Journal of Theoretical and Applied Vibration and Acoustics 10(1) 13-28 (2024)

I S A V

Journal of Theoretical and Applied Vibration and Acoustics

journal homepage: http://tava.isav.ir

Hybrid Koopman-neural network approach for robust parameter estimation and prediction in duffing oscillators

Yassin Riyazi [a], **Navidreza Ghanbari** [a], **Arash Bahrami** [b,*]

[a] *M.Sc. Student, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran*

[b] *Assistant Professor, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran*

Research Article

ARTICLE INFO

Article history:
Received 17 May 2024

Received in revised form 31 August 2024

Accepted 11 October 2024

Available online 19 November 2024

Keywords:
Koopman operator

parameter estimation

nonlinear dynamical systems

neural networks

period-doubling

ABSTRACT

Nonlinear dynamical system research is essential in science and engineering because it could be potentially employed to represent real-world phenomena. Traditional methods rely on pre-defined models or computationally expensive simulations, limiting their applicability to only numerical data. In the present research, without any prior knowledge of the system, we suggest a novel way to build a linear representation of the Duffing oscillator by fusing the capabilities of deep neural networks with the Koopman operator. This recently established methodology makes it easier to estimate system parameters effectively and accurately predict the oscillator's future behavior. Our approach incorporates a modified training procedure that restricts the Koopman operator to a single linear layer within the neural network, improving interpretability and potentially reducing training complexity. This methodology not only simplifies nonlinear system analysis but also paves the way for advancements in predictive modeling across various fields. Notably, our method yields distinctive eigenvalues of the Koopman generator matrix, enabling the Koopman operator to exhibit robustness against noise and capture a spectrum of Duffing equation behaviors. This includes the precise prediction of periodic oscillations and the capturing of period-doubling bifurcation, all while maintaining tractability within the neural network framework.

© 2024 Iranian Society of Acoustics and Vibration, All rights reserved.

1. Introduction

Dynamical systems theory has evolved by integrating three distinct viewpoints: Newton's classical approach, the state space representation paradigm, and the modern Operator Theoretic framework.

* Corresponding author:
E-mail address: arash.bahrami@ut.ac.ir (A. Bahrami)

http://dx.doi.org/10.22064/tava.2024.2029513.1241

Each perspective offers unique insights into the behavior of dynamic systems, contributing to our understanding of complex phenomena across diverse scientific disciplines.

Newton's methodology, rooted in classical mechanics, relies on formulating physical laws as second-order ordinary differential equations. These equations describe the evolution of system variables over time, enabling predictions of future states based on initial conditions. Through this lens, system behavior manifests as trajectories in a state space defined by state variables plotted against time, emphasizing the role of forces and accelerations in shaping system dynamics.

However, Newton's technique encountered limitations, notably highlighted by Henri Poincaré when grappling with the three-body mass problem. Pioneered by mathematicians such as Poincaré, the state space representation paradigm has proven invaluable for analyzing complex systems with nonlinear dynamics, unveiling phenomena like chaos and bifurcations. Unlike Newton's approach, this paradigm maps the entire trajectory of system evolution in a multidimensional space, treating time as another dimension. Poincaré's work laid the foundation for modern dynamical systems theory by emphasizing qualitative analysis of system dynamics through geometric properties of trajectories in the state space.

Despite the advancements facilitated by the state space representation, challenges persist in understanding and predicting the behavior of high-dimensional systems and managing vast amounts of data from modern scientific experiments. In response, the Operator Theoretic approach has emerged as a powerful framework. Rooted in linear operator theory, this approach offers a systematic methodology for extracting meaningful information from observational data, enabling the construction of predictive models capturing system dynamics. By representing dynamical systems as operators acting on function spaces, it provides a unified perspective on diverse systems, facilitating the integration of data-driven techniques like machine learning and deep learning into dynamical systems analysis.

The Koopman operator [1], one of the family of operator theory, with its inherent structure involving a mapping to a higher dimension, a linear transformation, and an inverse mapping, resembles the structure of autoencoders [2]. Bethany Lusch et al. in [3] developed a neural network deliberately tailored for conciseness and interpretability, aiming to encapsulate system dynamics within a low-dimensional manifold. They introduced nonlinear coordinates, which are identified through a modified auto-encoder, and showed under these coordinates, the dynamics of the system exhibit a global linearity property. Furthermore, they expanded the concept of Koopman representations to encompass systems with continuous spectra, introducing an auxiliary network to efficiently parameterize continuous frequencies. This innovative approach establishes a bridge between deep learning models and decades of asymptotic research, offering a fusion of the advantages of deep learning with the capacity for physical interpretability provided by Koopman embedding.

In a recent study, Kathleen Champion et al. [4] addressed the challenge of discovering governing equations from scientific data in data-rich fields lacking well-characterized quantitative descriptions. They utilized advances in sparse regression to identify both the structure and parameters of nonlinear dynamical systems from data, resulting in models that balance simplicity with descriptive power. To achieve this, they introduced a custom deep autoencoder network designed to uncover a coordinate transformation into a reduced space, simplifying the representation of dynamics. This innovative approach allows for simultaneous learning of governing equations and the associated coordinate system. They applied this technique to various

examples of high-dimensional systems exhibiting low-dimensional behavior, creating a modeling framework that combines the flexibility of deep neural networks with the economical nature of sparse identification of nonlinear dynamics (SINDY) [5].

While the sole use of a neural network proves accurate and meets our requirements, it does not guarantee the exclusive confinement of the Koopman operator to a designated linear layer. Consequently, the entire network structure incorporates elements of mapping, linear transformation, and inverse mapping simultaneously. This challenges the utility of using the linear layer weight to represent the system, as they only encapsulate a portion of the Koopman operator. To overcome these challenges, we introduce a novel approach that capitalizes on the synergy between the Koopman operator and Deep Neural Networks [6-9]. This groundbreaking fusion is aimed at converting the Duffing equation into a linearized representation, offering promising solutions to the intricacies encountered in traditional methods. Our approach enables a more accurate Koopman linearized representation of the system behavior. This is done by harnessing the computational power of deep learning and the Koopman operator's capability to provide a linear representation of nonlinear systems. Central to this endeavor lies the Duffing oscillator, embodying nonlinear dynamical system archetypes. Its versatility allows the emulation of diverse behaviors, making it pertinent for various physical phenomena, from mechanical vibrations to biological oscillations. Traditionally, analytical techniques and numerical simulations analyze and predict Duffing oscillator behavior. However, these methods encounter challenges with nonlinear complexities and uncertainties, prompting our innovative fusion of Koopman operator theory and deep neural networks.

The behavior of duffing oscillators has historically been analyzed and predicted using analytical techniques in conjunction with numerical simulations. While these methods provide insightful information, they often encounter difficulties when attempting to handle nonlinear complexities in real-world situations appropriately. Analytical solutions can be algebraically complex or even elusive, especially for highly nonlinear or higher dimensional systems. On the other hand, despite their powerful nature, numerical simulations require a lot of computational resources and struggle to produce long-term forecasts due to inherent numerical errors and uncertainties.

The following sections discuss the foundational principles of Koopman operator theory and the adaptability of deep neural networks. Furthermore, it is shown how their fusion forms a compelling framework for analyzing and predicting the behavior of Duffing oscillators. Next, the process of transforming Duffing oscillator dynamics into a linear representation is outlined. The discussion follows by introducing a modified loss function designed to enhance the generality of the Koopman linear representation of the dynamical system within this context. Through numerical simulations and comparisons with traditional methods, the efficacy of our approach in providing accurate predictions for the future behavior of the Duffing oscillator is demonstrated. Finally, the theoretical framework developed here enriches our understanding of nonlinear dynamics and offers a powerful tool with transformative potential across scientific, engineering, and practical applications.

For access to the computer codes and resources, one may refer to the GitHub repository available at: https://github.com/yriyazi/Leveraging-Koopman-operator-and-Deep-Neural-Networks-for-Parameter-Estimation-and-Future-Prediction

2. Koopman operator

The Koopman operator is a mathematical tool that has recently gained popularity in engineering; it offers a novel perspective on dynamic systems. Based on functional analysis, the Koopman operator presents a paradigm shift by shifting the attention from the state space to the space of observable functions. Even with intrinsically nonlinear systems, this enables us to see the system evolution through a linear lens. This viewpoint provides us with a fresh perspective through which to see the dynamics of intricate systems.

We first transform the existing state of the system into a high-dimensional Hilbert space and then identify a linear representation for it. Next, we locate a map that restores data to the system's initial space.

2.1. Dynamical system representation

We consider a dynamical system described by a set of state variables $x(t)$, which evolve over time t. Mathematically, it can be represented as:

$$\frac{dx}{dt} = f(x,t) \tag{1}$$

where

- n is the state vector dimension.
- $x \in R^n$ is the state vector representing the system's state variables.
- $t \in R^+$ is time.
- $f(x): R^n \to R^n$ is a vector-valued function describing how the state variables change over time.

2.2. Koopman operator transformation

The Koopman operator, denoted as K, is an infinite-dimensional linear operator that acts on observables or functions of the state variables. Assume $g(x)$ is such an observable that maps original data to high-dimensional space. The Koopman operator maps this observable from the state space to a higher-dimensional space:

$$K \times g(x_{t_o}) \equiv g\left(f(x_{t_o})\right) \equiv g(x_{t_{o+1}}) \tag{2}$$

where:

- $K(.): C^m \to C^m$ or $K(.): R^{2m} \to R^{2m}$ is the Koopman operator Generator. We use the real representation for simplicity.
- $2m$ is supposed to be infinite-dimensional, but in numerical approximation, a value, for example, 200 or 600, is assigned.
- t_o is an arbitrary value of time and t_{o+1} is a one-time step forward.
- $g(x): R^n \to C^m$ or $g(x): R^n \to R^{2m}$ is an observable or function defined on the state space.
- n is the state vector dimension.

- $g\big(f(x)\big)$ represents the observable after the system evolves according to $f(x)$.
- Koopman operator in Discrete Time
- In discrete-time dynamical systems, the Koopman operator is applied at discrete time steps. For these systems, (2) may be represented as:

$$K \times g(x_k) = g(x_{k+1}) \tag{3}$$

where:

- x_k represents the state of the system at the time k.
- x_{k+1} represents the state of the system at the next time step $k+1$.

As observed in Figure 1, the important relationship is apparent: the transformation of data x_t from its low-dimensional representation in Euclidean space to an infinite-dimensional Hilbert space is facilitated through the utilization of Koopman observables $g(x_t)$. Leveraging this Koopman transformation allows for the mapping of y_t to y_{t+1} via a linear matrix transformation. Furthermore, employing the inverse Koopman observable mapping $g^{-1}(y_{t+1})$, enables the derivation of x_{t+1}. With the help of the Koopman operator $f(x_t) = g^{-1}\big(K \times g(x_t)\big) = x_{t+1}$.

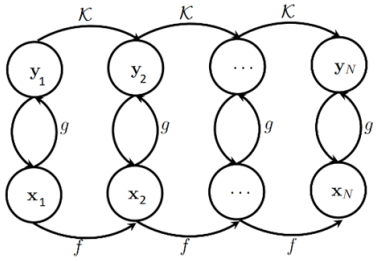

Fig.1. Koopman operator Evolution and a discrete dynamical system.

3. Coupling Koopman operators with Deep Neural Networks

Deep neural networks have demonstrated remarkable abilities in approximating intricate functions and mastering complex patterns from data. One of the key challenges encountered in the realm of Koopman operators is the identification of suitable observable functions. In methods such as Dynamic Mode Decomposition (DMD) [10, 11], the observable function is typically the identity function, and in extended DMD (EDMD) [12], observable functions take the form of polynomials or trigonometric functions. While these approaches are straightforward and accurate, they do not exhibit resiliency to noise.

In this study, we combine a deep neural network with the Koopman operator, thereby generating a linearized representation of the Duffing equation. This neural network effectively learns the intricate relationship between system parameters and observed behaviors, facilitating efficient parameter estimation. Moreover, the neural network undergoes training to predict the future trajectory of the Duffing oscillator, thereby equipping us with a valuable tool for forecasting system behavior. Neural networks do not necessarily always yield the best result, but neural networks do offer a flexible alternative that works well in many cases.

3.1. Case study

Our investigation centers on the Duffing oscillator, notable for its mechanical analogies, particularly when Galerkin's first mode approximation principle is applied to magnetoelastic systems [13, 14]. Beyond its tangible real-world applications, the Duffing equation exhibits diverse nonlinear behaviors, rendering it a compelling candidate for approximation using neural networks. Here, we approach the Duffing oscillator as a mathematical equation to unravel its intricate dynamics and explore its predictive potential.

3.2. Data acquisition

The non-dimensional Duffing oscillator is a dynamical system described by the following second-order differential equation [15]:

$$\frac{d^2x}{dt^2} + \delta\frac{dx}{dt} + \alpha x + \beta x^3 = \gamma\, cos\,(\omega t) \tag{4}$$

where:

- α is the linear stiffness constant, which equals 1.
- β characterizes the nonlinearity in the system, which equals -1.
- δ denotes the damping constant, which equals 0.1.
- ω is the angular frequency of the driving force, which equals 1.4.
- γ is the amplitude of the external driving force.

The Duffing oscillator numerical solution is generated using the Runge-Kuta method [16].

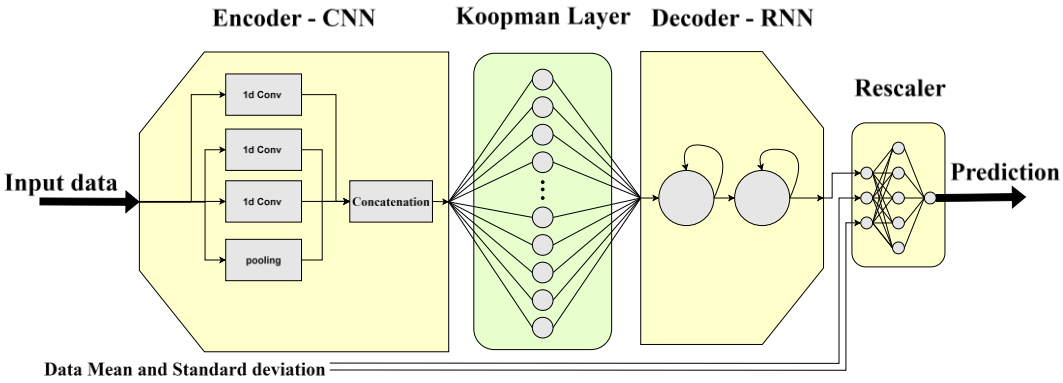

Fig. 2. Neural Network Diagram

3.3. Data normalization

Given the nature of regression, it is advisable to normalize the data before training the neural network. This normalization is crucial because even small variations in the input data can lead to significant changes in the backpropagation stage, potentially reducing the model's robustness against changes in input conditions. Our approach applies data normalization before feeding it into the convolutional neural network (CNN) architecture. During this process, we pass the statistical properties of the data, such as its mean and variance, through the network.

After the decoder block, the data is remapped to its original statistical properties. Then, it is passed to the Rescaler Block. The purpose of the Rescaler Block is to further reduce the variation in the output, ultimately leading to a more stable and controlled model response.

3.4. Structure

The model's architecture is shown in Figure 2, and its hyperparameters are presented in Table 1. It is structured around an Encoder-Decoder paradigm, effectively capturing the essence of complex dynamics. Specifically, the Encoder component is meticulously designed, featuring a sequence of Inception Blocks [17] in a convolutional neural network (CNN) [18]. These Inception Blocks are robust feature extractors, enabling the model to discern intricate patterns and relevant features from the input data.

Following the Inception-based Encoder, a pivotal transformation occurs through a linear layer. This linear layer assumes a distinct role within the architecture, embodying the essence of the Koopman operator evolution function K. It is important to note that this linear layer operates without an activation function and bias, preserving the linear nature of the Koopman operator transformation.

Transitioning from the Koopman operator layer, the architecture takes an intriguing turn with the integration of a two-layer Long Short-Term Memory (LSTM) [19] network. This LSTM component acts as the Decoder, expertly leveraging its sequential memory to unravel the transformed linearized representation. This sequence-to-sequence modeling approach facilitates the reconstruction of the system's temporal evolution, a crucial aspect in capturing its complicated behaviors.

3.5. Training

A custom loss function and a particular method of network data feeding are essential for improving the interpretability of neural networks. Although the network could be trained end-to-end without any modification, it would be more like a black box.

In a common Neural network training, data will be shuffled and fed to the neural network multiple times. However, in future time series predictions, the network can be over-trained. Also, the evolution function of the Koopman operator does not remain confined solely to the Koopman part; instead, it spreads throughout the network.

3.5.1. Data loader

The initial step involves generating a vector representing the solution of the Duffing oscillator. Subsequently, this vector is segmented into distinct batches based on pre-defined network input and output dimensions. The first batch is the input for an autoregressive prediction of the system's future behavior. The network input dimension specifies the number of samples required for the prediction, while the network output dimension determines the number of future steps to be forecast.

Table 1. CNN parameters

	Sequential Blocks	In Channels	Out Channels	kernel size	padding
Branch 1	1dConv	1	Out*	1	0
	ReLU	-	-	-	-
Branch 2	Conv 1	1	Out*	1	0
	ReLU	-	-	-	-
	1dConv	Out*	Out*	7	3
	ReLU	-	-	-	-
Branch 3	1dConv	1	Out*	1	0
	ReLU	-	-	-	-
	1dConv	Out*	Out*		15
	ReLU	-	-	-	-
Branch 4	1dMaxPool	1	1	3	1
	Linear	-	-	-	-
	Conv 1	1	Out*	1	0
	ReLU	-	-	-	-

*Out hyperparameter is 20

3.5.2. Custom loss function

A two-stage training algorithm has been proposed to restrict the Koopman generator matrix exclusively to the Koopman linear layer. In this algorithm, after each optimization step:

1. The weights of all layers except the Koopman Linear Layer $W_{Koopman}$ are frozen.
2. The output of the Koopman Linear Layer $g(x_{n_0}) \times (W_{Koopman}^{n})^{T}$ is calculated for time steps n to n_{KPH} . Since computing the backpropagation of a matrix would be more costly, 20 was selected as the Koopman Horizon Hyperparameters (KPH).
3. The weights of the Koopman Linear Layer are updated based on the linearity property. This update aims to minimize the prediction error of the *n*th output.

$$\sum_{n}^{KPH} L\left(g(x_{n_0}) \times (W_{Koopman}^{n})^{T}, g(x_n) \times (W_{Koopman})^{T}\right) \tag{5}$$

By implementing this two-stage training process, we ensure that the Koopman generator matrix is confined and utilized specifically within the Koopman linear layer, enhancing the network's predictive accuracy and control.

The hyperparameters for the optimizer and loss parameters are presented in Table 2 for reference.

Table 2. Optimizer and loss parameters

		Stage 1	Stage 2
Optimizer	Type	SGD	SGD
	Learning Rate	5.00E-02	5.00E-04
	momentum	0.9	-
	weight decay	1.00E-04	-
loss	Type	MSE	MSE

4. Results and discussion

In this section, the results of the study are presented, detailing the development and evaluation of our proposed methodology.

Duffing equation parameters are initially fixed and fed into the neural network for training. This step enables us to understand the system's behavior and evaluate the neural network's proficiency in learning and predicting based on fixed parameters.

Subsequently, we conduct experiments to assess the neural network's capability to capture a range of behaviors exhibited by the Duffing oscillator. By systematically varying one parameter of the Duffing equation (γ or excitation amplitude in physical setup), while keeping other parameters constant, we observe how the network adapted to different conditions and whether it could accurately predict the corresponding oscillatory patterns. This analysis provides valuable insights into the model's robustness and generalization capacity across diverse scenarios. Furthermore, by varying the value of γ, we investigate how the neural network adapted to different behavioral regimes, shedding light on the system's sensitivity to parameter changes and the network's proficiency in capturing complex dynamics.

Throughout the neural network training process, we employ an input horizon of 200 previous samples to extract relevant features, with data sampled at 10 Hz for the first case and 600 samples at 30 Hz for the second case. This ensures the model can access sufficient historical data to make accurate predictions. Additionally, in all cases, the future state prediction is limited to a confined portion of data obtained by numerically solving the Duffing equation to 150000 steps, equivalent to 15000 seconds. We keep the last 5000 data samples and feed them to neural network.

In all cases, we confine future state predictions to a specific portion of the data. This portion is obtained by solving the Duffing equation numerically for 150,000 steps, equivalent to 15,000 seconds, and retaining the last 15000 data samples for input to the neural network.

4.1. Periodic responses

In this section, the behavior of the Duffing oscillator is investigated under the condition that all parameters are kept fixed. Our primary focus is to examine how the neural network performs when trained on a dataset generated with these fixed parameters. To simulate real-world data challenges, we add a uniformly distributed noise in the range of [-0.5,0.5] to the data during the neural network training process. An input horizon of 200 prior samples is employed to extract essential features, with data sampled at 10 Hz. In addition, the initial conditions are set to $x(0) = 1.5$ and $\dot{x}(0) = -1.5$.

Initial observations reveal a remarkable resilience of the neural network to noise, even in scenarios where the noise-to-data ratio is as high as 1. This robustness is underscored by the consistent and accurate predictions made by the network, as illustrated in Figures 3a and 3b.

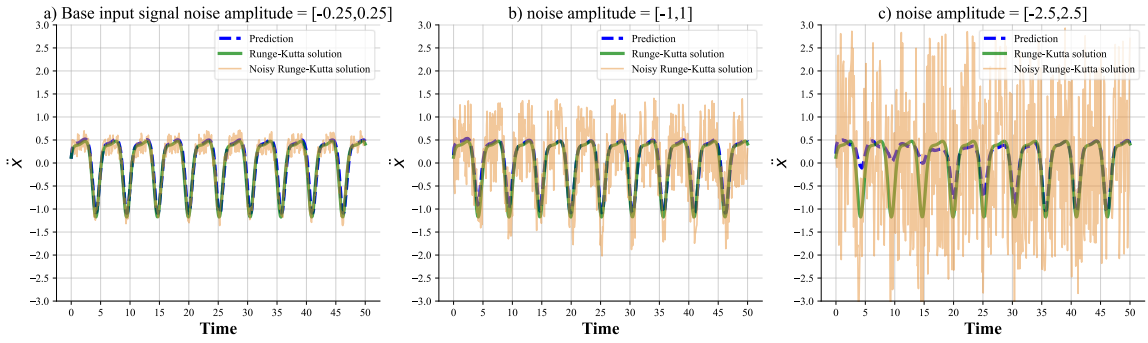

Fig. 3a. Robustness against noise $\gamma = 0.2$.

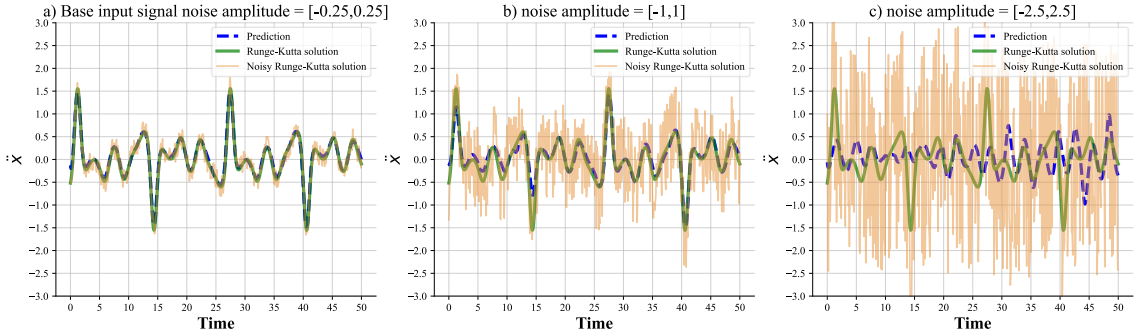

Fig. 3b. Robustness against noise $\gamma = 0.37$

In Figures 3a and 3b, each subplot contains three lines: the exact numerical solution of the Duffing equation, the numerical solution with added noise of the same amplitude as when training the neural network, and the prediction by our model. Plotting the future states produced auto-regressively from the neural network's output. The network is fed 200 prior samples of the noisy signal to generate these figures. Subplot (a) illustrates the neural network's performance under

normal training conditions; (b) presents a similar scenario with noise levels doubled compared to the training phase; and (c) represents a worst-case situation where noise completely overwhelms the data, resulting in the network's inability to provide accurate predictions. The neural network generally exhibits robustness against noise within a range spanning [-1, 1].

Furthermore, the success of our methodology prompts an intriguing question: can this approach be extended to capture a more diverse range of behaviors exhibited by the Duffing oscillator? This question forms the basis of our subsequent investigations, as we explore the neural network's adaptability and generalization capabilities across a broader spectrum of system dynamics.

4.2. Period-doubling bifurcations

In this section, the behavior of the Duffing oscillator with varying excitation amplitude γ is explored when ω =1.4. A uniform noise in the range of [-0.25, 0.25] is added to simulate real-world data challenges during the neural network training process. An input horizon of 600 is employed before samples to extract essential features, with data sampled at 30 Hz. Additionally, initial conditions are generated randomly to capture both stable and unstable solutions. Using 40 instances of Duffing equation solutions, a mean squared error (MSE) of 0.078 is observed. Feeding data in small, balanced batches and allowing the network to train on data points multiple times have proven crucial for optimizing performance.

As mentioned before, the control (bifurcation parameter) γ is modified while the excitation frequency ω is held unchanged. By changing the control parameter, a period-doubling bifurcation occurs around γ = 0.3085, as depicted in Figure 4. Our analysis of the neural network's predictive capabilities reveals intriguing insights. In particular, neural network can capture the underlying oscillatory structures, including periodic responses and period-doubling bifurcations. This will be discussed in more detail in the remaining parts of the article.

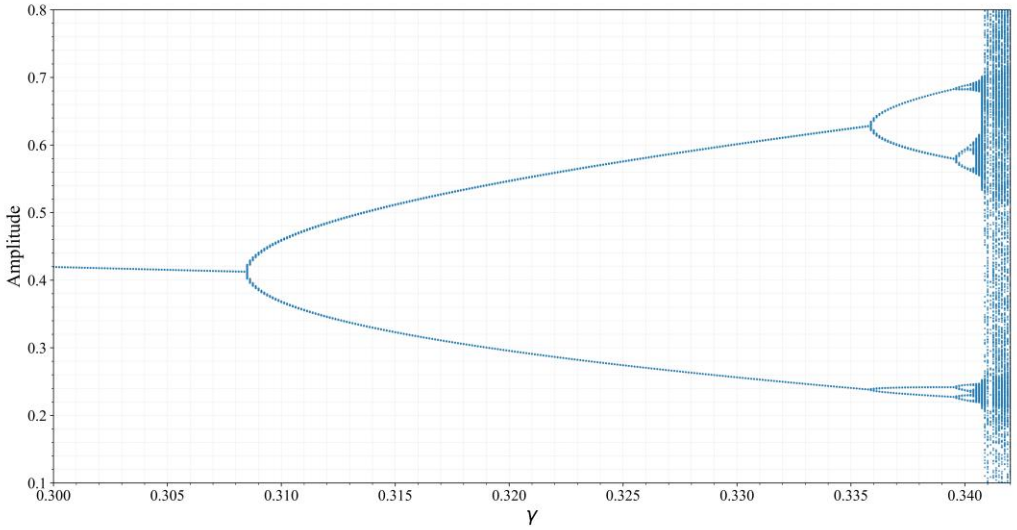

Fig. 4. Bifurcation diagram of the Duffing oscillator with $\omega = 1.4$, illustrating the variation of the response amplitude versus excitation amplitude. The analysis is confined to the domain of γ from 0.3 to 0.34.

In Figure 5, a model is trained to predict the variable x of the Duffing equation. It is observed that the model's performance is acceptable in both periodic and period-doubling scenarios. Although the behavior of variable x may appear similar in all cases of Figure 5, there are differences in their $\ddot{x}$ values, as demonstrated in Figure 6. A model is trained to predict the future state of $\ddot{x}$ in the Duffing equation as shown in Figure 6, maintaining consistent conditions and parameters. Similar predictive capabilities are observed across all scenarios. However, we encountered a significant challenge during our efforts to capture both stable and unstable solutions for the Duffing oscillator. The data exhibited vastly different amplitudes, leading to an optimization problem that favored cases with higher amplitudes due to their larger loss. All cases are mapped into the [-1, 1] domain before the rescaler block to address this issue.

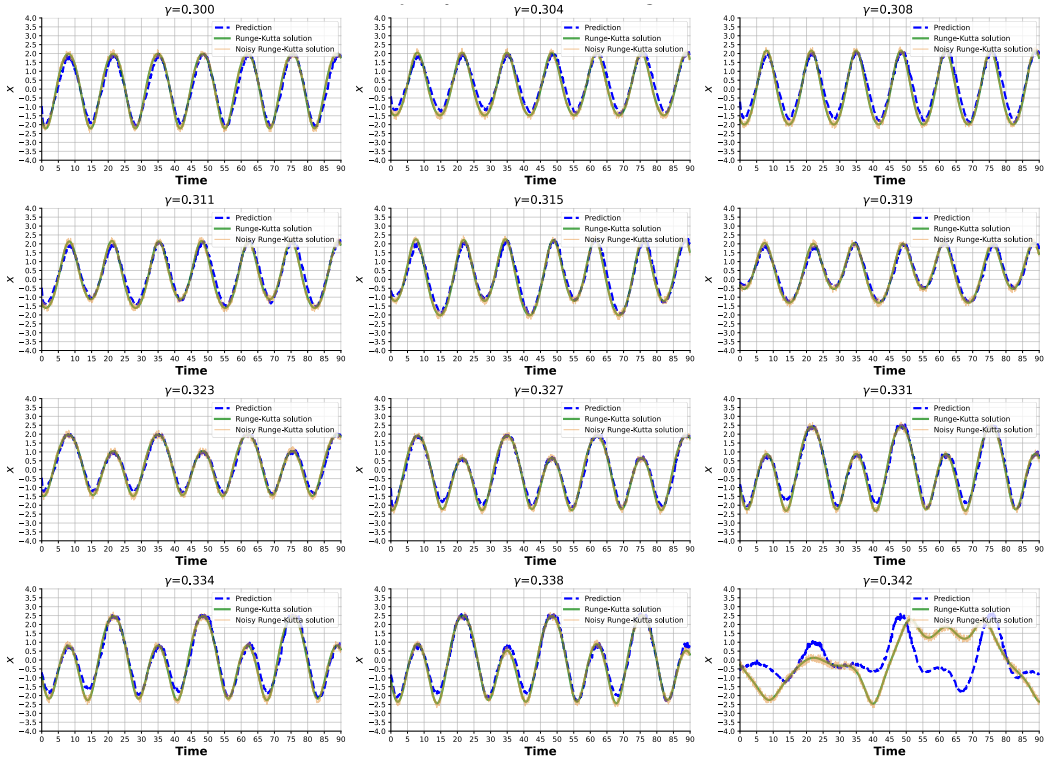

Fig. 5. Demonstration of the neural network's capability in capturing versatile periodic behaviors of the Duffing oscillator. The variable x has been predicted.

Essentially, two separate rescaling are applied to the data: one inside the neural network and one in the post-processing section.

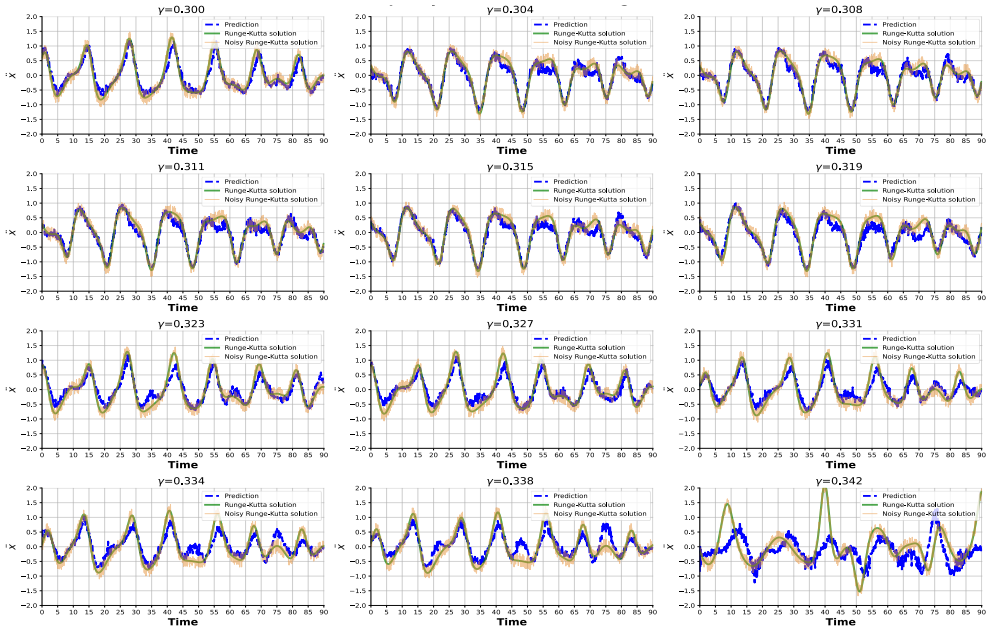

Fig. 6. $\ddot{x}$ versus time for different excitation amplitudes and excitation frequency ω = 1.4 within the moderate noise scenario

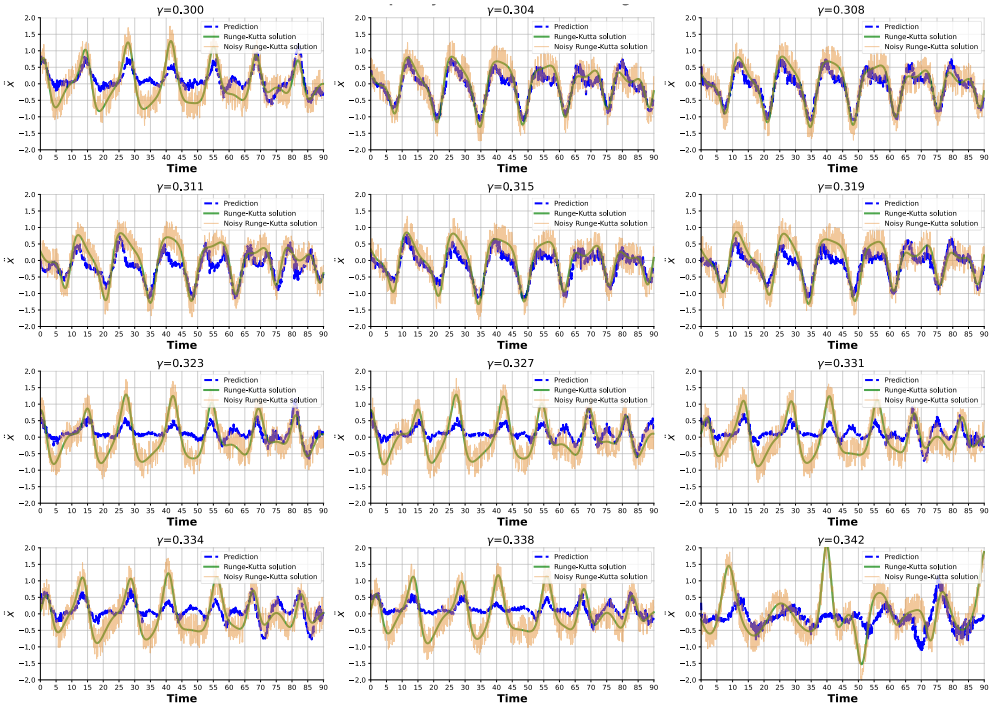

Fig. 7. $\ddot{x}$ versus time for different excitation amplitudes and excitation frequency ω = 1.4 within the heavy noise scenario

It should be noted that the model struggles to capture the chaotic regime, likely due to the absence of recognizable patterns within the 600 previous samples. Despite this, we demonstrated the model's effectiveness in capturing periodic behavior and period doubling phenomenon, proving its

versatility and efficiency. Notably, the model is not restricted to predicting x or $\ddot{x}$; it can be adapted for various outputs.

Moreover, the model's generalizability under different noise levels is demonstrated in Figures 6 and 7. The model is trained on data with noise in the range of [-0.25, 0.25], and its performance is examined with noise levels in the ranges of [-0.5, 0.5] and [-1, 1]. While a loss of noise resiliency is observed, a trade-off is necessary to gain generalizability for varying force parameters. One may suspect that higher noise levels may confuse the network, and the expectation of a wide variety of solutions reduces the model's robustness.

4.3. Koopman eigenvalues

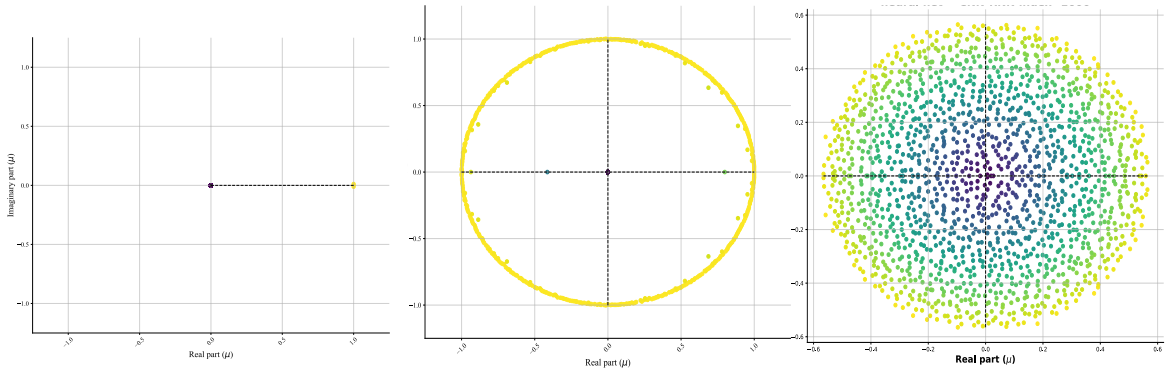

Fig. 8. Eigenvalues of left) DMD, Middle) Hankel-DMD methods. Right) Our approach Koopman Layer Eigenvalues

A critical aspect of our study is analyzing the Koopman operator's eigenvalues derived from our neural network approach compared to traditional methods. Figure 8 contrasts the eigenvalues obtained via DMD, Hankel-DMD, and our proposed method. Traditional methods like DMD exhibit limitations in capturing a diverse set of eigenvalues due to the small number of system state variables. While Hankel-DMD improves upon this, most of its Koopman evolution eigenvalues reside on the unit circle, leading to long-term instability, as shown in Figure 9.

An earlier work [20] suggests approximating the Koopman operator evolution matrix using radial basis functions to lessen this constraint. Even so, the eigenvalue spread in the unit circle, their quantity remains small. In contrast, our approach yields a rich distribution of eigenvalues within a 0.6-radius circle centered at (0,0), as shown in Figure 8 right image. This distribution is directly linked to enhanced noise stability and robustness.

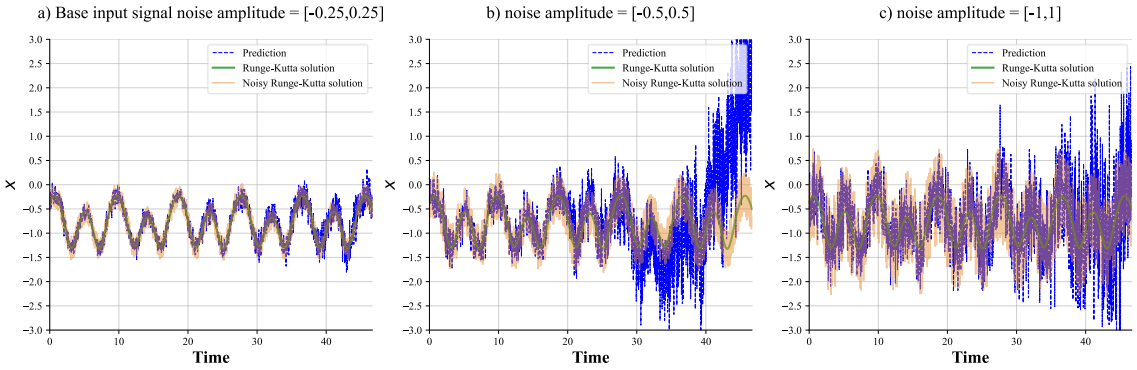

Figure 9. Result of the Hankel-DMD

5. Conclusion

This research mainly focused on implementing an Autoregressive Self-Supervising Autoencoder through a two-stage training process. Our primary objectives are to constrain the Koopman operator within a specific layer of the neural network, evaluate its robustness against noise, and predict the steady-state behavior of the system while ensuring resilience across various initial conditions and maintaining consistency with the Duffing coefficient.

Our methodology yielded unique eigenvalues of the Koopman generator matrix, enabling the Koopman operator to capture various Duffing equation behaviors and be noise-resistant. This entails capturing period-doubling bifurcations and properly predicting periodic oscillations while providing a more interpreted neural network framework. Although our method could accurately predict the Duffing oscillator's future behavior, chaotic behavior could not be predicted.

Furthermore, compared to conventional methods such as Dynamic Mode Decomposition (DMD) and Extended DMD (EDMD), our approach exhibited superior performance, particularly evident in the increased diversity of Koopman eigenvalues. Notably, our methodology surpassed radial basis function (RBF) approaches in this regard, highlighting its efficacy in capturing the intricate dynamics of nonlinear systems.

In conclusion, our research contributes to the advancement of predictive modeling techniques for nonlinear dynamical systems, offering insights into the robustness and versatility of autoregressive neural network architectures. While challenges remain, our findings pave the way for future developments in leveraging deep learning methodologies for nonlinear system analysis and prediction.

References

[1] B.O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proceedings of the National Academy of Sciences, 17 (1931) 315-318.
[2] G.E. Hinton, R.R. Salakhutdinov, Reducing the dimensionality of data with neural networks, science, 313 (2006) 504-507.
[3] B. Lusch, J.N. Kutz, S.L. Brunton, Deep learning for universal linear embeddings of nonlinear dynamics, Nature communications, 9 (2018) 4950.

[4] K. Champion, B. Lusch, J.N. Kutz, S.L. Brunton, Data-driven discovery of coordinates and governing equations, Proceedings of the National Academy of Sciences, 116 (2019) 22445-22451.
[5] E. Kaiser, J.N. Kutz, S.L. Brunton, Sparse identification of nonlinear dynamics for model predictive control in the low-data limit, Proceedings of the Royal Society A, 474 (2018) 20180335.
[6] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Communications of the ACM, 60 (2017) 84-90.
[7] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, (2014).
[8] K.He, X.Zhang, S.Ren,J.Sun,Deep residual Learning for image recognition, in: Proceeding of the IEEE conference on computer vision and pattern recognition,(2016)770-778.
[9] S.L. Brunton, M. Budišić, E. Kaiser, J.N. Kutz, Modern Koopman Theory for Dynamical Systems, SIAM Review, 64 (2022) 229-340.
[10] C.W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, D.S. Henningson, Spectral analysis of nonlinear flows, Journal of fluid mechanics, 641 (2009) 115-127.
[11] P.J. Schmid, L. Li, M.P. Juniper, O. Pust, Applications of the dynamic mode decomposition, Theoretical and computational fluid dynamics, 25 (2011) 249-259.
[12] M.O. Williams, I.G. Kevrekidis, C.W. Rowley, A data–driven approximation of the koopman operator: Extending dynamic mode decomposition, Journal of Nonlinear Science, 25 (2015) 1307-1346.
[13] F. Moon, P.J. Holmes, A magnetoelastic strange attractor, Journal of Sound and Vibration, 65 (1979) 275-296.
[14] K.A. Kumar, S.F. Ali, A. Arockiarajan, Magneto-elastic oscillator: Modeling and analysis with nonlinear magnetic interaction, Journal of Sound and Vibration, 393 (2017) 265-284.
[15] G. Duffing, Ingenieur: Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung. Sammlung Vieweg. Heft 41/42, Braunschweig 1918. VI+ 134 S, Zamm-zeitschrift Fur Angewandte Mathematik Und Mechanik, 1 (1918) 72.
[16] C. Runge, Über die numerische Auflösung von Differentialgleichungen, Mathematische Annalen, 46 (1895) 167-178.
[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1-9.
[18] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86 (1998) 2278-2324.
[19] J. Schmidhuber, S. Hochreiter, Long short-term memory, Neural Comput, 9 (1997) 1735-1780.
[20] Q. Li, F. Dietrich, E.M. Bollt, I.G. Kevrekidis, Extended dynamic mode decomposition with dictionary learning: A data-driven adaptive spectral decomposition of the Koopman operator, Chaos: An Interdisciplinary Journal of Nonlinear Science, 27 (2017).