

Real-Time State Estimation Using Recurrent Neural Network and Topological Data Analysis

Arman Razmarashooli^a, Daniel A. Salazar Martinez^a, Yang Kang Chua^b, Simon Laflamme^{a,c},
and Chao Hu^b

^aDepartment of Civil, Construction, and Environmental Engineering, Iowa State University,
Ames, IA, 50010, USA

^bDepartment of Mechanical Engineering, University of Connecticut, Storrs, CT 06269, USA

^cDepartment of Electrical and Computer Engineering, Iowa State University, Ames, IA, 50010,
USA

ABSTRACT

High-rate systems are defined as physical systems that undergo large perturbations, often exceeding 100 g 's, over very short durations, often less than 100 milliseconds. Examples include blast mitigation mechanisms and advanced weaponry. The use of control feedback to empower high-rate systems requires the capability to estimate system states of interest in the realm of microseconds. However, due to the dynamics of these high-rate systems being highly nonlinear and nonstationary, it is challenging to predict their behavior using conventional state estimation methods. To address this issue, we conduct a study that explores the integration of topological data analysis (TDA) and recurrent neural network (RNN) to improve predictive capabilities for high-rate systems. Here, TDA features are used as the input to a machine learning algorithm to determine the state of a high-rate system. We conduct practical evaluations using laboratory datasets from experiments in the dynamic reproduction of projectiles in ballistic environments for advanced research (DROPBEAR), focusing on localizing fast-changing boundary conditions on a cantilever beam. The study demonstrates the ability of the method to classify and predict a system's fundamental frequencies. This approach helps understand the structure of the underlying high-rate dynamics, leading to improved accuracy and precision in state estimation and prediction.

Keywords: Structural health monitoring, high-rate systems, nonlinear time series, topological data analysis, recurrent neural network, ensemble learning

1. INTRODUCTION

Algebraic topology and topological data analysis (TDA) have emerged as powerful tools in understanding complex dynamic systems by extraction of pseudo-states and shape-like structures in data.¹ The significance of these methods lies in their ability to represent the essential characteristics of the system by utilizing Taken's embedding theorem, which facilitates the mapping of pseudo-states to dynamic states through a machine learning algorithm.² This approach improves our understanding of complex problems across multiple fields, ranging from the analysis of 3D shapes to the classification of electrocardiogram signals and time series data. The integration of TDA with machine learning, particularly in analyzing non-linear datasets in noisy environments, has shown promise in improving the accuracy and efficiency of classification algorithms.

High-rate systems are defined as those experiencing dynamic events with amplitudes exceeding 100 g_n over durations of less than 100 ms, such as blast mitigation systems, advanced weaponry, automotive airbag deployment mechanisms, and hypersonic vehicles.^{3,4} The primary motivation behind conducting real-time state estimation for these systems lies in enabling feedback mechanisms to enhance operational performance and safety. Yet, this task is difficult, because these systems comprise large uncertainties in external loads, high levels of non-stationarity and heavy disturbances, and unmodeled dynamics resulting from changes in system configuration. Physics-based techniques have been proposed to conduct high-rate structural health monitoring, including a

Further author information: (Send correspondence to A. R.)

A.R.: E-mail: shooli@iastate.edu, Telephone: 1 515 815 0366

sliding mode observer-based algorithm for real-time state estimation based on a physical representation,⁵ and a model reference adaptive system-based algorithm that can achieve computation speed in the sub-millisecond time-frame.⁶ Data-based techniques have also been explored. A key challenge in data-based formulations is in data scarcity arising, because high-rate experiments are expensive to conduct and may only include limited dynamic responses.⁷ To address this issue, several studies proposed the use of transfer learning to train algorithms.^{8,9} Of interest, recurrent neural networks are promising for modeling and predicting time series due to their sequential organization, but they face the challenge of vanishing gradient issues. Gated neural networks, such as long short-term memory (LSTM) architectures, can overcome these issues.¹⁰ Nonlinear time series modeling has been effectively implemented using RNNs with LSTM cells.¹¹ In particular, the authors have previously shown promise in forecasting non-stationary time series in high-rate systems by decomposing the nonstationary system into stationary systems and combining predictions from an ensemble of RNNs using a weighted sum.⁷

The focus of this paper is the real-time state estimation of dynamic systems using data-based techniques to map time series measurements to structural states using TDA features, emphasizing the novel application of these features as input to recurrent neural network (RNN) algorithms. The objective of this approach is to improve the predictive capabilities of the RNNs using physical insights provided by TDA. This paper investigates the potential of TDA features extracted from point clouds from time series data formed by delay vectors as input features for RNN algorithms constructed with LSTM cells. The performance of the method is evaluated numerically using laboratory datasets obtained from the dynamic reproduction of projectiles in ballistic environments for advanced research (DROPBEAR) testbed. This dataset is selected due to the wide availability of experimental datasets and the suitability of DROPBEAR datasets for validating and benchmarking the performance of high-rate algorithms.⁷

The rest of the paper is organized as follows. Section 2 provides background information on time series embedding and computational topology. Section 3 describes the datasets and the overall methodology used in validating the algorithm. Section 4 presents and discusses the results. Section 5 presents our conclusion and recommendations for future work.

2. BACKGROUND

In computational topology, we characterize the topological space by evaluating homology groups in a point cloud formed by a finite number of points.¹² The computation of homology groups can be done through a variety of methods, but typically starts with the construction of simplicial complexes that create spaces using simplices.¹³ To determine topological invariants of point cloud data, a common approach is to assign a persistent simplicial complex to the data and define its invariants. The simplicial complex associated with the point cloud is dependent on the scale parameter ϵ . There are two major methods to associate a simplicial complex with the point cloud data, namely the Cech Complex and Vietoris-Rips complex.² In this paper, our method utilizes the Vietoris-Rips complex due to its computation efficiency for large datasets. With the Vietoris-Rips method, a group of points span a k -simplex if and only if the pairwise distances between the vertices are all less than or equal to ϵ .¹³

In practical applications, determining the value of ϵ can be challenging as it often remains unknown. To address this challenge, one can use a flexible approach that consists of keeping a record of various ϵ values, and for each ϵ generate simplicial complexes. As the ϵ parameter varies, the homology groups change, and new topological features may appear or disappear. Persistent homology tracks these changes by associating each topological feature with a birth and death time. The output of persistent homology is a collection of all topological features that can be represented by birth and death times in the form of a persistence diagram.¹⁴ These diagrams can be used to extract several TDA features, including bottleneck and Wasserstein Distance, the persistence landscape and silhouette, and the number of off-diagonal points.¹⁵ To illustrate, Figure 1 shows the Filtration of Rips complexes obtained from a point cloud consisting of 20 randomly scattered points in space, each surrounded by expanding balls of radius ϵ . Initially, there are 20 connected components, one for each point, forming feature H_0 (Figure 1(a)). As the radius of the balls increases, they create a simplicial complex, here at $\epsilon = 0.85$. Under that radius, the last H_0 dies and H_1 is born from the creation of a hole (Figure 1(b)). The hole H_1 dies at $\epsilon \geq 1.86$ (Figure 1(c)). The associated persistence diagram that tracks the evolution of H_0 and H_1 is plotted in Figure 1(d).

In practice, data are often collected using sensors, which are then represented as time series. To perform topological data analysis this time series must be converted into a point cloud format. This conversion is effectively conducted using the framework established by Taken's embedding theorem and its extensions.^{2,16,17} It consists of the transformation of time series data $\mathbf{x}(t) = [x_1, x_2, \dots, x_m]$ into delay vectors $\chi(t) = [\mathbf{x}(t), \mathbf{x}(t - \tau), \mathbf{x}(t - 2\tau), \dots, \mathbf{x}(t - (d - 1)\tau)]$. Based on Taken's embedding theorem, choosing the appropriate embedding dimension and time delay can create an embedding that is topologically equivalent to the original dynamic system. This technique can be utilized to calculate topological invariants of the system. Typically, the embedding dimension is determined using the false nearest neighbor test and time delay through mutual information.⁷ It is important to note that the Embedding theorem, regardless of its extension, does not apply to non-stationary systems.

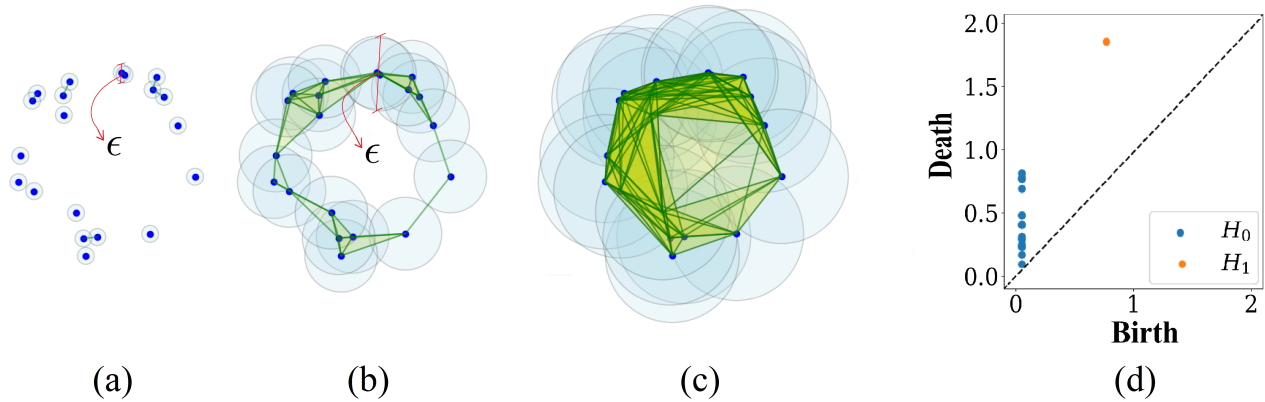


Figure 1: The filtration of Rips Complex. There are 20 points and 20 individual connected components in (a) (feature H_0). As the balls of radius ϵ expand, they connect and merge into one connected component. Eventually, all of the points connect, after which H_0 dies and a hole (feature H_1) is initially born within the connected components for $\epsilon = 0.85$ (b). The hole dies for $\epsilon \geq 1.86$ as shown in (c). The associated persistence diagram is plotted in (d).

3. METHODOLOGY

The DROPBEAR testbed was developed by the Air Force Research Laboratory to validate state estimation algorithms for high-rate systems.⁵ As shown in Figure 2, this testbed consists of a cantilever beam subjected to a fast-moving boundary condition using a movable cart. The dynamics of the beam can be altered by the cart to simulate gradual changes in stiffness and/or excited by a impact hammer to simulate an impact load. DROPBEAR can be simplified as a beam with a variable effective length, supported by fixed and roller supports at its ends. The beam typically experiences free vibration, and in this context, its response $x(t)$ can be simplified to a single harmonic time series:

$$x(t) = A \cos(\omega t) = \cos(2\pi f t) \quad (1)$$

where A is the amplitude, $x(t)$ is the vertical acceleration, f is the natural (first) cyclic frequency in Hz, and t is time in seconds. The fundamental frequency of the beam can be shown to be inversely proportional to the square of its effective length.¹⁸

Previous studies have shown that the optimal embedding dimension for a single harmonic signal is two.¹⁹ However, this study considers a dimension of three to account for noise, and thus gives rise to a 3-dimensional point cloud with its related persistence diagram features including connected components (H_0), loops (H_1), and voids (H_2). In noise-free data, the maximum persistence of H_0 represents the maximum Euclidean distance between points in the point cloud and it changes with varying frequency. The number of H_0 points in the persistence diagram represents the sampling rate of the system. Under high sampling rates, the value of H_0

converges to zero. Even though the number of H_0 points is a non-informative feature, under a fixed sampling rate, the normalized maximum persistence of H_0 can be a reasonable indicator of the system's frequency.

The maximum persistence of H_1 represents the evolution of loops in the point cloud, an indication of periodicity in the signal. As the frequency increases, the hole persists longer. Under this scenario, the maximum persistence of H_1 is born when the maximum persistence of H_0 dies. To remove the effect of sampling rate on the maximum persistence of H_1 , we introduced a new definition for maximum persistence of H_1 considering zero birth time which provides a better understanding of the persistence of topological features compared to normalized maximum persistence of H_0 . Using a time delay parameter $\tau = 0.25/f$, the harmonic signals form a unit circle in phase-space. Below this τ , an ellipse forms, where its minor to major axis ratio affects the maximum persistence of H_1 so more circular ellipses yield higher persistence. Thus, the frequency of the harmonic signal, under a specific τ , influences the maximum persistence of H_1 by affecting the ellipse's shape. It is important to note that in noise-free scenarios H_2 is absent because of the repetitive nature of periodic functions. When dealing with a simple harmonic signal, we can easily correlate most of the TDA features to the maximum persistence values within each homology group.²⁰

To analyze our dynamics of interest, we had to address the nonstationarity caused by the moving boundary condition. To do so, we applied two sliding windows over the dataset to extract local TDA features assuming our system within these windows is stationary. The first window has a size of $w = 1/f_{\min} + \tau$, with $f_{\min}$ being the minimum frequency of the system to ensure that the point cloud will form a complete loop, representing H_1 . We define a second window that is smaller in size ($w = 1/f_{\max} + \tau$), which ensures that the points do not overlap because as the frequency of the system increases, loops are likely to overlap, which can affect the value of the maximum persistence of H_0 .

4. RESULTS AND DISCUSSION

We study the performance of our windowing method to extract the physically meaningful TDA feature H_0 and H_1 for a non-stationary harmonic excitation. This study is conducted over two types of datasets: 1) a synthetic noise-free harmonic signal; and 2) experimental data from DROPBEAR.

4.1 Synthetic Harmonic Signal

The synthetic harmonic signal is taken as

$$x(t) = \cos(2\pi f(t)t) \quad (2)$$

with $f(t)$ varying between 20 and 50 Hz. The excitation is plotted in Figure 3 (grey solid line). In this excitation, the frequency remains constant at 20 Hz for the first two seconds (0-2 seconds) before increasing to 50 Hz over the next two seconds (2-4 seconds), after which it remains at 50 Hz for another 2 seconds (4-6 seconds), and then decreases to 20 Hz and remains constant for another 2 seconds (6-8 seconds). The size of the moving windows is $w_1 = 1/f_{\min} + \tau = 0.05 + 0.001 = 0.051$ seconds for plotting H_1 , and $w_2 = 1/f_{\max} + \tau = 0.02 + 0.001 = 0.021$ seconds for plotting H_0 . Data are embedded using $\tau = 0.001$ seconds for both windows ($0.25/f_{\max} = 0.005$ seconds). Figure 3 plots the evolution of maximum persistence of H_0 and H_1 compared to frequency signal. We can observe that the normalized values of the maximum persistences of H_0 and H_1 change proportionally with the frequency of the system, but with a delay compared to system's change in frequency. This delay is caused by the use of sliding windows take measurements in the past, and the magnitude of that delay depends on the size of sliding window and is greater for H_1 due to the larger window size in use.

4.2 DROPBEAR

The investigation is pursued using realistic datasets taken from DROPBEAR experimental testbed. The first two datasets correspond to Dataset-6, where one experiment does not involve the use of an impact hammer (test 9) and the other does (test 11).²¹ In these tests, the cart is initially located 50 mm from the clamp, moves at 200 mm from the clamp, and comes back to its original position. The temporal location of the cart is plotted in Figures 5 and 6 (black line). The maximum velocities and accelerations for both tests are 120

mm/s and 250 mm/s^2 , respectively. The system's fundamental frequency varies between 17.7 Hz (at 50 mm) and 31.0 Hz (at 200 mm). Accelerometer data was acquired at 25 kHz. Data is embedded using $\tau = 0.004$ seconds ($0.25/f_{\max} = 0.008$ seconds) and the size of the moving windows are taken as $1/f_{\min} + \tau = 0.06$ and $1/f_{\max} + \tau = 0.04$ seconds. The window slides every 0.001 seconds to lessen computational demands. Figures 5 and 6 plot the maximum persistence for one homology group (H_1) for tests 9 and 11, respectively. Results are compared against those obtained using a short-time Fourier transform (STFT) conducted under a window size of 4096 samples (equivalent to 0.164 seconds) and a window length overlap of 0.008 seconds. In these tests, the initial part of the excitation produces a very noisy TDA feature due to the low acceleration of the signal when the cart starts moving. As the cart begins to move, the maximum persistence of H_1 correlates with the cart's location. In the middle and at the end of the excitation when the cart stays at one location, the signal becomes less noisy due to the presence of free vibrations, leading to clearer TDA features. In test 11, due to the hammer impact, we observe a richer signal and less noisy TDA features.

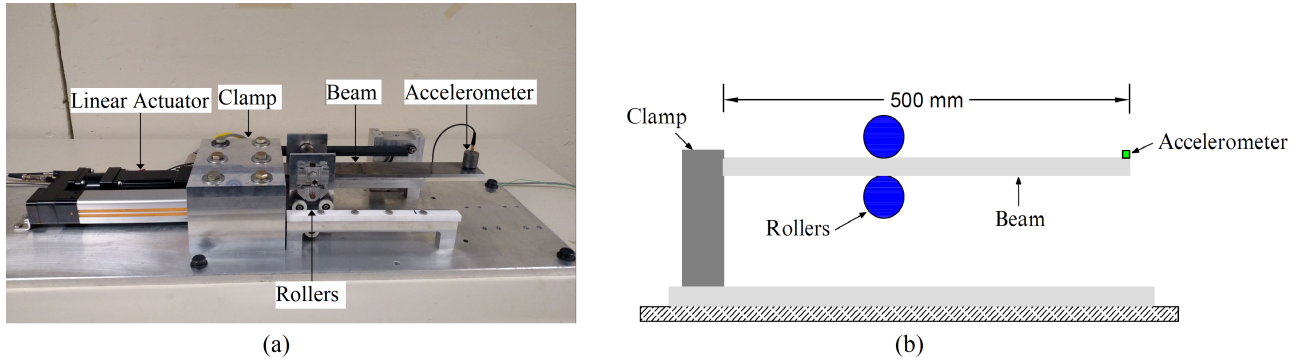


Figure 2: DROPBEAR testbed: (a) picture of the testbed; and (b) its schematic diagram.

4.3 Recurrent Neural Network Architecture

In this study, we employed an LSTM-based RNN architecture to model complex dynamics between cart location and the maximum persistence of H_1 . Our model consisted of a sequential stack of LSTM cells of 512 units, integrated with batch normalization and a dropout rate of 0.2 to enhance model generalization and avoid overfitting, followed by another LSTM layer of identical configuration to refine the features extracted. The model was compiled with Adam optimizer and mean squared error loss function. Training was conducted over 550 epochs with a batch size of 128, leveraging a split of 70% training and 30% validation data to monitor and prevent overfitting.

Our experimental setup has distinct datasets for training and testing phase to evaluate the model's predictive capability. The algorithm was trained using data from test 9, which does not have the impact hammer. Subsequently, the model's generalization were tested using data from test 11 which has three hammer impacts.

The aim of this research is to determine if TDA features, specifically the maximum persistence of H_1 , can be utilized in RNN to predict the dynamics of high-rate systems. The findings from DROPBEAR demonstrate that the maximum persistence of H_1 can be linked to the cart's location. To assess the outcomes of the RNN, results from the STFT are post-processed to map to the cart location by assuming a quadratic (physical) relationship. Regression parameters are estimated using a least square estimator (LSE), which are listed in Table 1.

Next, we investigate how well the maximum persistence of H_1 can be utilized to find the right cart location. We evaluate its performance through three performance metrics (J_1 , J_2 , and J_3) on the parts signal that remain after removing the initial noisy section.

Metric J_1 is the mean absolute error between the actual cart location and the mapped cart location for the entire excitation or section of the excitation. Metric $J_{2,i}$ is the number of instances over the signal that the estimation does not remain below a given distance i (in mm, here taken as $i = 5, 10$, and 20 mm, with 20 mm arbitrarily taken as an acceptable target). Metric $J_{3,i}$ is the mean absolute error of the cart location when the estimated cart position is within the given distance i (in mm). Table 2 presents results for these performance

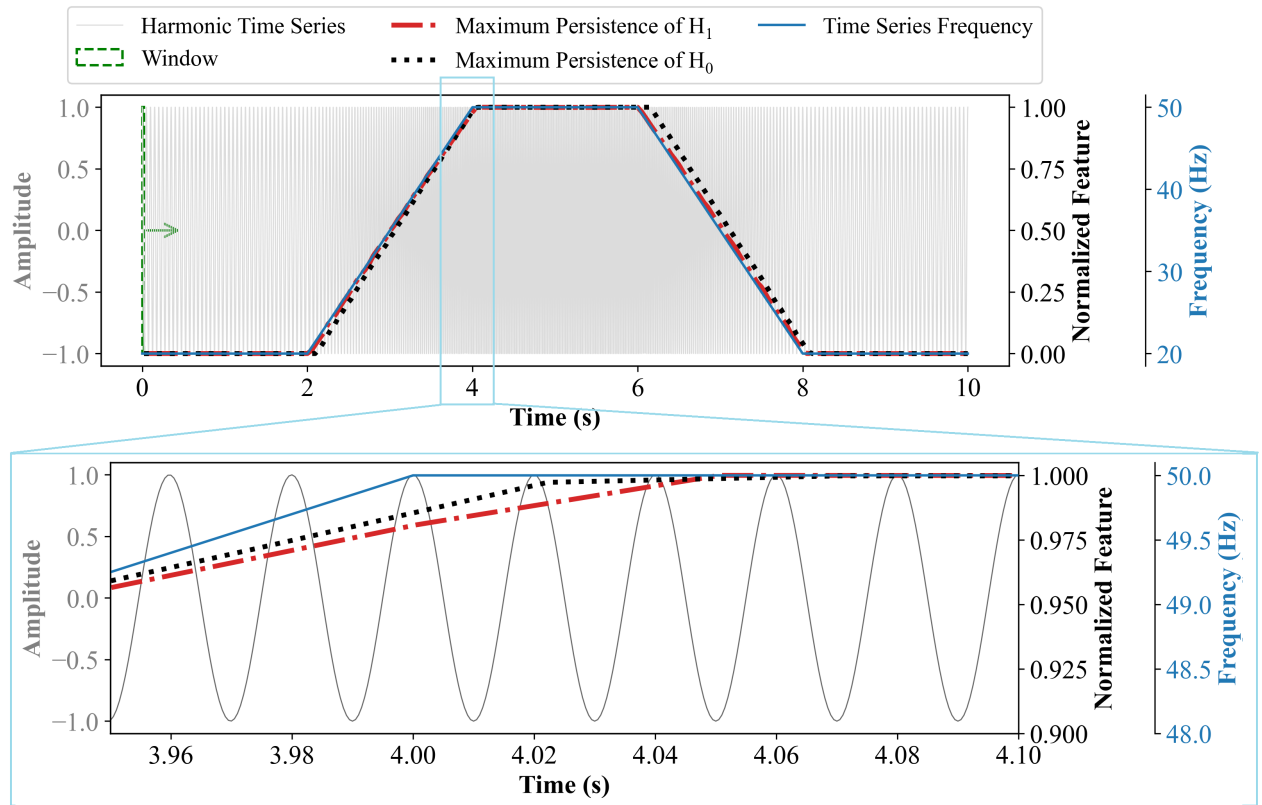


Figure 3: Evolution of TDA features constructed from H_0 and H_1 over a synthetic, noise-free harmonic signal.

metrics across the test. Metric J_1 shows that our TDA-based method remains within acceptable ranges (under 20 mm), and that it tends to perform similarly to the STFT over very rapid changes in movements. Metric J_2 shows that our TDA-based method maps within 20 mm of the correct cart location usually above 86% of the time. Metric J_3 shows that our TDA-based method performed at least similarly to the STFT-based method, if not better ($J_{3,20}$).

Table 1: Results from the LSE

Test	Feature	Intercept	Slope	R^2
Test 9	STFT	-0.13	1.070	0.98
Test 11	STFT	-0.11	1.059	0.96

Table 2: Performance results for cart localization

Test	Feature	J_1 (mm)	$J_{2,5}$ (%)	$J_{3,5}$ (mm)	$J_{2,10}$ (%)	$J_{3,10}$ (mm)	$J_{2,20}$ (%)	$J_{3,20}$ (mm)
Test 9 (training dataset)	H_1	14.0	88	2.4	42	6.8	14	8.7
	STFT	11.9	81	2.8	48	5.6	14	10.6
Test 11 (testing dataset)	H_1	15.1	75	2.9	47	4.8	10	9.7
	STFT	12.1	84	2.4	52	5.6	14	10.8

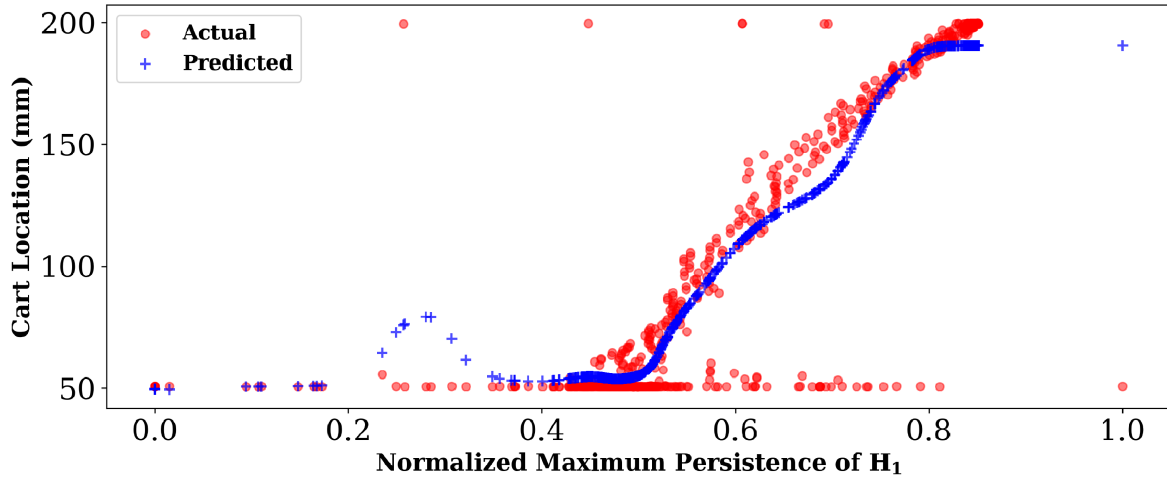


Figure 4: Actual vs. predicted cart Location as a function of normalized H_1 : RNN model accuracy for test 9

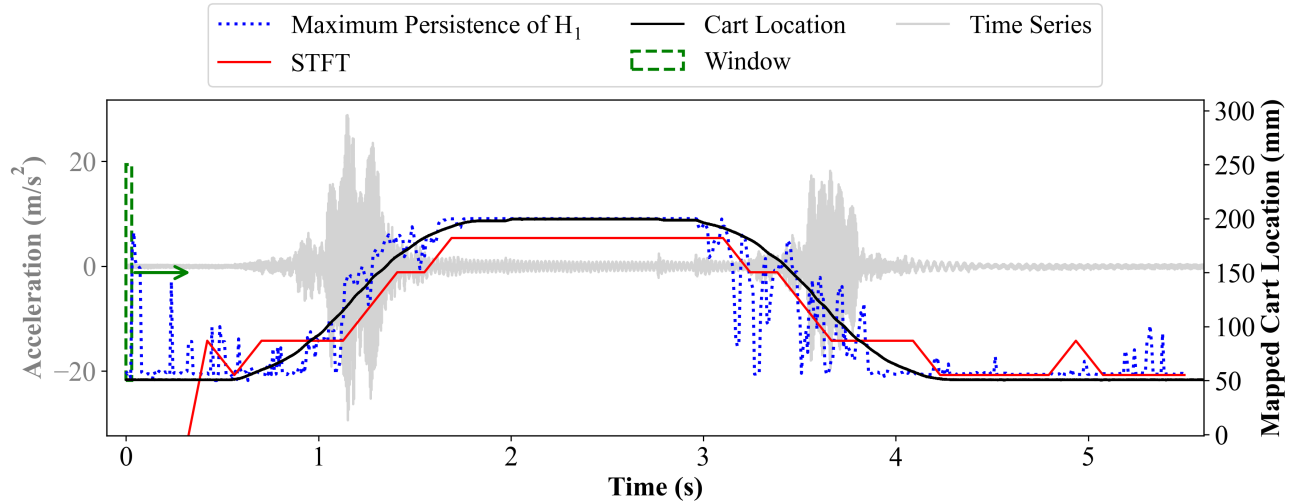


Figure 5: Evolution of TDA features over test 9 (training dataset)

5. CONCLUSION

This research explored the effectiveness of employing topological data analysis (TDA) features as input to a recurrent neural network for real-time estimation of high-rate systems. Specifically, the study focused on detecting a rapidly moving cart using acceleration measurements. Initially, we simplified the problem by linking TDA features to the frequency of a single harmonic signal. We investigated key TDA features from a physical perspective for their applicability to high-rate systems and introduced a novel definition for the maximum persistence of H_1 , which represents the largest void in the data. To address the nonstationary nature of these systems, we proposed a method that utilizes multiple resolutions and sliding windows, along with a technique for converting time series data into a point cloud. We tested this approach on synthetic data and demonstrated its effectiveness by comparing the signal's frequency to the maximum persistence of H_1 and H_0 . The method was applied to real data from DROPBEAR, focusing on the maximum persistence of H_1 as input in an RNN with one LSTM and multiple layers. Our results indicated that RNNs trained solely on the maximum persistence of H_1 could accurately determine the cart's location. This observation remained true even under conditions of higher cart acceleration and confirmed that trained RNNs could outperform the short-time Fourier transform

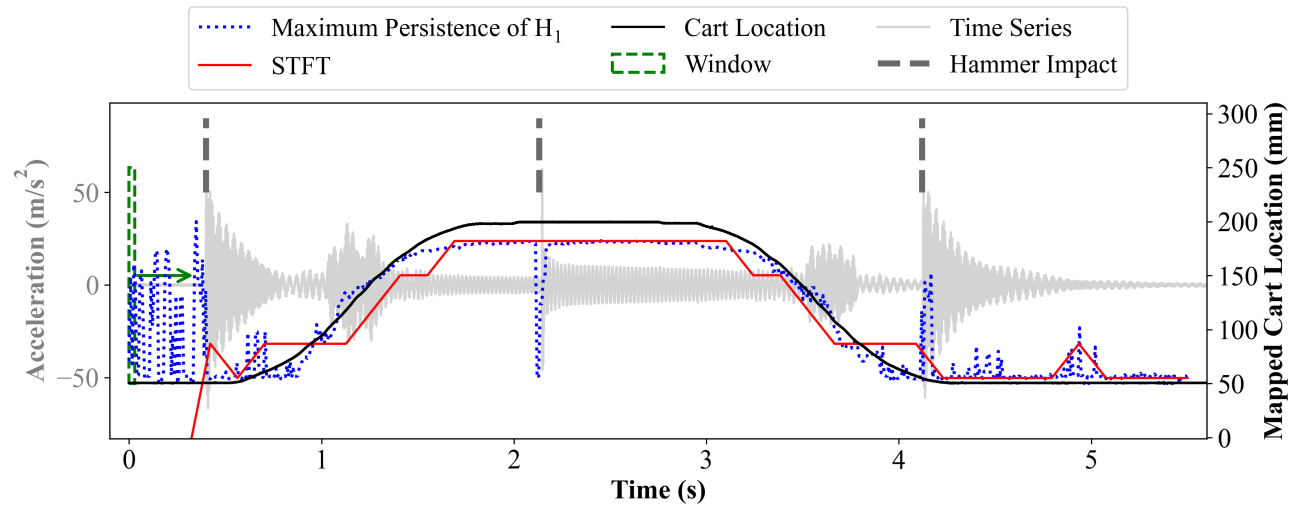


Figure 6: Evolution of TDA features over test 11 (testing dataset)

(STFT) in some cases.

Overall, our findings suggest that TDA features are effective for monitoring dynamic systems with a dominating, nonstationary fundamental frequency, and they can outperform traditional methods such as STFT under certain scenarios. Future work will extend the study to multi-harmonic systems and explore the formulation of ensemble RNNs to better link TDA features to the system's states.

ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support from the Defense Established Program to Stimulate Competitive Research (DEPSCoR) award number FA9550-22-1-0303, the Air Force Office of Scientific Research (AFOSR) award number FA9550-23-1-0033, and the National Science Foundation awards numbers 1937460 & 2234919. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the United States Air Force.

REFERENCES

- [1] Schrader, P. T., "Topological multimodal sensor data analytics for target recognition and information exploitation in contested environments," in *[Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII]*, **12547**, 114–143, SPIE (2023).
- [2] Takens, F., "Detecting strange attractors in turbulence," in *[Lecture Notes in Mathematics]*, 366–381, Springer Berlin Heidelberg (1981).
- [3] Hong, J., Laflamme, S., Dodson, J., and Joyce, B., "Introduction to state estimation of high-rate system dynamics," *Sensors* **18**, 217 (Jan. 2018).
- [4] Dodson, J., Downey, A., Laflamme, S., Todd, M. D., Moura, A. G., Wang, Y., Mao, Z., Avitabile, P., and Blasch, E., "High-rate structural health monitoring and prognostics: An overview," in *[Data Science in Engineering, Volume 9]*, 213–217, Springer International Publishing (Oct. 2021).
- [5] Joyce, B., Dodson, J., Laflamme, S., and Hong, J., "An experimental test bed for developing high-rate structural health monitoring methods," *Shock and Vibration* **2018**, 1–10 (jun 2018).
- [6] Yan, J., Laflamme, S., Hong, J., and Dodson, J., "Online parameter estimation under non-persistent excitations for high-rate dynamic systems," *Mechanical Systems and Signal Processing* **161**, 107960 (Dec. 2021).
- [7] Barzegar, V., Laflamme, S., Hu, C., and Dodson, J., "Ensemble of recurrent neural networks with long short-term memory cells for high-rate structural health monitoring," *Mechanical Systems and Signal Processing* **164**, 108201 (Feb. 2022).

- [8] Bengio, Y., “Deep learning of representations for unsupervised and transfer learning,” in [*Proceedings of ICML workshop on unsupervised and transfer learning*], 17–36, JMLR Workshop and Conference Proceedings (2012).
- [9] Raghu, M., Zhang, C., Kleinberg, J., and Bengio, S., “Transfusion: Understanding transfer learning for medical imaging,” *Advances in neural information processing systems* **32** (2019).
- [10] Hu, Y., Huber, A., Anumula, J., and Liu, S.-C., “Overcoming the vanishing gradient problem in plain recurrent networks,” *arXiv preprint arXiv:1801.06105* (2018).
- [11] Yeo, K. and Melnyk, I., “Deep learning algorithm for data-driven simulation of noisy dynamical system,” *Journal of Computational Physics* **376**, 1212–1231 (2019).
- [12] Alexander, Z., *A topology-based approach for nonlinear time series with applications in computer performance analysis*, PhD thesis, University of Colorado at Boulder (2012).
- [13] Carlsson, G. and Vejdemo-Johansson, M., [*Topological data analysis with applications*], Cambridge University Press (2021).
- [14] Edelsbrunner, H., Letscher, D., and Zomorodian, A., “Topological persistence and simplification,” in [*Proceedings 41st Annual Symposium on Foundations of Computer Science*], 454–463 (2000).
- [15] Hatcher, A., [*Algebraic topology*], Cambridge University Press, Cambridge (2002).
- [16] Stark, J., “Delay embeddings for forced systems. i. deterministic forcing,” *Journal of Nonlinear Science* **9**, 255–332 (1999).
- [17] Stark, J., Broomhead, D. S., Davies, M. E., and Huke, J., “Delay embeddings for forced systems. ii. stochastic forcing,” *Journal of Nonlinear Science* **13**, 519–577 (2003).
- [18] Kausel, E., [*Advanced Structural Dynamics*], Cambridge University Press (2017).
- [19] Small, M., [*Applied nonlinear time series analysis: applications in physics, physiology and finance*], vol. 52, World Scientific (2005).
- [20] Razmarashooli, A., Chua, Y. K., Barzegar, V., LIU, H., Laflamme, S., Hu, C., Downey, A. R., and Dodson, J., “Topological data analysis for real-time extraction of time series features,” *STRUCTURAL HEALTH MONITORING 2023* (2023).
- [21] Nelson, M., Laflamme, S., Hu, C., G.Moura, A., Hong, J., Downey, A., Lander, P., Wang, Y., Blasch, E., and Dodson, J., “Generated datasets from dynamic reproduction of projectiles in ballistic environments for advanced research (DROPBEAR),” *Data in Brief* (2022).