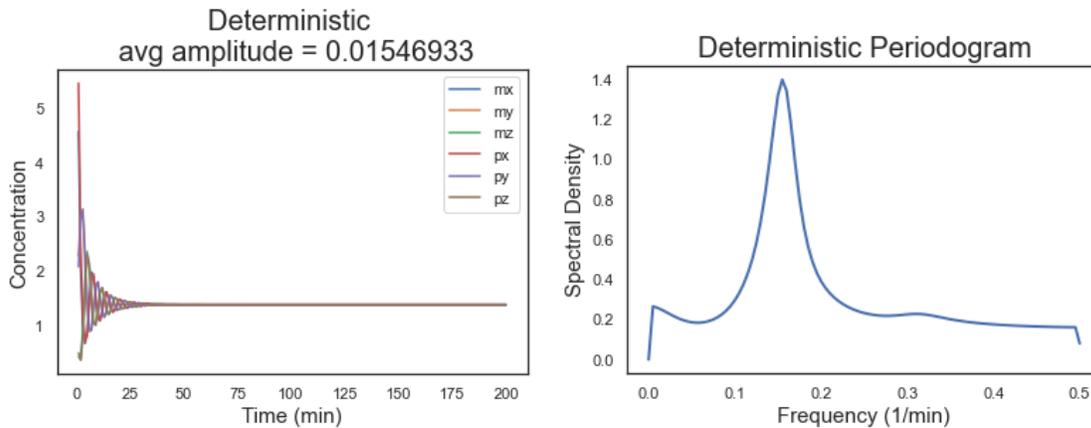


Figure 16: A time-domain trajectory plot (left) and corresponding periodogram plot (right) for the deterministic model of another example parameter set ($\alpha = 4$ and $\beta = 5$), which is predicted to have no oscillation from the analytic solution since it is to the left of the line in Figure 14.



Figure 15 is one example of a system which is inside the oscillation parameter space predicted by the analytical solution, but does not exhibit visual oscillation, and another example which is outside the analytic oscillation parameter space. Notice that compared to the first system, the second system dampens to zero quicker and the periodogram has a larger height at low values. This helps us better distinguish this subset of the parameter space which we identified as not visually oscillatory using the amplitude but which is oscillatory according to the analytical solution.

We next use this classification method to analyze the stochastic model parameter space. In the stochastic model, we implement the peak-finding algorithm on the trajectory of the average over all replicas to create our set of amplitudes. This set of amplitudes is then averaged to obtain a specific value of α and β .

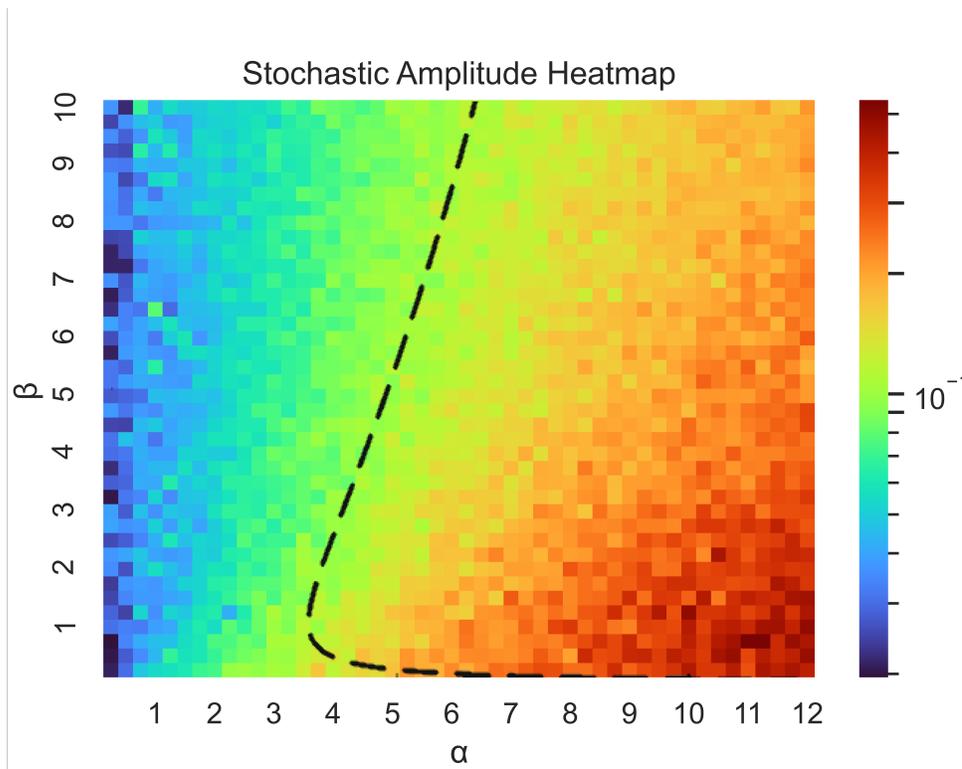


Figure 17: A log plot of average amplitude for stochastic model, using the previously described peak-finding method of determining amplitude. The parameter space has α increasing from 0 to 12 along the x-axis and β increasing from 0 to 10 along the y-axis, both in increments of 0.25. Throughout all runs, these variables are kept constant: $\alpha_0 = 0$, $n = 2$, Replica = 100, Volume = $10U$. Note that each square shows the average amplitude of oscillation from 100 simulation replicas.

First notice that across the entire parameter space, the stochastic model exhibits a smaller range of oscillation than the deterministic, indicated by the color scale on the right. There is also more noise in this plot, as expected by the nature of the stochastic models.



To compare to the plot from Figure 13, we again note the similarities. Increasing β has less effect on changing the system's classification, since the difference between values of high β and low β are less than in the deterministic model. We notice that the region before which exhibits no visual oscillation in the deterministic model, where α and β are both large, does exhibit prominent oscillation. We see that the range of parameters which permit oscillation is broader in this stochastic plot when compared to the deterministic plot from Figure 14. This also corroborates with our observations from Figures 11 and 13, where the range of oscillation in the stochastic models was larger than in the deterministic models.

4 Discussion and Conclusion

4.1 Volume in Stochastic Systems

Looking at the effect of the volume parameter from Figure 4, first notice how a system with lower volume has a more sporadic trajectory, exhibiting larger amplitude variation. In Figure 5, this correlates with a lower periodogram amplitude relative to the deterministic model, which confirms that a stochastic system with low volume has relatively more stochastic noise and thus will have less consistent oscillation. This frequency amplitude is also dependent on the rate of transcription and translation, and a higher rate of transcription or rate of translation will produce more apparent oscillations in the stochastic model.

In our stochastic models, we find that a stochastic system is more likely to oscillate than a deterministic system when the volume is small, as showcased in Figure 13 and Figure 17. This implies that in an oscillatory system, stochastic fluctuations are more likely to perpetuate oscillation rather than dampen it. We also find that in this oscillatory repressilator system, because the stochastic model has broader conditions for oscillation than the deterministic model, a physical system which is simulated with a deterministic model should be expected to oscillate more than predicted, assuming that the parameter set is near the boundary between “oscillatory” and “non-oscillatory” parameter sets. This broader condition for oscillation of the stochastic model means that it is closer to the analytical solution than the deterministic model. We predict that this is because a system which satisfies the conditions for oscillation given by the analytical solution may have stochastic fluctuations induce oscillation.

4.2 Divide Point

We can see two interesting trends from our divide points; the first trend is seen in Figure 6. When the value of α is incrementally increased from 0.5 to 7.5 and the value β is held constant at 4.5, Figure 6 shows a peculiar trend of the deterministic and stochastic system when plotted in the frequency domain. Initially, the deterministic graph is at a higher maximum amplitude. As α is increased, the stochastic and deterministic system both increase their maximum amplitude, with the stochastic rising faster. The stochastic graph slowly rises to match the graph of the deterministic, leading to an increased



alignment between the two graphs; the closest alignment is reached at the divide point. After this divide point, we note that although the deterministic and stochastic model continue to increase their maximum amplitude, the deterministic is now faster. We believe the divide point is where noise becomes relevant in the stochastic system, limiting the prevalence of certain frequencies. Hence, this causes the maximum amplitude of the stochastic system to drop off from the deterministic system for later values of α as the stochastic system is unable to maintain its original speed of increase seen in the graphs before the divide point. In biological terms, as α , the base rate of mRNA production, increases, more particles are produced in a system. A greater rate of production means there are more particles in a system, giving greater variation and noise. Eventually, this noise becomes prominent enough that the stochastic system loses its ability to match a deterministic system in terms of oscillatory behavior.

For our second trend, we compare Figure 7 to Figures 11 and 14. We notice that Figure 7 shows that divide points fall near the line dividing “no oscillation” and “constant oscillation.” This is consistent with our findings as divide points occur at low amplitudes of the frequency domain (as seen in Figure 6), showing how the system may lack a clear and consistent oscillation at divide points. Investigating this phenomenon further, we find that divide points belonged to graphs that gradually dampen. An example can be found in the left graph of Figure 8. This gradual behavior falls between two other types of graphs we notice: one where graphs quickly dampen to “no oscillation” and one where graphs dampen to a constant oscillation. In Figure 7, the area of gradually damping oscillations center between the area of quickly damping graphs on the left and the constant oscillation graphs on the right.

4.3 Oscillatory Behavior

In our classification of “no oscillation” and “constant oscillation,” we find that a periodogram classification of the deterministic graph in Figure 11 matches our amplitude-based classification in Figure 14. Both classifications differ from the analytical solution modeled in [1], which we map onto Figure 14. This creates two different dividing lines for oscillatory behavior: analytical oscillation and observable oscillation. We believe the discrepancy is caused by a difference between observable oscillation and analytical oscillation. As shown in Figures 15 and 16, there are many parameter sets which oscillate from an analytical perspective but which oscillate with an amplitude that is insignificant for physical models. This oscillation discrepancy is also present in other numerical solving methods and techniques. The reason for this discrepancy between the analytical and numerical oscillation state is something which we hope to further study in future work, but is outside the scope of this paper.

The amplitude of the oscillations from the deterministic and stochastic models is also notable. When a system experiences oscillation, a stochastic model will predict smaller oscillation than a deterministic one. However, when a system experiences damped or no oscillation, a stochastic model will predict small oscillation while a deterministic one will predict no visible oscillations. This allows us to conclude that deterministic models have



a larger range of oscillation amplitudes, as shown in Figures 14 and 17. As expected, the deterministic periodograms are stronger than the stochastic periodograms, indicating that stochastic fluctuations disrupt the oscillation frequency.

One observation we find is that the period of oscillation predicted by the deterministic models is shorter than those predicted by a correlating stochastic model. This can be seen in Figure 6 as the peak of the stochastic model is shifted to the left of the deterministic one, indicating the deterministic model has a higher frequency of oscillation and thus lower period. This effect is magnified when reaction rates α and β are high and the volume of the system is small. For similar oscillating systems, stochastic models are assumed to be more accurate, thus a deterministic model will tend to predict a faster turnover rate than the system may exhibit. We believe this arises when stochastic events permit the transcription of mRNA with a high repressor concentration. A high repressor concentration lowers the mRNA concentration, which amplifies the effects of stochastic events. Translation of mRNA in the presence of a high repressor concentration will delay the repressilator feedback loop, since the concentration of the repressor will remain high as mRNA is transcribed due to stochastic events. This leads to the repressor staying at a high concentrations for longer, which increases the period of oscillation.

4.4 Conclusion

In conclusion, our comparisons yield valuable insight into the differences between the behavior of stochastic and deterministic models of oscillatory systems. In the future, we hope to do further investigations on the effect of homogeneity of the system and modeling more complex oscillatory systems. We want to investigate inhomogeneous models and compare them to our current homogeneous models. Beyond our repressilator system, we want to investigate fluctuations and oscillatory behavior in other gene networks and expression systems.

Acknowledgments

We would like to thank Dr. Elizabeth Read for her mentorship and expertise on this paper.

References

- [1] M.B. Elowitz, S. Leibler, A synthetic oscillatory network of transcriptional regulators, *Nature*, **403** (2000) 335–338.
- [2] D.T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81** (1977) 2340–2361.
- [3] D.T. Gillespie, Stochastic Simulation of Chemical Kinetics, *Annu. Rev. Phys. Chem.*, **58** (2007) 35–55.
- [4] A. Lipshtat, A. Loinger, N.Q. Balaban, O. Biham, Genetic toggle switch without cooperative binding, *Phys. Rev. Lett.*, **96** (2006).
- [5] A. Loinger, B. Ofer, Stochastic Simulations of the Repressilator Circuit, *J. Phys. Rev. E*, **76** (2007).



- [6] C.F. Lopez, J.L. Muhlich, J.A. Bachman, P.K. Sorger, *Mol. Syst. Biol.*, **9** (2013) 646.
- [7] R. Ma, J. Wang, Z. Hou, H. Liu, Small-number effects, *Phys. Rev. Lett.*, **109** (2012).
- [8] M. Mackey, L. Glass, Oscillation and Chaos in Physiological Control Systems, *Science*, **197** (1977) 287–289.
- [9] N. Maheshri, E.K. O’Shea, Living with noisy genes: how cells function reliably with inherent variability in gene expression, *Annu. Rev. Biophys. Biomol. Struct.*, **36** (2007), 413–434.
- [10] B. Novák, J. Tyson, Design principles of biochemical oscillators, *Nat. Rev. Mol. Cell. Biol.*, **9** (2000) 981–991.
- [11] D. Schultz, A.M. Walczak, J.N. Onuchic, P.G. Wolynes, Extinction and resurrection in gene networks, *Proc. Natl. Acad. Sci.*, **105** (2008), 19165–19170.
- [12] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, J.K. Millman, N. Mayorov, A.R.J., Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, I. Polat, Y. Feng, E.W. Moore, J. Vanderplas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. Mulbregt, S. Tygier, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nat. Methods*, **17** (2000) 261–272).

Howard Ji

University High School
4771 Campus Drive
Irvine, CA 92612
E-mail: 24jihoward@gmail.com

Jackson Solheid

University of California, Irvine
280 Aldrich Hall
Irvine, CA 92697
E-mail: jsolheid@uci.edu

Junhao Gu

University of California, Irvine
280 Aldrich Hall
Irvine, CA 92697
E-mail: junhag2@uci.edu

Received: November 22, 2022 **Accepted:** August 10, 2023
Communicated by Alex Chen

