INFERRING CAUSAL CONNECTIVITY IN EPILEPTOGENIC ZONE USING DIRECTED INFORMATION

Rakesh Malladi[⋆] *Giridhar P Kalamangalam*[†] *Nitin Tandon*[‡] *Behnaam Aazhang*[⋆]

[⋆] Department of Electrical and Computer Engineering, Rice University, Houston, USA
[†] Department of Neurology, University of Texas Health Sciences Center, Houston, USA
[‡] Department of Neurosurgery, University of Texas Health Sciences Center, Houston, USA

ABSTRACT

Directed information, an information theoretic quantity, is developed in this paper to infer the causal connectivity from electrocorticography (ECoG) recordings of an epileptic patient. The causal connectivity can be used to infer the optimal electrodes for electrical stimulation based treatments of epilepsy. A parametric estimator for directed information between two ECoG signals is also proposed. The estimator estimates entropy and causally conditioned entropy and their difference is the estimate of DI. The estimator is then applied to ECoG data recorded from the electrodes in the epileptogenic zone (EZ) in two patients with focal epilepsy to learn the changes in causal connectivity during seizures.

***Index Terms*—** Directed information, epilepsy, causal connectivity, ECoG, entropy

1. INTRODUCTION

Epilepsy is a common neurological disease affecting about three million patients in the United States alone. Buoyed by the success of electrical stimulation in treating movement disorders like Parkinson's disease, neuromodulation via electrical stimulation is considered a promising approach to treat epilepsy in patients where current treatment options are not effective [1]. Learning the location of optimal spatial locations for stimulation is a major challenge in this endeavour [2, 3]. Our conjecture is that these locations can be inferred from the changes in causal connectivity during seizures. This paper focusses on developing novel techniques to learn the causal connectivity in the epileptogenic zone (EZ). The EZ is defined as the minimum area of the brain that must be resected during surgery to achieve seizure freedom [4].

We develop directed information (DI), an information theoretic quantity, to infer the causal connectivity between the discrete-time, continuous-valued signals recorded at different electrocorticographic (ECoG) electrodes implanted in the brain of an epileptic patient in this paper. Directed information was first introduced in [5] and was further developed for discrete-time, discrete-valued time-series in [6, 7] and was also used to quantify the capacity of wireless channels with feedback [8]. DI between two time-series is the amount of uncertainty in one time-series causally explained by the other time-series [9]. DI is also successfully used in many other applications - gene networks, portfolio theory and hypothesis testing, to name a few [10–12].

Causal connectivity is inferred predominantly using techniques based on Granger causality [13–15] and more recently using techniques based on transfer entropy [16] and dynamic causal modeling [17, 18]. The main contribution of this paper is developing a novel quantitative technique to infer the causal connectivity from ECoG by broadening the definition of directed information to the class of discrete-time, continuous-valued processes like ECoG recordings. This ensures that DI is a very broad concept applicable to a large class of electrophysiological measurements from the brain - from spike trains [9] to ECoG data. This is the main adavantage of DI over other techniques based on Granger causality [13, 16, 19, 20] to estimate causal connectivity. In addition, we propose a parametric estimator to estimate the DI from one ECoG signal to another. The proposed estimator first estimates entropy and causally conditioned entropy from ECoG signals and their difference is the estimate of directed information. The estimator is validated on simulated time-series. The proposed DI estimator is applied to ECoG data to infer causal connectivity in the EZ of two patients with focal epilepsy. The results of our analysis show that optimal spatial locations for electrical stimulation can be identified in two steps - first use entropy estimates to select a small group of electrodes and then use estimated DI between them to isolate the spatial locations that are drivers of epileptic activity.

The outline of the rest of the paper is as follows. Section 2 defines DI and proposes a parametric estimator for DI between two time-series. The performance of the proposed estimator is characterized in section 3. In section 4, the causal connectivity is inferred using DI from ECoG recordings. Concluding remarks are given in section 5.

This work is funded in part by a NSF grant and Texas Instruments. The authors can be reached at {Rakesh.Malladi, aaz}@rice.edu and {Giridhar.P.Kalamangalam, Nitin.Tandon}@uth.tmc.edu.

2. DIRECTED INFORMATION

Consider the recordings from a ECoG electrode array with multiple channels. The causal influence between two channels is quantified by directed information between them. In this paper, we solve the problem of estimating DI between all pairs of channels to learn the causal connectivity between them. Let $\mathbf{X}^N = (x_1, x_2, \cdots, x_N)^{\mathrm{T}}$ and $\mathbf{Y}^N = (y_1, y_2, \cdots, y_N)^{\mathrm{T}}$ denote recordings from two electrodes (or channels) of the ECoG array. Here N is the number of samples recorded from each electrode. The ECoG data used in this paper is collected from 154 electrodes at a sampling rate of 1 KHz. For notational convenience the variables with non-positive subscripts are empty sets and the subscripts are not shown when equal to 1. The following sections define DI and provide an estimator of DI between two ECoG signals.

2.1. DI definition

The DI from N samples of time-series $\mathbf{X}$ to N samples of time-series $\mathbf{Y}$, denoted by $I\left(\mathbf{X}^N \to \mathbf{Y}^N\right)$, is defined as

$$I\left(\mathbf{X}^N \to \mathbf{Y}^N\right) = h\left(\mathbf{Y}^N\right) - h\left(\mathbf{Y}^N \| \mathbf{X}^N\right), \qquad (1)$$

where $h\left(\mathbf{Y}^N\right)$ and $h\left(\mathbf{Y}^N \| \mathbf{X}^N\right)$ are respectively the differential entropy [21] of the N dimensional continuous random vector $\mathbf{Y}^N$ and the causally conditioned differential entropy of $\mathbf{Y}^N$, causally conditioned on $\mathbf{X}^N$. The causally conditioned differential entropy is defined as

$$h\left(\mathbf{Y}^N \| \mathbf{X}^N\right) = \sum_{n=1}^{N} h\left(y_n | \mathbf{Y}^{n-1}, \mathbf{X}^n\right). \qquad (2)$$

The definitions of DI and causally conditioned differential entropy in (1) and (2) are obtained by broadening the definitions of the same quantities from discrete-time, discrete-valued random processes [6, 7] to discrete-time, continuous-valued processes. One of the main differences between discrete-valued and continuous-valued random processes is that the entropy of a discrete-valued process is always non-negative, whereas the differential entropy of a continuous-valued process can be negative [21]. However DI is always non-negative since conditioning cannot increase differential entropy [21], i.e., $0 \leq I\left(\mathbf{X}^N \to \mathbf{Y}^N\right)$. Directed information, unlike other causality metrics like Granger causality, can be interpreted as the number of bits of uncertainty in one process that is causally explained away by the other process. If $I\left(\mathbf{X}^N \to \mathbf{Y}^N\right) = 0$, then there is no causal influence from $\mathbf{X}$ to $\mathbf{Y}$. Now, the DI between the time-series $\mathbf{X}$ and $\mathbf{Y}$ is defined as

$$\begin{aligned} I\big(\mathbf{X} \to \mathbf{Y}\big) &= \lim_{N\to\infty} \tfrac{1}{N} I\big(\mathbf{X}^N \to \mathbf{Y}^N\big) \\ &= \lim_{N\to\infty} \tfrac{1}{N} h\big(\mathbf{Y}^N\big) - \lim_{N\to\infty} \tfrac{1}{N} h\big(\mathbf{Y}^N \| \mathbf{X}^N\big) \\ &= h\big(\mathbf{Y}\big) - h\big(\mathbf{Y} \| \mathbf{X}\big), \end{aligned} \qquad (3)$$

provided the limits exist. In (3), $h\left(\mathbf{Y}\right)$ and $h\left(\mathbf{Y} \| \mathbf{X}\right)$ are respectively the differential entropy of $\mathbf{Y}$ and the causally conditioned differential entropy of $\mathbf{Y}$ given $\mathbf{X}$. The DI from $\mathbf{Y}$ to $\mathbf{X}$ is also similarly defined. Note that the DI is not a symmetric metric, i.e., $I\left(\mathbf{X}^N \to \mathbf{Y}^N\right) \neq I\left(\mathbf{Y}^N \to \mathbf{X}^N\right)$.

2.2. DI Estimator

Let us focus on estimating $I\left(\mathbf{X} \to \mathbf{Y}\right)$. The estimator assumes the time-series $\mathbf{X}$ and $\mathbf{Y}$ are stationary, ergodic and Markovian in the observed time-window. These are reasonable assumptions to model ECoG data [13, 15, 22]. An implicit assumption in the problem of estimating the causal connectivity is that the true causal connectivity does not vary in the ECoG recording window, which is mathematically captured by stationarity. Ergodicity is required to ensure that the estimates from long enough recording windows converge to the true value. Finally, the Markovian assumption only formalizes the duration of the past activity that influences the current activity at different electrodes and does not explicitly model the dependence. Let the current sample of the time-series $\mathbf{Y}$ depend on the past J and past K samples of the time-series $\mathbf{Y}$ and $\mathbf{X}$ respectively. Also $\mathrm{P}\left(\mathbf{Y}^N \| \mathbf{X}^N\right) = \prod_{n=1}^{N} \log \mathrm{P}\big(y_n | \mathbf{Y}_{n-J}^{n-1}, \mathbf{X}_{n-K+1}^{n}\big)$ denotes the causal likelihood of the N samples of $\mathbf{Y}$, conditioned on the causal past of $\mathbf{Y}$ and $\mathbf{X}$. In addition, let us focus on the distributions for which the differential entropy exists. Under these assumptions, an estimator for the causally conditioned entropy of $\mathbf{Y}$ given $\mathbf{X}$ is:

$$\hat{h}\left(\mathbf{Y} \| \mathbf{X}\right) = \tfrac{1}{N} \sum_{n=1}^{N} \left(- \log \mathrm{P}\left(y_n | \mathbf{Y}_{n-J}^{n-1}, \mathbf{X}_{n-(K-1)}^{n}\right)\right). \qquad (4)$$

An estimator for $\hat{h}\left(\mathbf{Y}\right)$ can be easily obtained by setting $K = 0$ in (4). The estimators proposed here are inspired by the ideas from [9, 21]. We propose the following estimator for DI from time-series $\mathbf{X}$ to $\mathbf{Y}$:

$$\hat{I}\left(\mathbf{X} \to \mathbf{Y}\right) = \hat{h}\left(\mathbf{Y}\right) - \hat{h}\left(\mathbf{Y} \| \mathbf{X}\right). \qquad (5)$$

The performance of this estimator is tested on simulated time-series in the following section and applied to real ECoG data in the subsequent section.

3. PERFORMANCE ON SIMULATED DATA

The accuracy of DI estimator proposed in the earlier section is characterized using data simulated from a multivariate autoregressive (MVAR) model. Consider N samples of two independent time-series $\mathbf{X}^N$ and $\mathbf{Z}^N$, where the x_n, z_n are i.i.d. zero mean Gaussian with variances σ_x^2 and σ_z^2 respectively. Without loss of generality, let $\sigma_x^2 = 1$. The time-series $\mathbf{Y}$ is generated according to the following MVAR model:

$$y_n = \beta_1 x_n + \beta_2 x_{n-1} + z_n, n = 1, 2, \cdots, N. \qquad (6)$$

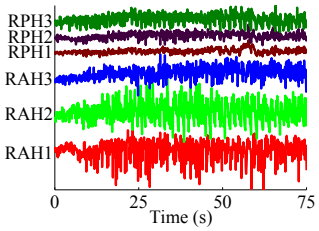


Fig. 1. A 75 second snapshot of ECoG recordings from six channels of patient P1.

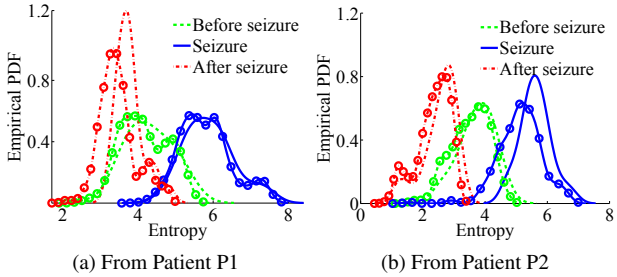


Fig. 2. Empirical PDFs of entropy (curves without circles) and causal conditional entropy (curves with circles) estimates from three segments - before seizure, during seizure and after seizure.

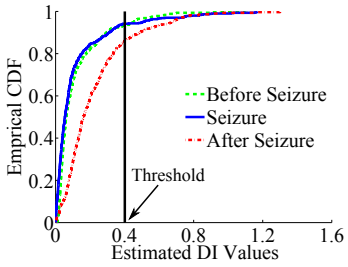


Fig. 3. CDFs of DI estimates of patient P1 from a segment before, during and after a seizure.

Let us first look at the true value of DI in both directions for the model given by (6). When $\beta_1 = 1, \beta_2 = 0$, (6) reduces to $y_n = x_n + z_n$, and it is obvious that both $\mathbf{X}$ and $\mathbf{Y}$ have equal causal information about each other. It is easy to see that $I(\mathbf{X} \to \mathbf{Y}) = I(\mathbf{Y} \to \mathbf{X}) = I(\mathbf{X}; \mathbf{Y}) = C$, where $I(\mathbf{X}; \mathbf{Y})$ is the mutual information between $\mathbf{X}$ and $\mathbf{Y}$ and $C = \frac{1}{2}\log\left(1 + \frac{1}{\sigma_z^2}\right)$. The other extreme case of $\beta_1 = 0, \beta_2 = 1$ is also interesting and (6) reduces to $y_n = x_{n-1} + z_n$. In this case, $\mathbf{X}$ has causal information about $\mathbf{Y}$, while $\mathbf{Y}$ has no causal information about $\mathbf{X}$. More precisely, $I(\mathbf{X} \to \mathbf{Y}) = I(\mathbf{X}; \mathbf{Y}) = C$ and $I(\mathbf{Y} \to \mathbf{X}) = 0$. For the remaining case of non-zero β_1, β_2, the analytical expressions for DI are:

$$I(\mathbf{X} \to \mathbf{Y}) = \frac{1}{2}\log\left(\frac{|\beta_1\beta_2|}{\sigma_z^2}\right) + \frac{1}{2}\cosh^{-1}\left(\frac{\beta_1^2 + \beta_2^2 + \sigma_z^2}{2|\beta_1\beta_2|}\right),$$
$$I(\mathbf{Y} \to \mathbf{X}) = \frac{1}{2}\log\left(1 + \frac{\beta_1^2}{\sigma_z^2}\right). \quad (7)$$

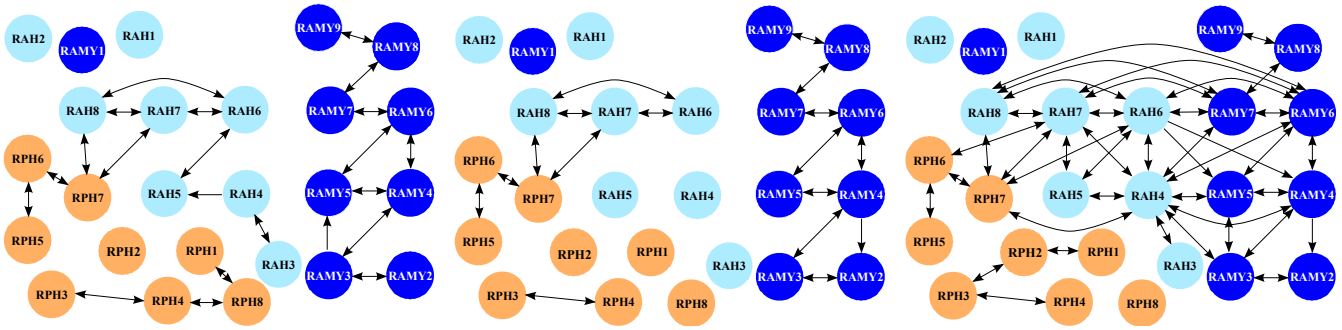
The derivation of (7) uses tridiagonal matrix determinant from [23] and is omitted due to limited space. Note that from (7), DI from $\mathbf{Y}$ to $\mathbf{X}$ does not depend on β_2. This is because the uncertainty in the current sample of $\mathbf{X}$ does not depend on β_2, when causally conditioned on the past of $\mathbf{X}$ and $\mathbf{Y}$.

Now using the estimator in (5), DI is estimated from 10^5 samples of the two time-series $\mathbf{X}$ and $\mathbf{Y}$, generated with unit variances from (6). The parameters, $\beta_1, \beta_2, \sigma_z^2$, are estimated from data. Let us focus on estimating $\hat{I}(\mathbf{X} \to \mathbf{Y})$. Estimating $\hat{I}(\mathbf{X} \to \mathbf{Y})$ using (5) requires estimating the causal conditional entropy of time-series $\mathbf{Y}$ given $\mathbf{X}$ and the entropy of $\mathbf{Y}$. The causal conditional entropy can be estimated using (4) and it requires estimating the distribution of samples on $\mathbf{Y}$, conditioned on the past of $\mathbf{Y}$ and $\mathbf{X}$. This distribution is Gaussian from (6). The parameters of this distribution $\beta_1, \beta_2, \sigma_z^2$, along with the model orders are estimated using maximum-likelihood (ML) estimation with minimum description length (MDL) penalty [9, 24]. $\hat{h}(\mathbf{Y}\|\mathbf{X})$ is obtained by substituting the parameters in (4). Entropy of $\mathbf{Y}$ is estimated using similar procedure. The resultant estimates are substituted into (5) to find $\hat{I}(\mathbf{X} \to \mathbf{Y})$. $\hat{I}(\mathbf{Y} \to \mathbf{X})$ can be estimated similarly.

The accuracy of the DI estimator is evaluated by the normalized root mean-square error (NRMSE) between the true value from (7) and the estimate from (5) for 100 different values of the ordered pair (β_1, β_2), where $0 \leq \beta_1, \beta_2 \leq 1$. The NRMSE between the estimate and the true value is 0.29% for DI from $\mathbf{X}$ to $\mathbf{Y}$. The NRMSE after a similar analysis for DI from $\mathbf{Y}$ to $\mathbf{X}$ is 0.61%. This validates the accuracy of the proposed DI estimator.

4. CAUSAL CONNECTIVITY FROM ECOG DATA

Directed information is now applied to infer the causal connectivity from the ECoG recordings of two patients with focal epilepsy under the treatment of our coauthors. We analyzed multiple recordings from both patients, recorded when the patients are awake, asleep and are having a seizure. Here we present the results from a connectivity analysis restricted to the 25 electrodes in epileptogenic zone (EZ) in these patients, due to pragmatic space constraints. The journal version of this article to be submitted soon will have more results. The epileptogenic zone (EZ) of the patient P1 was localized to the right amygdala (RAMY), right anterior and posterior hippocampus (RAH and RPH respectively), based on typical clinical criterion for localizing seizure onsets. The EZ of the patient P2 was localized to the left anterior and posterior hippocampus (LAH and LPH respectively). Each seizure record (approximately 10 minutes long) is divided into multiple overlapping segments of varying duration (between 30 and 100 seconds). A snapshot of ECoG recordings in six channels from patient P1 is plotted in Fig. 1. The ECoG signals in each segment are modeled using MVAR processes. The parameters of the MVAR model, namely, the model orders, auto-regressive parameters and the noise variances are estimated using ML with MDL penalty, as described in section 3. After some trial and error, we restricted the search space for the optimal MVAR process model orders to $[75, 125]$ to reduce the complexity of the parameter estimation. A down sampling factor of 10 (i.e. every 10^{th} sample is used from the past) is used to model the past dependence. Once all the parameters are estimated, DI from one ECoG



(a) From a segment before the seizure (b) From a segment during the seizure (c) From a segment after the seizure

Fig. 4. Causal connectivity in the epileptogenic zone of patient P1 from three segments - one before seizure, one during seizure and one after seizure. Twenty five electrodes, identified by the label inside each circle, are chosen in the epileptogenic zone.

channel to another is the difference between the estimates of the entropy and causal conditional entropy. The estimation algorithm is described in section 2.2. This process is repeated for every pair of ECoG channels to learn the causal connectivity graphs during seizures in both these patients. The properties of the entropy and DI estimates from our analysis of six seizure records, three from each patient, are presented in the remainder of this section.

To begin with, let us focus on the entropy estimates of 25 channels from a segment. The distribution of entropy values of the ECoG channels changes within a seizure record. Fig. 2a, 2b plot the empirical probability density function (PDF) of the estimates of entropy and the causal conditional entropy in three different segments within a seizure record from patient P1 and P2 respectively. The dashed line, solid line, dash-dot lines in both these figures correspond to the PDF estimates from a segment before seizure, during seizure and after seizure respectively. The curves with circles correspond to the PDF of the causal conditional entropy and without circles correspond to entropy. It is clear from both these figures that entropy increases during a seizure and falls back down after the seizure. This is expected, since the abnormal excessive activity in the brain during seizures leads to increase in variance of the signals leading to more entropy. Also the peak of the entropy distribution after a seizure is slightly to the left of entropy distribution before seizures, indicating that activity returns to almost normal levels immediately after a seizure. Estimated causal conditional entropy between 600 directed links present between 25 ECoG electrodes is used to calculate its empirical PDF. As expected, the PDFs of the causal conditional entropy are left-shifted with respect to the entropy PDF from the same segment, since causal conditional entropy is less than entropy. Directed information is estimated by subtracting the causal conditional entropy from entropy.

Fig. 3 plots the CDF of the estimated DI in three segments, before a seizure, during a seizure and after a seizure, recorded from 25 electrodes in the EZ of patient P1. The inferred causal connectivity graphs in these three segments are plotted in Fig. 4. The seizure record from which the three segments in Fig. 3, 4 are selected is different from the one in Fig. 2a. Each electrode is represented by a circle with its name in these figures. Due to limited space, we are not presenting the results from a similar analysis done for other seizure recordings of patients P1 and P2. Only those connections whose DI estimate exceeds a threshold are depicted in the connectivity graphs in Fig. 4. The threshold is patient-specific and is equal to the DI value corresponding to the cumulative distribution function (CDF) of 0.9 over all the segments analyzed. For the connectivity graphs in Fig. 4, the threshold is 0.4 and it is depicted by a vertical black line in Fig. 3. From Fig. 3, it is clear that the number of DI estimates exceeding the threshold are more in the segment after seizure when compared with the other two segments. This explains the increase in number of causal connections after a seizure when compared with segments before and during a seizure observed in Fig. 4. The analysis presented here can be further extended to learn the optimal spatial locations for electrical stimulation. Increase in entropy during seizures can be used as feature to first select a subset of electrodes from 154 electrodes. The causal connectivity between these subset of electrodes can be further analyzed to narrow down and identify the electrodes responsible for sustaining the seizure network.

5. CONCLUSIONS

In this paper, we develop directed information to infer and present the causal connectivity in the EZ from the ECoG recordings of two patients with focal epilepsy. Our analysis showed that entropy increases during seizure and this can be used as a feature to select a small group of electrodes. Causal connectivity between these electrodes can then be further analyzed to identify the electrodes driving the seizure activity. These electrodes could potentially be the optimal spatial locations for electrical stimulation. This should be confirmed by analyzing a larger patient cohort in future work.

6. REFERENCES

[1] Casey H Halpern, Uzma Samadani, Brian Litt, Jurg L Jaggi, and Gordon H Baltuch, "Deep brain stimulation for epilepsy," *Neurotherapeutics*, vol. 5, no. 1, 2008.

[2] "NIH fact sheet - epilepsy," 2010.

[3] Robert S Fisher, "Neurostimulation for epilepsy: do we know the best stimulation parameters?," *Epilepsy Currents*, vol. 11, no. 6, 2011.

[4] Hans O Luders, Imad Najm, Dileep Nair, Peter Widdess-Walsh, and William Bingman, "The epileptogenic zone: general principles," *Epileptic Disorders*, vol. 8, 2006.

[5] Hans Marko, "The bidirectional communication theory–a generalization of information theory," *IEEE Transactions on Communications*, vol. 21, no. 12, 1973.

[6] J Massey, "Causality, feedback and directed information," in *International Symposium on Information Theory Applications (ISITA)*, 1990.

[7] Gerhard Kramer, *Directed information for channels with feedback*, Ph.D. thesis, ETH Zürich, 1998.

[8] Sekhar Tatikonda and Sanjoy Mitter, "The capacity of channels with feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 1, 2009.

[9] Christopher J Quinn, Todd P Coleman, Negar Kiyavash, and Nicholas G Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *Journal of Computational Neuroscience*, vol. 30, no. 1, 2011.

[10] Arvind Rao, Alfred O Hero, David J States, and James Douglas Engel, "Inference of biologically relevant gene influence networks using the directed information criterion," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2006, vol. 2.

[11] Haim H Permuter, Young-Han Kim, and Tsachy Weissman, "Interpretations of directed information in portfolio theory, data compression, and hypothesis testing," *IEEE Transactions on Information Theory*, vol. 57, no. 6, 2011.

[12] Tsachy Weissman, Young-Han Kim, and Haim H Permuter, "Directed information, causal estimation, and communication in continuous time," *IEEE Transactions on Information Theory*, vol. 59, no. 3, 2013.

[13] Clive WJ Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, 1969.

[14] Mario Chávez, Jacques Martinerie, and Michel Le Van Quyen, "Statistical assessment of nonlinear causality: application to epileptic eeg signals," *Journal of Neuroscience Methods*, vol. 124, no. 2, pp. 113–128, 2003.

[15] Evgenia Sitnikova, Taras Dikanev, Dmitry Smirnov, Boris Bezruchko, and Gilles Van Luijtelaar, "Granger causality: cortico-thalamic interdependencies during absence seizures in wag/rij rats," *Journal of neuroscience methods*, vol. 170, no. 2, pp. 245–254, 2008.

[16] Thomas Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, 2000.

[17] Karl J Friston, Lee Harrison, and Will Penny, "Dynamic causal modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, 2003.

[18] Stefan J Kiebel, Marta I Garrido, Rosalyn Moran, Chun-Chuan Chen, and Karl J Friston, "Dynamic causal modeling for eeg and meg," *Human brain mapping*, vol. 30, no. 6, pp. 1866–1876, 2009.

[19] Meng-Hung Wu, Richard E Frye, and George Zouridakis, "A comparison of multivariate causality based measures of effective connectivity," *Computers in Biology and Medicine*, vol. 41, no. 12, 2011.

[20] Lionel Barnett and Anil K Seth, "The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference," *Journal of Neuroscience Methods*, vol. 223, 2014.

[21] TM Cover and JA Thomas, *Elements of information theory*, Wiley, 2006.

[22] Maciej Kamiński, Mingzhou Ding, Wilson A Truccolo, and Steven L Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biological Cybernetics*, vol. 85, no. 2, 2001.

[23] GY Hu and RF O'Connell, "Analytical inversion of symmetric tridiagonal matrices," *Journal of Physics A: Mathematical and General*, vol. 29, no. 7, 1996.

[24] Peter D Grünwald, *The minimum description length principle*, MIT press, 2007.