cCCM-Based Quantification of Directed Brain Information Transfer for Early Detection of Mild Cognitive Impairment

Boxin Sun[1] Jinxian Deng[1] Ming Gu[1] Voyko Kavcic[2,3,5] Jian Ren[1] Bruno Giordani[4,5] Tongtong Li[1,5]
[1]Department of Electrical and Computer Engineering, Michigan State University, East Lansing, USA
[2]Institute of Gerontology, Wayne State University, Detroit, MI, USA
[3]International Institute of Applied Gerontology, Ljubljana, Slovenia
[4]Departments of Psychiatry, University of Michigan, Ann Arbor, MI, USA
[5]Michigan Alzheimer's Disease Research Center, Ann Arbor, MI, USA

***Abstract*—Directed information transfer in the brain network—also referred to as effective connectivity or brain causality—has been a central topic in cognitive neuroscience. In this paper, we quantify the effective connectivity from one brain region to another by applying causalized convergent cross mapping (cCCM) to EEG data. Our study demonstrates that cCCM-based effective connectivity can capture directed information transfer between brain regions that may not be detected by Pearson correlation, and that region pairs with low functional connectivity may still exhibit strong effective connectivity. We further apply cCCM to analyze differences between individuals with normal cognition (NC) and those with mild cognitive impairment (MCI). Our results show that machine learning models trained on cCCM features achieve high accuracy in the early detection and prediction of MCI, which is often considered a transitional stage between NC and Alzheimer's disease (AD). The proposed approach characterizes the overall physiological status of resting-state brain activity and shows promising capability for individualized cognitive health assessment in older adults.**

***Index Terms*—Directed information, brain causality, causalized convergent cross mapping, mild cognitive impairment, Alzheimer's disease**

I. Introduction

Alzheimer's disease (AD) is a degenerative brain disorder, characterized by progressive deterioration of nerve cells, eventually leading to cell death. When the neuronal damage extends to parts of the brain that enable basic bodily functions such as breathing and swallowing, it becomes fatal [1]. Unfortunately, there is currently no effective cure for the irreversible brain damage caused by AD. Therefore, early detection and prediction of pathological brain changes—long before AD symptoms are evident—is the key to prevent or delay the onset of AD.

Mild cognitive impairment (MCI) is a condition in which people show a slight, but noticeable decline in cognitive capabilities beyond what is considered to be normal for their age, and is often regarded as a transitional stage between normal cognition (NC) and AD. For this reason, in this study, we aim to develop an innovative and non-invasive techinque for early detection and prediction of MCI.

Existing research suggests that the abnormal brain functions in AD and MCI are closely related to the weakening or loss of connectivity among critical brain regions [2], since the most vulnerable neurons in neurodegeneration are the association neurons with long projections that formulate the communication channels between the brain regions. In literature, brain connectivity is generally modeled using both functional and effective connectivity. Functional connectivity characterizes the mutual dependence between brain regions, and is often measured using the Pearson correlation or mutual information. Effective connectivity (also known as causality) measures the causal influences that one neural unit exerts over another and seeks to understand the directed information transfer in the brain network [3]–[5]. Note that two time series which are weakly correlated may still have strong unidirectional or bidirectional causal coupling between them [6]. Effective connectivity, therefore, is expected to provide more insight on the directed information transfer in the brain network than functional connectivity.

Recently, causalized convergent cross mapping (cCCM) [4]—a state space reconstruction approach rooted in dynamic systems theory [7]—has shown to be conditionally equivalent to the directed information framework, and demonstrated to be a robust, effective, and easy–to–implement tool in quantifying directed information transfer and identifying causal relationships in the brain network [4], [5].

Motivated by the obseravations above, in this paper, we quantify dynamic interactions between brain regions by applying cCCM to resting-state EEG data. Our findings show that cCCM-based effective connectivity can capture directed information transfer that may not be identifiable through traditional Pearson correlation, revealing that even region pairs with weak functional connectivity can exhibit strong effective connectivity. We further apply cCCM to compare individuals with normal cognition (NC) and mild cognitive impairment (MCI). Our results demonstrate that machine learning models trained on dynamic cCCM features achieve high accuracy in detecting and predicting MCI. Overall, the proposed approach characterizes the physiological status of resting-state brain activity and offers a promising framework for individualized cognitive health assessment in older adults.

979-8-3315-7111-5/26/$31.00 ©2026 IEEE

II. Causalized Convergent Cross-Mapping and Directed Information

In this section, we briefly revisit causalized convergent cross mapping (cCCM) [4], [5], [7] and discuss its relationship with directed information (DI) [8].

A. Causalized Convergent Cross-Mapping (cCCM)

Consider two dynamically coupled variables X and Y which share the same attractor manifold $\mathbf{M}$. Let $\mathbf{X}^n = [X_1, X_2, \cdots, X_n]$ and $\mathbf{Y}^n = [Y_1, Y_2, \cdots, Y_n]$ be the time series consisting of samples of X and Y, respectively. The cCCM alorithm can be summarized as:

- *Step 1*: Construct the shadow manifolds with respect to $\mathbf{X}^n$ and $\mathbf{Y}^n$.

$$\mathbf{M}_x = \{\mathbf{x}_t \mid \mathbf{x}_t = [X_t, X_{t-\tau}, \cdots, X_{t-(E-1)\tau}], \quad t = 1 + (E-1)\tau, \cdots, n\} \tag{1}$$

$$\mathbf{M}_y = \{\mathbf{y}_t \mid \mathbf{y}_t = [Y_t, Y_{t-\tau}, \cdots, Y_{t-(E-1)\tau}], \quad t = 1 + (E-1)\tau, \cdots, n\} \tag{2}$$

- *Step 2*: For each vector $\mathbf{x}_t$, find its $E+1$ nearest neighbors in $\mathbf{M}_x$ with an index smaller than t and denote the time indices (from closest to farthest) of the $E+1$ nearest neighbors of $\mathbf{x}_t$ by $t_1, ...t_{E+1}$.
- *Step 3*: If the two signals X and Y are dynamically coupled, then the nearest neighbors of $\mathbf{x}_t$ in $\mathbf{M}_x$ would be mapped to the nearby points of Y_t on manifold $\mathbf{M}_y$. The estimated Y_t based on $\mathbf{M}_x$, or say the cross mapping from X to Y, is defined as:

$$\hat{Y}_t|\mathbf{M}_x = \sum_{i=1}^{E+1} w_i Y_{t_i} \tag{3}$$

where $t_i \leq t$,

$$w_i = \frac{u_i}{\sum_{j=1}^{E+1} u_j}, \quad \text{with} \quad u_i = exp\{-\frac{d(\mathbf{x}_t, \mathbf{x}_{t_i})}{d(\mathbf{x}_t, \mathbf{x}_{t_1})}\},$$

and d denoting the Euclidean distance between two vectors. The cross mapping from Y to X can be defined in a similar way. As n increases, it is expected that $\hat{X}_t|\mathbf{M}_y$ and $\hat{Y}_t|\mathbf{M}_x$ would converge to X_t and Y_t, respectively.

- *Step 4*: The cross mapping correlations are defined as

$$\rho_{\text{cCCM}}(X \to Y) = \rho(\mathbf{Y}^n, \hat{\mathbf{Y}}^n) \quad \text{and}$$
$$\rho_{\text{cCCM}}(Y \to X) = \rho(\mathbf{X}^n, \hat{\mathbf{X}}^n) \tag{4}$$

where ρ denotes the Pearson correlation.

B. Conditional Equivalence of cCCM and DI

The DI from $\mathbf{X}^n$ to $\mathbf{Y}^n$ is defined as [8]:

$$I(\mathbf{X}^n \to \mathbf{Y}^n) = \sum_{i=1}^{n} [H(Y_i|\mathbf{Y}^{i-1}) - H(Y_i|\mathbf{Y}^{i-1}, \mathbf{X}^i)] = \sum_{i=1}^{n} I(\mathbf{X}^i; Y_i|\mathbf{Y}^{i-1})$$

The average DI from X to Y, measured in bits per sample is defined as

$$\bar{I}_n(X \to Y) = \frac{I(\mathbf{X}^n \to \mathbf{Y}^n)}{n}$$

As we demosntrated in our previous work, based on Takens' theorem [9] and Gel'fand's result [10], and the Shannon–McMillan–Breiman theorem [11], cCCM is approximately equivalent to DI. More specifically, if (i) X and Y are dynamically coupled, zero-mean Gaussian random variables and their joint distribution is bivariate Gaussian, and (ii) $\mathbf{X}^n$ and $\mathbf{Y}^n$ are stationary ergodic Gaussian random processes, then when n is sufficiently large, we have

$$\bar{I}_n(X;Y) \approx -\frac{1}{2}\log(1 - \rho^2_{\text{cCCM}}(X,Y)). \tag{5}$$

III. Data Description

In this study, we apply cCCM to quantify directed information transfer in the brain network using EEG data. This section describes the participant cohort, data collection and preprocessing procedures, and the selection of regions of interest.

A. Demographic Information

We recruited 137 community-dwelling American older adults (122 females, 15 males), aged 60 to 90 years, from the greater Detroit area. All participants were diagnosed through the Michigan Alzheimer's Disease Research Center (MADRC) consensus conference utilizing the National Alzheimer's Coordinating Center (NACC) Unified Data Set (UDS)–84 of them being NC and 53 with MCI (42 amnestic MCI (aMCI) and 11 non-amnestic MCI (naMCI)). Because of the small number of naMCI participants, all MCI participants were combined into one group. All participants provided written informed consent. There were no significant differences among NC, aMCI, and naMCI participants in terms of education, and as expected, the average age of the MCI group (74.62$\pm$7.05) is slightly higher than that of the NC group (72.19$\pm$6.17).

B. EEG Data Collection and Pre-Processing

Brain Vision (Brain Vision, Inc.) equipment was used to record scalp EEG activity with a non-invasive high-density actiCap (64 active electrodes), modified according to the International 10-20 System. As part of a larger study on event-related potential (ERP) brain activity, each participant received a 3-minute, eye-closed resting-state EEG recording. In this article, we focus on resting-state EEG—which requires no training or active responses of the participants—due to its higher scalability, reproducibility, and clinical operability.

Analyzer 2 (Brain Vision, Inc.) was used to preprocess the baseline EEG data. Offline inspection was performed to identify and remove EEG segments contaminated by excessive noise, signal saturation, or lack of EEG activity. After preprocessing, we kept the first 100s of the eye-closed, resting-state EEG signals from all 64 channels for further analysis.

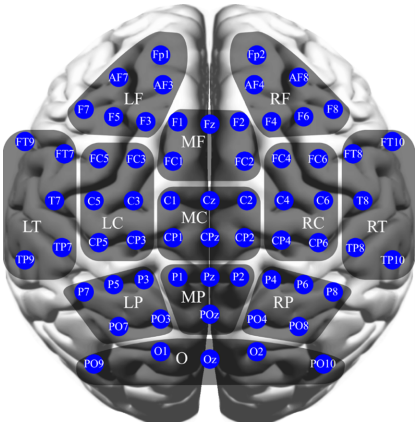


Fig. 1. Selection of regions of interest.

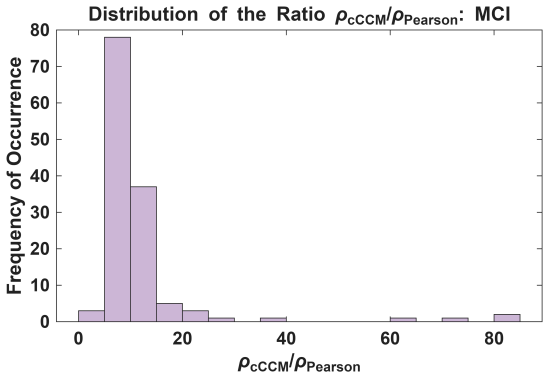


Fig. 2. The distribution of the ratio $\rho_{cCCM}/\rho_{pearson}$ for all directed ROI pairs, averaged over all window blocks and all MCI participants. Similar results were observed in the NC group as well.

C. Selection of Regions of Interest

As illustrated in Fig.1, we selected a total of twelve regions of interest (ROIs): Right frontal– RF, Medium frontal–MF, Left frontal–LF, Left temporal–LT, Left central–LC, Medial central–MC, Right central–RC, Right temporal–RT, Left parietal–LP, Medial parietal–MP, Right parietal– RP, and Occipital–Occ.

Before further analysis, we calculated the current source density (CSD) or the Laplacian (second spatial derivative) of the scalp voltage from the EEG signal for all the ROIs using the CSD Toolbox in MATLAB.

IV. A Comparison of cCCM-Derived Effective Connectivity and Pearson Correlation–Based Functional Connectivity

In this study, a comparative analysis was performed between cCCM and Pearson correlation values across all possible ROI pairs. Notably, many region pairs exhibited cCCM values markedly higher than their corresponding Pearson correlations in both the NC and MCI groups. As illustrated in Fig. 2, 51 out of the 132 directional region pairs yielded a cCCM to Pearson correlation ratio ($\rho_{cCCM}/\rho_{pearson}$) exceeding 10, while 9 pairs exceeded a ratio of 20, with one instance even reaching as high as 83.63. This aligns with our previous findings [6].

From an information-theoretic perspective, the underlying argument is that Pearson correlation characterizes the instantaneous information transfer between the current values of X and Y (i.e., X_n and Y_n). cCCM, on the other hand, considers the overall impact of both the current and previous values of X (i.e., $X_n, X_{n-1}, \cdots, X_{n-L}$) on Y_n. When the delay in information transmission from one brain region to the other is larger than one sampling period, the impact of the previous values of X on Y_n (measured by the cCCM value) would be much larger than the impact of the current value of X on Y_n (measured by the Pearson correlation).

V. Application of cCCM-Based Brain Effective Connectivity Analysis for the Early Detection of Mild Cognitive Impairment

In this section, we analyzed the differences in cCCM-based dynamic effective connectivity between NC and MCI, and extracted features to build machine learning models for soft early detection of MCI.

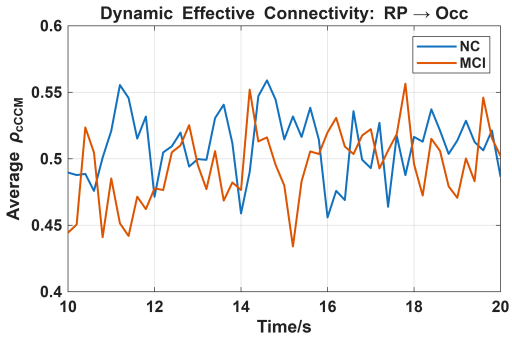
A. Differences in cCCM-Based Dynamic Effective Connectivity Between NC and MCI

To evaluate the dynamic brain effective connectivity, the sliding window approach was integrated with the cCCM algorithm. For each participant, first, the CSD sequence of each ROI was partitioned into non-overlapping blocks (windows). Effective connectivity across all directed ROI pairs was then computed using cCCM for every block, yielding a dynamic cCCM sequence. For each directed ROI pair, the mean dynamic effective connectivity for each group (NC or MCI) was obtained by averaging the corresponding dynamic cCCM sequences across all participants in that group.

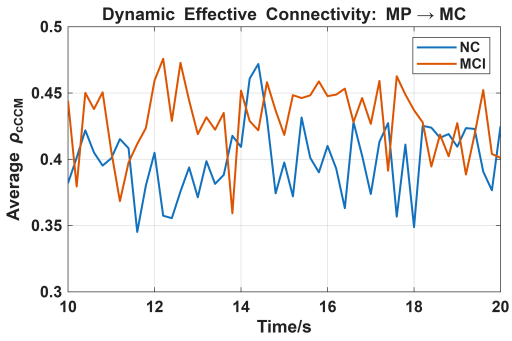
The differences in cCCM-based dynamic effective connectivity between the NC and MCI groups are illustrated in Fig. 3 (here the window size = 200 ms). Our results indicate that: (i) the effective connectivity of both NC and MCI groups varies significantly over time; and (ii) at a time period, MCI may exhibit lower effective connectivity than NC in some region pairs, but concurrently shows higher effective connectivity in other region pairs. The latter may reflect the diversity and compensatory mechanisms of the brain's communication network.

B. Multiscale Analysis of Dynamic Effective Connectivity and Discrimination of NC and MCI

In most of the existing work involving the sliding window approach, a fixed window size is used, which limits the analysis of brain dynamics to one particular scale, leading to loss of important features and observation flexibility. To overcome these drawbacks, in this study, instead of using a fixed observation window size, we chose to use a whole set of different window sizes ($L = 100, 200, 400, 800$ ms),



(a)



(b)

Fig. 3. Dynamic cCCM-based effective connectivity, NC versus MCI. (a) From right parietal (RP) to occipital (Occ). (b) From medial parietal (MP) to medial centel (MC).

therefore capture the dynamics of the functional connectivity at different frequency resolutions, which we referred to as "multiscale analysis". The reconfigurable machine learning model for NC/MCI discrimination was built in three steps:

(i) Feature selection: We performed time-frequency analysis on each cCCM vector using the continuous Wavelet Toolbox in MATLAB and took the mean, minimum and maximum of the wavelet coefficients as primitive features. For each fixed window size L, all the primitive features corresponding to NC and MCI were evaluated using the t-test, and only the M primitive features with the smallest p-values were selected to formulate the feature vector. In this research, the window size $L = 100, 200, 400, 800$ ms and the number of selected features $M = 5, 10, 15, 20, 25, 30, 35, 40, 45$.

(ii) Dimension reduction: Using a regularized Linear discriminant analysis (LDA), we mapped the feature vector of each subject to a point in a one-dimensional subspace or axis, where the differences between NC and MCI subjects are maximized. Here regularized LDA was used to reduce the noise effect caused by both biological variability and measurement errors.

(iii) Classification: Finally, we constructed decision trees based on the LDA output and performed classification using the AdaBoost classifier. By tuning the window size and the group of features used, we obtained a series of approaches, where each setting defines a unique classifier and reflects the differences between NC and MCI from a unique perspective. That is, a reconfigurable NC/MCI discrimination model was built using cCCM and resting-state EEG (Fig. 4).

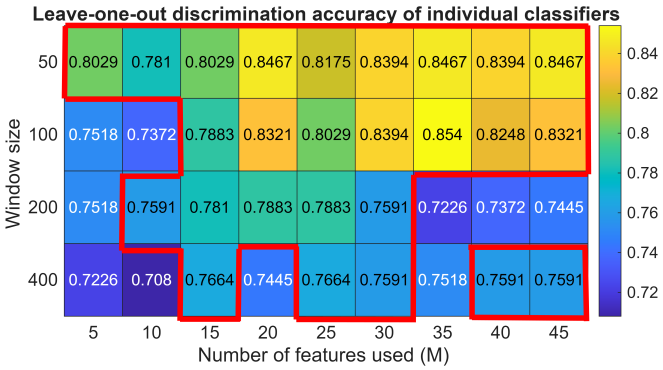


Fig. 4. Accuracy of the reconfigurable NC/MCI discrimination model based on resting-state effective connectivity.

C. Soft Detection and Prediction of MCI

Unlike existing work which generally relies only on an individual classifier, we developed a soft discrimination model [12] for NC and MCI through weighted majority voting of a selected group of reliable discrimination approaches.

More specifically, for $n = 1, 2, \ldots, N$, let the vote of voter n be v_n, where $v_n = +1$ if the decision is NC and $v_n = -1$ if the decision is MCI. Let a_n denote the accuracy of voter n. For each subject, in addition to the binary classification (NC or MCI), we defined a soft discrimination score

$$s = \frac{1}{N} \sum_{n=1}^{N} a_n v_n,$$

which is a connectivity-based score of the participant obtained from the EEG data and reflects the status of the overall physiological brain activity at the resting state. Potentially, it can be used to predict how likely the participant would progress from NC to MCI.

In this study, we selected $N = 26$ individual classifiers with the highest accuracy in leave-one-out (LOO) cross-validation to serve as voters, as marked by the red boxes in Fig. 4. The same group of voters was used for the soft discrimination of NC and MCI for 5-fold cross-validation as well,and the results are shown in Fig. 5. In 5-fold cross-validation, random shuffles were applied to enhance the dataset cardinality (i.e., the number of participants in that dataset). Specifically, the results represent the combination of 10 independent 5-fold cross-validation runs, with the test group randomly generated in each run.

The soft discrimination score for each participant is shown in Fig. 6 (a). As can be seen, each participant receives a score s within $[-1, +1]$, where a positive s implies that the decision is NC, and a negative s implies that the decision is MCI. The soft score s also serves as an indicator of the cognitive status level in the sense that the larger the s, the better the cognitive health. For example, a score of 0.8 would indicate that the participant is in a very optimistic NC condition, and a score of 0.08 would indicate that, although the participant is classified as

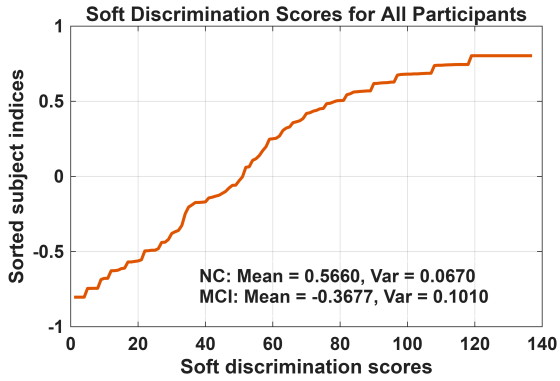
Leave-one-out cross-validation		Predicted Class		Recall (or Sensitivity)
		NC	MCI	
Actual Class	NC	80	4	95.23%
	MCI	6	47	88.68%
Precision		93.02%	92.16%	ACC=92.70%

(a)

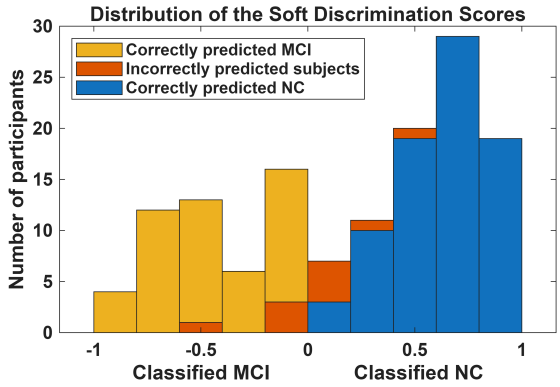
5-fold cross-validation		Predicted Class		Recall (or Sensitivity)
		NC	MCI	
Actual Class	NC	798	42	95.00%
	MCI	78	452	85.28%
Precision		91.10%	91.49%	ACC=91.24%

(b)

Fig. 5. Confusion matrices of the classification result. (a) Leave-one-out (LOO) cross-validation. (b) 5-fold cross-validation, test repeated for 10 random runs.



(a)



(b)

Fig. 6. Soft discrimination of NC and MCI based on leave-one-out cross-validation. (a) Soft discrimination scores in ascending order for cCCM classifications. (b) Distribution of the soft discrimination scores.

NC, however, there is a trend that the participant may progress to the MCI condition. A similar interpretation applies to the negative scores. In other words, the soft discrimination score may predict whether an NC participant is likely to progress to MCI, which is critical for timely intervention to prevent further cognitive decline. The distribution of soft scores (Fig. 6(b)) also showed that incorrect predictions occurred primarily (70%) when the soft discrimination score fell within the range $[-0.2,\ 0.2]$, where differences between low-scoring NC and high-scoring MCI participants were not significant.

Based on the longitudinal data of a subset sample of 50 participants, we showed that the 9-18 months progression trend prediction accuracy based on the soft score is 82%. This implies that EEG-based machine learning models can be used to predict the individualized progression of cognitive health in older adults.

VI. Conclusions

In this paper, we quantified directed brain connectivity by applying cCCM to EEG data. Our analysis indicated that cCCM may capture directional interactions which are not reflected by Pearson correlation, and region pairs with low functional connectivity can still exhibit strong effective connectivity. We then applied cCCM to analyze the differences between NC and MCI. Our results demonstrated that machine learning models trained on dynamic EEG-cCCM features can achieve high accuracy in the early detection MCI, and shows promising capability in proactive prediction of people at risk of MCI before clinical symptoms may occur.

Acknowledgements

This research is partially supported by NSF under award 2032709/Li, and NIH under awards R21-AG046637/Kavcic, R01-AG054484/Kavcic, P30AG072931/Paulson and P30AG024824/Yung.

References

[1] Alzheimer's Association, "2025 Alzheimer's disease facts and figures," Tech. Rep., 2025.

[2] H. Fu, J. Hardy, and K. E. Duff, "Selective vulnerability in neurodegenerative diseases," *Nature Neuroscience*, vol. 21, no. 10, pp. 1350–1358, sep 24 2018.

[3] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, pp. 1273–1302, 8 2003.

[4] J. Deng, B. Sun, N. Scheel, A. Renli, D. Zhu, D. Zhu, J. Ren, T. Li, and R. Zhang, "Causalized convergent cross-mapping and its approximate equivalence with directed information in causality analysis," *PNAS Nexus*, vol. 3, 2024.

[5] B. Sun, J. Deng, N. Scheel, D. C. Zhu, J. Ren, R. Zhang, and T. Li, "Causalized convergent cross mapping and its implementation in causality analysis," *Entropy*, vol. 26, p. 539, 6 2024.

[6] A. Renli, B. Sun, M. Gu, T. Martin, V. Kavcic, T. Li, and B. Giordani, "EEG-based information transfer paths during motion discrimination: Detectable differences between normal cognition and MCI," *Alzheimer's & Dementia*, vol. 21, p. e70868, 11 2025.

[7] G. Sugihara, R. May, H. Ye, C.-h. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting Causality in Complex Ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, 10 2012.

[8] J. Massey, "Causality, feedback, and directed information," in *The International Symposium on Information Theory and Its Applications*, 11 1990, pp. 303–305.

[9] D. Ruelle and F. Takens, "On the nature of turbulence," *Communications in Mathematical Physics*, vol. 20, no. 3, pp. 167–192, 9 1971.

[10] I. Gel'fand and Y. A, "Calculation of amount of information about a random function contained in another such function," in *American Mathematical Society translations*, 12 1957, pp. 199–246.

[11] P. H. Algoet and T. M. Cover, "A sandwich proof of the Shannon-McMillan-Breiman theorem," *The Annals of Probability*, vol. 16, 4 1988.

[12] J. Deng, B. Sun, V. Kavcic, M. Liu, B. Giordani, and T. Li, "Novel methodology for detection and prediction of mild cognitive impairment using resting-state EEG," *Alzheimer's and Dementia*, vol. 20, 2024.