



# Matlab Open Source Code: Noise-Assisted Multivariate Empirical Mode Decomposition Based Causal Decomposition for Causality Inference of Bivariate Time Series

Yi Zhang<sup>1,2,3†</sup>, Guan Wang<sup>1†</sup>, Ziwen Li<sup>4†</sup>, Mingjun Xie<sup>1</sup>, Branko Celler<sup>5</sup>, Steven Su<sup>6</sup>, Peng Xu<sup>2,3\*</sup> and Dezhong Yao<sup>2,3</sup>

<sup>1</sup> School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu, China, <sup>2</sup> Key Laboratory for Neuro Information of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu, China, <sup>3</sup> Center for Information in BioMedicine, University of Electronic Science and Technology of China, Chengdu, China, <sup>4</sup> Glasgow College, University of Electronic Science and Technology of China, Chengdu, China, <sup>5</sup> Biomedical Systems Laboratory, University of New South Wales, Sydney, NSW, Australia, <sup>6</sup> Center for Health Technologies, Faculty of Engineering and Information Technology, University of Technology, Sydney, NSW, Australia

## OPEN ACCESS

### Edited by:

Stavros I. Dimitriadis,  
Greek Association of Alzheimer's  
Disease and Related Disorders,  
Greece

### Reviewed by:

Yuri Antonacci,  
University of Palermo, Italy  
Danilo Mandic,  
Imperial College London,  
United Kingdom

### \*Correspondence:

Peng Xu  
xupeng@uestc.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

**Received:** 10 January 2022

**Accepted:** 04 May 2022

**Published:** 16 June 2022

### Citation:

Zhang Y, Wang G, Li Z, Xie M,  
Celler B, Su S, Xu P and Yao D (2022)  
Matlab Open Source Code:  
Noise-Assisted Multivariate Empirical  
Mode Decomposition Based Causal  
Decomposition for Causality Inference  
of Bivariate Time Series.  
Front. Neuroinform. 16:851645.  
doi: 10.3389/fninf.2022.851645

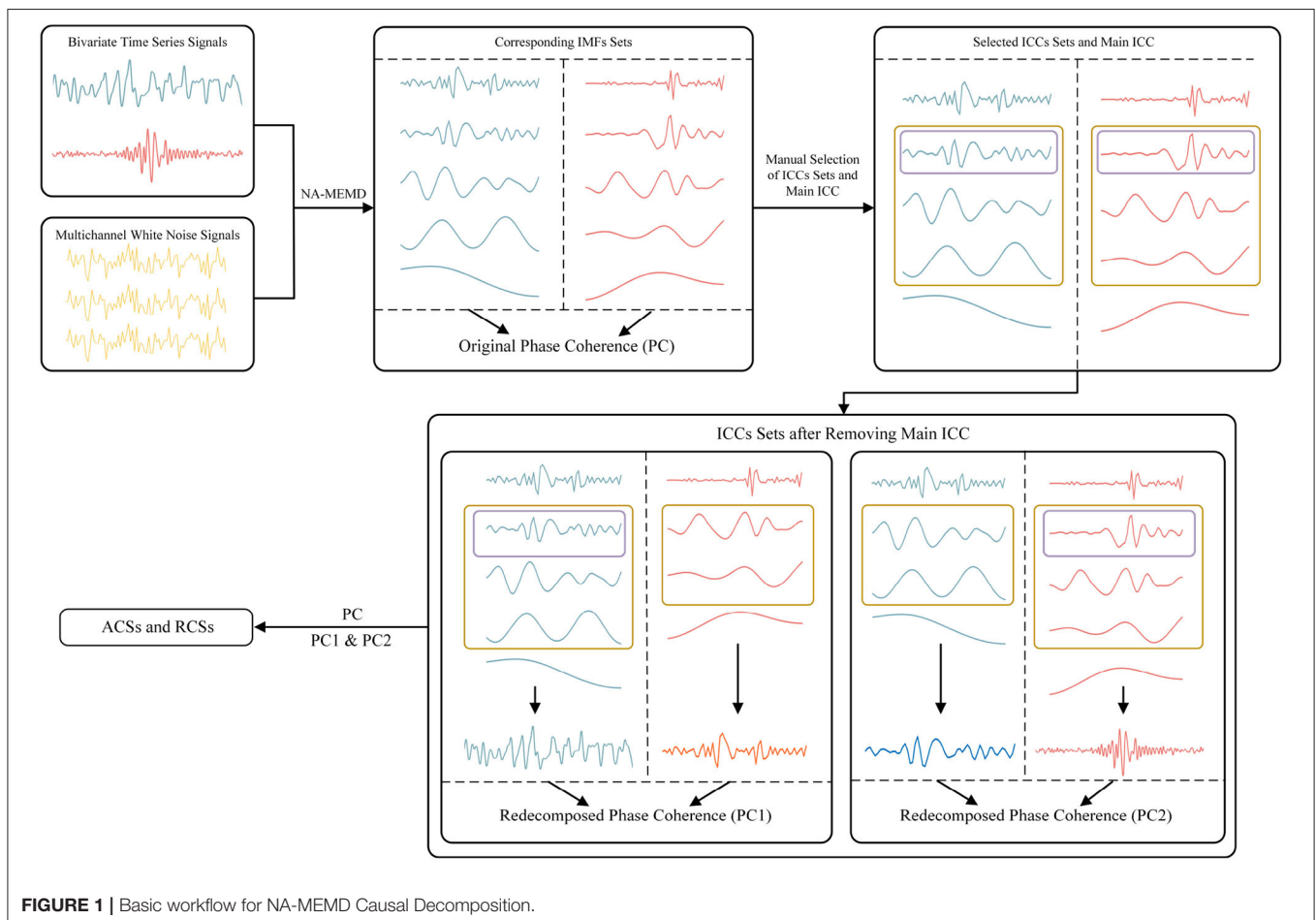
Causality inference has arrested much attention in academic studies. Currently, multiple methods such as Granger causality, Convergent Cross Mapping (CCM), and Noise-assisted Multivariate Empirical Mode Decomposition (NA-MEMD) are introduced to solve the problem. Motivated by the researchers who uploaded the open-source code for causality inference, we hereby present the Matlab code of NA-MEMD Causal Decomposition to help users implement the algorithm in multiple scenarios. The code is developed on Matlab2020 and is mainly divided into three subfunctions: *na\_memd*, *Plseries*, and *cd\_na\_memd*. *na\_memd* is called in the main function to generate the matrix of Intrinsic Mode Functions (IMFs) and *Plseries* can display the average frequency and phase difference of IMFs of the same order in a matrix which can be used for the selection of the main Intrinsic Causal Component (ICC) and ICCs set. *cd\_na\_memd* is called to perform causal redecomposition after removing the main ICC from the original time series and output the result of NA-MEMD Causal Decomposition. The performance of the code is evaluated from the perspective of executing time, robustness, and validity. With the data amount enlarging, the executing time increases linearly with it and the value of causal strength oscillates in an ideally small interval which represents the relatively high robustness of the code. The validity is verified based on the open-access predator-prey data (wolf-moose bivariate time series from Isle Royale National Park in Michigan, USA) and our result is aligned with that of Causal Decomposition.

**Keywords:** NA-MEMD Causal Decomposition, empirical mode decomposition, Granger causality, CCM, Causal Decomposition, causality inference, bivariate time series

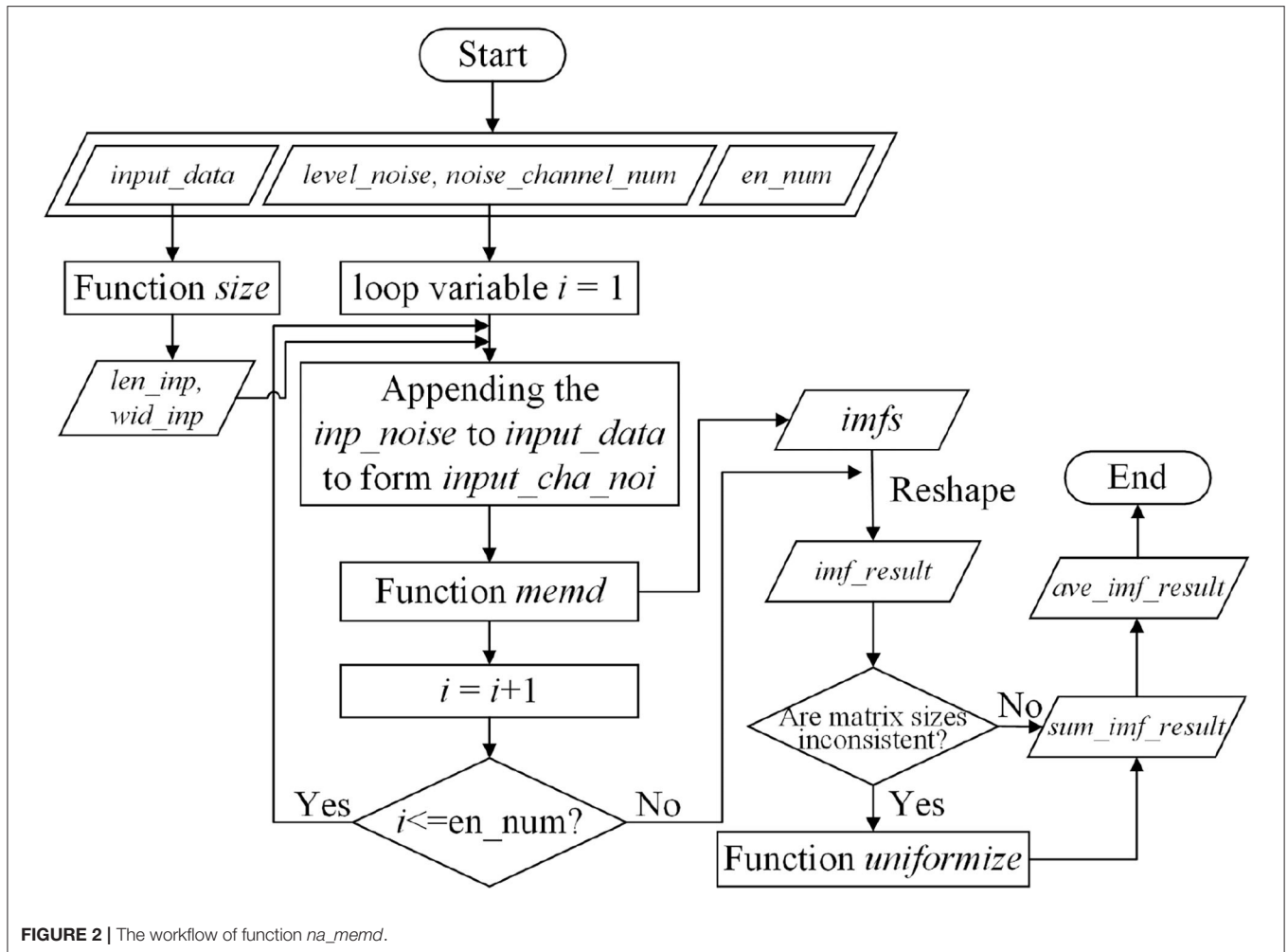
## 1. INTRODUCTION

In early studies, the definition of Noise-assisted Multivariate Empirical Mode Decomposition (NA-MEMD) based Causal Decomposition was given by Ur Rehman and Mandic (2011), She et al. (2017), and Zhang et al. (2017, 2021), dealing with the cause-effect relationship observed across from two time series signals in the real-world situations (Adarsh and Janga Reddy, 2019). The study provides Matlab Open Source Code of NA-MEMD Causal Decomposition and helps scholars and researchers (particularly in function connectivity Mueller et al., 2013; Meshi et al., 2016, signal detection and processing Abbate et al., 1997; Song and Que, 2006, and Statistical Causality Cox, 1992; Cox and Wermuth, 2004) to determine the causality occurring at stochastic, deterministic and complex dynamical (nonlinear deterministic) processes on the basis of time series (Small, 2008). The development of such a theoretical framework has been arising since the publication of Granger causality (Granger, 1969; Kamiński et al., 2001; Seth, 2007). It defines that variable  $X$  is considered to be the Granger cause of variable  $Y$  if  $X$  helps to explain future changes in  $Y$ . The test of Granger causality introduces F-test to quantify the autoregressive property between  $X$  and  $Y$  by solving a best

least-square problem. Granger causality was fundamentally based on the hypothesis of uncoupling cause-effect and, thus, would be applicable to stochastic processes. For multivariate time series, Looney et al. (2018) propose a novel multivariate sample entropy that can handle the analysis of within- and cross-channel dynamics. It is also the only method to identify synchronized regularity dynamics. In addition, inspired by Takens' Embedding Theorem (Noakes, 1991), Sugihara et al. proposed Convergent Cross Mapping (CCM) (Sugihara et al., 2012; Krakovská et al., 2015), which held that if the projections of  $X$  and  $Y$  in a certain dimension, i.e.,  $X'$  and  $Y'$  respectively, existed the causal relationship, then  $X$  and  $Y$  would have CCM causality. It was assumed that a cause-effect relation was embedded in a complex dynamical process which was also likely to be a linear/nonlinear deterministic system. CCM accommodated the inseparability/coupling of causal effects. Not surprisingly, Yang et al. (2018) established Causal Decomposition and further confirmed the utility of Hilbert-Huang Transform (HHT) (Huang, 2014a,b) in causality inference. Cause-effect mutual information was assumed to be carried over the instantaneous phase of the observed time series. However, standard methods such as Standard Fourier, wavelet, and Hilbert may encounter problems in analyzing real-world data. First, the approaches

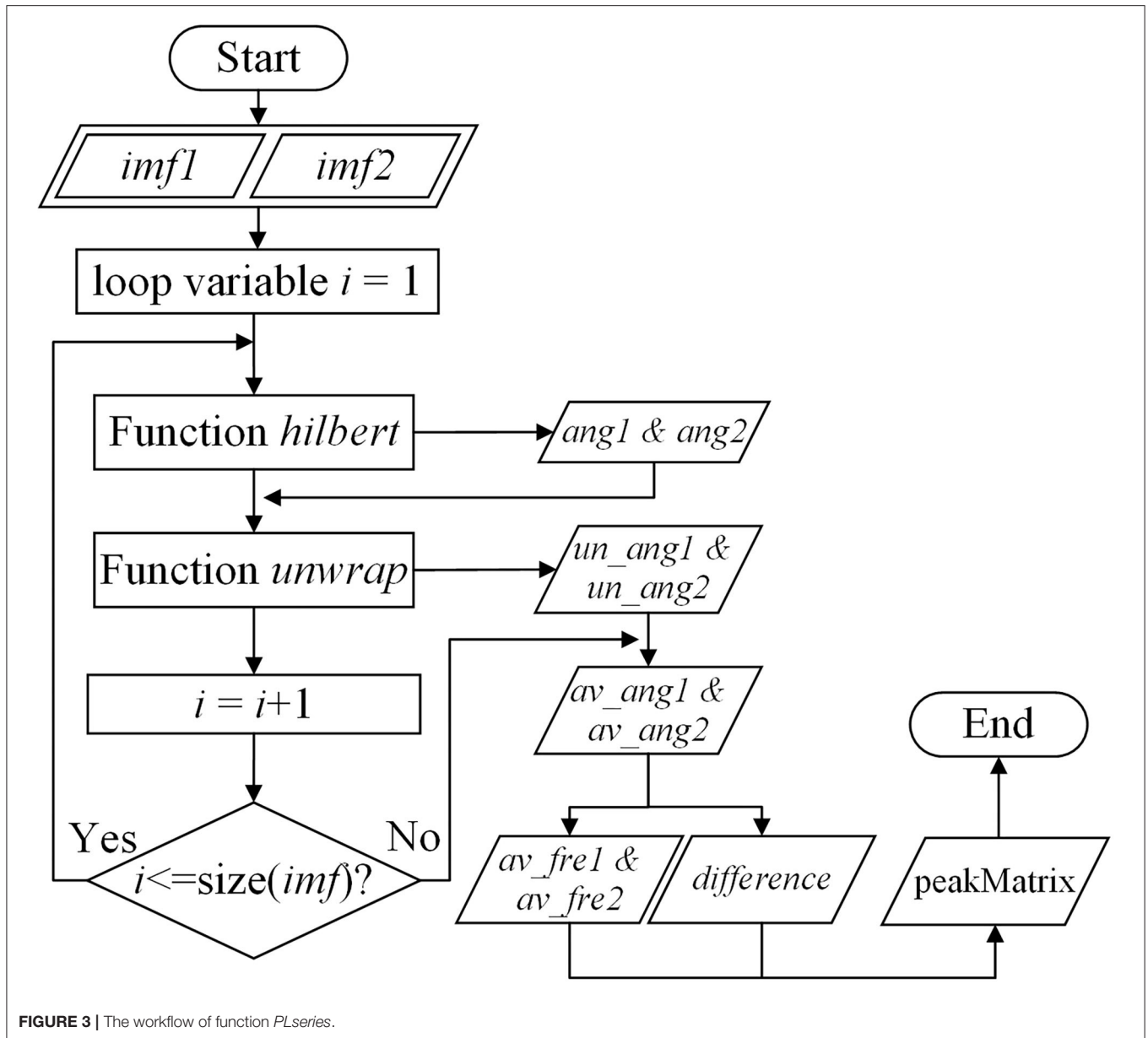


**FIGURE 1** | Basic workflow for NA-MEMD Causal Decomposition.



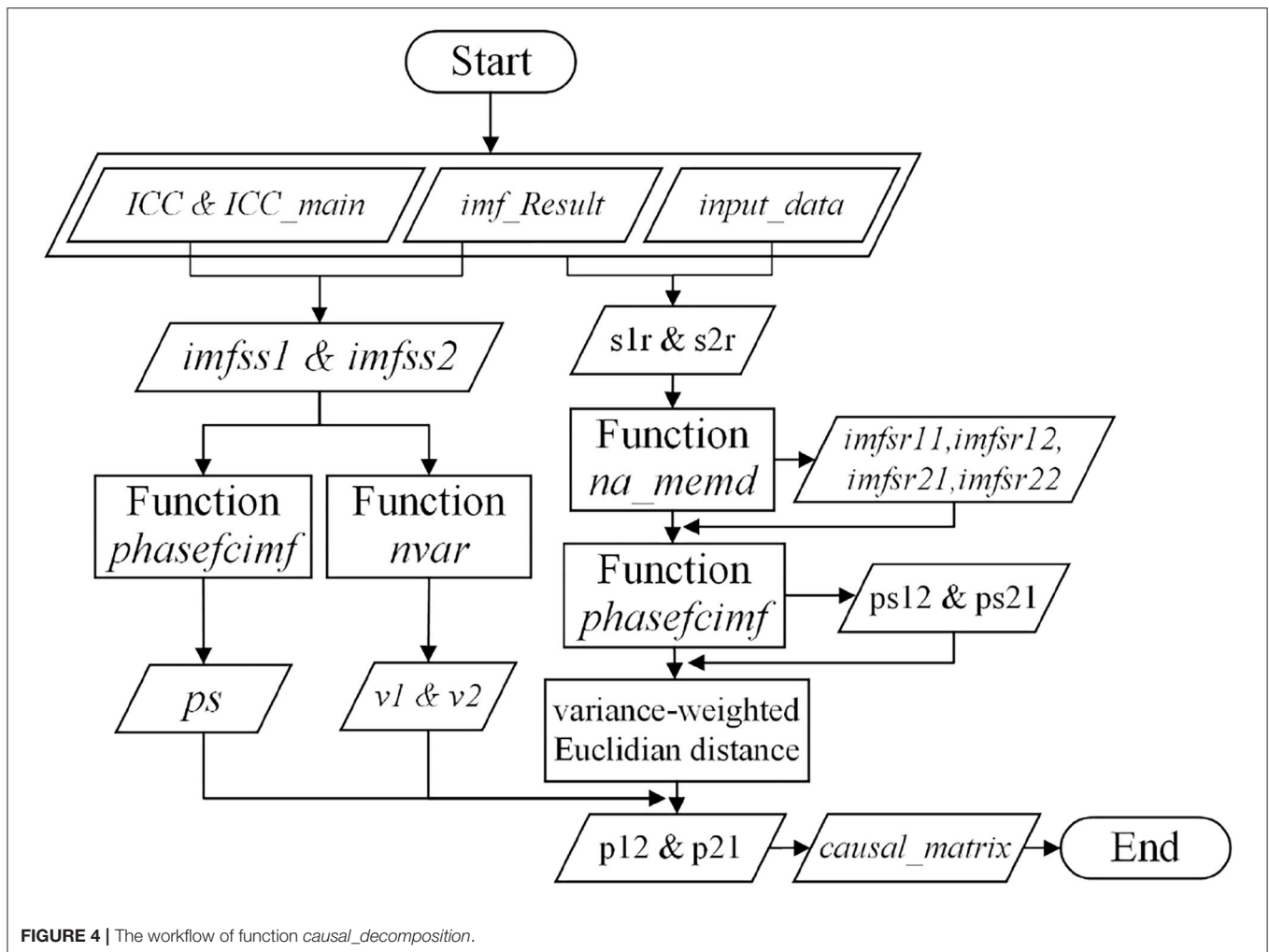
depending on predefined function heavily depend on the data length and stationarity and the real-world data which are often short and intermittent may hinder the analysis process. Second, among the standard methods, integral transforms pursue more frequency concepts than the temporal concept which may increase the frequency resolution but lose information in the time domain. Third, for nonmonocomponent data, the direct implementation of analytic signal representation results in negative IFs which have no physical justification. Next, accurate observation and measurement should be conducted on the corresponding scales rather than globally. However, correlation, coherence, and synchrony measure cannot be performed on certain scales (Mandic et al., 2013). Causal Decomposition was based on phase dependency, which was distinguished from the prediction paradigm such as Granger causality, CCM, Mutual Information from Mixed Embedding (Kugiumtzis, 2013; Jia et al., 2019), Dynamic Causal Modeling (Friston et al., 2003, 2013), and Transfer Entropy (Staniek and Lehnertz, 2008; Bossomaier et al., 2016). It was suitable for complex dynamical processes, acquired in the manner of time series in certain similar time scales. The development of the MEMD theoretical framework

has been started by Altaf et al. (2007) who put forward a method that develops a mathematical approach to adapt EMD to both real and complex domains. The so-generated Intrinsic Mode Functions (IMFs) can be used in processing both real and complex signals. Another method proposed by Tanaka and Mandic (2007) also provides an insight into extending standard EMD to the complex domain. It takes advantage of both positive and negative frequency components of signals to generate complex-valued IMFs. Since the EMD is initially limited to real-valued time series, an extension of the EMD framework to bivariate time series is designed by extracting zero-mean rotating components (Rilling et al., 2007). In order to make EMD compatible with trivariate signals, ur Rehman and Mandic (2009) comes out with a theoretical framework that projects local mean envelop to multiple directions in three-dimensional spaces adapting the rotation property of quaternions. In order to handle the causality analysis in multiscale time series, like multiple physiological time series (e.g., EEG, EMG, and ECG) composed network (Bashan et al., 2012; Faes et al., 2017), we recently presented NA-MEMD Causal Decomposition (Zhang et al., 2021) and pointed out its potential to the causality



inference in a complex dynamical process. In Zhang et al. (2017), multichannel EMG signals are processed by EEMD, MEMD, and NA-MEMD, and their outcomes are quantitatively assessed by comparing their number of IMFs, mode-alignment, and mode-mixing. It has been justified that NA-MEMD has a relatively outstanding performance in processing brain-muscle signals compared with EEMD and MEMD. Many studies have contributed to algorithm implementations related to causality analyzes in open-source scenarios. C and C++ code of EMD using Matlab Coder™ was introduced in R2018a in MathWorks (Huang, 2022). Furthermore, Rehman and Mandic generated the algorithm of Multivariate empirical mode decomposition

(MEMD) (Rehman and Mandic, 2010). Then, Zhang et al. (2017) published a Matlab toolbox (Wen et al., 2022) of NA-MEMD in 2017. In the causality analysis, Mønster provided a Convergent Cross Mapping algorithm in MATLAB in 2018 (Jakubik, 2022). Yang published the proposed Causal Decomposition approach in GitHub (Yang, 2022), which then was exchanged to MathWorks in 2020. In the study, we hereby present the Matlab code package for NA-MEMD Causal Decomposition used in the preliminary article by Zhang et al. (2021), offering the configuration specification details required in the data analyzes and tests, and providing the functional specification of line-by-line codes in the workflow.



## 2. SPECIFICATION OF CALLING FUNCTION

### 2.1. Background Theory

The overall procedure for NA-MEMD Causal Decomposition is illustrated in **Figure 1** by 5 steps. It first adds multichannel auxiliary white noise (shown in step **a-2**) to original bivariate time series signals (shown in step **a-1**), followed by NA-MEMD to obtain two corresponding IMFs sets (shown in steps **b-1** and **b-2**) as well as their phase coherence. Next, the main ICC (framed by the purple rectangles in steps **c** and **d**) and ICCs sets (framed by the yellow rectangles in steps **c** and **d**) are selected according to average frequency and phase difference. Only the IMFs with identical frequency scales and small phase differences can be regarded as ICCs. The main ICC is usually selected as the first ICC in the set. Removing the main ICC from IMFs sets and adding the remaining signals in each IMFs setup, two pairs of signals are formed by one modified signal without the main ICC and the other original one (shown in step **d**). The phase coherence of the two pairs is calculated. Combined with the value of the phase coherences of two redecomposed signal pairs and the original signal pair, the value of Absolute Causal Strengths (ACSs) and Relative Causal Strengths (RCSs)

can be obtained by NA-MEMD Causal Decomposition (shown in step **e**).

### 2.2. Function *na\_memd*

The procedure for function *na\_memd* is illustrated in **Figure 2**. The function enables to form a matrix, giving Intrinsic Mode Functions (IMFs) decomposed from NA-MEMD with causal-effect time series defined by input matrix *input\_data*. Relevant variables about function *na\_memd* are listed in **Table 1**. The function *na\_memd* is invoked when variables *input\_data*, *ave\_noise*, *level\_noise*, *noise\_channel\_num*, and *en\_num* are set. Then, function *size* returns the number of row elements of matrix *input\_data* to variable *len\_inp*, and the number of column elements of matrix *input\_data* to variable *wid\_inp*. *level\_noise* and *noise\_channel\_num* are used to generate the random Gaussian-White noise time series. After then, function *memd* is run if vector/matrix *inp\_noise* is appended to matrix *input\_data*, which is named matrix *input\_cha\_noi*. It repeats for constant *en\_num* times (*i* is the loop variable). For each repetition, a three-dimension array *imfs* is returned which contains all IMFs decomposed from matrix *input\_cha\_noi*. In

**TABLE 1** | Parameter configurations for function *na\_memd*.

Parameter names	Parameter specification	Configurations
<i>input_data</i>	Data type: <b>Matrix</b> IMFs decomposed from NA-MEMD	Two time series data on identical length.
<i>level_noise</i>	Data type: <b>Number</b> Used to determine the intensity of alternating component of noise in <i>na_memd</i>	With it increasing, the represented cause-effect relationship may be attenuated.
<i>noise_channel_num</i>	Data type: <b>Number</b> Used to determine the channel number of appended noise signals in <i>na_memd</i>	With it increasing, the represented cause-effect relationship may be attenuated.
<i>en_num</i>	Data type: <b>Number</b> Used to determine the cycle index of noise-assisted calculation in <i>na_memd</i>	With it increasing, the running time will be prolonged dramatically.

**TABLE 2** | Parameter configurations for function *PLseries*.

Parameter names	Parameter specification	Configurations
<i>imf_1</i> , <i>imf_2</i>	Data type: <b>Matrix</b> IMFs decomposed from NA-MEMD	Two time series data on identical length.
<i>ave_fre1</i> , <i>ave_fre2</i>	Data type: <b>Vectors</b> Represent the average instantaneous frequency of two IMF sets	Null
<i>difference</i>	Data type: <b>Vector</b> Represent the average phase difference of two IMF sets	Null
<i>peakMatrix</i>	Data type: <b>Matrix</b> A matrix to illustrate <i>av_freq1</i> , <i>av_freq2</i> and <i>difference</i> in different columns	The first two columns refer to <i>av_freq1</i> and <i>av_freq2</i> . The third column represents <i>difference</i> .

**TABLE 3** | Parameter configurations for function *causal\_decomposition*.

Parameter names	Parameter specification	Configurations
<i>ICC_main</i>	Data type: <b>Number</b> The index of selected main ICC	The main ICC is the one with the minimal phase difference.
<i>input</i>	Data type: <b>Matrix</b> The raw data to be analyzed	Null
<i>imf_result</i>	Data type: <b>Matrix</b> The result from <i>na_memd</i> in <i>na_memd_causal_decomposition</i>	Null
<i>level_noise</i>	Data type: <b>Number</b> Used to determine the intensity of alternating component of noise in <i>na_memd</i>	With it increasing, the represented cause-effect relationship may be attenuated.
<i>noise_channel_num</i>	Data type: <b>Number</b> Used to determine the channel number of appended noise signals in <i>na_memd</i>	With it increasing, the represented cause-effect relationship may be attenuated.
<i>en_num</i>	Data type: <b>Number</b> Used to determine the cycle index of noise-assisted calculation in <i>na_memd</i>	With it increasing, the running time will be prolonged dramatically.
<i>causal_matrix</i>	Data type: <b>Matrix</b> Used to output the relative causal strengths and the absolute strengths.	The first two columns represent the relative causal strengths. The latter two columns represent the absolute causal strengths.

order to facilitate the ensemble process of NA-MEMD, array *imfs* is reshaped into cell *imf\_result*. Considering repetitions in-between, if the matrix sizes in cell *imf\_result* perhaps are inconsistent, the residual is embedded with a pseudo-monotone variation. In order to guarantee the consistency of the size in cell *imf\_result*, the residuals can be ignored together. Matrix elements of *imf\_result*, *sum\_imf\_result*, and *imf\_result* are first

initialized to be zero, which stand for reshaped IMF data, the summation of the reshaped IMF, and the average value of the summation respectively. The noise data are generated by random alternating noise components determined by parameter *level\_noise*. This process is repeated by times of *en\_num*. For each loop, the noise is appended to matrix *input\_data* which is matrix *input\_cha\_noi*. Then matrix *imfs* is calculated by performing

**TABLE 4** | The relationship between *noise\_channel\_num* and executing time.

<i>noise_channel_num</i>	3	4	5	6	7	8	9	10
Executing time(s)	24.50	25.68	26.86	28.01	31.73	33.58	34.80	37.64

MEMD on *input\_cha\_noi* with function *memd*. Next, *imfs* is reshaped into *imf\_result* for further summation which data is stored in *sum\_imf\_result*. The average value is also calculated and assigned to *ave\_imf\_result*.

### 2.3. Function *PLseries*

The workflow for function *PLseries* is shown in **Figure 3**. The function outputs a matrix that shows the average frequency and their phase difference of IMFs decomposed from NA-MEMD by matrices *imf1* and *imf2* (**Table 2**), respectively. The row of *imf1* and *imf2* stand for the index of IMFs and the column of *imf1* and *imf2* refers to the length of the time series. The output matrix is defined as matrix *peakMatrix*. The function *PLseries* is called when matrix *imf1* and *imf2* are input. The elements in vectors *av\_fre1*, *av\_fre2*, and *difference* are initialized to be zero and their length of the row is also set to be consistent with that of *imf1* and *imf2* (their rows are in the same length). After that, for each IMF, Hilbert Transform is used to calculate its instantaneous phase. To alleviate the effect of phase winding, the function *unwarp* is called to obtain the real phase which is defined as vectors *un\_ang1* and *un\_ang2*. Then the average value of the unwrapped phase is calculated and stored in vectors *av\_ang1* and *av\_ang2*, respectively. Their difference is stored in vector *difference*. According to vectors *un\_ang1* and *un\_ang2*, the phase information is converted to frequency information which is stored in vectors *fre1* and *fre2*, respectively. Their average values are calculated and are stored in vectors *av\_fre1* and *av\_fre2*. The final output matrix *peakMatrix* is the combination of vectors *av\_fre1*, *av\_fre2*, and *difference* which are in the first, the second, and the third column, respectively.

### 2.4. Function *causal\_decomposition*

The workflow for function *causal\_decomposition* is demonstrated in **Figure 4**. The function generates a matrix *causal\_matrix* to provide values of RCSs and ACSs according to the times series inputted by matrix *input* referred to in **Table 3**. The raw data of bivariate time series are row-by-row loaded to input which is subsequently defined as vectors *s1* and *s2*. Those two decomposed IMF sets are then assigned as matrices *imfs1* and *imfs2*, respectively. Intrinsic Causal Components (ICCs) sets, vector *ICC*, are manually reviewed and selected by function *plot\_and\_pick*. Based on *ICC*, the number *ICC\_main* is confirmed to indicate RCSs and ACSs of raw bivariate time series of input. Main *ICC* is selected manually by function *plot\_and\_pick*, and is defined as vector *imfs1* and *imfs2*. Their phase coherence and variance are calculated by function *phasefcimf* and function *nvar*, respectively. After obtaining the main *ICC*, it is removed from the corresponding inputs *s1* and *s2*, and the rest are defined as vectors *s1r* and *s2r*. They are used to replace the corresponding signal in input to form two new input

**TABLE 5** | The relationship between *en\_num* and executing time.

<i>en_num</i>	5	10	15	20	25	30	35	40
Executing time(s)	41.26	78.94	117.09	155.09	194.79	217.20	245.20	288.27

**TABLE 6** | The relationship between data length and executing time.

Data length	250	500	750	1,000	1,250	1,500	1,750	2,000
Executing time(s)	16.11	22.46	31.49	37.65	44.14	50.87	62.34	70.41

matrices *input1* and *input2*. Then they are re-decomposed by calling function *na\_memd* to calculate two IMFs pairs (matrices *imfsr11*, *imfsr12* and matrices *imfsr21*, *imfsr22*, respectively). The phase coherence between paired IMFs sets and the distance between the phase coherences of the original IMFs pairs and the re-decomposed ones are calculated and stored in (*ps12*, *ps21*) and (*p12*, *p21*), respectively. The final causal strengths are outputted by the matrix *causal\_matrix*.

## 3. RESULT

### 3.1. Test for Time Series Data

To guarantee the experimental efficiency in the real application, variable *noise\_channel\_num* and *en\_num* and the data length should be modified. The input data is the Gaussian-White noise with a mean value of 0 and variance of 1, and the experiment is performed on Matlab 2020b on the fully powered laptop.

#### 3.1.1. Variable *noise\_channel\_num*

The relationship between *noise\_channel\_num* and executing time is demonstrated below in **Table 4**.

In this test, *en\_num* and data length are determined to be 3 and 500, respectively. It is clear that executing time has a linearly increasing trend with *noise\_channel\_num* from the recorded data in **Table 4**.

#### 3.1.2. Variable *en\_num*

The relationship between *en\_num* and executing time is demonstrated below in **Table 5**.

In this test, *noise\_channel\_num* and data length are fixed to be 3 and 500, respectively, and with *en\_num* increasing, executing time increases linearly along with *en\_num* according to **Table 5**.

#### 3.1.3. Data Length

The relationship between *data length* and executing time is demonstrated below in **Table 6**.

In this test, *noise\_channel\_num* and *en\_num* are both fixed to be 3, and with *data length* increasing, executing time increases linearly along with *data length* referred to in the data in **Table 6**.

### 3.2. Robustness and Validity Test

Since the random noise is involved in NA-MEMD, it is likely that the results of different execution will be slightly different from one another. Therefore, the consistency of the output should be tested. In this test, *level\_noise*, *noise\_channel\_num*, *en\_num*,

**TABLE 7** | Result of robustness and validity test.

Test number	RCSs of X-to-Y ( $C_{xy}$ )	RCSs Y-to-X ( $C_{yx}$ )	ACSs of X-to-Y ( $C_{xy}$ )	ACSs Y-to-X ( $C_{yx}$ )
1	0.7168	0.2832	0.2237	0.0884
2	0.5374	0.4626	0.1927	0.1659
3	0.6804	0.3196	0.2003	0.0941
4	0.7594	0.2306	0.1810	0.0542
5	0.6364	0.3636	0.2048	0.1170
6	0.6042	0.3958	0.1971	0.1291
7	0.6398	0.3602	0.2038	0.1148
8	0.6585	0.3415	0.1974	0.1024
9	0.5943	0.4057	0.1923	0.1313
10	0.5473	0.4527	0.1979	0.1637

and **data length** are set to be 0.001, 3, 5, and 61, respectively, and the open-access predator-prey data is provided by Vucetich and Peterson (2012). The result is shown in **Table 7**.

In **Table 7**, since the X-to-Y causality value and Y-to-X causality value sum to one, only the X-to-Y causal value is considered in the stability discussion.

From the statistical point of view, the mean value of  $C_{xy}$  is 0.6384 and the variance is 0.004712 which is less volatile. Combined with the actual causal process demonstrated by the causality values by Yang et al. (2018), all of them indicate strong and valid causality since our results are close to that of Yang et al. (2018), which demonstrates the outcome of the procedure is consistent with the algorithm.

## 4. DISCUSSION

### 4.1. The Approach to Exert Noise in Original Signals in *na\_memd*

Generally, input parameters *level\_noise*, *en\_num*, and *noise\_channel\_num* are used to determine the total amount of noise appended to the original signal. The description of the relevant variables can be checked in **Table 3**. Before setting the parameters, the noise level in the original input data should be evaluated. If the input has already been overwhelmed by noise, both the two parameters are supposed to be set higher to better remove the noise in the original data. However, with the Gaussian-White noise increasing, the effective data in the input is also likely to be attenuated by the appended noise. As a result, it is of significance to choose a proper value of *level\_noise* and *noise\_channel\_num* to avoid excessive Gaussian-White noise. One method of removing the original noise without doing damage to the effective data is to reduce the product of *level\_noise* and *noise\_channel\_num* and to increase *en\_num*.

### 4.2. The Method of Selecting ICCs Set From IMFs

For ICCs, the ICCs set and the main ICC need to be selected for further operation. Among all the IMFs, ICCs are the ones whose average frequencies and phases are on the same scale. IMFs in different frequency scales or with large phase differences should be excluded from the ICCs set. In ICCs set, the main ICC is the one with the smallest average phase difference. Besides,

the phase diagram of IMFs generated by Matlab can also be used to discriminate ICCs from all IMFs. Assist from both data in *peakMatrix* and visual aid of signal diagram is necessary for ICCs discrimination.

### 4.3. Calculation of RCSs and ACSs

In this code, it is recommended that only RCSs are used to detect the causal-effect relationship between bivariate time series signals.

Absolute Causal Strengths is defined as the variance weighted Euclidean distance (De Leeuw and Pruzansky, 1978) between the re-decomposed IMFs set and the original IMFs set. Its value can reflect the relative cause-effect relationship between multiple signals. The signals with higher ACS can be regarded as the cause and the ones with smaller ACS can be considered as the effect, respectively. Specifically, if only two signals are considered, RCS can also be calculated to illustrate their cause-effect relationship. It normalizes the ACS into an interval between zero and one, and we use the value of RCS to interfere with the causal-effect relationship. When the value of RCS is larger than 0.5, it represents the relationship of cause, and when the value of RCS is smaller than 0.5, it represents the relationship of effect. When the value is equal to 0.5, the relationship can be either reciprocal causation or an irrelevant relationship. When the ACS of two candidate signals is negligible, RCS will converge to 0.5. The threshold value is normally set as 0.05 (Yang et al., 2018).

### 4.4. Further Analysis of the Result

#### 4.4.1. The Reason for Linear Increase in Executing Time

In this code, the ensemble process is designed as simple loops for multiple noise appending rather than nested loops or complex function iteration. According to ur Rehman et al. (2013), compared with EEMD, NA-MEMD applies an ensemble algorithm that massively reduces the reconstruction error. As a result, the impact of the number of noise channels on creating error can be negligible. However, ur Rehman et al. (2013) recommend the number to be 2 which is suitable for the test situation and can be adjusted accordingly. Therefore, this experiment can change the *noise\_channel\_num* and observe differences. As a result, the executing time is in direct proportion to the total data size. When *noise\_channel\_num*, *en\_num*, and *data length* increase, respectively, the total data size inflates



linearly. Therefore, the executing time would have a tendency of linear increase with those variables.

#### 4.4.2. The Determination for System Robustness

For any existing causality pairs, their cause-effect relationship is stable. RCS between them is considered a constant. As a result, it is appropriate to use the variance of relative causality data to represent the robustness of the system. If the variance has a small value which means RSC is less likely to fluctuate between the ideal causality strength, the system has relatively high robustness. If the variance is measured high, then it proves that the system is more susceptible to the external or internal environment.

## 5. CONCLUSION

A code package of NA-MEMD has been proposed to facilitate the potential users to apply the algorithm in an effective way. The users need to select the ICCs set and main ICC to check the causal-effect relationship according to the printed frequency-phase matrix and the signal figures. Based on MEMD, the crucial step is to append multi-channel random Gaussian-White noise to the original signal repeatedly to attenuate self-carrying noise. By adjusting the number of noise channels, the intensity of appended noise per channel, and the cycle index, it has been shown that the users can alleviate the influence of original noise without damaging the original data. As a result, NA-MEMD Causal Decomposition has potential for applications in analyzing bivariate and multiscale time series signals.

## 6. DATA AVAILABILITY STATEMENT

The code of NA-MEMD Causal Decomposition for causality inference of bivariate time series will be available through open-source platform github, [https://github.com/AaronLi43/ginkgo\\_glasgow](https://github.com/AaronLi43/ginkgo_glasgow). The demo data used in section RESULT is Wolf and moose field data which are available online at the United States Isle Royale National, <https://isleroyalewolf.org/data/data/home.html>. Further inquiries can be directed to the corresponding author.

## REFERENCES

- Abbate, A., Koay, J., Frankel, J., Schroeder, S. C., and Das, P. (1997). Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection. *IEEE Trans. Ultrasonics Ferroelectrics Frequency Control* 44, 14–26. doi: 10.1109/58.585186
- Adarsh, S., and Janga Reddy, M. (2019). Evaluation of trends and predictability of short-term droughts in three meteorological subdivisions of india using multivariate emd-based hybrid modelling. *Hydrol. Process.* 33, 130–143. doi: 10.1002/hyp.13316
- Altaf, M. U. B., Gautama, T., Tanaka, T., and Mandic, D. P. (2007). "Rotation invariant complex empirical mode decomposition," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07, Vol. 3* (Honolulu, HI: IEEE), III-1009.
- Bashan, A., Bartsch, R. P., Kantelhardt, J. W., Havlin, S., and Ivanov, P. C. (2012). Network physiology reveals relations between network topology and physiological function. *Nat. Commun.* 3, 1–9. doi: 10.1038/ncomm1705

## AUTHOR CONTRIBUTIONS

YZ, GW, and ZL found the potential utility of the theoretical framework of NA-MEMD Causal Decomposition in function connectivity, signal detection and processing, Statistical Causality, implemented Matlab open-source code for NA-MEMD Causal Decomposition, tested the efficiency, robustness, validity of the code, carried out the results and discussions, and drafted and revised the manuscript. MX participated in Section Result in part. SS verified and proved the theoretical part of NA-MEMD Causal Decomposition. DY and PX supervised the project and checked the paper quality. PX advised the submission procedure and the manuscript quality linked toward its extensions to fMRI studies. All authors contributed to the article and approved the submitted version.

## FUNDING

This study was supported by the National Natural Science Foundation of China (#61961160705, #U19A2082, #62103085, and #61801094), Sichuan Science and Technology Program (#2021YFG0126), Major Science and Technology Special Projects in Sichuan (#2020YFG0469), and Fundamental Research Funds for the Central Universities China (#ZYGX2019J086).

## ACKNOWLEDGMENTS

The authors are thankful for the support from the School of Aeronautics and Astronautics [University of Electronic Science and Technology of China (UESTC)], Key Laboratory for NeuroInformation of Ministry of Education, School of Life Science and Technology (UESTC), Center for Information in BioMedicine (UESTC), Glasgow College (UESTC), Biomedical Systems Laboratory (University of New South Wales, Sydney, Australia), and Center for Health Technologies (University of Technology Sydney, Australia).

- Bossomaier, T., Barnett, L., Harr9, M., and Lizier, J. T. (2016). "Transfer entropy," in *An Introduction to Transfer Entropy* (Cham: Springer), 65–95.
- Cox, D. R. (1992). Causality: some statistical aspects. *J. R. Stat. Soc. A* 155, 291–301. doi: 10.2307/2982962
- Cox, D. R., and Wermuth, N. (2004). Causality: a statistical view. *Int. Stat. Rev.* 72, 285–305. doi: 10.1111/j.1751-5823.2004.tb00237.x
- De Leeuw, J., and Pruzansky, S. (1978). A new computational method to fit the weighted euclidean distance model. *Psychometrika* 43, 479–490. doi: 10.1007/BF02293809
- Faes, L., Nollo, G., Stramaglia, S., and Marinazzo, D. (2017). Multiscale granger causality. *Phys. Rev. E* 96, 042150. doi: 10.1103/PhysRevE.96.042150
- Friston, K., Moran, R., and Seth, A. K. (2013). Analysing connectivity with granger causality and dynamic causal modelling. *Curr. Opin. Neurobiol.* 23, 172–178. doi: 10.1016/j.conb.2012.11.010
- Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage*. 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7
- Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424–438. doi: 10.2307/1912791

- Huang, N.orden E., and Z. S. (2022). *Empirical Mode Decomposition*. Available online at: [https://www.mathworks.com/help/signal/ref/emd.html?s\\_tid=srchtitle\\_EMD\\_1](https://www.mathworks.com/help/signal/ref/emd.html?s_tid=srchtitle_EMD_1) (accessed November 1, 2021).
- Huang, N. E. (2014a). Hilbert-Huang transform and its applications, *World Scientific* 16. doi: 10.1142/8804
- Huang, N. E. (2014b). "Introduction to the hilbert-huang transform and its related mathematical problems," in *Hilbert-Huang Transform and Its Applications* (Tok Tuck Link: World Scientific), 1–26.
- Jakubik, J. (2022). *Convergent Cross Mapping*. Available online at: <https://www.mathworks.com/matlabcentral/fileexchange/52964-convergent-cross-mapping> (accessed November 1, 2021).
- Jia, Z., Lin, Y., Jiao, Z., Ma, Y., and Wang, J. (2019). Detecting causality in multivariate time series via non-uniform embedding. *Entropy* 21, 1233. doi: 10.3390/e21121233
- Kamiński, M., Ding, M., Truccolo, W. A., and Bressler, S. L. (2001). Evaluating causal relations in neural systems: granger causality, directed transfer function and statistical assessment of significance. *Biol. Cybern.* 85, 145–157. doi: 10.1007/s004220000235
- Krakovská, A., Jakubík, J., Budáčová, H., and Holeciová, M. (2015). Causality studied in reconstructed state space. examples of unidirectionally connected chaotic systems. *arXiv preprint arXiv:1511.00505*. doi: 10.48550/arXiv.1511.00505
- Kugiumtzis, D. (2013). Direct-coupling information measure from nonuniform embedding. *Phys. Rev. E* 87, 062918. doi: 10.1103/PhysRevE.87.062918
- Looney, D., Adjei, T., and Mandic, D. P. (2018). A novel multivariate sample entropy algorithm for modeling time series synchronization. *Entropy* 20, 82. doi: 10.3390/e20020082
- Mandic, D. P., ur Rehman, N., Wu, Z., and Huang, N. E. (2013). Empirical mode decomposition-based time-frequency analysis of multivariate signals: the power of adaptive data analysis. *IEEE Signal Process Mag.* 30, 74–86. doi: 10.1109/MSP.2013.2267931
- Meshi, D., Mamerow, L., Kirilina, E., Morawetz, C., Margulies, D. S., and Heekeren, H. R. (2016). Sharing self-related information is associated with intrinsic functional connectivity of cortical midline brain regions. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep22491
- Mueller, S., Wang, D., Fox, M. D., Yeo, B. T., Sepulcre, J., Sabuncu, M. R., et al. (2013). Individual variability in functional connectivity architecture of the human brain. *Neuron* 77, 586–595. doi: 10.1016/j.neuron.2012.12.028
- Noakes, L. (1991). The takens embedding theorem. *Int. J. Bifurcat. Chaos* 1, 867–872. doi: 10.1142/S0218127491000634
- Rehman, N., and Mandic, D. P. (2010). Multivariate empirical mode decomposition. *Proc. R. Soc. A* 466, 1291–1302. doi: 10.1098/rspa.2009.0502
- Rilling, G., Flandrin, P., Gonçalves, P., and Lilly, J. M. (2007). Bivariate empirical mode decomposition. *IEEE Signal Process. Lett.* 14, 936–939. doi: 10.1109/LSP.2007.904710
- Seth, A. (2007). Granger causality. *Scholarpedia* 2, 1667. doi: 10.4249/scholarpedia.1667
- She, Q.-S., Ma, Y.-L., Meng, M., Xi, X.-G., and Luo, Z.-Z. (2017). Noise-assisted memd based relevant imfs identification and eeg classification. *J. Central South Univer.* 24, 599–608. doi: 10.1007/s11771-017-3461-5
- Small, M. (2008). *The Taxonomy of Time Series*. Jönköping: New Nonlinear Phenomena Research.
- Song, S.-P. and Que, P.-W. (2006). Wavelet based noise suppression technique and its application to ultrasonic flaw detection. *Ultrasonics* 44, 188–193. doi: 10.1016/j.ultras.2005.10.004
- Staniek, M., and Lehnertz, K. (2008). Symbolic transfer entropy. *Phys. Rev. Lett.* 100, 158101. doi: 10.1103/PhysRevLett.100.158101
- Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., et al. (2012). Detecting causality in complex ecosystems. *Science* 338, 496–500. doi: 10.1126/science.1227079
- Tanaka, T., and Mandic, D. P. (2007). Complex empirical mode decomposition. *IEEE Signal Process. Lett.* 14, 101–104. doi: 10.1109/LSP.2006.882107
- ur Rehman, N., and Mandic, D. P. (2009). Empirical mode decomposition for trivariate signals. *IEEE Trans. Signal Process.* 58, 1059–1068. doi: 10.1109/TSP.2009.2033730
- Ur Rehman, N., and Mandic, D. P. (2011). Filter bank property of multivariate empirical mode decomposition. *IEEE Trans. Signal Process.* 59, 2421–2426. doi: 10.1109/TSP.2011.2106779
- ur Rehman, N., Park, C., Huang, N. E., and Mandic, D. P. (2013). Emd via memd: multivariate noise-aided computation of standard emd. *Adv. Adapt. Data Anal.* 5, 1350007. doi: 10.1142/S1793536913500076
- Vucetich, J. A., and Peterson, R. (2012). *The Population Biology of Isle Royale Wolves and Moose: An Overview*. Available online at: <https://isleroyalewolf.org/data/data/home.html> (accessed November 20, 2021).
- Wen, Y., Zhang, Y., Xu, P., and Yao, D. (2022). *Noise-Assisted Multivariate Empirical Mode Decomposition (na-memd)*. Available online at: <https://neuro.uestc.edu.cn/content/324>.
- Yang, A. (2022). *Causal-Decomposition-Analysis*. Available online at: <https://github.com/acccyang/causal-decomposition-analysis> (accessed November 1, 2021).
- Yang, A. C., Peng, C.-K., and Huang, N. E. (2018). Causal decomposition in the mutual causation system. *Nat. Commun.* 9, 1–10. doi: 10.1038/s41467-018-05845-7
- Zhang, Y., Xu, P., Li, P., Duan, K., Wen, Y., Yang, Q., et al. (2017). Noise-assisted multivariate empirical mode decomposition for multichannel emg signals. *Biomed. Eng. Online* 16, 1–17. doi: 10.1186/s12938-017-0397-9
- Zhang, Y., Yang, Q., Zhang, L., Ran, Y., Wang, G., Celler, B., et al. (2021). Noise-assisted multivariate empirical mode decomposition based causal decomposition for brain-physiological network in bivariate and multiscale time series. *J. Neural Eng.* 18, 046018. doi: 10.1088/1741-2552/abecf2

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Wang, Li, Xie, Celler, Su, Xu and Yao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.