

Review Article

Mathematical Issues in the Inference of Causal Interactions among Multichannel Neural Signals

Young-Jin Jung,^{1,2} Kyung Hwan Kim,³ and Chang-Hwan Im¹

¹ *Department of Biomedical Engineering, Hanyang University, Seoul 133-791, Republic of Korea*

² *Research Institute of Industrial Science, Hanyang University, Seoul 133-791, Republic of Korea*

³ *Department of Biomedical Engineering, Yonsei University, Wonju 220-710, Republic of Korea*

Correspondence should be addressed to Chang-Hwan Im, ich@hanyang.ac.kr

Received 28 October 2011; Accepted 16 November 2011

Academic Editor: Kiwoon Kwon

Copyright © 2012 Young-Jin Jung et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Within the last few decades, attempts have been made to characterize the underlying mechanisms of brain activity by analyzing neural signals recorded, directly or indirectly, from the human brain. Accordingly, inference of functional connectivity among neural signals has become an indispensable research tool in modern neuroscience studies aiming to explore how different brain areas are interacting with each other. Indeed, remarkable advances in computational sciences and applied mathematics even allow the estimation of causal interactions among multichannel neural signals. Here, we introduce the brief mathematical background of the use of causality inference in neuroscience and discuss the relevant mathematical issues, with the ultimate goal of providing applied mathematicians with the current state-of-the-art knowledge on this promising multidisciplinary topic.

1. Introduction and Background

Traditional functional neuroimaging studies have focused on the functional specification of brain areas. However, only a limited amount of information regarding the underlying neuronal mechanisms can be obtained when such spatial specification is studied. Recently, research interests have shifted toward describing how different brain areas interact with each other, with the hope of better understanding the functional organization of the cortical network [1–7]. Correlation [1, 2], coherence [3], phase locking value [4], mean phase coherence [5], and mutual information [6, 7] have been used to estimate functional interaction between multiple neural assemblies. These methods have been applied to signals obtained via many different functional neuroimaging modalities such as electroencephalography (EEG), local field potential (LFP), intracranial EEG (iEEG), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI). Recent advances in neural signal analysis

have also enabled the estimation of direction of information flow between different cortical areas [8–12] even beyond conventional correlation-based functional connectivity analyses.

Over the past few decades, a number of measures for “directional” coupling between neural activities have been developed [8–13] and applied to various fields in both basic and clinical neuroscience [14–24]. Although a variety of causality estimators have been widely used for characterizing the mechanisms of neuronal networks, notable limitations and issues still exist that require intervention by applied mathematicians. For example, multivariate autoregressive (MVAR) model-based causality estimators do not accurately infer information flow between nonstationary and/or highly nonlinear neural signals. The determination of model order and the dependency on the analysis sample size are other issues that should be addressed in future studies. Furthermore, most non-MVAR-based causality estimators can only be used to infer causality between two signals, and thus need to be extended to the case of multichannel (≥ 3) signal analyses [12, 13, 25].

Here, we introduce several mathematical signal analysis methods for estimating directional coupling between neural activities, all of which have been widely used in basic and applied neuroscience. Additionally, this paper attempts to illustrate the important mathematical issues that need to be addressed to improve the conventional causality estimators, with the aim to stimulate interest in this imperative multidisciplinary research topic among applied mathematicians.

2. MVAR-Based Causality Estimators

Recently, a number of causality estimation techniques have been developed to infer causality among multiple neural signal generators. The MVAR model—a linear multivariate time series model with a long history of application in econometrics [8]—has been frequently applied for causality estimations. The MVAR model is an extended version of the autoregressive (AR) model, a simple approach to time series characterization that assumes that for any given univariate time series, its consecutive measurements contain information regarding the process that generated it. The AR model can be implemented by modeling the current value of any variable as the weighted linear sum of its previous values. In the AR model, the value of a time series x at time t , x_t can be estimated using:

$$x_t = \alpha_0 + \sum_{k=1}^p \alpha_k x_{t-k} + e_t, \quad (2.1)$$

where α , p , and e_t represent AR-matrix coefficients, the model order, and the uncorrelated Gaussian random process with a zero mean, respectively.

2.1. Granger Causality

Granger causality [8] has been proposed in the field of econometrics to quantify the causal relationship between two different time series. Specifically, this simple technique uses an MVAR model to linearly predict the future values of \mathbf{x} and \mathbf{y} , vectors of deterministic variables. The MVAR model attempts to estimate the value of x_t using:

$$x_t = \alpha_0 + \sum_{k=1}^p \alpha_k x_{t-k} + \sum_{k=1}^p \beta_k y_{t-k} + w_t, \quad (2.2)$$

where α and β represent the AR-matrix coefficients and w_t the uncorrelated multivariate Gaussian random process with a zero mean. In contrast to (2.1), where the current value of a time series is estimated as the weighted sum of its previous values, the current value x_t in (2.2) is estimated using the previous values of two signal vectors x and y . We can judge whether there exists the Granger causality from y to x by inspecting whether the past information from both time series significantly improves the prediction of the future of x , rather than using the past information from x alone. In other words, if the prediction error for the MVAR model (w_t) is smaller than that for the AR model (e_t), it can be concluded that y causes x . In this way, Granger causality can be evaluated using

$$F \equiv \frac{(RSS_0 - RSS_1)/M}{RSS_1/(T - 2M - 1)}, \quad (2.3)$$

where

$$\begin{aligned} RSS_0 &= \sum_{i=1}^T e_i^2, \\ RSS_1 &= \sum_{i=1}^T w_i^2, \end{aligned} \quad (2.4)$$

where T represents the number of observations. To assess the statistical significance of the estimated Granger causality, the F -test with the null hypothesis, $H_0: \beta_k = 0$ (i.e., y_t does not influence the generation of x_t) is generally used. If $\beta_k = 0$ for all $k = 1, 2, \dots, p$, the Granger causality value F becomes zero as RSS_0 equals RSS_1 . Conversely, if the null hypothesis is rejected, that is, F is sufficiently large, it can be concluded that y_t causes x_t .

To test this hypothesis, a traditional F -test derived from an ordinary least squared regression for each equation can be used. To test the statistical significance of F , the cumulative F distribution is first estimated, after which the probability of the F value can be calculated by $P_{GC} = 1 - \text{CDF}(F)$, where CDF represents the cumulative distribution function and P_{GC} represents the probability of Granger causality. For example, $P_{GC} = 1$ would indicate that no causal interaction exists between two time series, while $P_{GC} = 0$ would signal a strong directional influence ($y_t \rightarrow x_t$).

However, the MVAR model is problematic when estimating the appropriate model order p . Basically, most model order estimation methods are based on the maximum likelihood principle, which allows the determination of the highest possible model order in MVAR signal modeling. Akaike information criterion (AIC) [26] is also based on this concept and was the earliest method to estimate MVAR model orders. As AIC generally chooses larger than optimal model orders, the Bayesian information criterion (BIC) [27]—which is based on the Bayes estimator—was developed by Schwarz. The BIC generally penalizes free parameters more strongly than the AIC, and thus provides more accurate estimates of MVAR model orders. Although several modifications of the AIC and BIC have been recently developed [28–35], the estimation of accurate and reliable model orders remains an important issue.

2.2. Directed Transfer Function

Directed transfer function (DTF) is a widely used tool in identifying information flow between multichannel neural signals. Even though both Granger causality and DTF are based on MVAR modeling, the DTF procedure differs slightly from Granger causality. As described

above, Granger causality uses the variance of prediction errors to estimate the causal interaction, while DTF uses a matrix transfer function derived from MVAR model coefficients [9, 36]. In the framework of the MVAR model, a multivariate process of DTF can be described as a data vector X of N source signals: $X_t = (X_1(t), X_2(t), \dots, X_N(t))^T$. The MVAR model can then be constructed as

$$X_t = \sum_{k=1}^p A_k X_{t-k} + E_t, \quad (2.5)$$

where E_t represents a vector composed of white noise values at time t , A_k is an $N \times N$ matrix composed of the model coefficients, and p is the model order of MVAR. Note that (2.1) is a special case of (2.5) when $N = 1$. The MVAR model is then transformed into the frequency domain as follows:

$$X(f) = A^{-1}(f)E(f) = H(f)E(f), \quad (2.6)$$

where f denotes a specific frequency and the $H(f)$ matrix represents the so-called transfer matrix, which is defined as

$$H(f) = A^{-1}(f) = \left(\sum_{k=0}^p A_k e^{-i2\pi f k \Delta t} \right)^{-1}, \quad A_0 = -I, \quad (2.7)$$

where I is the identity matrix.

The DTF can then be defined in terms of the elements of the transfer matrix H_{ij} as

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^k |H_{im}(f)|^2}, \quad (2.8)$$

where $\gamma_{ij}^2(f)$ denotes the ratio between inflow from signal j to signal i and all inflows to signal i and k represents the number of signals. The DTF ratio $\gamma_{ij}^2(f)$ ranges from 0 to 1, with values approaching to 1, suggesting that signal i is caused by signal j , whereas values approaching to 0 indicating that no information flow from signal j to signal i exists at a specific frequency.

2.3. Partial Directed Coherence

Partial directed coherence (PDC) was proposed by Baccalá and Sameshima as a frequency domain counterpart to Granger causality [11] and is based on a spectral representation of (2.5), defined as

$$A(f) = \sum_{k=0}^p A_k e^{-i2\pi f k \Delta t}, \quad (2.9)$$

$$\bar{A}(f) = I - A(f),$$

where I is an identity matrix [37]. The estimate of PDC from X_m to X_n is defined as

$$\text{PDC}_{X_n \leftarrow X_m}(f) = \frac{\overline{A}_{n,m}(f)}{\sqrt{\sum_{k=1}^Q |A_{n,k}(f)|^2}}, \quad (2.10)$$

where $\overline{A}_{n,m}$ is the (n, m) -th element of \overline{A} .

2.4. Modified MVAR-Based Estimators

Although the MVAR-based causality estimators described above have been shown to be useful in many neuroscience problems, they are not applicable for all types of neural signals. For example, signals which have severely unbalanced model residual variances are not appropriate for PDC. Accordingly, several additional modified causality estimators, including generalized partial directed coherence [38, 39], Geweke's Granger causality [40, 41], Wiener Granger causality [42], and direct directed transfer function [43], were subsequently developed to broaden application.

The generalized partial directed coherence (gPDC) was first proposed by Baccalá et al. to circumvent the numerical problem associated with time series scaling [38], by which a variance stabilization of the frequency domain representation of lagged causality could be achieved [39]. The gPDC estimator is defined as

$$\pi_{ij}^{(w)}(f) = \frac{(1/\sigma_i)\overline{A}_{ij}(f)}{\sqrt{\sum_{k=1}^N (1/\sigma_k^2)\overline{A}_{kj}(f)\overline{A}_{kj}^*(f)}}, \quad (2.11)$$

where σ_i represents the variance of the i th input process. gPDC was modified from PDC to improve the identification of causal interactions between signals with severely unbalanced model residual variances [39].

The Geweke's Granger causality is derived from Geweke's formulation [40, 41] and is defined as

$$I_{k \rightarrow l}(f) = \frac{(Z_{kk} - (Z_{lk}^2/Z_{ll}))|H_{lk}|^2}{|S_{ll}(f)|}, \quad (2.12)$$

$$I_{l \rightarrow k}(f) = \frac{(Z_{ll} - (Z_{kl}^2/Z_{kk}))|H_{kl}|^2}{|S_{kk}(f)|},$$

where $S_{kk}(f)$ and $S_{ll}(f)$ represent the individual power spectra of sites k and l , respectively, and the expressions for H_{lk} can be found in (2.6). Z_{kk} , Z_{ll} , Z_{lk} , and Z_{kl} are elements of the covariance matrix Z for the noise vector of the bivariate model. Geweke's Granger causality at frequency f is expressed as the fraction of the total power at the frequency at one site that can be explained by the causal influence from the other. As seen in (2.12) and (15), Geweke's Granger causality can be evaluated solely using the bivariate model. Recently, Bressler and Seth [42] introduced Wiener-Granger causality and discussed its merits and limitations in various neuroscience applications [42].

The direct directed transfer function (dDTF) was proposed by Korzeniewska [43] for the analysis of direct information transfer among brain structures using local field potentials. To calculate the dDTF, partial coherence χ_{ij} and full-band frequency DTF (ffDTF) η_{ij} were independently defined as

$$\begin{aligned}\chi_{ij}^2(f) &= \frac{M_{ij}^2(f)}{M_{jj}(f)M_{ii}(f)}, \\ \eta_{ij}^2(f) &= \frac{|H_{ij}(f)|^2}{\sum_f \sum_{m=1}^k |H_{im}(f)|^2},\end{aligned}\tag{2.13}$$

where M_{ij} represents the minor produced by removing the i th row and j th column of a spectral matrix S . In multivariate signals, partial coherences may provide more specific information regarding causal interactions among signals than ordinary coherences [44]. The value of dDTF is defined as the product of the above two variables and can be expressed as

$$\delta_{ij}(f) = \chi_{ij}^2(f)\eta_{ij}^2(f).\tag{2.14}$$

The dDTF method was proposed to circumvent some problems associated with DTF, specifically its inability to differentiate between the direct and indirect connections [43].

2.5. Examples of Practical Applications

To date, Granger causality has been extensively applied to the analysis of neural signals [45–58]. For examples, Hesse et al. used Granger causality to assess directed interdependencies between neural signal generators related to the Stroop task [45]. Seth also demonstrated that Granger causality may represent a useful tool for determining how interregional directional coupling is modulated by behavior [46]. Moreover, Sato et al. proposed a wavelet-based Granger causality, which they applied to fMRI signals [47]. Tang et al. applied both a blind source separation algorithm and Granger causality to the analysis of a high-density scalp EEG dataset and assessed the top-down and bottom-up influences [48]. Gow et al. demonstrated the potential value of combining Granger causality analyses with multimodal imaging to explore the functional architecture of cognition [49].

The DTF algorithm has also been extensively applied to various aspects within neuroscience, particularly to the analyses of electrophysiological signals such as EEG, MEG, and iEEG, because frequency-domain analysis is generally required in these modalities. Franaszczuk et al. first applied the DTF algorithm to the localization of ictal onset zones in temporal lobe epilepsy patients [59, 60]. Astolfi et al. demonstrated that the DTF algorithm could be used to assess the time-varying functional connectivity patterns from noninvasive EEG recordings in human [61]. Babiloni et al. investigated cortical causal interactions from combined high-resolution EEG and fMRI data and showed that DTF was able to unveil the direction of the information flow between the cortical regions of interest [62]. Kuś et al. attempted to characterize EEG activity propagation patterns in beta and gamma bands during finger movements, demonstrating that short-time DTF can successfully identify frequency selective information from EEGs [63]. Ding et al. and Wilke et al. also applied the DTF algorithm to EEG and iEEG signals acquired from intractable partial epilepsy patients

in order to better describe the structure of seizures in terms of space, time, and frequency [64, 65]. Kim et al. applied a source localization technique called FINEs in addition to DTF to iEEG signals and verified that directional connectivity analysis could be a useful tool to identify epileptogenic sources located outside of the iEEG electrodes [66].

As with DTF, PDC analyses have recently been applied to a variety of practical applications: Sun et al. demonstrated that PDC is a useful tool for evaluating changes in cortical interdependences in the context of different psychotic or mental states and can also be used to diagnose affective disorders [21]. Similarly, Zhang et al. used PDC to estimate cortical interactive networks during the mental rotation of Chinese characters, demonstrating different changes in cortical networks according to task difficulty [18]. Furthermore, Zhu et al. studied the effects of brain development and aging on cortical interactive network pattern, demonstrating that the PDC analysis of EEG is a powerful approach for characterizing brain development and aging [24].

3. Non-MVAR-Based Causality Estimators

3.1. Transfer Entropy

Information theoretic measures have widely been utilized to quantify mutual dependence between time series. Although standard time-delayed mutual information can estimate mutual dependence between neural signals, it is not able to distinguish information flow [12]. To circumvent this issue, Schreiber developed a new causality estimator named transfer entropy (TE), on the basis of the entropy rate,

$$h_x = \langle -\log[P(x_{n+1} | x_n)] \rangle, \quad (3.1)$$

where $\langle \cdot \rangle$ denotes an expectation value, $P(x)$ represent the probability of x , $P(x_{n+1} | x_n)$ is the conditional probability of x_{n+1} given x_n , and n is the time sample position. To estimate the information flow, the conditional entropy rate of x_{n+1} given both y_n and x_n

$$h_{x|y} = \langle -\log[P(x_{n+1} | x_n, y_n)] \rangle \quad (3.2)$$

has to be introduced. This indicated the average uncertainty about the future state ($= x_{n+1}$) of $x_{(t)}$, conditional on the current state ($= y_n$) of $y_{(t)}$ as well as on its own current state ($= x_n$). The transfer entropy can be defined as the difference between h_x and $h_{x|y}$ [67], in the following form:

$$T_{y \rightarrow x} = \sum P(x_{n+1}, x_n, y_n) \log \frac{P(x_{n+1} | x_n, y_n)}{P(x_{n+1} | x_n)}, \quad (3.3)$$

where $P(x_{n+1}, x_n, y_n)$ is the joint probability, evaluated by the sum of all available realizations of (x_{n+1}, x_n, y_n) in time series.

Many researchers now apply the TE algorithm to the field of neuroscience [67–71], as TE has been demonstrated to be more sensitive to nonlinear signal properties than the conventional MVAR-based causality estimators [69]. However, TE analyses are restricted to bivariate situations and require substantially more data samples than MVAR-based causality estimators.

3.2. Phase Slope Index

To robustly estimate the direction of information flow in multivariate time series, Nolte proposed a new causality estimator called phase slope index (PSI) [13], basic assumption of which states that mixing does not affect the imaginary part of the complex coherency of a multivariate times series [72]. Measured data Y_t are assumed to be a superposition of two sources X_t and additive noise E_t

$$Y_t = X_t + BE_t, \quad (3.4)$$

where B represents a mixing matrix that merges the additive noise into the measurement channels. The measured data are then divided into K segments and used to calculate the cross-spectral density as follows:

$$S_{ij}(f) = \frac{1}{K} \sum_k z_i(f, k) z_j^*(f, k), \quad (3.5)$$

where $z_i(f, k)$ represents the Fourier transform of the i th channel data and k th segment and S_{ij} is the cross-spectral matrix between i th and j th time series. PSI is defined as

$$\tilde{\Psi}_{ij} = \Im \left(\sum_{f \in F} C_{ij}^*(f) C_{ij}(f + \delta f) \right), \quad (3.6)$$

where

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f) S_{jj}(f)}} \quad (3.7)$$

is the complex coherency, δf is the specific frequency resolution, F is the frequency band of interest, and $\Im(\cdot)$ denotes the imaginary part. Finally, the PSI is normalized using its standard deviation and is expressed as

$$\Psi = \frac{\tilde{\Psi}}{\text{std}(\tilde{\Psi})}, \quad (3.8)$$

Nolte et al. presented several computer simulations, via which the relative performances of Granger causality and PSI were compared. In these simulations, PSI was found to perform better than Granger causality in inferring causal relationship between signals with nonlinear interactions. As the PSI is a nonparametric approach, it has several key advantages over conventional parametric approaches represented by the MVAR models. For instance, the PSI not only requires a lower computational load than the MVAR-based approaches, but it is also independent from the signal's stationarity. However, the PSI has a limitation in that it is also a pairwise metric of directional interactions and is thereby vulnerable to the ambiguity between direct and indirect influences [25].

3.3. Nonlinear Granger Causality

To estimate causal interactions between the nonlinear bivariate neural signals, nonlinear Granger causality (NGC) was developed [73, 74]. The basic concept of NGC is similar to TE in that NGC concludes that $y_{(t)}$ does not cause $x_{(t)}$ if the value of h_x in (3.1) is comparable to $h_{x|y}$ in (3.2). Gourévitch [37] defined the nonlinear Granger causality as follows:

$$\text{NGC}_{x \leftarrow y} = \frac{C^2(x_{n+1}, x_n, y_n)}{C^2(x_n, y_n)} - \frac{C^2(x_{n+1}, x_n)}{C^2(x_n)}, \quad (3.9)$$

where C^2 is the correlation integral of order 2. This correlation integral was proposed by Grassberger [75]. For any given vectorial signal dimension L and length of signal T , the correlation integral of order q is defined as

$$C^q(X) = \left(\frac{1}{T-L} \sum_{t=L+1}^T \left(\frac{1}{(T-L-1)} \sum_{s=L+1, s \neq t}^T 1_{\{\|X(t)-X(s)\| < r\}} \right)^{q-1} \right)^{1/(q-1)}, \quad (3.10)$$

where $\|\cdot\|$ represents the maximum norm, 1_A is 1 in a set A , 0 otherwise, and r is a positive scalar. The bivariate version for two signals X and Y of the same dimension L and the same length T is expressed as

$$C^2(X, Y) = \frac{1}{(T-L)(T-L-1)} \sum_{t=L+1}^T \sum_{s=L+1, s \neq t}^T 1_{\{\|X(t)-X(s)\| < r\}} 1_{\{\|Y(t)-Y(s)\| < r\}}. \quad (3.11)$$

3.4. Partial Nonlinear Granger Causality

Recently, Gourévitch et al. proposed a new method for estimating nonlinear causal interactions [37], termed partial nonlinear Granger causality (PNGC). The PNGC algorithm is able to estimate direct causality from X_m to X_n when Q signals are considered. PNGC is defined as

$$\text{PNGC}_{x_n \leftarrow x_m} = \frac{C^2(X_n^f, X_1^p, \dots, X_Q^p)}{C^2(X_1^p, \dots, X_Q^p)} - \frac{C^2(X_n^f, X_1^p, \dots, X_{m-1}^p, X_{m+1}^p, X_Q^p)}{C^2(X_1^p, \dots, X_{m-1}^p, X_{m+1}^p, X_Q^p)}, \quad (3.12)$$

Although PNGC showed promising results when applied to complex systems, it is still dependent on model order and scale [37]. Consequently, if nonlinearity is suspected, PNGC should be used only as a complementary tool.

4. Mathematical Issues in Causality Inference

4.1. Issues in MVAR-Based Causality Inference

The most popular causality estimators—GC [8], DTF [9], and PDC [11]—as well as their modifications are based on MVAR modeling of neural signals. The MVAR modeling is highly

dependent on the selection of model orders: too low order may not provide an exact expression of the signal feature, while too high model order may result in overfitting. Thus, the correct choice of an MVAR model orders is critically important for precise causality inference. Although several methods have been proposed to estimate proper model orders (like AIC [26], BIC [27], deviance information criterion (DIC) [33], minimum description length (MDL) [30], focused information criterion (FIC) [34], minimum message length (MML) [28], and others [29, 31, 33]) and some investigators have attempted to compare performances of different model order determination criteria [35, 76], no golden rule exists for the model order selection, and further research is clearly needed.

Moreover, MAVR-based causality estimators guarantee accurate causality inference only when datasets (signals) satisfy stationary conditions [9, 77], whereby their multivariate probability distribution is not affected by timeshift. At the very least, the mean, variance, and autocorrelation of multivariate time series should not vary over time. Unfortunately, these conditions cannot be satisfied in most cases, and thus, some mathematical transformations are often required to make the time series become roughly stationary. Nolte et al. demonstrated that MVAR-based approaches typically fail to estimate causal interactions between neuronal signals that are not stationary [13]. Although several stationary tests (e.g., unit root test [78] and Sargan and Bhargava test [79]) have been introduced to assess whether or not a time series is stationary, most were not verified in practical neural signals. In neuroscience applications, issues of stationarity also have to be carefully dealt with by considering an empirical appraisal of the participants' behavioral states [11].

Another critical limitation affecting the reliability of causality estimators is the linear modeling of neural signals [80]. Neural time-series signals can take several forms; for example, spikes, noisy signal, and highly correlated signals, may have a nonlinear form [37]. Accordingly, it is imperative to develop techniques for causality analysis that accommodate nonlinear time series, as most current studies on the causal network inference do not verify signal linearity, nor do they account for nonlinearity. Specifically, many MVAR-based models (such as PDC) are not robust to simple nonlinear linkage [37].

Generally, MVAR-based causality estimators require the appropriate selection of signal sample number. In one study, Schlögl assessed the dependency of several MVAR algorithms on the number of time samples, demonstrating that sufficient numbers of samples are required to obtain a reliable estimate of causal interactions among neural signals [81]. Moreover, Schlögl also showed causality inference to be highly dependent on both MVAR estimation methods as well as model order in cases with the same number of time samples. As the number of time samples is generally limited in most practical examples, a more systematic approach to reliably determine the number of time samples and appropriate MVAR estimators should be developed in future studies.

4.2. Issues in Non-MVAR-Based Causality Inference

While most non-MVAR-based causality estimators, such as PNGC, nonlinear Granger causality, TE, and PSI, were introduced to circumvent the well-described problems of MVAR-based causality estimators, many can only be applied to causality inferences of bivariate neural signals. As such, further research is required to extend bivariate causality inferences to include multivariate (more than three) causality inference. Furthermore, a method for determining the proper model order in PNGC remains an ongoing problem [37], as with MVAR-based estimators.

5. Conclusion

Here, we summarized the mathematical techniques used in causality estimation, all of which have been extensively applied to infer causal relationships among multichannel neural signals. We also described the limitations of current methods and presented several ongoing problems, some of which may be of interest to applied mathematicians. We hope that this paper will serve as a useful guide for researchers in the field of applied mathematics and helps raise awareness of this important research topic.

Author Contribution

Y.-J. Jung and K.-H. Kim are co-first authors and contributed equally to this work.

Acknowledgments

This work was supported by the Original Technology Research Program for Brain Science through the National Research Foundation of Korea (NRF) Grant funded by the Ministry of Education, Science, and Technology (no. 2010-0018840).

References

- [1] A. Aarabi, F. Wallois, and R. Grebe, "Does spatiotemporal synchronization of EEG change prior to absence seizures?" *Brain Research*, vol. 1188, no. 1, pp. 207–221, 2008.
- [2] G. Thut, D. Veniero, V. Romei, C. Miniussi, P. Schyns, and J. Gross, "Rhythmic TMS causes local entrainment of natural oscillatory signatures," *Current Biology*, vol. 21, no. 14, pp. 1176–1185, 2011.
- [3] G. Varotto, E. Visani, S. Franceschetti, G. Sparacino, and F. Panzica, "Spectral and coherence analysis of EEG during intermittent photic stimulation in patients with photosensitive epilepsy," *International Journal of Bioelectromagnetism*, vol. 11, no. 4, pp. 189–193, 2009.
- [4] J. P. Lachaux, E. Rodriguez, J. Martinerie, and F. J. Varela, "Measuring phase synchrony in brain signals," *Human Brain Mapping*, vol. 8, no. 4, pp. 194–208, 1999.
- [5] F. Mormann, K. Lehnertz, P. David, and C. E. Elger, "Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3, pp. 358–369, 2000.
- [6] S. Baillet, L. Garnero, G. Marin, and J. P. Hugonin, "Combined MEG and EEG source imaging by minimization of mutual information," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 5, pp. 522–534, 1999.
- [7] J. Jeong, J. C. Gore, and B. S. Peterson, "Mutual information analysis of the EEG in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 112, no. 5, pp. 827–835, 2001.
- [8] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.
- [9] M. J. 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.
- [10] K. Sameshima and L. A. Baccalá, "Using partial directed coherence to describe neuronal ensemble interactions," *Journal of Neuroscience Methods*, vol. 94, no. 1, pp. 93–103, 1999.
- [11] 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.
- [12] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, 2000.
- [13] G. Nolte, A. Ziehe, V. V. Nikulin et al., "Robustly estimating the flow direction of information in complex physical systems," *Physical Review Letters*, vol. 100, no. 23, Article ID 234101, 2008.
- [14] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy et al., "Slowing and loss of complexity in Alzheimer's EEG: two sides of the same coin?" *International Journal of Alzheimer's Disease*, pp. 1–10, 2011.

- [15] M. Grosse-Wentrup, B. Schölkopf, and J. Hill, "Causal influence of gamma oscillations on the sensorimotor rhythm," *NeuroImage*, vol. 56, no. 2, pp. 837–842, 2011.
- [16] K. Blinowska, R. Kus, M. Kaminski, and J. Janiszewska, "Transmission of brain activity during cognitive task," *Brain Topography*, vol. 23, no. 2, pp. 205–213, 2010.
- [17] A. Korzeniewska, C. M. Crainiceanu, R. Kuś, P. J. Franaszczuk, and N. E. Crone, "Dynamics of event-related causality in brain electrical activity," *Human Brain Mapping*, vol. 29, no. 10, pp. 1170–1192, 2008.
- [18] H. Zhang, Y. Sun, J. Yan et al., "Cortical interactive network during mental rotation of Chinese character," *Neuroscience Letters*, vol. 461, no. 2, pp. 185–189, 2009.
- [19] A. Korzeniewska, P. J. Franaszczuk, C. M. Crainiceanu, R. Kuś, and N. E. Crone, "Dynamics of large-scale cortical interactions at high gamma frequencies during word production: event related causality (ERC) analysis of human electrocorticography (ECoG)," *NeuroImage*, vol. 56, no. 4, pp. 2218–2237, 2011.
- [20] E. Başar, "The theory of the whole-brain-work," *International Journal of Psychophysiology*, vol. 60, no. 2, pp. 133–138, 2006.
- [21] Y. Sun, Y. Li, Y. Zhu, X. Chen, and S. Tong, "Electroencephalographic differences between depressed and control subjects: an aspect of interdependence analysis," *Brain Research Bulletin*, vol. 76, no. 6, pp. 559–564, 2008.
- [22] D. W. Gow and J. A. Segawa, "Articulatory mediation of speech perception: a causal analysis of multi-modal imaging data," *Cognition*, vol. 110, no. 2, pp. 222–236, 2009.
- [23] A. Sinai and A. Korzeniewska, "Event-related dynamics of brain oscillations," *Brain*, vol. 159, no. 6, pp. 275–295, 2006.
- [24] C. Zhu, X. Guo, Z. Jin et al., "Influences of brain development and ageing on cortical interactive networks," *Clinical Neurophysiology*, vol. 122, no. 2, pp. 278–283, 2011.
- [25] P. Rana, *Seizure Detection and Analysis Using Phase-Slope Index*, University Of Wisconsin, 2011.
- [26] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, 1974.
- [27] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461–464, 1978.
- [28] C. S. Wallace and D. M. Boulton, "An information measure for classification," *The Computer Journal*, vol. 11, no. 2, pp. 185–194, 1968.
- [29] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problemy Peredachi Informatsii*, vol. 9, no. 3, pp. 3–11, 1973.
- [30] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [31] E. J. Hannan and B. G. Quinn, "The determination of the order of an autoregression," *Journal of the Royal Statistical Society*, vol. 41, no. 2, pp. 190–195, 1979.
- [32] E. I. George, "The variable selection problem," *Journal of the American Statistical Association*, vol. 95, no. 452, pp. 1304–1308, 2000.
- [33] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. van der Linde, "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society*, vol. 64, no. 4, pp. 583–639, 2002.
- [34] G. Claeskens and N. L. Hjort, "The focused information criterion," *Journal of the American Statistical Association*, vol. 98, no. 464, pp. 900–916, 2003.
- [35] O. I. Shittu and M. J. Asemota, "Comparison of criteria for estimating the order of autoregressive process: a Monte Carlo approach," *European Journal of Scientific Research*, vol. 30, no. 3, pp. 409–416, 2009.
- [36] M. Eichler, "On the evaluation of information flow in multivariate systems by the directed transfer function," *Biological Cybernetics*, vol. 94, no. 6, pp. 469–482, 2006.
- [37] B. Gourévitch, R. L. Bouquin-Jeannès, and G. Faucon, "Linear and nonlinear causality between signals: methods, examples and neurophysiological applications," *Biological Cybernetics*, vol. 95, no. 4, pp. 349–369, 2006.
- [38] L. A. Baccalá, K. Sameshima, and D. Y. Takahashi, "Generalized partial directed coherence," in *Proceedings of the 15th International Conference on Digital Signal Processing (DSP '07)*, no. 1–4, pp. 163–166, July 2007.
- [39] L. Faes and G. Nollo, "Extended causal modeling to assess partial directed coherence in multiple time series with significant instantaneous interactions," *Biological Cybernetics*, vol. 103, no. 5, pp. 387–400, 2010.
- [40] J. F. Geweke, "Measures of conditional linear dependence and feedback between time series," *Journal of the American Statistical Association*, vol. 79, no. 388, pp. 907–915, 1984.

- [41] S. Bressler, C. Richter, and Y. Chen, "Cortical functional network organization from autoregressive modeling of local field potential oscillations," *Statistics in Medicine*, vol. 26, no. 21, pp. 3875–3885, 2007.
- [42] S. L. Bressler and A. K. Seth, "Wiener-granger causality: a well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, 2011.
- [43] A. Korzeniewska, M. Mańczakb, M. Kamiński, K. J. Blinowska, and S. Kasicki, "Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method," *Journal of Neuroscience Methods*, vol. 125, no. 1-2, pp. 195–207, 2003.
- [44] P. J. Franaszczuk, K. J. Blinowska, and M. Kowalczyk, "Biological cybernetics in the study of electrical brain activity," *Biological Cybernetics*, vol. 247, pp. 239–247, 1985.
- [45] W. Hesse, E. Möller, M. Arnold, and B. Schack, "The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies," *Journal of Neuroscience Methods*, vol. 124, no. 1, pp. 27–44, 2003.
- [46] A. K. Seth, "Causal connectivity of evolved neural networks during behavior," *Computation in Neural Systems*, vol. 16, no. 1, pp. 35–54, 2005.
- [47] J. R. Sato, E. A. Junior, D. Y. Takahashi, M. de Maria Felix, M. J. Brammer, and P. A. Morettin, "A method to produce evolving functional connectivity maps during the course of an fMRI experiment using wavelet-based time-varying Granger causality," *NeuroImage*, vol. 31, no. 1, pp. 187–196, 2006.
- [48] A. Tang, M. T. Sutherland, P. Sun et al., "Top-down versus bottom-up processing in the human brain: distinct directional influences revealed by integrating SOBI and granger causality," *Independent Component Analysis and Signal Separation*, vol. 4666, pp. 802–809, 2007.
- [49] D. W. Gow, J. A. Segawa, S. P. Ahlfors, and F.-H. Lin, "Lexical influences on speech perception: a Granger causality analysis of MEG and EEG source estimates," *NeuroImage*, vol. 43, no. 3, pp. 614–623, 2008.
- [50] M. X. Cohen et al., "Top-down-directed synchrony from medial frontal cortex to nucleus accumbens during reward anticipation," *Human Brain Mapping*, vol. 33, no. 1, pp. 246–252, 2012.
- [51] J. Szaflarski, M. DiFrancesco, T. Hirschauer et al., "Cortical and subcortical contributions to absence seizure onset examined with EEG/fMRI," *Epilepsy and Behavior*, vol. 18, no. 4, pp. 404–413, 2010.
- [52] A. Özkaya and M. Kortürek, "Estimating short-run and long-run interaction mechanisms in interictal state," *Journal of Computational Neuroscience*, vol. 28, no. 2, pp. 177–192, 2010.
- [53] C. Ligges, M. Ungureanu, M. Ligges, and B. Blanz, "Understanding the time variant connectivity of the language network in developmental dyslexia: new insights using Granger causality," *Journal of Neural Transmission*, vol. 117, no. 4, pp. 529–543, 2010.
- [54] L. Pollonini, S. Pophale, N. Situ et al., "Information communication networks in severe traumatic brain injury," *Brain Topography*, vol. 23, no. 2, pp. 221–226, 2010.
- [55] E. Sitnikova, "Thalamo-cortical mechanisms of sleep spindles and spike-wave discharges in rat model of absence epilepsy (a review)," *Epilepsy Research*, vol. 89, no. 1, pp. 17–26, 2010.
- [56] K. L. Anderson, R. Rajagovindan, G. A. Ghacibeh, K. J. Meador, and M. Ding, "Theta oscillations mediate interaction between prefrontal cortex and medial temporal lobe in human memory," *Cerebral Cortex*, vol. 20, no. 7, pp. 1604–1612, 2010.
- [57] L. Astolfi, J. Toppi, F. De Vico Fallani et al., "Neuroelectrical hyperscanning measures simultaneous brain activity in humans," *Brain Topography*, vol. 23, no. 3, pp. 243–256, 2010.
- [58] D. W. Gow, C. J. Keller, E. Eskandar, N. Meng, and S. S. Cash, "Parallel versus serial processing dependencies in the perisylvian speech network: a Granger analysis of intracranial EEG data," *Brain and Language*, vol. 110, no. 1, pp. 43–48, 2009.
- [59] P. J. Franaszczuk, G. K. Bergey, and M. J. Kaminski, "Analysis of mesial temporal seizure onset and propagation using the directed transfer function method," *Electroencephalography and Clinical Neurophysiology*, vol. 91, no. 6, pp. 413–427, 1994.
- [60] P. J. Franaszczuk and G. K. Bergey, "Application of the directed transfer function method to mesial and lateral onset temporal lobe seizures," *Brain Topography*, vol. 11, no. 1, pp. 13–21, 1998.
- [61] L. Astolfi, F. Cincotti, D. Mattia et al., "Assessing cortical functional connectivity by linear inverse estimation and directed transfer function: simulations and application to real data," *Clinical Neurophysiology*, vol. 116, no. 4, pp. 920–932, 2005.
- [62] F. Babiloni, F. Cincotti, C. Babiloni et al., "Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function," *NeuroImage*, vol. 24, no. 1, pp. 118–131, 2005.
- [63] R. Kuś, J. S. Ginter, and K. J. Blinowska, "Propagation of EEG activity during finger movement and its imagination," *Acta Neurobiologiae Experimentalis*, vol. 66, no. 3, pp. 195–206, 2006.

- [64] L. Ding, G. A. Worrell, T. D. Lagerlund, and B. He, "Ictal source analysis: localization and imaging of causal interactions in humans," *NeuroImage*, vol. 34, no. 2, pp. 575–586, 2007.
- [65] C. Wilke, W. van Drongelen, M. Kohrman, and B. He, "Neocortical seizure foci localization by means of a directed transfer function method," *Epilepsia*, vol. 51, no. 4, pp. 564–572, 2010.
- [66] J. S. Kim, C. H. Im, Y. J. Jung, E. Y. Kim, S. K. Lee, and C. K. Chung, "Localization and propagation analysis of ictal source rhythm by electrocorticography," *NeuroImage*, vol. 52, no. 4, pp. 1279–1288, 2010.
- [67] C. J. Honey and O. Sporns, "Dynamical consequences of lesions in cortical networks," *Human Brain Mapping*, vol. 29, no. 7, pp. 802–809, 2008.
- [68] C. J. Honey, R. Kötter, M. Breakspear, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 24, pp. 10240–10245, 2007.
- [69] M. Lungarella and O. Sporns, "Mapping information flow in sensorimotor networks," *PLoS Computational Biology*, vol. 2, no. 10, pp. 1301–1312, 2006.
- [70] L. Barnett and A. Barrett, "Granger causality and transfer entropy Are equivalent for gaussian variables," *Physical Review Letters*, vol. 103, no. 23, pp. 1–10, 2009.
- [71] G. Boris and J. Eggermont, "Evaluating information transfer between auditory cortical neurons," *Journal of Neurophysiology*, vol. 97, no. 3, pp. 2533–2543, 2007.
- [72] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, "Identifying true brain interaction from EEG data using the imaginary part of coherency," *Clinical Neurophysiology*, vol. 115, no. 10, pp. 2292–2307, 2004.
- [73] E. Baek and W. A. Brock, "A nonparametric test for independence of a multivariate time series," *Statistica Sinica*, vol. 2, no. 1, pp. 137–156, 1992.
- [74] C. Hiemstra and J. D. Jones, "Testing for linear and nonlinear granger causality in the stock price-volume relation," *Journal of Finance*, vol. 49, no. 5, pp. 1639–1664, 1994.
- [75] P. Grassberger, "Estimation of the Kolmogorov entropy from a chaotic signal," *Physical Review A: General Physics*, vol. 28, no. 4, pp. 2591–2593, 1983.
- [76] C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, "Choice of multivariate autoregressive model order affecting real network functional connectivity estimate," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 436–448, 2009.
- [77] J. D. Hamilton, *Time Series Analysis*, vol. 48, no. 1, Princeton University Press, Princeton, NJ, USA, 1994.
- [78] D. A. Dickey and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.
- [79] J. D. Sargan and A. Bhargava, "Testing residuals from least squares regression for being generated by the Gaussian random walk," *Econometrica: Journal of the Econometric Society*, vol. 51, no. 1, pp. 153–174, 1983.
- [80] D. Marinazzo, W. Liao, H. Chen, and S. Stramaglia, "Nonlinear connectivity by Granger causality," *NeuroImage*, vol. 58, no. 2, pp. 330–338, 2011.
- [81] A. Schlögl, "A comparison of multivariate autoregressive estimators," *Signal Processing*, vol. 86, no. 9, pp. 2426–2429, 2006.