SIKKIM UNIVERSITY
Gangtok, Sikkim

Methods for Extracting Governing Equations and Dominant Patterns from Data

By
Arindam Saikia

A thesis submitted to the Department of Physics, Sikkim University
In Partial Fulfillment of the Requirements for the Degree of Master of Science

Supervised by :
Dr. Rupak Mukherjee

11 July, 2024

6 माइल, सामदुर, तादोंग -737102
गंगटोक, सिक्किम, भारत
फोन-03592-251212, 251415, 251656
टेलीफैक्स -251067
वेबसाइट - www.cus.ac.in

6th Mile, Samdur,Tadong -737102
Gangtok, Sikkim, India
Ph. 03592-251212, 251415, 251656
Telefax: 251067
Website: www.cus.ac.in

सिक्किम विश्वविद्यालय
SIKKIM UNIVERSITY

(भारत के संसद के अधिनियम द्वारा वर्ष 2007 में स्थापित और नैक (एनएएसी) द्वारा वर्ष 2015 में प्रत्यायित केंद्रीय विश्वविद्यालय)
(A central university established by an Act of Parliament of India in 2007 and accredited by NAAC in 2015)

CERTIFICATE

This is to certify that this dissertation entitled "**Methods for Extracting Governing Equations and Dominant Patterns from Data**" submitted to **Sikkim University** in partial fulfilment of the requirement for the award of the degree of **Master of Science in Physics** is a research done by **Arindam Saikia** during the end semester 2022-2024 in the **Department of Physics** and that this dissertation has not formed the basis for the award of any Degree/Diploma/Associateship/Fellowship and any other similar title.

Dr. Ajay Tripathi
11/07/24

अध्यक्ष
Head
भौतिकी विभाग
Department of Physics
सिक्किम विश्वविद्यालय
Sikkim University

Head of the Department

Dr. Rupak Mukherjee
11-July-2024

Supervisor

Acknowledgments

First and foremost, I want to express my utmost gratitude to my advisor, Dr. Rupak Mukherjee, for his invaluable guidance, insightful feedback, and unwavering support throughout this project. His expertise and encouragement have been crucial in the completion of this thesis.

I would also like to extend my sincere thanks to the faculty members of the Department of Physics at Sikkim University: Dr.Ajay Tripathi, Prof.Subir Mukhopadhyay, Dr.Amitabha Bhattacharyya, Dr.Hemam Dinesh Singh, and Dr.Dhurba Rai, for their guidance and motivation in fostering my deep interest in this field.

I am deeply grateful to my colleagues and friends at Sikkim University, whose camaraderie and intellectual discussions provided a stimulating and enjoyable environment for research.

Lastly, I dedicate this thesis to everyone who believed in me and supported me through this journey. Your faith in me has been my greatest motivation.

Arindam Saikia
July 2024

Declaration

I hereby declare that the work presented in this thesis is my own effort and has not been submitted elsewhere for any other degree or qualification. I certify that the thesis is entirely my own work and does not contain any unacknowledged work from any other sources.

Arindam Saikia

Arindam Saikia
Physics Department
Sikkim University
11 July, 2024

Contents

Abstract

This thesis presents a comprehensive exploration of various methods to distill or closely approximate the governing equations behind a diverse array of dynamical systems. The systems considered range from those governed by simple linear equations to more complex non-linear systems, also including the Kuramoto model and the Kuramoto model with periodically varying amplitudes.

We begin by employing a simple linear regression algorithm to obtain initial close approximations of the governing equations. This foundational method serves as a baseline for comparison with more advanced techniques. Following this, we introduce and apply L1 regularization (lasso regression), which helps in enhancing the sparsity of the model by penalizing the absolute values of the coefficients. This method is particularly useful for identifying the most relevant variables and interactions in the system.

To assess the efficacy of each method, we compare the loss functions, providing a quantitative measure of the accuracy and robustness of the approximations. The comparison of loss functions helps in evaluating which method yields the best performance for different types of systems and data conditions.

Furthermore, we delve into a range of decomposition methods such as Singular Value Decomposition (SVD), Proper Orthogonal Decomposition (POD) and Dynamic Mode Decomposition (DMD). These techniques are instrumental in extracting the dominant modes from the data, which are crucial for understanding the primary dynamics of the system. By focusing on these dominant modes, we are able to derive simplified yet accurate governing equations that capture the essential behavior of the system.

Chapter 1

Introduction

Understanding and describing behavior in a concise mathematical manner involves articulating the underlying equations that govern its evolution over time. These systems can vary widely in nature, from simple oscillator-like systems with linear equations to more complex and chaotic systems described by nonlinear equations. To understand and predict their future behavior and states, it is crucial to comprehend these governing equations.

The traditional approach to this problem relies on using fundamental first principles to derive equations that best model observed phenomena. While this method has been effective for many systems and has been instrumental in discovering the principles behind numerous physical phenomena, it has limitations. It can be challenging to apply when systems become unpredictable or complex. Fortunately, modern advancements in numerical methods and mathematics are addressing this issue. We are now seeing the development of algorithms designed to tackle the challenge of equation discovery, such as pySINDY by Steve L. Brunton et al., which have proven successful in identifying governing equations from observational data. Large datasets and sophisticated algorithms are being utilized to uncover the relationships and interactions driving these systems.

In this thesis, we explore and build upon some of the techniques and methods for extracting such equations from provided data. Our goal is to develop a robust and reliable algorithm that can take observational data and predict the underlying equations. These predictions can then be verified through experiments and cross-testing.

1.1 Background

A vast number of systems observed in nature are inherently dynamical, evolving over time according to specific relationships dictated by the system's nature and its constituents. Deciphering these relationships and the driving equations governing these systems is of paramount importance in understanding their behavior. Classical models, such as Newton's laws of motion, have provided comprehensive explanations for a wide range of phenomena, from the motion of planets within our solar system to the execution of space missions reaching distant regions. Similarly, the Navier-Stokes equations have enabled an in-depth exploration of fluid dynamics, elucidating the complex physics governing fluid behavior.

However, despite the success of these classical approaches, there are scenarios where traditional theoretical models fall short. Certain systems possess underlying mechanisms that remain poorly understood or are too intricate to be captured analytically. In such contexts, data-driven methods emerge as powerful alternatives. These approaches aim to extract governing relationships directly from observational data without relying on predefined theoretical frameworks, offering novel insights into complex systems.

Data-driven equation discovery algorithms have demonstrated their capability to handle real-world noisy and potentially incomplete data, making them indispensable tools for analyzing a wide array of dynamical systems. By leveraging these methods, researchers can uncover the underlying dynamics of systems where classical models may not be applicable or sufficient. This versatility positions data-driven approaches as critical components in the toolkit for modern scientific investigation.

In contrast, data-driven methods aim to extract these governing equations directly from data, without relying heavily on predefined theoretical models. This is particularly useful for complex systems where the underlying mechanisms are not fully understood or are too intricate to model analytically. Data-driven approaches can also accommodate noisy and incomplete data, making them versatile tools for real-world applications.

1.2 Research Problem

This thesis primarily addresses the challenge of developing algorithms that effectively minimize the loss function. The loss function quantifies the accuracy of a given set of equations in recreating observed data by measuring the discrepancies between the predicted and actual data. Minimizing this function is crucial for ensuring that the derived equations accurately represent the underlying dynamics of the system.

Another significant challenge tackled in this thesis is the extraction of the dominant and most important features from the data while filtering out noisy disturbances. These disturbances often arise from natural interferences and can obscure the true underlying patterns. Effective data-driven methods must be capable of distinguishing between meaningful signals and noise, ensuring that the derived models focus on the essential dynamics of the system without being misled by extraneous fluctuations.

Addressing these challenges involves the development and application of sophisticated algorithms that can balance the trade-off between accuracy and complexity. The ability to minimize the loss function while effectively handling noisy data is a critical aspect of advancing the field of data-driven equation discovery. This thesis aims to contribute to this advancement by proposing novel methodologies and demonstrating their effectiveness through various applications.

1.3 Objectives

The specific objectives of this research are:

1. To employ simple linear regression algorithms to obtain approximations of governing equations.

2. To apply L1 regularization (lasso regression) to enhance model sparsity and identify the most relevant variables of the system.

3. To compare the effect of sparsity coefficient on the final model.

4. To utilize decomposition methods, such as Proper Orthogonal Decomposition (POD) and Dynamic Mode Decomposition (DMD), to remove noise, extract dominant modes and derive simplified governing equations for those particular modes..

1.4 Thesis Organization

The structure of this thesis is as follows:

- **Chapter 2: Theoretical Framework** - This chapter reviews briefly overviews some of the theoretical necessary to precede the upcoming discussions and terms used around them.
- **Chapter 3: Methodology** - This chapter details the methodologies employed in this research, including the theoretical foundations and implementation of linear regression, L1 regularization, and decomposition methods like POD and DMD.
- **Chapter 4: Results and Discussion** - This chapter applies the different methods at hand to data sets and presents the results. Each problem introducing some new complexity level to be solved.

By structuring the thesis in this manner, we provide a comprehensive exploration of data-driven methods for extracting governing equations from dynamical system data, offering both theoretical insights and practical applications.

Chapter 2

Theoretical Framework

2.1 Linear Regression

Linear regression is a foundational statistical method used to model the relationship between a dependent variable and one or more independent variables. In the context of this work, linear regression can be employed to approximate the governing equations by fitting linearly the columns in the provided "library matrix" of possible terms, to the observed data.

The primary objective of linear regression is to find the coefficient vector "b" that minimizes the following loss function:

$$\text{Loss Function} = \text{Minimize}\|y - Xb\|^2 \tag{2.1}$$

Here "y" represents the Data Matrix, and "X" represents the library matrix.

To apply linear regression, we first collect time-series data from the dynamical system. This data is then divided into training and testing sets. The training set is used to fit the model, while the testing set is used to validate its predictive power. The coefficients β_i are estimated using the least squares method, which minimizes the sum of the squared residuals:

$$\text{Minimize} \quad \sum_{i=1}^{m} \left(y_i - \beta_0 - \sum_{j=1}^{n} \beta_j x_{ij} \right)^2 \tag{2.2}$$

2.2 L1 Regularization

L1 Regularization, also known as Lasso (Least Absolute Shrinkage and Selection Operator), is a regression technique that adds a penalty equal to the absolute value of the magnitude of coefficients to the loss function. This method helps in feature selection by shrinking some coefficients to zero, effectively reducing the number of variables.

The objective function for Lasso regression is given by:

$$\text{Loss Function} = \|y - Xb\|^2 + \lambda \sum_{i=1}^{n} |b_i| \tag{2.3}$$

where λ is the regularization parameter (or Sparsity Strength) that controls the amount of shrinkage applied to the coefficients. Hence, making the resulting model sparse in its usage of terms with fewer terms and fewer degrees of freedom, which in turn makes it much more easily interpretable.

To apply Lasso regression, we need to select an appropriate value for λ. This can be achieved through cross-validation, where the data is divided into multiple folds, and the model is trained and validated on different subsets of the data. The value of λ that minimizes the cross-validation error is chosen as the optimal parameter.

Then, regression is performed as before.

2.3 Decomposition Methods

Matrix Decomposition methods are used to break down complex systems into simpler, more understandable components. These methods are particularly useful in identifying dominant modes in the data.

2.3.1 Proper Orthogonal Decomposition (POD)

POD (Proper Orthogonal Decomposition) is a powerful mathematical tool that can be employed in a range of situations where there is a need for analyzing big data sets, extracting essential information from complex dynamics, reducing computation cost of algorithms, eliminating noise from data etc.

The mathematical foundation for POD was established in statistics where Principal Component Analysis or PCA was formulated. PCA aimed at reducing the dimensionality of data while preserving the most significant features. In statistical terms, the principal components represent the directions of maximum variance, while in POD, the modes capture dominant structures or patterns in the data. Thus, POD comes in handy whenever there are large and complex datasets with high dimensionality and a more concise representation is desired to gain insights into the underlying dynamics of the system.

Theory

POD involves finding a set of orthogonal basis functions that capture the dominant patterns in the data. The data matrix $\mathbf{X}$ is decomposed as:

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T \tag{2.4}$$

where $\mathbf{U}$ contains the orthogonal modes, $\boldsymbol{\Sigma}$ is a diagonal matrix of singular values, and $\mathbf{V}$ contains the temporal coefficients.

2.3.2 Dynamic Mode Decomposition (DMD)

Dynamic Mode Decomposition (DMD) is a data-driven technique used to extract dynamic modes from time-series data. It is particularly useful for systems with periodic or quasi-periodic behavior.

Theory

DMD decomposes the data into a set of modes, each associated with a specific frequency and growth/decay rate. The data matrix $\mathbf{X}$ and its shifted version $\mathbf{X}'$ are related as:

$$\mathbf{X}' = \mathbf{A}\mathbf{X} \tag{2.5}$$

where $\mathbf{A}$ is the dynamic matrix. The eigenvalues and eigenvectors of $\mathbf{A}$ provide the DMD modes and their corresponding dynamics.

Chapter 3

Methodology

This chapter outlines the approach taken to ensure the algorithms function seamlessly with the prepared or recorded data.

3.1 Data Preparation

Five distinct datasets, generated using known governing equations, were prepared. These datasets were then formatted into matrices, with each column representing a time snapshot in the case of dynamical systems.

Data preparation methodology for all the problem cases are stated below :

3.1.1 Equation with Linear terms

Preparation of this set of data is pretty straightforward. We only have to know the following equation :

$$u = 11x + 5y \tag{3.1}$$

and store the resulting data into matrix format.

3.1.2 Equation with Non-linear terms

Again, we take a similar approach as the linear equation before and use the following equation :

$$u = 11xy + 5y^2 \tag{3.2}$$

3.1.3 Kuramoto Model Order Parameter

The Kuramoto model is a mathematical model used to describe the synchronization of a large set of coupled oscillators. The model is governed by the Kuramoto order parameter, which quantifies the degree of synchronization in the system.

Kuramoto Model Description

The Kuramoto model consists of N oscillators, each described by a phase angle $\theta_i(t)$ that evolves according to the following differential equation:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i)$$

where ω_i is the natural frequency of the i-th oscillator, and K is the coupling strength between oscillators.

Order Parameter Definition

The order parameter $r(t)$ is a measure of the average phase coherence of the system and is defined as:

$$r(t)e^{i\psi(t)} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j(t)}$$

where $r(t)$ represents the magnitude of the order parameter, and $\psi(t)$ is the average phase of the oscillators.

Order Parameter Interpretation

The magnitude of the order parameter $r(t)$ ranges from 0 to 1:

- When $r(t) \approx 0$, the oscillators are in a desynchronized state.
- When $r(t) \approx 1$, the oscillators are perfectly synchronized.

The Kuramoto model and its order parameter are crucial for understanding the dynamics of collective behavior in systems of coupled oscillators, such as the synchronization of neurons or the coherence of fireflies.

3.2 Library Construction

To analyze the data effectively, a set of basis functions was constructed for the regression model. The library included a range of terms from varying powers of time "t", to trigonometric terms with different frequencies and Gaussian functions to capture various aspects of the data.

3.3 Data Normalization

Normalization was performed to standardize the features in the library and the target data, ensuring that all terms were on a comparable scale.

3.4 Lasso Regression Model

Lasso Regression was applied to the normalized data using L1 regularization to identify significant terms and constrain the magnitude of model coefficients.

This method was opted for in place of linear regression for increasingly complex systems due to the following problem of parsimonious modelling and avoiding over-fitting.

3.4.1 Parsimonious Modelling

As shown in Figure 3.1, we see that as we increase the complexity of the governing equations, we expect a decrease error between the predicted data and and the original data. With the most complex model given by the linear regression algorithm. But this creates a problem of over-fitting and doesn't portray a true representation of the terms involved in the system's dynamics. So, in order to counter this problem we need to introduce some kind of method that has sparsity promotion built into it that we can control, and hence affect the complexity of the resulting model directly. This allows us to select our target point on the Pareto Curve and choose a Sparsity strength which brings the model to the target fit point.

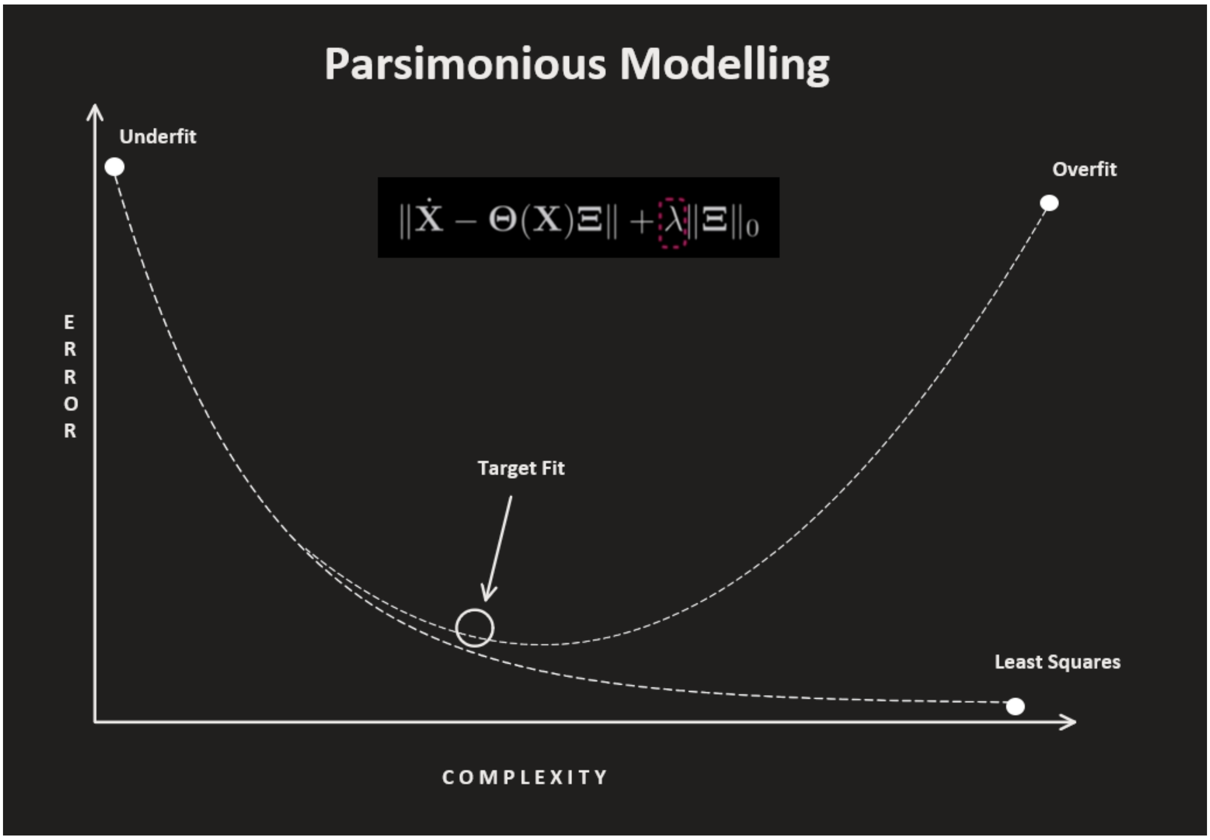

Figure 3.1: Pareto Principle for Parsimonious Modelling

3.5 Coefficient Analysis and Visualization

The coefficients obtained from the Regression were analyzed to interpret the contribution of each term in the library to the data. Visualization was performed to display the magnitude of the coefficients.

Chapter 4

Results and Discussion

4.1 Linear Regression

4.1.1 Problem 1 : $u = 11x + 5y$ data

In this section, we utilize the linear regression algorithm on data generated from the following equation : $u = 11x + 5y$ as an example. The following figures illustrate the data and results obtained.

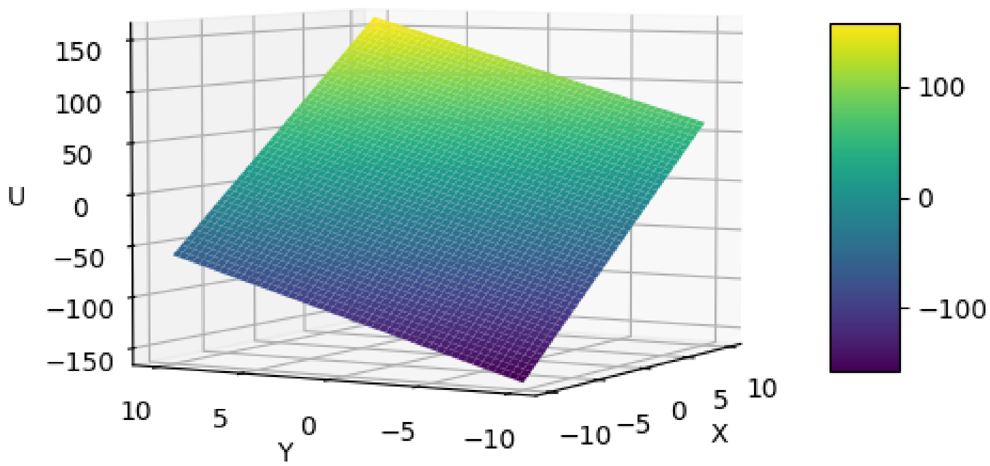

Figure 4.1: Surface Plot of the Equation

The above figure depicts the surface made by the given equation, which appears flat due to absence of non linear terms like y^2 or xy etc.

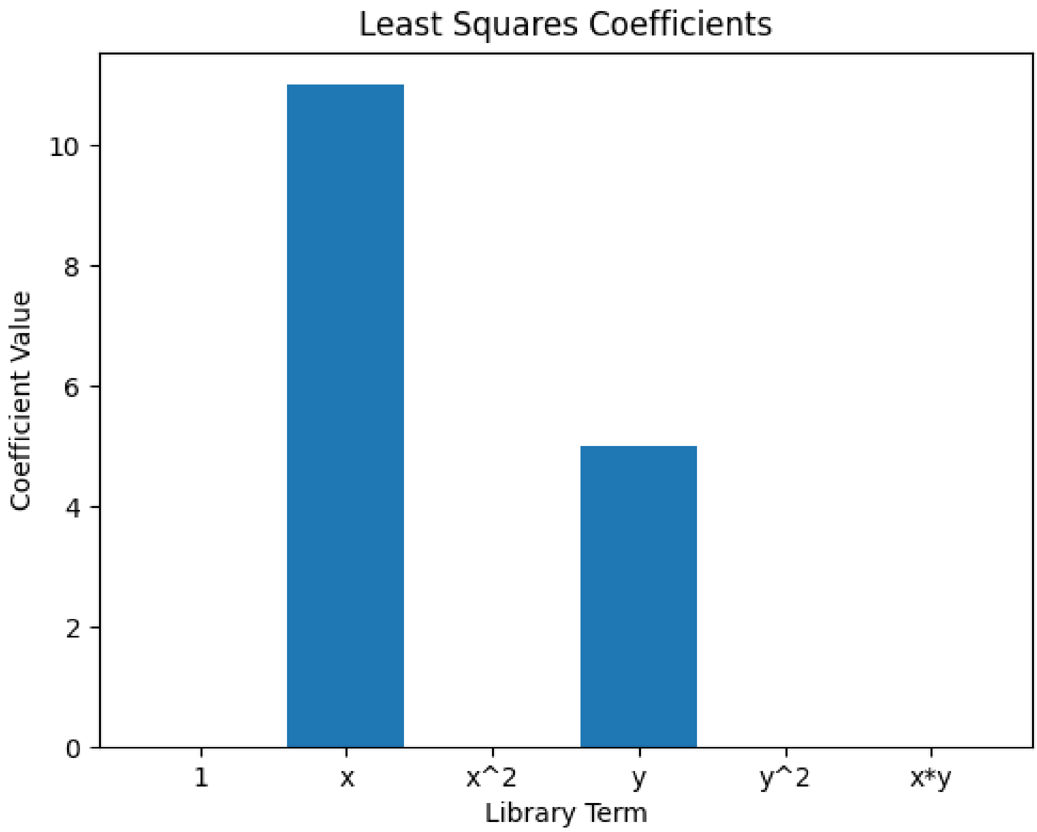

Figure 4.2: Scatter plot of the data points and the fitted regression line.

Figure 4.2 shows the scatter plot of the data points with the fitted linear regression line. The line represents the equation $u = 11x + 5y$, which was obtained using the least squares method.

The regression analysis indicates that the coefficients for x and y are approximated near 11 and 5 respectively.

4.1.2 Problem 2 : $u = 11xy + 5y^2$ data

In this section, we examine the linear regression case where $u = 11xy+5y^2$. The following figures illustrate the data and results obtained.

4.3 depicts the curved surface made by the given equation, which is a result of the presence of terms like y^2 and xy this time.

Figure 4.4 shows the plot of the data points with the fitted regression surface. The surface represents the equation $u = 11xy + 5y^2$, which was obtained using the nonlinear least squares method.

The regression analysis indicates that the coefficients for y^2 and xy are approximated near 5 and 11 respectively.

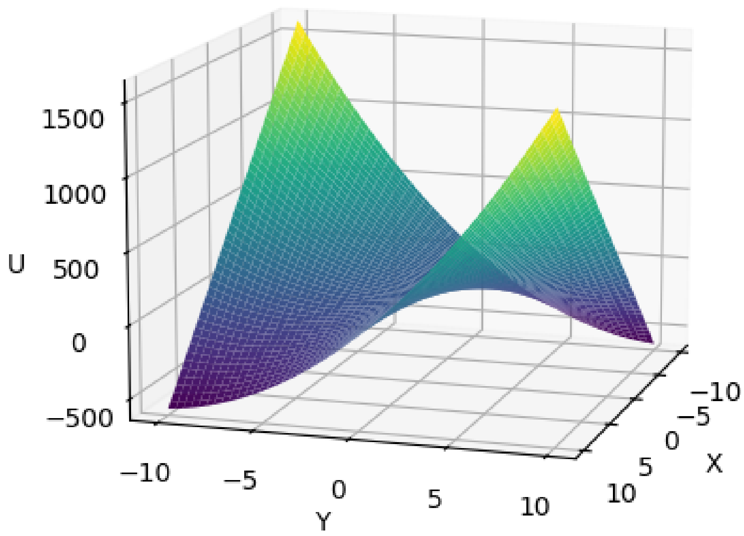

Figure 4.3: Surface Plot of the Equation

4.2 Lasso Regression

4.2.1 Problem 3 : Phase Coupled Kuramoto Model's Order Parameter

As discussed earlier, In this section, we discuss the analysis of the Kuramoto order parameter data using the L1(Lasso) algorithm. The following figures illustrate the results obtained.

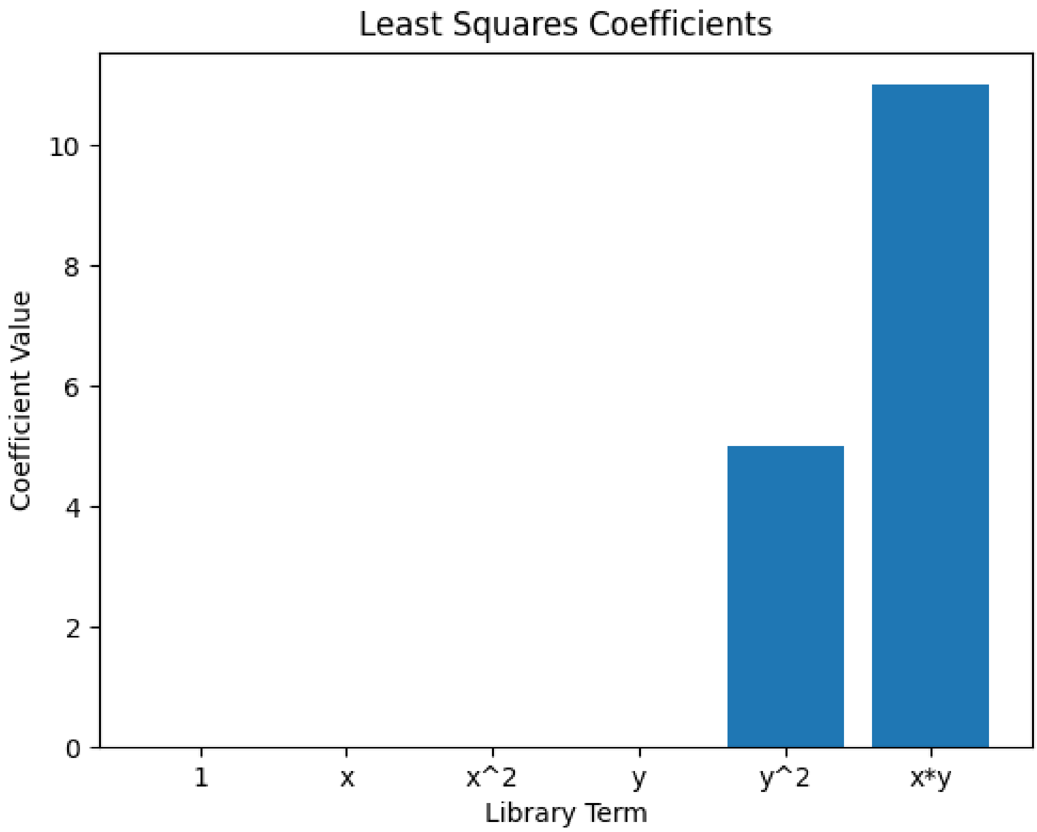

Figure 4.4: Plot of the data points and the fitted regression surface.

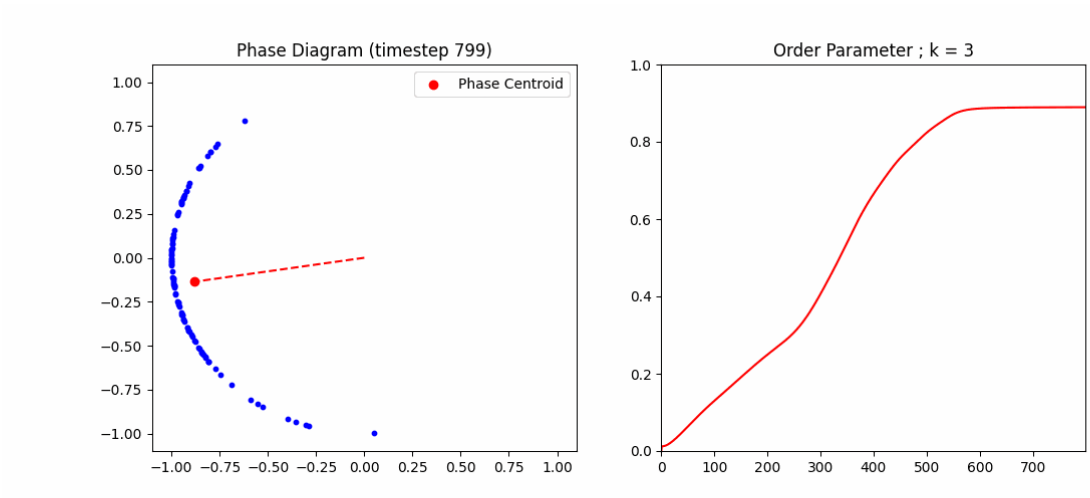

Figure 4.5: Order parameter trend

Figure 4.6 shows the plot of the Kuramoto order parameter over time. The L1 algorithm was applied to analyze the synchronization behavior of coupled oscillators.

4.8 shows the coefficients discerned by the Lasso algorithm.

The analysis reveals that the Kuramoto order parameter reaches a stable value, indicating synchronization among the oscillators. The L1 algorithm effectively captures the transitions and stable states in the synchronization process.

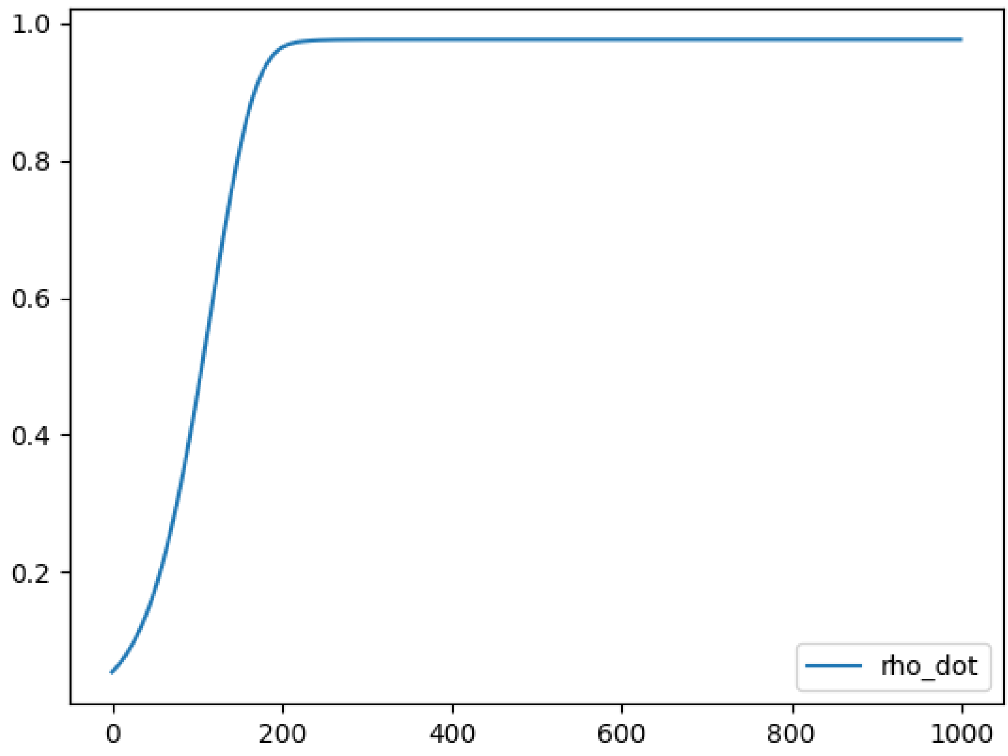

Figure 4.6: Kuramoto order parameter over time.

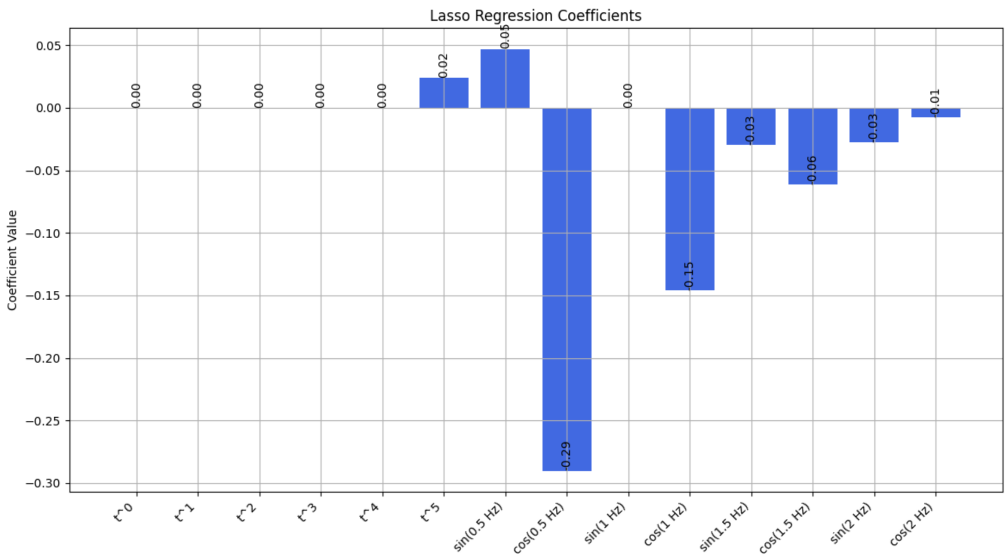

Figure 4.7: Weightage of terms obtained from Lasso Regression

4.2.2 Problem 4 : Amplitude-Variated Phase coupled Kuramoto Model's Order Parameter

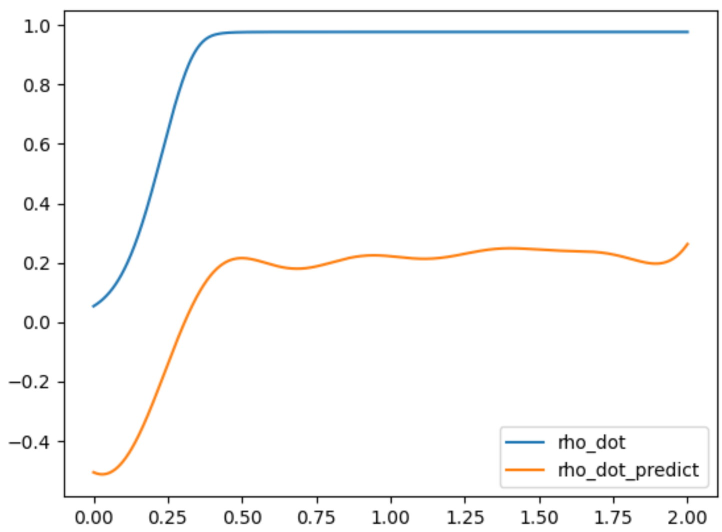

Figure 4.8: Reconstruction of Data from obtained terms

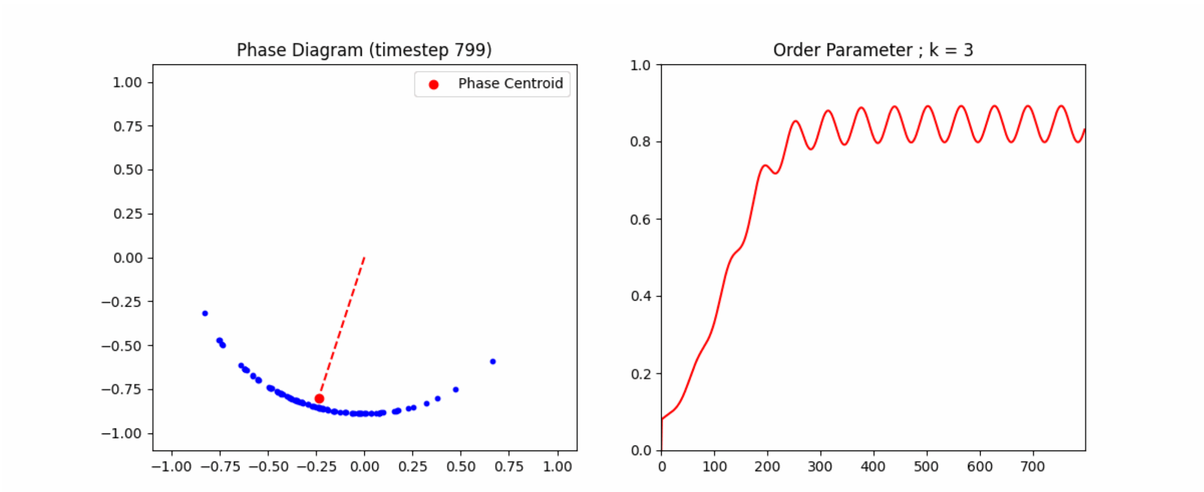

Figure 4.9: Order parameter trend

This section presents the analysis of amplitude-varied Kuramoto data using the L1 algorithm. The following figures illustrate the results obtained.

Figure 4.8 shows the plot of amplitude variations in the Kuramoto data. The L1 algorithm was used to analyze the impact of amplitude changes on the synchronization behavior.

The analysis demonstrates that varying the amplitude affects the synchronization dynamics. The L1 algorithm helps in identifying critical points where amplitude changes lead to desynchronization or resynchronization among oscillators.

Figures 4.10 to 4.19 shows the changes in the coefficient amplitudes when sparsity strength of different values are used in the algorithm.

Also, Figures 4.20 to 4.29 shows the resulting recreation of the order parameter curve using these terms respectively for different strengths.

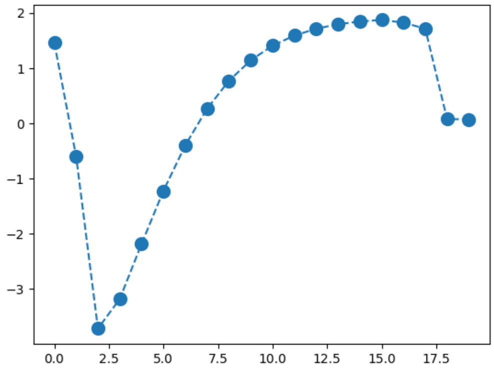

Figure 4.10: Sparsity 0

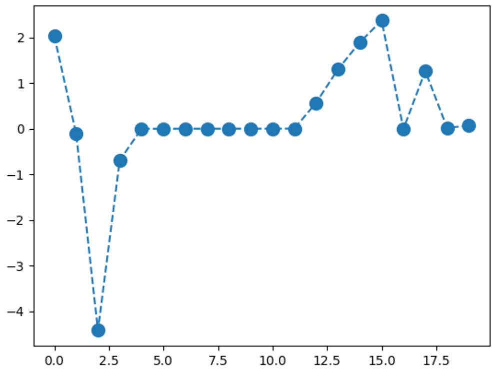

Figure 4.11: Sparsity 0.2

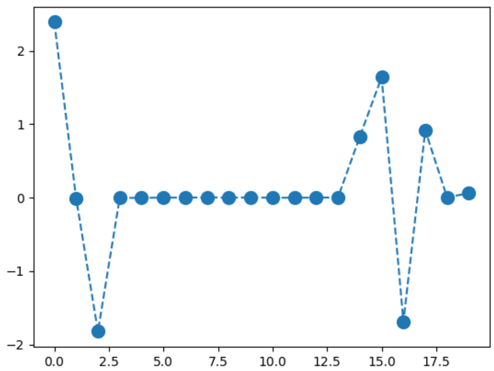

Figure 4.12: Sparsity 0.5

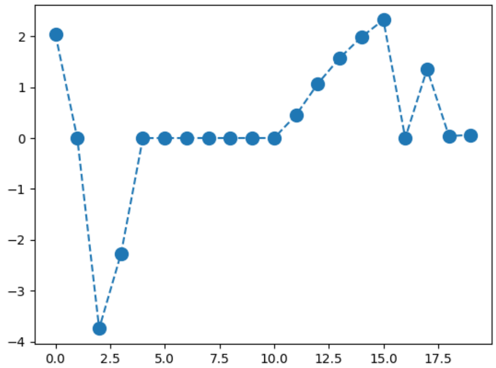

Figure 4.13: Sparsity 0.1

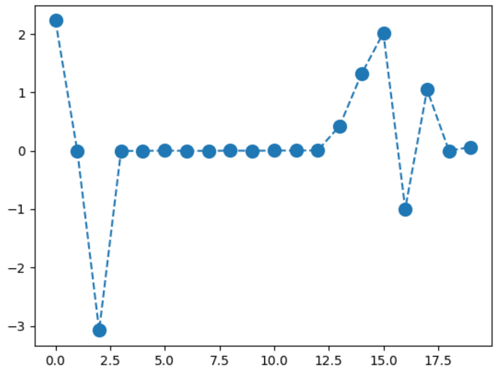

Figure 4.14: Sparsity 0.4

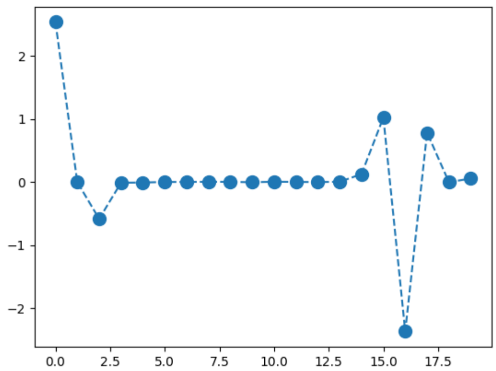

Figure 4.15: Sparsity 0.6

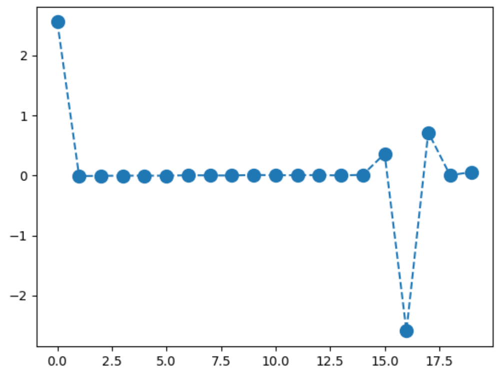

Figure 4.16: Sparsity 0.7

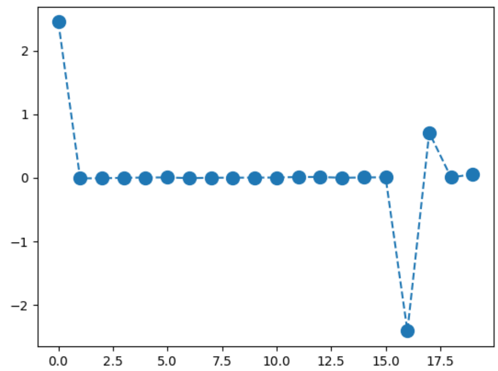

Figure 4.18: Sparsity 0.8

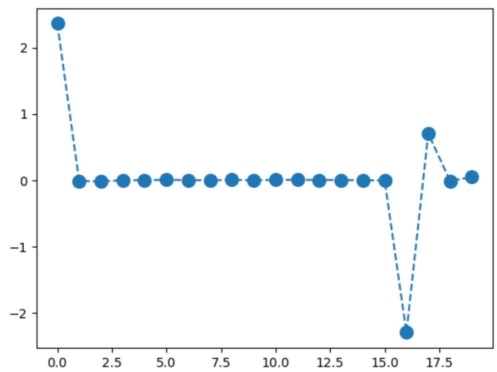

Figure 4.17: Sparsity 0.9

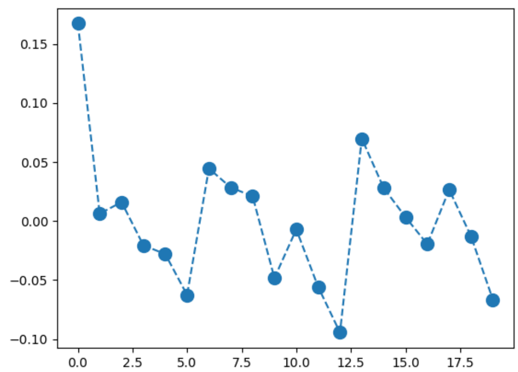

Figure 4.19: Sparsity 10

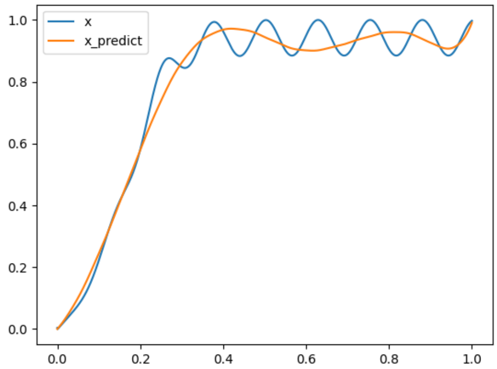

Figure 4.20: Sparsity 0

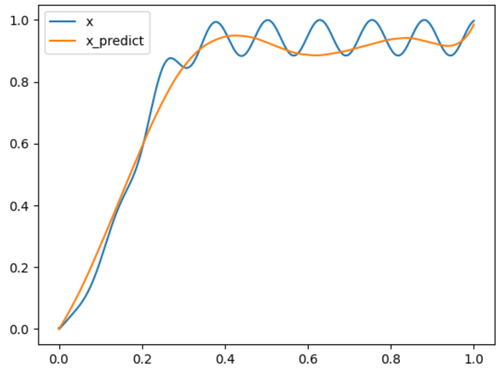

Figure 4.23: Sparsity 0.1

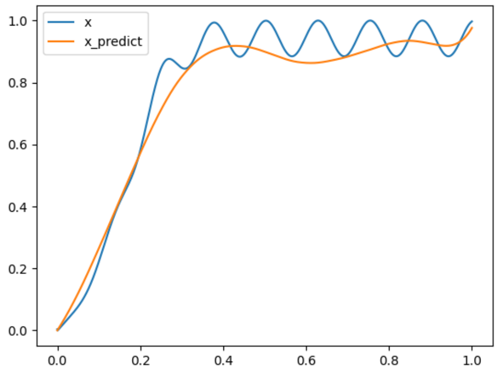

Figure 4.21: Sparsity 0.2

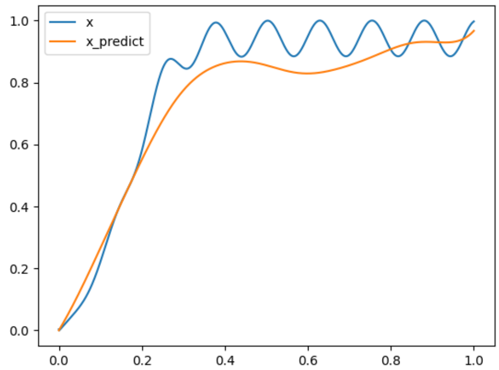

Figure 4.24: Sparsity 0.4

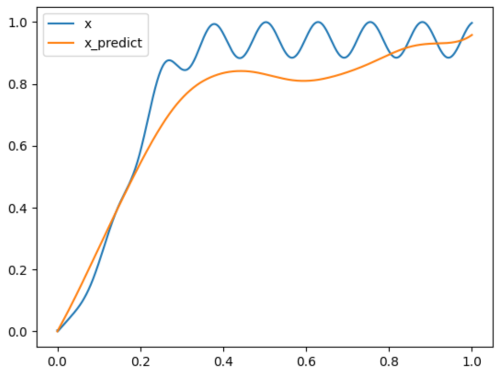

Figure 4.22: Sparsity 0.5

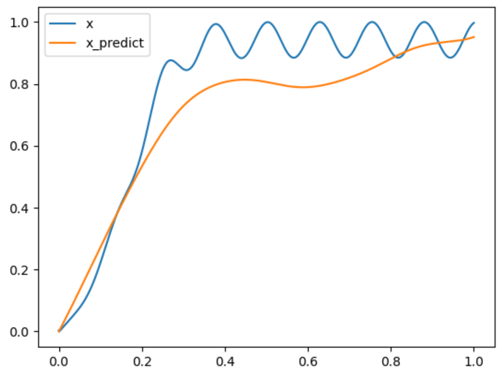

Figure 4.25: Sparsity 0.6

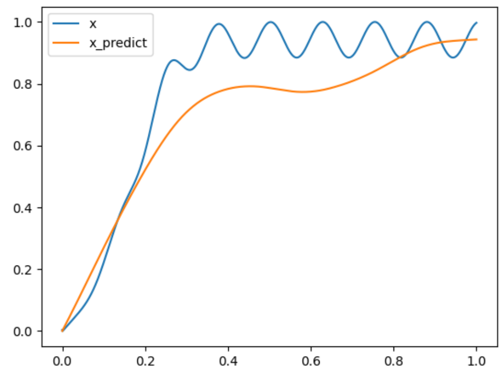

Figure 4.26: Sparsity 0.7

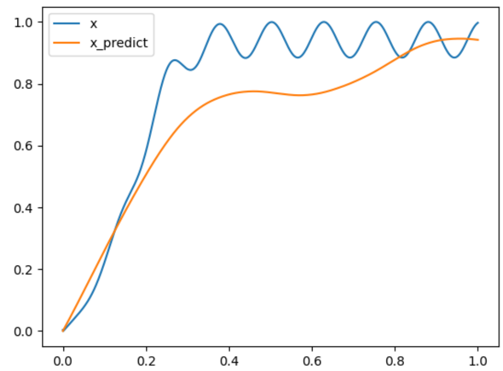

Figure 4.28: Sparsity 0.8

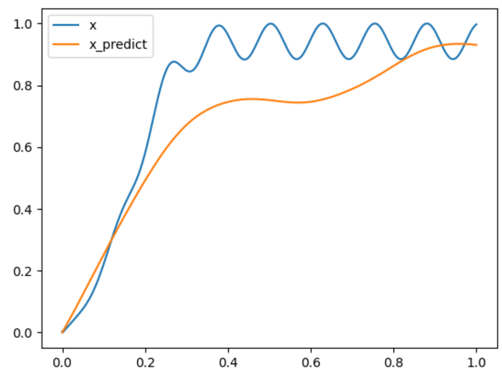

Figure 4.27: Sparsity 0.9

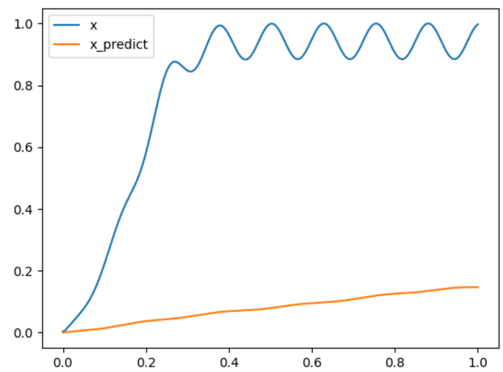

Figure 4.29: Sparsity 10

4.3 POD Application

4.3.1 Problem 5 : Gaussian + Sine profile with Noise data

This subsection presents the application of the Proper Orthogonal Decomposition (POD) method and the Dynamic Mode Decomposition (DMD) method to separate out the dominant modes in the data and remove noise. The following figures illustrate the results obtained.

Figures 4.30 to 4.34 shows the POD modes obtained using the L1 algorithm. The plots demonstrate the dominant modes and their contributions to the overall data variability.

The analysis shows that the L1 algorithm effectively identifies the dominant modes, capturing the essential features of the data. The POD method, combined with the L1 algorithm, provides a powerful tool for data dimensionality reduction and feature extraction.

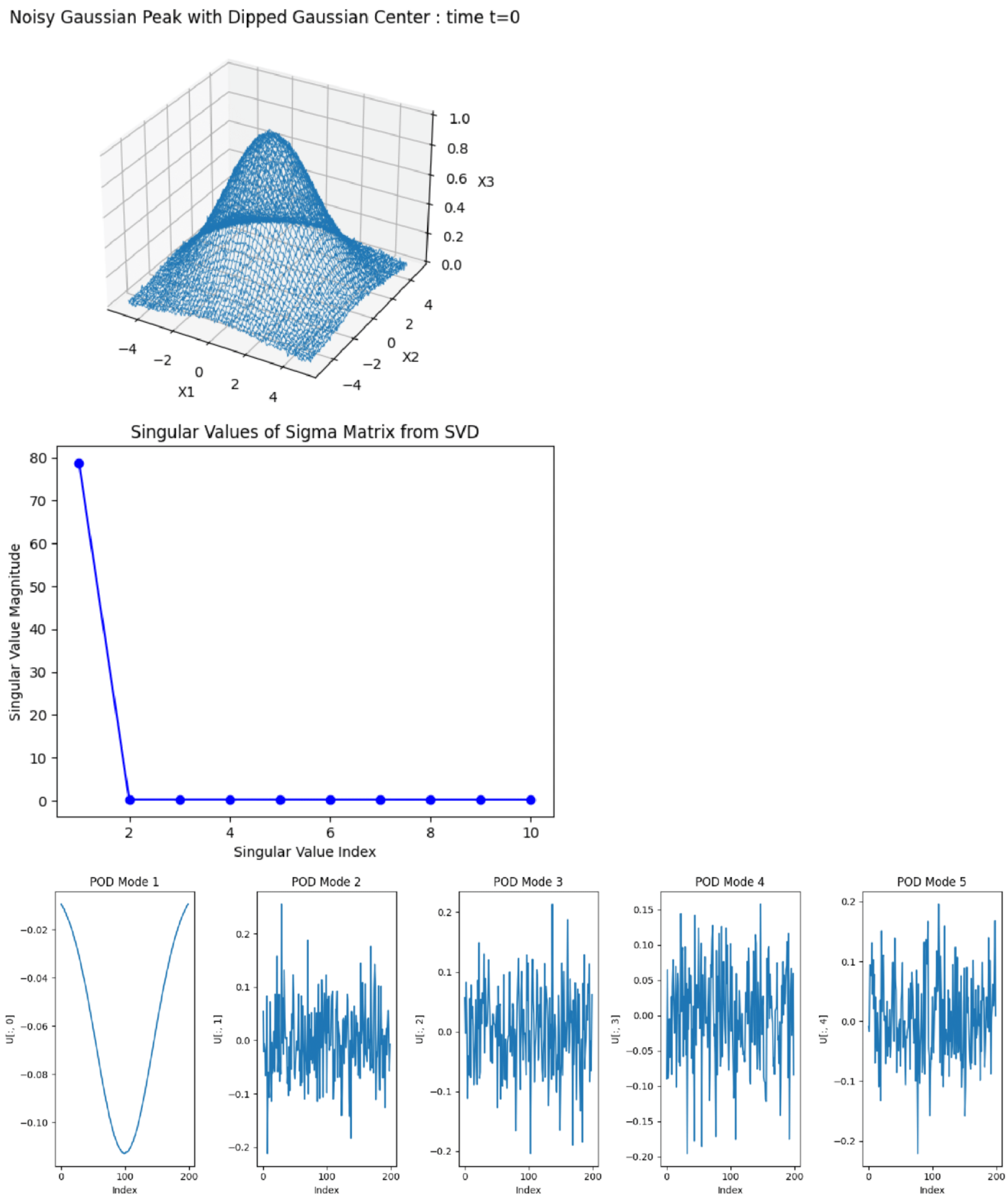

Figure 4.30: Output Plots for t=0

4.4 DMD Application

This section presents the application of the Dynamic Mode Decomposition (DMD) method to run modes through the L1 algorithm. The following figures illustrate the results obtained.

Figure 4.35 shows the DMD modes obtained using the L1 algorithm. The plot illustrates the temporal evolution of the dominant modes.

The analysis reveals that the DMD method, when combined with the L1 algorithm, effectively captures the dynamic behavior of the system. The identified modes represent the key dynamics and provide insights into the underlying processes driving the observed data patterns.

We can thus truncate the modes upto the second place and recreate successfully noiseless data. (As shown in 4.35

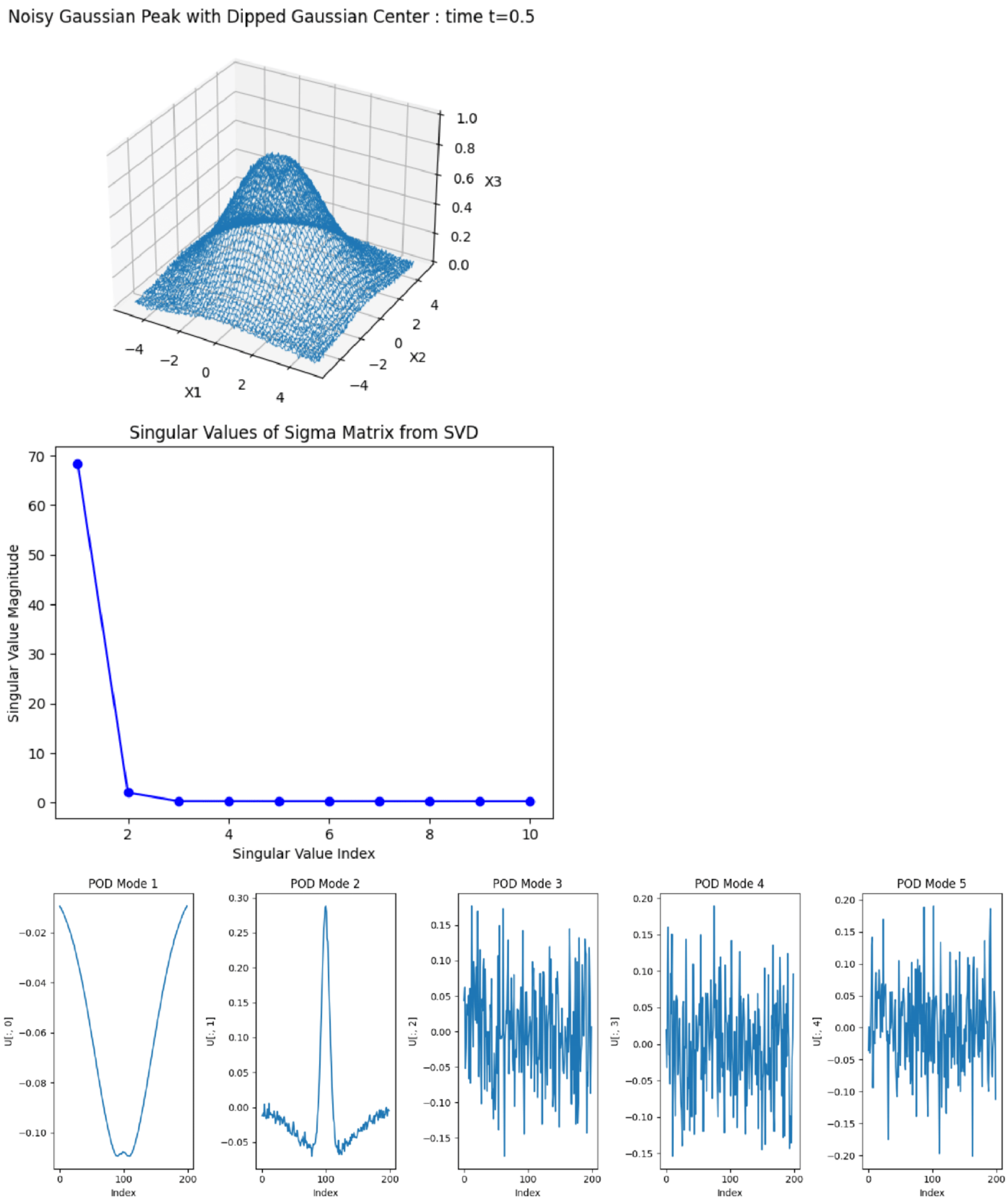

Figure 4.31: Output Plots for t=1

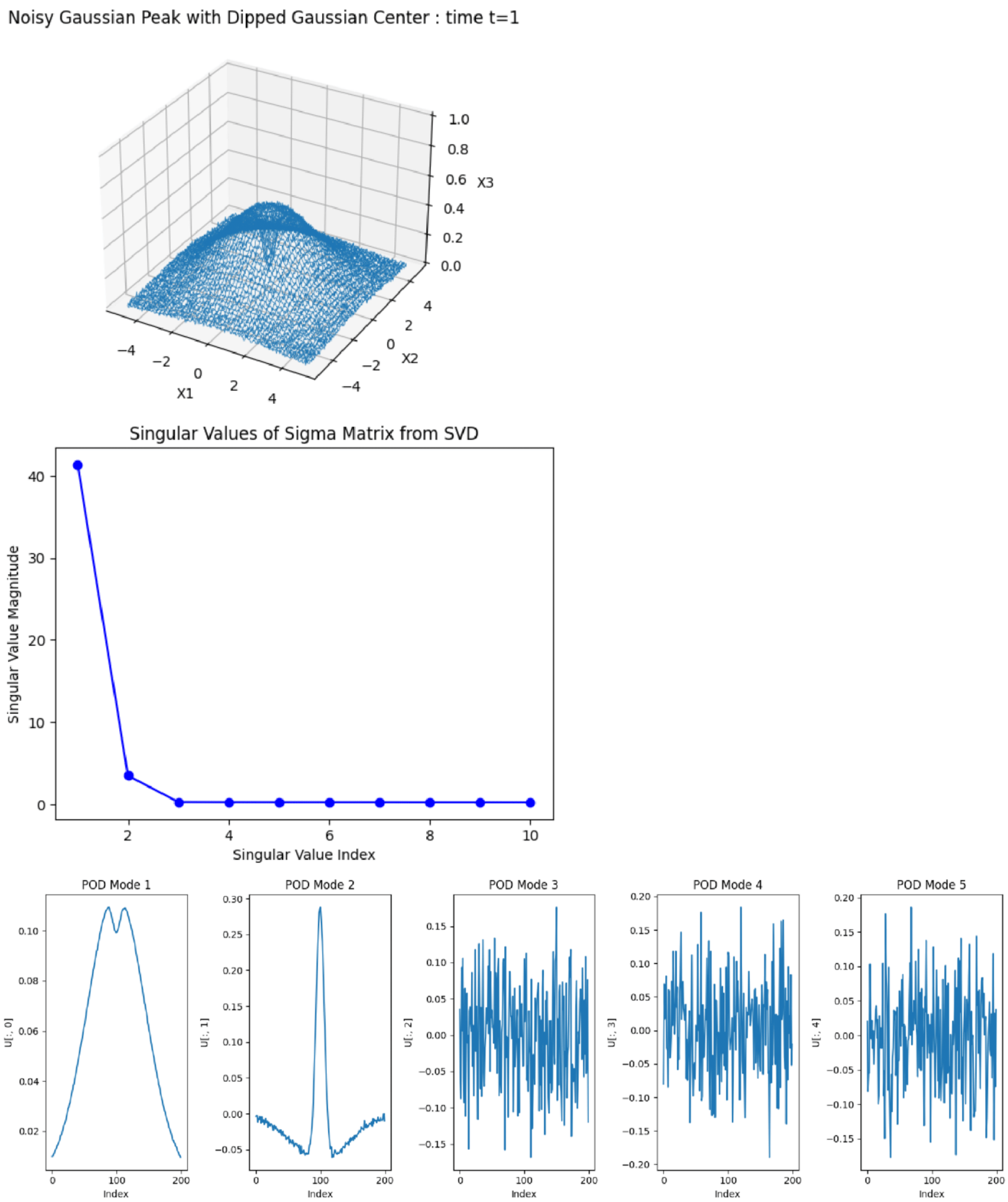

Figure 4.32: Output Plots for t=1.5

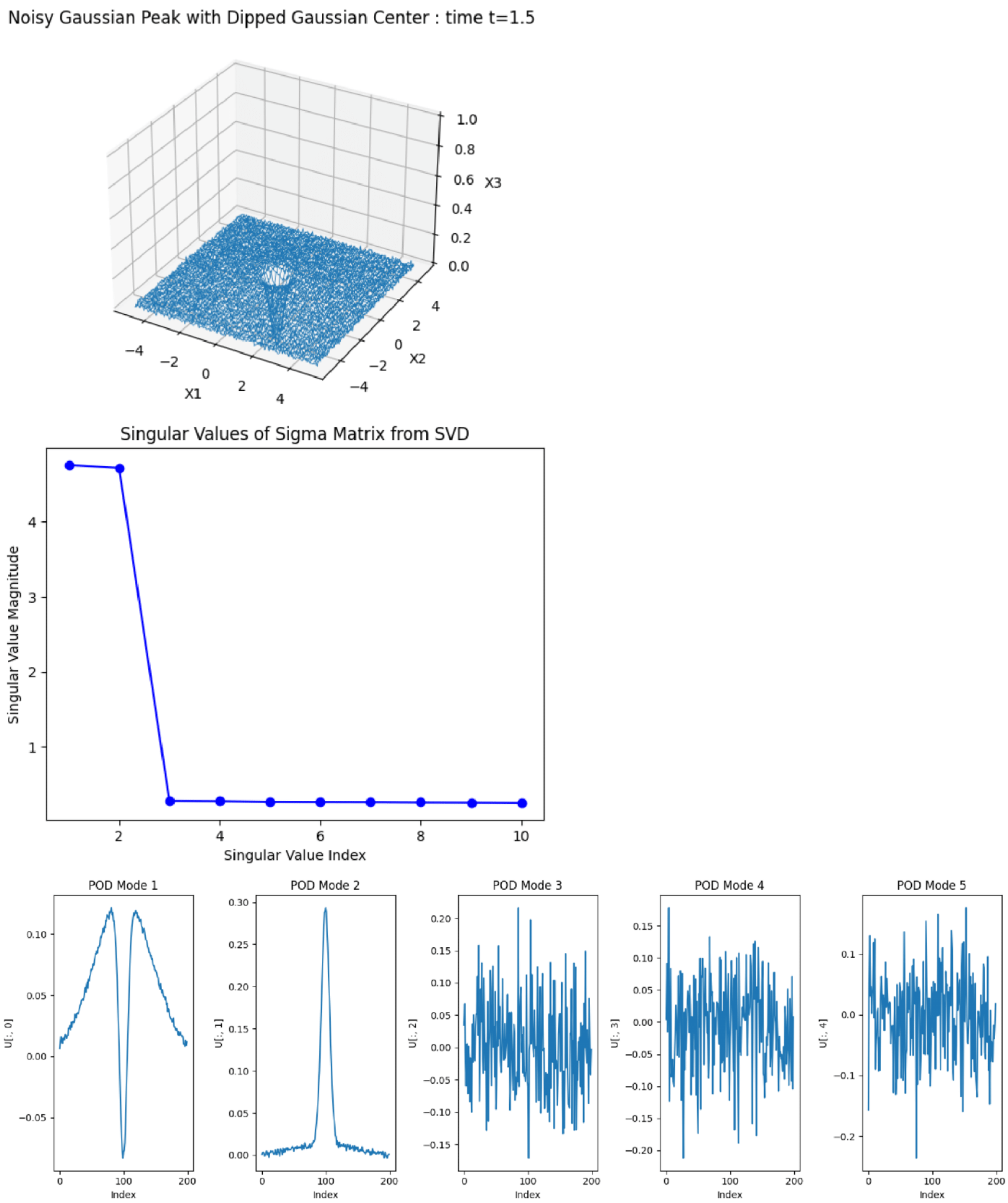

Figure 4.33: Output Plots for t=0

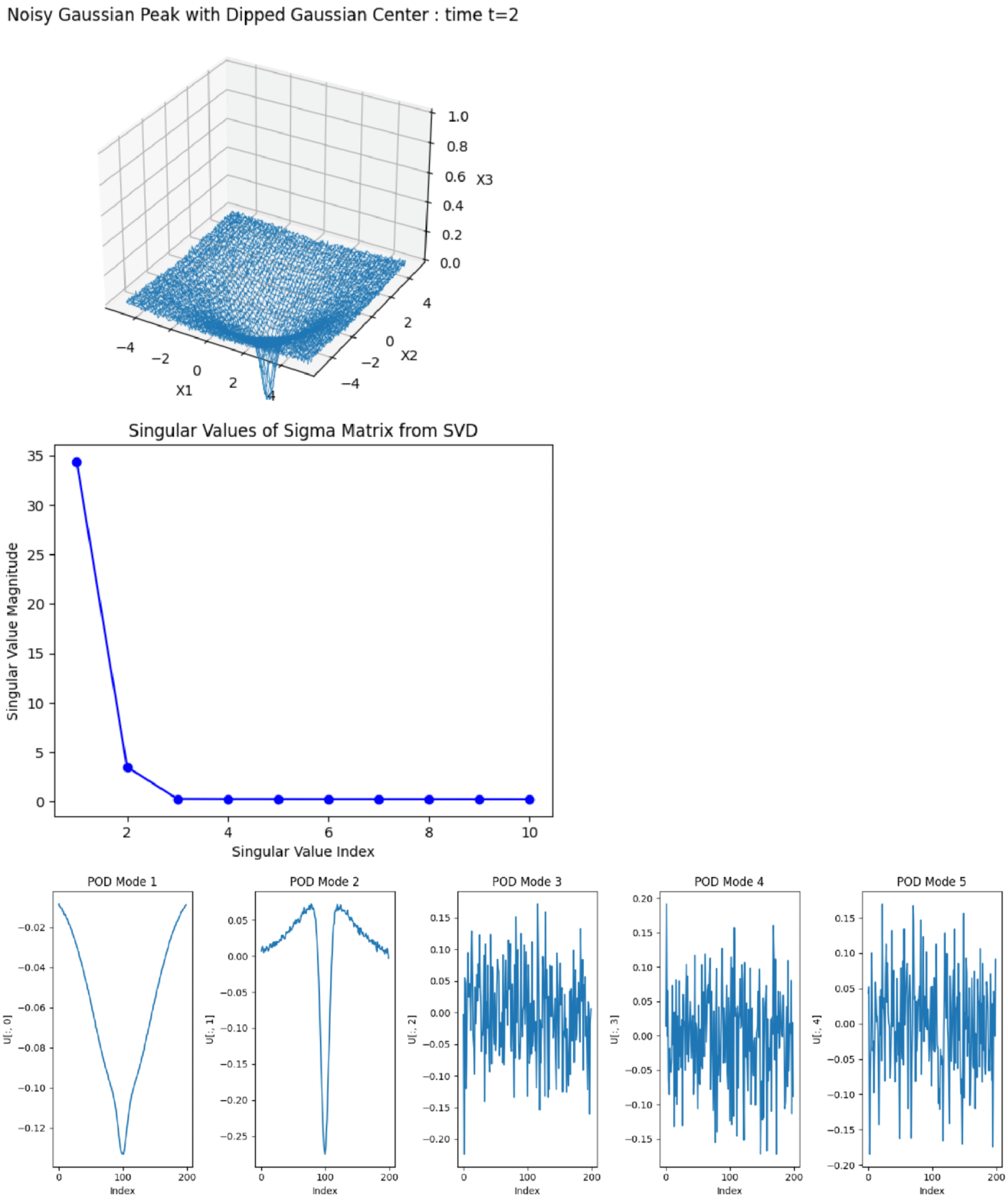

Figure 4.34: Output Plots for t=2

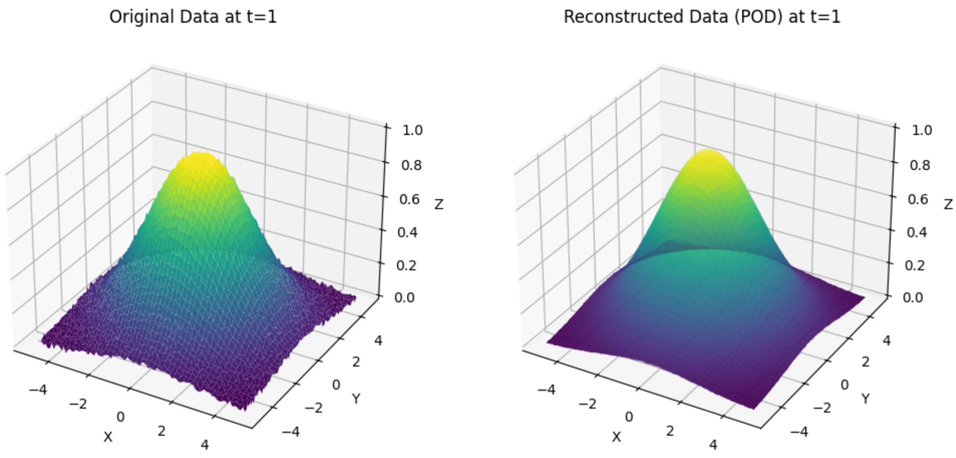

Figure 4.35: Reconstructed plot demonstrating the removal of noise in the data.

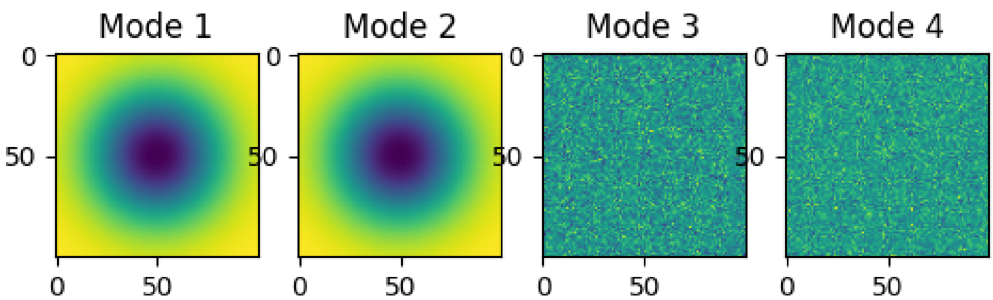

Figure 4.36: Reconstructed plot demonstrating the removal of noise in the data.

Bibliography

[1] Kaptanoglu, A. A., Morgan, K. D., Hansen, C. J., & Brunton, S. L. (2020). *Characterizing magnetized plasmas with dynamic mode decomposition*, *Physics of Plasmas, 27*(3), 032108, https://doi.org/10.1063/1.5138932

[2] Michael Schmidt and Hod Lipson, *Distilling Free-Form Natural Laws from Experimental Data*, Science, 324 (5923), DOI: 10.1126/science.1165893 Brookhaven National Lab on June 22, 2023

[3] Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, *Proceedings of the National Academy of Sciences, 113*(15), 3932-3937, https://doi.org/10.1073/pnas.1517384113

[4] Bongard, J., & Lipson, H. (2007). *Automated reverse engineering of nonlinear dynamical systems*, *Proceedings of the National Academy of Sciences, 104*(24), 9943-9948, https://doi.org/10.1073/pnas.0609476104

[5] De Silva, B., Champion, K., Quade, M., Loiseau, J.-C., Kutz, J. N., & Brunton, S. L. (2020). *PySINDy: A Python package for the sparse identification of nonlinear dynamical systems from data*, *Journal of Open Source Software, 5*(49), 2104, https://doi.org/10.21105/joss.02104

[6] Kaptanoglu, A. A., de Silva, B. M., Fasel, U., Kaheman, K., Goldschmidt, A. J., Callaham, J., Delahunt, C. B., Nicolaou, Z. G., Champion, K., Loiseau, J.-C., Kutz, J. N., & Brunton, S. L. (2022). *PySINDy: A comprehensive Python package for robust sparse system identification*, *Journal of Open Source Software, 7*(69), 3994, https://doi.org/10.21105/joss.03994

[7] Sen, A., Ahalpara, D. P., Thyagaraja, A., & Krishnaswami, G. S. (2014). *A KdV-like advection-dispersion equation with some remarkable properties*, (arXiv:1109.3745v3 [nlin.PS]), https://doi.org/10.48550/arXiv.1109.3745