Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

A Novel Propagation Path Identification Framework for Faults in Industrial Processes ⋆

Liang Ma*,, Kaixiang Peng**, Jie Dong****

** Shunde Graduate School of University of Science and Technology Beijing, Foshan, 528399, China*
*** Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, 100083, China*

(e-mail: liangma@ustb.edu.cn)

Abstract: In modern industry, timely and accurate fault diagnosis plays an important role in satisfying the demands of production safety and stability of production quality. This paper dedicates on propagation path identification of faults in industrial processes, which will offer a feasible technology or solution to take corrective and timely maintenance measures for field engineers. Specifically, a recurrent neural networks-based Granger causality analysis approach is developed, which has sufficiently considered the nonlinear and dynamic relationships among time series after faults happen. Finally, we validate our approach on a typical industrial process, finishing mill process, to demonstrate the efficiency of the proposed scheme.

Keywords: Propagation path identification, root cause diagnosis, Granger causality analysis, recurrent neural networks, industrial process

1. INTRODUCTION

Modern industrial processes, like steel-making and chemical, are progressing towards large-scale, continuous and automation. Since there are too many coupled control loops and processes in the whole production line, an abnormal operation may cause widespread alarms because of coupling relationships and propagation characteristics of faults, which seriously affects the production performance and final product quality. Associated with these trends, process monitoring and fault diagnosis play an important role in satisfying the demands of production safety and stability of production quality.

As the core technologies in process monitoring and fault diagnosis, faut detection, diagnosis and classification have been paid extensive attention in academic research and industrial application areas (Yin et al., 2014; Chen et al., 2019; Ge et al., 2018; Ma et al., 2020a). However, in contrast with the achievements in above areas, limited attentions have been focus on propagation path identification approaches and their applications. Although some data-based causality analysis means, like cross-correlation function (CCF) (Bauer et al., 2008) and Granger causality (GC) analysis (Yuan et al., 2014; Landman et al., 2014), have obtained satisfactory results, most of them assume linear dynamics. As well known, time series may turn into non-normally distributed or non-stationary after faults happen (Ma et al., 2020b), which make those approaches be impossible to achieve satisfactory causality analysis results. In such cases, several methods, such as dynamic time warping (DTW) (Li et al., 2016) and transfer entropy (TE) (Schreiber, 2000), have been developed. However, issues on higher computational complexity still have for industrial processes, which may limit further promotions and applications. More importantly, most of current methods are focused on fault diagnosis, while propagation path identification problems have not been in-depth studies (Landman et al., 2019; Ma et al., 2018; Ahmed et al., 2017), especially for faults in industrial processes.

Motivated by above observations, in this paper, the propagation path identification issue on faults is investigated from a new perspective, which combines recurrent neural networks (RNNs) with GC analysis method. The proposed scheme has fully considered the nonlinear and dynamic relationships among time series after faults happen, which will provide a feasible technology or solution to take corrective and timely maintenance measures for field engineers.

The remainder of this work is organized as follows. In Section 2, the basic ideas of GC analysis and problem formulation are presented. Then, Section 3 is focused on the developed method for faults in industrial processes. Next, a case study on finishing mill process (FMP) is given in Section 4. Finally, conclusions are made in Section 5.

⋆ This work was supported by the Natural Science Foundation of China (NSFC) under Grants (61873024, 61773053), the Project funded by China Postdoctoral Science Foundation (2019M660464), the Guangdong Basic and Applied Basic Research Foundation (2019A1515110991), the Fundamental Research Funds for the China Central Universities of USTB (FRF-TP-19-049A1Z), PR China and the National Key R&D Program of China (No.2017YFB0306403).

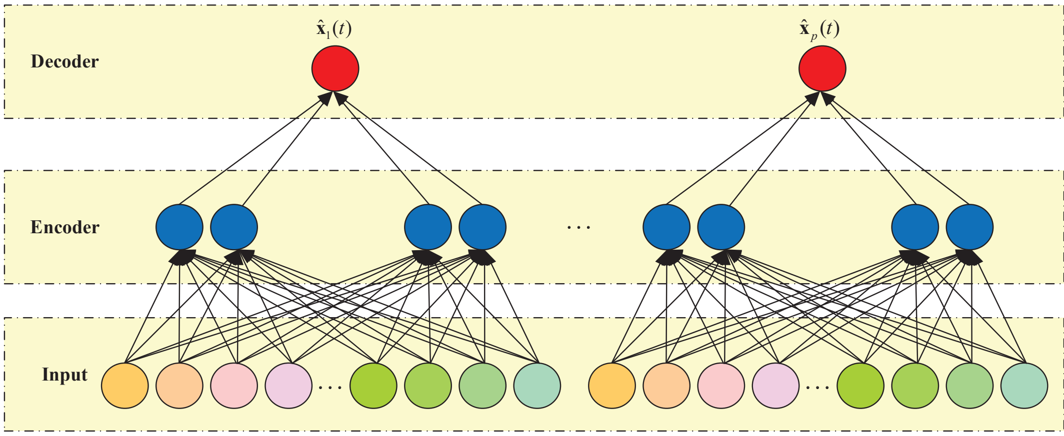

Fig. 1. Illustration of modeling GCs using MLPs

2. PROBLEM FORMULATION

In this section, the basic ideas of GC analysis method are described in brief. Then, the problem formulation is given.

Let $\mathbf{x}(t) \in \mathbf{R}^p$ for $t = 1, \cdots, T$ denotes a p-dimensional time series, then time series analysis based on GC can be defined as:

$$\mathbf{x}(t) = \sum_{r=1}^{R} \mathbf{A}^{(r)}\mathbf{x}(t-r) + \mathbf{e}(t) \tag{1}$$

where $\mathbf{A}^{(r)} \in \mathbf{R}^{l\times l}$ represents linear dependency relationships of $\mathbf{x}(t)$ to $\mathbf{x}(t-r)$ up to lag R, $\mathbf{e}(t)$ denotes white noise (Bressler et al., 2010).

In above model, if $\forall r$, $A_{ij}^{(r)} = 0$, then j does not have causality with i. Therefore, the above time series analysis can be determined by:

$$\min_{\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(R)}} \sum_{t=1}^{T} \left(\mathbf{x}(t) - \sum_{r=1}^{R} \mathbf{A}^{(r)}\mathbf{x}(t-r)\right)^2 + \alpha \sum_{ij} \parallel A_{ij}^{(1)}, \cdots, A_{ij}^{(R)} \parallel_2 \tag{2}$$

where $\parallel * \parallel_2$ is L_2 norm used for shrinking all values to 0, $\alpha > 0$ is a tuning parameter.

In industrial processes, once faults happen, the relevant time series may change into non-normally distributed and non-stationary, and the causalities may become nonlinear, which will influence causality analysis results. To address above issues, in this paper, a new causality analysis method is proposed, which benefits from classical RNNs and GC analysis methods, which has fully considered the nonlinear and dynamic relationships among time series after faults happen.

3. THE PROPOSED PROPAGATION PATH IDENTIFICATION METHOD

Due to the RNNs can compress the past of a time series into a hidden state, it is very suitable for modeling nonlinear time series. Inspired by the works of Tank et al (Tank et al., 2018a,b), based on above linear model, a nonlinear model can be constructed by:

$$\begin{aligned}\mathbf{x}(t) =& \widehat{\mathbf{x}}(t) + \mathbf{e}_t = \mathbf{g}\big(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-R)\big) + \mathbf{e}_t \\ =& \begin{bmatrix} g_1(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-R)) \\ \vdots \\ g_p(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-R)) \end{bmatrix} + \mathbf{e}_t\end{aligned} \tag{3}$$

where $g_i(\bullet)$ denotes how the past R lags influence i.

Then, the p time series can be modeled by p distinct multilayer perceptron (MLP) separately, as presented in Fig. 1. The value of GCs can be calculated by the magnitudes of encoder weights.

Assume that b denotes the number of hidden units in the 1st encoder layer, $\mathbf{A}^h \in \mathbf{R}^{pR\times b}$ denotes the encoder weights of the hth MLP module. It can be seen that if $A_{prb'}^h = 0$, then $\mathbf{x}_p(t-r)$ does not affect b'. Therefore, the hth MLP can be transformed into:

$$\min_{\mathbf{A}_{MLP}^h} \mathcal{L} = \min_{\mathbf{A}_{MLP}^h} \sum_{t=1}^{T} \big(\mathbf{x}_h(t) - g_h(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-R))\big)^2 + \alpha \sum_{b'=1}^{b}\sum_{r=1}^{R} \parallel A_{1rb'}^h, \cdots, A_{prb'}^h \parallel_2 \tag{4}$$

where $\mathbf{A}_{MLP}^h$ are weights of decoders and encoders for the hth MLP module.

Suppose that $\mathbf{y}_{t-1} \in \mathbf{R}^b$ is the bth hidden state at t, then the hidden state at $t-1$ can be updated by:

$$\mathbf{y}_t = f(\mathbf{x}_t, \mathbf{y}_{t-1}) \tag{5}$$

where $f(\star)$ is a nonlinear recurrent function.

The standard RNNs model takes the form:

$$f_t = \sigma(A^f\mathbf{x}_t + B^f\mathbf{y}_{(t-1)}) \tag{6}$$

$$i_t = \sigma(A^{in}\mathbf{x}_t + B^{in}\mathbf{y}_{(t-1)}) \tag{7}$$

$$o_t = \sigma(A^o\mathbf{x}_t + B^o\mathbf{y}_{(t-1)i}) \tag{8}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \sigma(A^c\mathbf{x}_t + B^c\mathbf{y}_{(t-1)}) \tag{9}$$

$$\mathbf{y}_t = o_t \odot \sigma(c_t) \tag{10}$$

where σ is sigmoid function, $\odot$ represents componentwise multiplication, i_t, f_t, o_t and c_t denote input, forget, output gates and state cell, respectively, of which c_t can be transformed into the hidden state used for prediction $\mathbf{y}_t$.

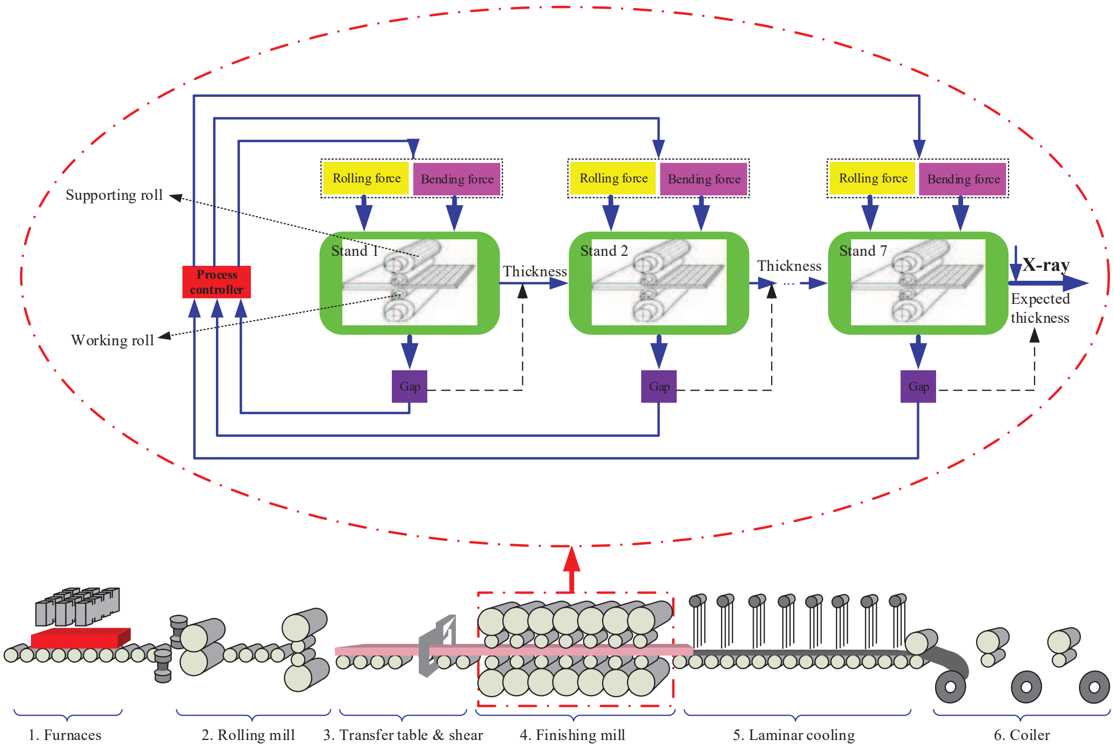

Fig. 2. Schematic layout of the HSMP

Thus, the set of input matrices can be obtained by:

$$\mathbf{A} = [(A^f)^T, (A^{in})^T, (A^o)^T, (A^c)^T]^T \tag{11}$$

which reflects how the past time series $\mathbf{x}_t$ affects the forget gates, input gates, output gates and cell updates.

A group lasso penalty across columns of $\mathbf{A}$ can be used for selecting which series cause series i during estimation:

$$\min_{\mathbf{A},\mathbf{B}} \sum_{t=2}^{T} (\mathbf{x}_{it} - g_h(\mathbf{x}(t-1), \cdots, \mathbf{x}(t-R)))^2 + \lambda \sum_{j=1}^{p} \|\mathbf{A}_{:j}\|_2 \tag{12}$$

where

$$\mathbf{B} = [(B^f)^T, (B^{in})^T, (B^o)^T, (B^c)^T]^T \tag{13}$$

If λ is enough large, then many columns of $\mathbf{A}$ will be zero, which results in a sparse set of Granger causal connections.

To this end, the GCs of all variable pairs $GC_{\mathbf{x}_i(t)\to\mathbf{x}_j(t)}$ can be built by $\mathbf{A} = [A^1, \cdots, A^p]$. Then, GCs can be calculated by:

$$\mathbf{GC} = \begin{bmatrix} GC_{\mathbf{x}_1(t)\to\mathbf{x}_1(t)} & \cdots & GC_{\mathbf{x}_p(t)\to\mathbf{x}_1(t)} \\ \vdots & \ddots & \vdots \\ GC_{\mathbf{x}_1(t)\to\mathbf{x}_p(t)} & \cdots & GC_{\mathbf{x}_p(t)\to\mathbf{x}_p(t)} \end{bmatrix}$$
$$= \begin{bmatrix} \sqrt{\sum_{b'=1}^{b}\sum_{r=1}^{R}(A^1_{1rb'})^2} & \cdots & \sqrt{\sum_{b'=1}^{b}\sum_{r=1}^{R}(A^1_{prb'})^2} \\ \vdots & \ddots & \vdots \\ \sqrt{\sum_{b'=1}^{b}\sum_{r=1}^{R}(A^p_{1rb'})^2} & \cdots & \sqrt{\sum_{b'=1}^{b}\sum_{r=1}^{R}(A^p_{prb'})^2} \end{bmatrix}. \tag{14}$$

Further, in order to get binary GCs, the following threshold must be set. That is, if

$$GC_{\mathbf{x}_i(t)\to\mathbf{x}_j(t)} \geq \gamma \max(\mathbf{GC}) \tag{15}$$

where γ is a proportionality coefficient determined by expert knowledge. If above formula satisfies, then relevant elements of $\mathbf{GC}$ can be appointed to 1, and 0 otherwise.

4. A CASE STUDY ON FMP

In this section, the new scheme will be used for FMP, and real datasets are for validating the performance of the developed propagation path identification method.

4.1 Process description

Hot strip mill process is a typical multi-stage, long-process industrial process, which involves complex physical and chemical changes from raw materials to final products. There are lots of coupled control loops and variables in the manufacturing line, an abnormal operation may cause widespread alarms, which seriously affects system safety, reliability and final product quality. As a result, it is a hot topic to ensure high-quality and high-efficiency operation nowadays by means of reasonable propagation path identification technology.

As shown in Fig. 2, the whole production line is composed of six subprocesses, where FMP is the crucial one, which will be used as background process in this paper. It can be seen that there are seven stands in the finishing mill group. In each stand, hydraulic cylinder is used for strip gauge control, and electromechanical system rotates the rolls to make strip steels move forward smoothly.

According to the analyses above, the developed framework will be applied for identifying the propagation path of faults in FMP. The measured variables of FMP are listed

Table 1. Description of process variables in FMP

Variable	Description	Unit
$1-7$	Average gap of the qth stand, $q=1,...,7$	mm
$8-14$	Total force of the qth stand, $q=1,...,7$	MN
$15-20$	Work roll bending force of the qth stand, $q=2,...,7$	MN

Table 2. GC-based causality matrix

Variable No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
10	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0	0	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

in Table 1. When the considered fault occur, it not only influences the rolling force in the associated stand and ambient temperature, but also makes subsequent stand cannot satisfy the reduction rate accurately, which finally affects the exit strip thickness or flatness.

4.2 Propagation path identification results

During simulations, based on Bayesian information criterion (BIC), the model order R was set as 5. Moreover, the tuning parameter α, the number of encoder layers, b and λ were chosen to be 0.13, 1, 32 and 0.1, respectively. Further, in order to get the binary GCs, the threshold was set as 0.33. The causality matrix is listed in Table 2, where the columns and rows are cause and response variables, respectively. Then, the propagation path identification results are shown in Fig. 3. It can be observed that variable 4 can be diagnosed as the root cause of the fault. After this fault happens, by means of feedback control, total forces of the 5th, the 7th and average gap in the subsequent stands are influenced by this abnormal event.

It can be seen that well identification performance can be got by the new method, which will provide a feasible technology or solution to take corrective and timely maintenance measures for field engineers.

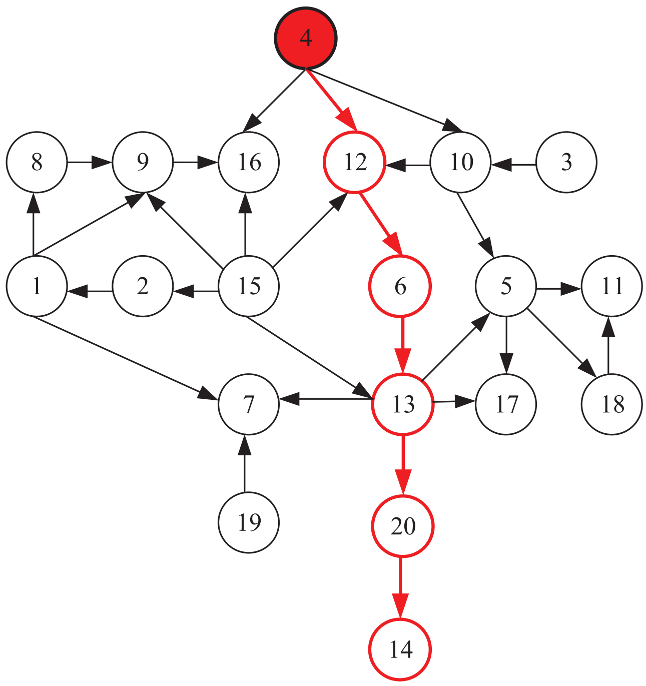

Fig. 3. Propagation path identification results

5. CONCLUSION

In this paper, an accurate propagation path identification method has been designed from a new perspective, which is crucial for supporting the field engineers' decision making. Based on the framework, the new RNNs-based GC analysis method was applied to construct causal topologies for identifying the propagation path, which has fully considered the nonlinear and dynamic relationships among time series after faults happen. Finally, the validity of the developed algorithm was verified with real FMP data, where well identification performance has been obtained.

REFERENCES

Yin, S., Ding, S. X., Xie, X. C., and Luo, H.(2014). A review on basic data-driven approaches for industrial process monitoring. IEEE Transactions on Industrial Electronics, 61, 6418-6428.

Chen, Z. W., Cao, Y., Ding, S. X., et al.(2019). A distributed canonical correlation analysis-based fault detection method for plant-wide process monitoring. IEEE Transactions on Industrial Electronics, 15(5), 2710-2720.

Ge, Z. Q.(2018). Distributed predictive modeling framework for prediction and diagnosis of key performance index in plant-wide processes. Journal of Process Control, 65, 107-117.

Ma, L., Dong, J., and Peng, K. X.(2020a). A novel hierarchical detection and isolation framework for quality-related multiple faults in large-scale processes. IEEE Transactions on Industrial Electronics, 67(2), 1316-1327.

Bauer, M. and Thornhill, N. F.(2008). A practical method for identifying the propagation path of plant-wide disturbances. Journal of Process Control, 18, 707-719.

Yuan, T. and Qin, S. J.(2014). Root cause diagnosis of plant-wide oscillations using Granger causality. Journal of Process Control, 24(2), 450-459.

Landman, R., Kortela, J., Sun, Q., and Jämsä-Jounela, S. L.(2014). Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity. Computers & Chemical Engineering, 71, 446-456.

Ma, L., Dong, J., and Peng, K. X.(2020b). A novel key performance indicator oriented hierarchical monitoring and propagation path identification framework for complex industrial processes. ISA Transactions, 96, 1-13.

Li, G., Qin, S. J., and Yuan, T.(2016). Data-driven root cause diagnosis of faults in process industries. Chemometrics & Intelligent Laboratory Systems, 159, 1-11.

Schreiber, T.(2000). Measuring information transfer. Physical Review Letters, 85, 461-464.

Landman, R. and Jämsä-Jounela, S.(2019). Fault propagation analysis by implementing nearest neighbors method using process connectivity. IEEE Transactions on Control Systems Technology, 27(5), 2058-2067.

Ma, L., Dong, J., and Peng, K. X.(2018). A practical propagation path identification scheme for quality-related faults based on nonlinear dynamic latent variable model and partitioned Bayesian network. Journal of the Franklin Institute, 355, 7570-7594.

Ahmed, U., Ha, D., Shin, S., et al(2017). Estimation of disturbance propagation path using principal component analysis (PCA) and multivariate granger causality (MVGC) techniques. Industrial & Engineering Chemistry Research, 56, 7260-7272.

Bressler, S. L. and Seth, A. K.(2010). Wiener-granger causality: A well established methodology. NeuroImage, 58, 323-329.

Tank, A., Covert, I., Foti, N., et al(2018a). An interpretable and sparse neural network model for nonlinear granger causality discovery. arXiv preprint arXiv: 1711.08160.

Tank, A., Covert, I., Foti, N., et al(2018b). Neural granger causality for nonlinear time series. arXiv preprint arXiv: 1802.05842.