A MODEL-FREE CAUSALITY MEASURE BASED ON MULTI-VARIATE DELAY EMBEDDING

Saba Emrani and Hamid Krim, Fellow, IEEE

North Carolina State University, Electrical and Computer Engineering Department, Raleigh, NC
{semrani, ahk}@ncsu.edu

ABSTRACT

We further delay embedding framework application to multiple time series for extraction of their potential causal interactions. We introduce a novel geometric model-free causality measure that can efficiently detect linear and nonlinear causal interactions between time series with no prior information or parameter estimation. Using multivariate delay embedding, we construct a point cloud from a set of time domain signals and propose the inverse of its fractal dimension as a causal interaction measure between the corresponding time series. Correlation dimension estimation is fully exploited as a fractal-based method for uncovering the dimensions of generated point clouds. Extensive simulation results are presented to substantiate the capabilities of the proposed approach.

***Index Terms*—** Causal interaction, multivariate delay embedding, time series analysis.

1. INTRODUCTION

Causal interaction among a set of signals is deduced if a significant potentially time-lagged influence between the corresponding time series is discovered. Causality detection has important applications in neuroscience [1–3], economics [4] and computational biology [5]. Various causality measures have been proposed for quantifying this influence, many of which have important limitations. These approaches are mainly formulated using autoregressive(AR) models such as Granger causality [6], the cross spectrum such as coherence and phase slope index [7], or directed transfer function [8], and partial directed coherence methods [9]. Most of these measures are rooted in linear regression modeling and cannot identify the nonlinear interactions between signals. For example Granger causality as a widely used tool in many valuable applications in neuroscience is based on a linear bivariate autoregressive model. It has also been recently extended based on a nonlinear autoregressive model by fitting a nonlinear polynomial model [10]. This however requires prior knowledge of the system model. Since in many applications such as brain data analysis, the interaction structures are unknown in advance, a model-free causality measure which accounts for nonlinear as well as linear causal interactions is highly desirable.

In this paper, we present a geometric, non-parametric and model-free causality measure based on manifold learning that can efficiently detect linear and nonlinear causal interactions between time series. This framework is based on the concept of multivariate delay embedding. The mathematical foundation of delay coordinate embedding method was first proposed by Takens [11] to capture the dynamics of a time series in a higher dimensional space. There is a large body of literature on the application of delay embedding to dynamical systems with chaotic attractors [12, 13]. We have used univariate time delay embedding as an efficient tool for identification of quasi-harmonic patterns in signals as well as estimating their spectral characteristics [14, 15]. In this study, we extend the delay embedding framework to more than one modality data by developing an approach for evaluating causal interactions using the dimension of multivariate delay embedding. We use fractal dimension as a measure of the intrinsic dimensionality of the point cloud.

The remainder of the paper is organized as follows. In Section 2, we describe in detail our proposed analysis framework. The multivarie delay embedding approach is first formulated in Section 2.1. Correlation dimension estimation is presented in Section 2.2 as a fractal-based method to determine the dimension of multivarie delay embedding point clouds. We present our causal interaction measure in Section 2.3 and validate the introduced approach using experimental results in Section 3. Several examples are provided to evaluate the capability of the proposed measure in detecting directed, linear and nonlinear causal interactions in both noise free and noisy cases. Finally, Section 4 concludes the paper.

2. CAUSALITY BY DELAY EMBEDDINGS

2.1. Multivariate Delay Embedding

Univariate delay embedding for a time series $\{x_n\}_{n=1}^{N}$ is defined as $\{X_n\} = (x_n, x_{n-\tau}, \ldots, x_{n-(m-1)\tau})$, where τ is the time delay, and m is the embedding dimension. With a similar intuition, multivariate delay embedding of p time series

Research supported by National Science Foundation EEC-1160483

$\{x_{i,n}\}_{n=1}^{N}, i = 1, 2, \ldots, p$ can be defined as

$$\{X_n\} = (x_{1,n}, x_{1,n-\tau}, \ldots, x_{1,n-(m_1-1)\tau}, x_{2,n}, x_{2,n-\tau}, \ldots, x_{2,n-(m_2-1)\tau}, \ldots, x_{p,n-(m_p-1)\tau}), \quad (1)$$

where $\boldsymbol{m} = (m_1, m_2, \ldots, m_p)$ is the embedding dimension vector, and determines the number of components included from each time series. The ambient dimension of the embedding space is therefore $M = \sum_{i=1}^{p} m_i$. In a more general case, different τ's for each time series can be used as

$$\{X_n\} = (x_{1,n}, x_{1,n-\tau_1}, \ldots, x_{1,n-(m_1-1)\tau_1}, x_{2,n}, x_{2,n-\tau_2}, \ldots, x_{2,n-(m_2-1)\tau_2}, \ldots, x_{p,n-(m_p-1)\tau_p}), \quad (2)$$

where $\tau = (\tau_1, \tau_2, \ldots, \tau_p)$ is the time delay vector for p time series. Non-uniform multivariate delay embedding which is the most general embedding technique can be defined for m_i varying delays denoted by $l_{ij}, j = 1, \ldots, m_i$ at each time series $\{x_{i,n}\}_{n=1}^{N}$ as $\{X_n\} = (x_{1,n-l_{11}}, x_{1,n-l_{12}}, \ldots, x_{1,n-l_{1m_1}}, x_{2,n-l_{21}}, \ldots, x_{p,n-l_pm_p})$.

2.2. Fractal-based Dimension Estimation

While the multivariate embedding point cloud is embedded in space $\mathbb{R}^M$, its intrinsic dimension is not necessarily M. The intrinsic dimension is defined as the minimum number of free variables needed to represent the data (degrees of freedom) with no loss of information. Equivalently, the intrinsic dimension of a dataset in $\mathbb{R}^M$ is equal to d if its elements lie totally within a d-dimensional subspace of $\mathbb{R}^M$, where $d < M$ [16]. The fractal dimension of a set of points is an important measure of the intrinsic dimensionality of the points. Fractal-based dimension estimation techniques are global methods that can provide a non-integer value as the intrinsic dimension of data. A variety of techniques have been previously proposed for fractal dimension estimation [17]. Correlation dimension [18] is a computationally efficient approach for estimating intrinsic dimension. Suppose $\boldsymbol{S} = \{s_1, s_2, ...s_n\}$ is a set of data points in $\mathbb{R}^M$. The correlation integral $C(n, r)$ is defined as

$$C(n, r) = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} I(\|s_j - s_i\| \leq r), \quad (3)$$

where $I(.)$ is an indicator function, i.e. $I(\lambda) = 1$ if and only if condition λ holds and zero otherwise. Also, $\|s_j - s_i\|$ denotes the Euclidean distance between data points s_j and s_i. $C(n, r)$ is basically the probability of a pair of points having a distance smaller than or equal to r. The correlation fractal dimension D_c of $\boldsymbol{S}$ is then defined as

$$D_c = \lim_{n\to\infty} \lim_{r\to 0} \frac{\ln(C(n, r))}{\ln r}. \quad (4)$$

D_c is estimated in practice by using the plot of $\ln C(n, r)$ with respect to r and by computing the slope of its linear part.

2.3. CIM: Causal Interaction Measure

In this section, we introduce our proposed causal interaction measure based on the intrinsic dimension of multivariate delay embeddings. Consider three time series $X = \{x_n\}_{n=1}^{N}$, $Y = \{y_n\}_{n=1}^{N}$ and $Z = \{z_n\}_{n=1}^{N}$ and their multivariate delay embedding vectors $(x_n, y_{n-\tau_1})$ and $(x_n, z_{n-\tau_2})$. The causal interaction between time series $X = \{x_i\}_{i=1}^{N}$ and $Y = \{y_i\}_{i=1}^{N}$ with delay τ_1 is higher than the causal interaction between time series $X = \{x_i\}_{i=1}^{N}$ and $Z = \{z_i\}_{i=1}^{N}$ with delay τ_2 if the dimension of the point cloud represented by $(x_n, y_{n-\tau_1})$ is lower than the dimension of the point cloud $(x_n, z_{n-\tau_2})$. (The theoretical proof can be found in [19].) In order to find the delay corresponding to the strongest causal interaction between time series, we simply replace Z with Y in the statement above and conclude the following: The causal interaction between time series $X = \{x_i\}_{i=1}^{N}$ and $Y = \{y_i\}_{i=1}^{N}$ with delay τ_1 is higher than the causal interaction between them with delay τ_2 if the dimension of the point cloud represented by $(x_n, y_{n-\tau_1})$ is lower than the dimension of the point cloud $(x_n, y_{n-\tau_2})$.

We denote the dimension of the delay embedding point cloud $(x_n, y_{n-\tau_1})$ as d_{XY} and propose $1/d_{XY}$ as a causal interaction measure (CIM) between time series $X = \{x_i\}_{i=1}^{N}$ and $Y = \{y_i\}_{i=1}^{N}$ with delay τ_1. In other words, there is a directed information flow from X to Y with delay τ_1 with strength $1/d_{XY}$. We exploit the correlation dimension estimation method to compute the dimensions of the point clouds as described in Section 2.2. The calculated fractal dimension and the final CIM values are therefore not necessarily integer numbers. For example, when we construct two-dimensional multivariate emebedding, the smallest dimension value corresponds to a curve and occurs when the highest interaction exists between two time series. Also, the largest possible dimension equals the dimension of the ambient space which corresponds to the case where there is no interaction between the two time series (2 in this case).

2.4. Efficient Estimation: A Computational Approach

As investigating all possible combinations of the whole time series is computationally complex, we use a progressive geometrical method to build the embedding vectors. First, we determine a maximum lag according to the existing practical estimations of maximum delays in the causal interaction between time series $\{x_{i,n}\}_{n=1}^{N}, i = 1, 2, \ldots, p$ as $L_i, i = 1, \ldots, p$. We will then construct a candidate embedding vector using all $L_1 + L_2 + \ldots + L_p$ elements as

$$\boldsymbol{B} = (x_{1,n}, x_{1,n-1}, \ldots, x_{1,n-L_1}, x_{2,n}, x_{2,n-1}, \ldots, x_{p,n-L_p}). \quad (5)$$

The embedding vector $\boldsymbol{b}$ to be chosen is a subset of $\boldsymbol{B}$ including the time series with causal interactions with the corresponding delays. Similar to the concept of Granger causality, we can infer that the selected embedding vector is to increase our knowledge about the future of the system in one or several steps ahead. Suppose that for time series $\{x_{i,n}\}_{n=1}^{N}$, T_i future steps need to be investigated. The future state of the system is then denoted by $X_F = (x_{i,n+1}, x_{i,n+2}, \dots, x_{i,n+T_1})$.

One possible solution for constructing the embedding vector $\boldsymbol{b}$ starts with an empty vector $\boldsymbol{b_0}$. Suppose that the selected embedding vector at step $(j-1)$ is $\boldsymbol{b_{j-1}} = (x_1^\star, x_2^\star, \dots, x_{j-1}^\star)$. At step j, the element $x_j^\star \in \boldsymbol{B} \setminus \boldsymbol{b_{j-1}}$ will be added to $\boldsymbol{b_{j-1}}$ to build a candidate vector for the next step $\boldsymbol{b_j}$. We will then analyze the point cloud of the delay embedding vector $\boldsymbol{b_j}$ with $x_1^\star, x_2^\star, \dots, x_{j-1}^\star$ and $x_j^\star$ as the corresponding coordinates. If the point cloud forms a manifold with an intrinsic dimension less than a predefined threshold, we conclude that there is significant interaction between the components of $\boldsymbol{b_j}$. As a result, we include $x_j^\star$ in the embedding vector. Otherwise, $x_j^\star$ will be skipped and the next element will be taken into consideration. This reconstruction scheme is based on the causality measure presented in Section 2.3, meaning the intrinsic dimension of the multivariate delay embedding is inversely related to the strength of causal interaction between the corresponding time series. The dimensionality estimation of the delay embedding point clouds constructed using real data sets, will be performed using correlation dimension estimation method. The threshold on the obtained fractal dimension can be experimentally determined and components that establish a dimension less than the selected threshold will be included in the embedding vector.

The approach described above is a general formulation and can describe a variety of models. Our previous framework of univariate delay embedding [20, 21] can be obtained when X_F and $\boldsymbol{B}$ include elements from the same time series. Cross modeling happens when X_F has elements from one time series and $\boldsymbol{B}$ from another one. The more general case is mixed modeling when X_F has elements from one time series and $\boldsymbol{B}$ from all the existing time series in the application. The most comprehensive condition is full modeling when X_F and $\boldsymbol{B}$ both have elements from all time series.

3. EXPERIMENTAL VALIDATION

In order to substantiate the ability of the proposed causal interaction measure to identify linear and nonlinear directed interactions within different delays, we analyze various synthetic data including an example with linear directed flow, an auto-regressive(AR) process and unidirectional Hénon map.

Example 1: In this example, we test the capability of our proposed method in detecting simple linear directed causal interaction by considering a process fully dependent on another one. Time series $\{y_i\}_{i=1}^{N}$ is causally affected by series $\{x_i\}_{i=1}^{N}$ with delay 1, while $\{x_i\}_{i=1}^{N}$ is independent, as expressed by the following equation.

$$\begin{aligned} x_i &= w_x, \\ y_i &= a x_{i-1}, \end{aligned} \tag{6}$$

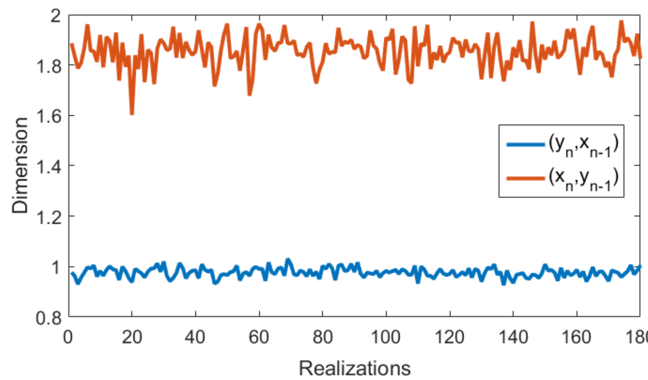

Fig. 1. Estimated dimensions of multivariate delay embeddings of the process presented in Example 1. Blue: dimension of (y_n, x_{n-1}), red: dimension of (x_n, y_{n-1}).

where w_x is zero mean white Gaussian noise with standard deviation of 1 and $a = 0.5$. We constructed multivariate delay embeddings (x_n, y_{n-1}) and (y_n, x_{n-1}) and calculated the dimensions of the two point clouds for 180 different realizations of the random noise. Figure 1 illustrates the estimated dimensions of the point clouds (x_n, y_{n-1}) and (y_n, x_{n-1}) in red and blue, respectively. Clearly, the dimension of (y_n, x_{n-1}) is much lower than (x_n, y_{n-1}), showing strong information flow from X to Y with delay 1. The average dimension over all realizations for (x_n, y_{n-1}) and (y_n, x_{n-1}) are 1.84 and 0.98, respectively. Therefore, the CIM value equals 0.54 for information flow from Y to X and 1.02 for information flow from X to Y, showing a strong flow from X to Y with delay 1. This example validates the ability of the proposed method in detecting the direction of linear causal interaction between two time series. Moreover, the average dimensions of the multivariate delay embeddings for values of delays other than 1 are also much greater than those for (y_n, x_{n-1}). Specifically, the average dimension for (x_n, y_{n-2}) and (y_n, x_{n-2}) are 1.85. This shows the capability of our method in identifying the correct delay in causal interaction.

Example 2: In this example, we validate the capability of the proposed method to identify directed linear causal interactions in a first order AR process, where the second time series is driven by the first one according to the following equation:

$$\begin{aligned} x_i &= 0.5 x_{i-1} + u_i, \\ y_i &= 0.2 y_{i-1} + 0.8 x_{i-1} + v_i, \end{aligned} \tag{7}$$

where u_i and v_i represent independent white Gaussian noise with standard deviation of 1 and 0.3, respectively. Each series consists of 200 samples, and 180 realizations of the process are implemented. We constructed two dimensional multivariate delay embeddings of $\{x_i\}_{i=1}^{N}$ and $\{y_i\}_{i=1}^{N}$. The dimension of (y_n, x_{n-1}) is always less than dimension of (x_n, y_{n-1}) as illustrated in Figure 2 for different realizations. The average dimension over all realizations for (x_n, y_{n-1}) and (y_n, x_{n-1})

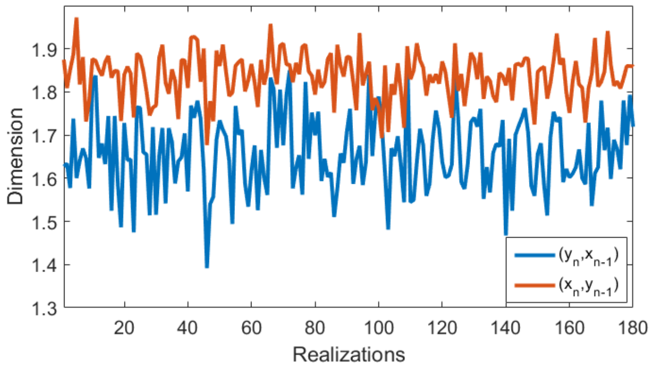

Fig. 2. Estimated dimensions of multivariate delay embeddings of the AR process in Example 2. Blue: dimension of (y_n, x_{n-1}), red: dimension of (x_n, y_{n-1}).

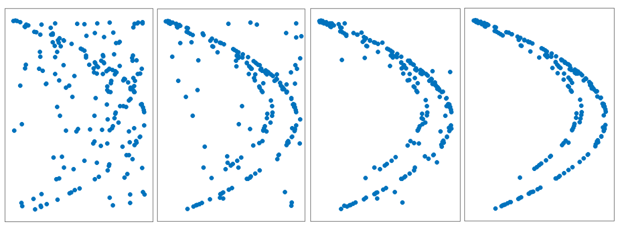

Fig. 3. Hénon system's point cloud as the coupling strength increases in the range $[0.01, 0.6]$.

are 1.83 and 1.65, respectively. Therefore, the CIM value equals 0.54 for information flow from Y to X and 0.61 for information flow from X to Y, confirming the existence of a stronger information flow from $\{x_i\}_{i=1}^N$ to $\{y_i\}_{i=1}^N$ with delay 1, validating the proposed method in identifying directed linear causal interaction.

Example 3: In this example, we validate the effectiveness of the proposed method in detecting nonlinear causal interactions and in identifying different interaction strengths using coupled Hénon system with various coupling strengths. The coupled Hénon map is given by

$$\begin{aligned} x_i &= 1.4 - x_{i-1}^2 + 0.3x_{i-2} \\ y_i &= 1.4 - (Cy_{i-1}x_{i-1} + (1-C)y_{i-1}^2) + 0.3y_{i-1}, \end{aligned} \tag{8}$$

where C is the coupling strength and in this experiment takes many values in the range $[0.01, 0.6]$. For each value of C, we construct two dimensional multivariate delay embeddings for time series of length 200 samples with delay 1 as (y_n, x_{n-1}). Four samples of the embedding point clouds are shown in Figure 3 as the coupling strength increases. The dimension of the point cloud for each value of coupling strength C is illustrated in Figure 4. Clearly, the dimension of the delay embedding decreases as the driving strength increases. These results demonstrate the capability of our method in determining the strength of causal interactions. Moreover, for comparison we have shown in Figure 4 the dimension of the point cloud (z_n, x_{n-1}) for an arbitrary independent nonlinear system given by $z_i = \sin(i) + 1.5\sin(z_{i-1})) + 0.6$. Since there is no interaction between $\{x_i\}_{i=1}^N$ and $\{z_i\}_{i=1}^N$, the dimension of (z_n, x_{n-1}) is higher than all dimensions obtained for Hénon system. We add white Gaussian observation noise with SNR $= 20dB$ to each time series and repeat the experiment. In the presence of noise, the dimension of (y_n, x_{n-1}) slightly increases but the results in detecting the causal interaction, and its strength are similar to the time series with no noise as shown in Figure 4.

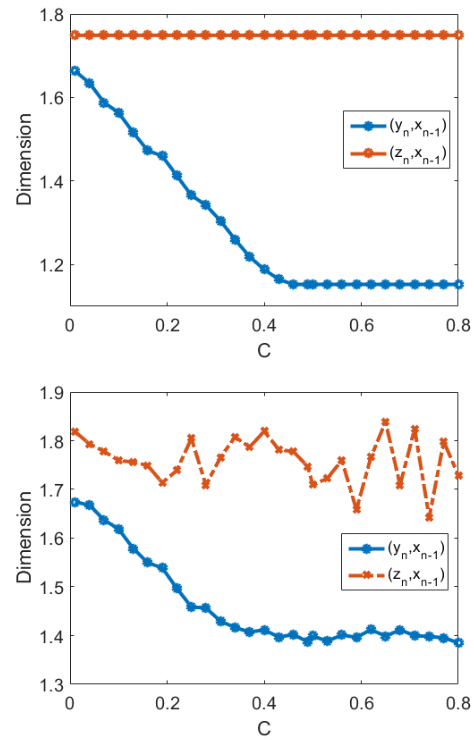

Fig. 4. Top: the dimension of multivariate delay embedding (y_n, x_{n-1}) for coupled Hénon map for different coupling strengths in $[0.01, 0.6]$, compared with that of (z_n, x_{n-1}), where $\{z_i\}_{i=1}^N$ and $\{x_i\}_{i=1}^N$ are independent. Bottom: same results in the presence of noise.

4. CONCLUSION

In this study, we proposed a geometric model-free causality measure based on fractal dimension of multivariate delay embedding. This measure demonstrates a capacity of efficiently detecting linear and nonlinear directed causal interactions between time series with no prior model of the system. Future work will focus on exploiting the proposed causality measure in electroencephalography(EEG) and magnetoencephalography (MEG) signal analysis to construct a comprehensive map of brain activity while processing different stimuli.

5. REFERENCES

[1] S. Haufe, R. Tomioka, G. Nolte, K.-R. Müller, and M. Kawanabe, "Modeling sparse connectivity between underlying brain sources for eeg/meg," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 8, pp. 1954–1963, 2010.

[2] Y. Zhao, S. A. Billings, H.-L. Wei, and P. G. Sarrigiannis, "A parametric method to measure time-varying linear and nonlinear causality with applications to eeg data," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3141–3148, 2013.

[3] B. He, L. Yang, C. Wilke, and H. Yuan, "Electrophysiological imaging of brain activity and connectivity-challenges and opportunities," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 7, pp. 1918–1931, 2011.

[4] U. Soytas and R. Sari, "Energy consumption and gdp: causality relationship in g-7 countries and emerging markets," *Energy economics*, vol. 25, no. 1, pp. 33–37, 2003.

[5] L.-Y. Lo, K.-S. Leung, and K.-H. Lee, "Inferring time-delayed causal gene network using time-series expression data," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, vol. 12, no. 5, pp. 1169–1182, 2015.

[6] C. W. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[7] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, and K.-R. Müller, "Robustly estimating the flow direction of information in complex physical systems," *Physical review letters*, vol. 100, no. 23, p. 234101, 2008.

[8] M. Kaminski and K. J. Blinowska, "A new method of the description of the information flow in the brain structures," *Biological cybernetics*, vol. 65, no. 3, pp. 203–210, 1991.

[9] L. A. Baccalá and K. Sameshima, "Partial directed coherence: a new concept in neural structure determination," *Biological cybernetics*, vol. 84, no. 6, pp. 463–474, 2001.

[10] Y. Zhao, S. A. Billings, H. Wei, F. He, and P. G. Sarrigiannis, "A new narx-based granger linear and nonlinear casual influence detection method with applications to eeg data," *Journal of neuroscience methods*, vol. 212, no. 1, pp. 79–86, 2013.

[11] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, vol. 898 of *Lecture Notes in Mathematics*, pp. 366–381, 1981.

[12] H. D. I. Abarbanel, R. Brown, J. J. Sidorowich, and L. Sh, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics*, vol. 65, pp. 1331–1392, Oct. 1993.

[13] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge University Press, 1997.

[14] S. Emrani, T. Gentimis, and H. Krim, "Persistent homology of delay embeddings and its application to wheeze detection," *IEEE Signal Processing Letters*, vol. 21, no. 4, pp. 459–463, 2014.

[15] S. Emrani and H. Krim, "Spectral estimation in highly transient data," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pp. 1721–1725, IEEE, 2015.

[16] P. R. Krishnaiah and L. N. Kanal, "Classification, pattern recognition, and reduction of dimensionality, volume 2 of handbook of statistics," *North4Holland Amsterdam*, 1982.

[17] B. Klinkenberg, "A review of methods used to determine the fractal dimension of linear features," *Mathematical Geology*, vol. 26, no. 1, pp. 23–46, 1994.

[18] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors," in *The Theory of Chaotic Attractors*, pp. 170–189, Springer, 2004.

[19] S. Emrani and H. Krim, "Effective connectivity-based neural decoding: A causal interaction-driven approach," *arXiv preprint arXiv:1607.07078*, 2016.

[20] S. Emrani, T. Gentimis, and H. Krim, "Persistent homology of delay embeddings and its application to wheeze detection," *Signal Processing Letters, IEEE*, vol. 21, pp. 459–463, April 2014.

[21] S. Emrani and H. Krim, "Spectral estimation in highly transient data," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, pp. 1721–1725, Aug 2015.