Tracking Epileptic Seizure Activity via Information Theoretic Graphs

Yonathan Murin, Jeremy Kim, and Andrea Goldsmith
Department of Electrical Engineering, Stanford University, Stanford CA, USA

***Abstract*—This work introduces an algorithm for localization of the seizure onset zone (SOZ) of epileptic patients based on electrocorticography (ECoG) recordings. The algorithm represents the set of electrodes using a directed graph in which nodes correspond to recording electrodes, while the edge weights are the pair-wise *causal influence*. This causal influence is quantified by estimating the pair-wise directed information functional. The SOZ is inferred from the graph by identifying nodes which act as *sources* of causal influence. Results based on several datasets show a close match between the inferred SOZ and the SOZ estimated by expert neurologists.**

I. Introduction

Epilepsy is one of the most common neurological disorders affecting about $0.5-1\%$ of the world population. It is characterized by recurrent unprovoked seizures, in which a sudden burst of uncontrolled electrical activity occurs within a group of neurons in the cerebral cortex, resulting in a variety of symptoms ranging from muscle stiffness to impaired consciousness [1]. Epileptic seizures can be divided into two groups, based on the location in the brain from which the seizures originate and how they propagate: Partial, or focal, seizures originate from a limited area in the brain, commonly referred to as the *seizure onset zone* (SOZ). In contrast, generalized epilepsy seizures begin with a widespread discharge that involves the entire brain. Despite the fact that the term "generalized seizure" is frequently used, [2] demonstrated that some epileptic seizures that appeared to be generalized actually had a focal onset. Moreover, according to [3], about 80% of the seizures are believed to be focal in nature. In this work we consider focal epilepsy and present an algorithm for SOZ localization, defined as determining the area in the brain where the focal seizure originates. A guiding hypothesis throughout this our work is that in focal seizures there is a singular focal point, from which the seizure originates.

The aim of epilepsy treatment is to suppress the seizures. In most cases, treatment with antiepileptic drugs is successful, yet, this treatment is not effective for about one third of the patients [1], and they are diagnosed with refractory epilepsy. The traditional treatment for refractory epilepsy is a surgical procedure, in which the SOZ is first located and then possibly removed (if it is not responsible for indispensable brain functions). The main tool for SOZ identification is invasive electroencephalogram, also known as electrocorticography (ECoG), in which grids or strips of electrodes are placed on the cortex. ECOG allow a direct measurement and recording of the brain's electrical activity (local field potentials). These recordings, together with video monitoring, are used by expert neurologists to approximate the electrodes associated with the area within which the SOZ lies. Another treatment approach for patients with refractory epilepsy is electrical stimulation [4], and in-particular closed-loop SOZ stimulation [5]. In both of these treatment options, *accurate localization of the SOZ is critical in the treatment of patients with refractory epilepsy*. The objective of the current work is to design an algorithm that localizes the SOZ based on the ECoG recordings. Such an automatic solution will increase the localization accuracy as well as save analysis time for the neurologists.

As reported in [6], focal seizures start in the SOZ and spread to surrounding areas in the brain. The algorithm proposed in this work builds upon this understanding: Just before a seizure starts, namely, in the pre-ictal phase, *signals recorded at electrodes close to the SOZ should have a relatively large causal influence on the rest of the recorded signals*. However, to fully quantify the statistical causal influence between two electrodes, one must evaluate this influence *when conditioning on the rest of the electrodes* [7]. For a large number of electrodes, this task is computationally demanding and requires a huge amount of data per each seizure. The algorithm we propose uses a practical approximation by considering the electrodes as nodes in a directed graph, where the edges weights are estimations of the *pair-wise causal influence*. Then the algorithm infers the SOZ based on the graph properties.

Detection and quantification of statistical causal influence among time-series is a long standing problem in signal processing [8]. Roughly speaking, causality measures can be divided into two groups: parametric and non-parametric, and can be estimated in the time or in the frequency domain. Parametric measures implicitly assume that the observed time-series follow a specific model, which makes estimation more efficient and simpler. On the other hand, a mismatch between the observed time-series and the assumed model usually leads to poor results. Examples for parametric causality measures are Granger causality in the time domain, and its counterpart in the frequency domain, direct transfer function (see [9] for a detailed review). As an accurate statistical model for ECoG recordings is not known, we propose to use the non-parametric causality measure of directed information (DI) [10], taken from the field of information theory [11]. To estimate the pairwise DI, we use a novel estimator based on the k-nearest-neighbor (k-NN) principle, which extends the best known mutual information estimator [12]. Our analysis indicates that this estimator is more accurate than other known non-

The work was partly supported by the NSF Center for Science of Information (CSoI) under grant NSF-CCF-0939370.

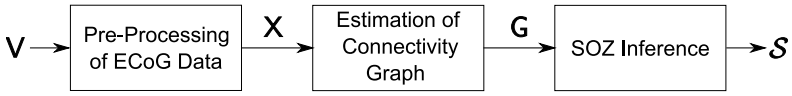


Fig. 1: High level description of the *per-block* processing. V is an $10 \cdot F_{\rm s} \times N$ matrix of the ECoG recordings. X is an $N_0 \times N, N_0 \leq 10 \cdot F_{\rm s}$ matrix, containing the inputs to the graph estimation. G is the estimated graph, and $\mathcal{S}$ is a set of electrodes corresponding to the SOZ.

parametric estimation approaches such as the k-NN estimator of [13], estimation of the causal conditional likelihood via kernel density estimation [14], or estimation via correlation integrals [15].

Numerus works studied the problem of SOZ localization based on ECoG recordings, see [1], [14] and references therein. Most of the algorithms proposed in these works follow a similar high-level structure: 1) Pre-processing of the ECoG recordings; 2) Estimation of a connectivity graph; and 3) Inference of the SOZ from the estimated graph. This structure is illustrated in Fig. 1. The algorithm proposed in the current work follows a similar approach, yet it implements each step differently. In contrast to most of the works mentioned in [1], our proposed algorithm uses DI to quantify the causal influence, thus avoiding the assumption of a parametric statistical model. Furthermore, this algorithm uses an improved estimator compared to the estimators proposed in [14] and [15]. This improvement requires fewer samples for estimation, thus, we can estimate DI from a time interval that is more stationary. Note that despite the recent advances in automated SOZ localization, the gold standard is still considered to be the localization performed by neurologists. Hence, to evaluate the performance of our proposed algorithm, we compare its inference to the inference made by expert neurologists and show that for three patients these two inferences match.

The rest of this paper is organized as follows: Section II formally defines the problem. Section III discusses the DI functional and proposes a method to estimate it. Section IV describes the proposed algorithm, and Section V presents the localization results.

***Notation*:** We denote random variables (RVs) by upper case letters, X, and their realizations with the corresponding lower case letters. We use the short-hand notation X_1^n to denote the sequence $\{X_1, X_2, \ldots, X_n\}$. We denote random processes using boldface letters, e.g., $\mathbf{X}$. Matrices are denoted by sans-serif font, e.g., V. We denote sets by calligraphic letters, e.g., $\mathcal{S}$, where $\mathcal{R}$ denotes the set of real numbers. Finally, $f_X(x)$ denotes the probability density function (PDF) of a continuous RV X on $\mathcal{R}$, and $\log(\cdot)$ denotes the natural basis logarithm.

II. Problem Definition

The input in the considered problem is a data-set, listed in the iEEG portal [16], corresponding to a patient with refractory epilepsy. Each data-set contains ECoG recordings, i.e., voltage traces, as well as annotations indicating which time intervals in the recordings correspond to seizures. The data-sets also include reports describing the spatial locations, on the cortex, of the electrodes, and comments by expert neurologists as to where the seizures originate from. An electrode that is highlighted in the comments is referred to as an *electrode of interest* (EOI). Note that the data-sets may contain recordings from several strips and grids. In such cases the algorithm analyzes the largest grid of electrodes.[1]

Based on the annotations in the data-sets, the algorithm uses ECoG signals from two types of 10 second intervals (blocks): Pre-ictal blocks and rest blocks. Pre-ictal blocks are recordings from a block right before a seizure. Intuitively, in pre-ictal blocks the seizure activity has not spread out across the brain yet, and therefore these blocks should give the clearest insights as to the SOZ location. Rest blocks are randomly sampled from intervals that exclude seizures, artifacts, and 10 second windows before and after seizures. In this time, the patient is either resting or awake and conscious. These blocks are used for significance testing.

Fig. 1 provides a general description of the processing applied for *each block*. Let the sampling rate used in recording the considered data-set be $F_{\rm s}$, and let the number of recorded electrodes be N. The input for the block processing is a $10{\cdot}F_{\rm s}{\times}N$ matrix V, in which the i^{th} column corresponds to the recordings from the i^{th} electrode. The algorithm applies pre-processing on the matrix V resulting in an $N_0{\times}N, N_0 \leq 10{\cdot}F_{\rm s}$ matrix X. We refer to this phase as pre-processing since it precedes the core of the algorithm, namely, estimating the directed graph. This estimation takes as input the matrix X and outputs a matrix G representing a directed graph. $[\mathsf{G}]_{i,j}$ is an estimation of the causal influence of the signal recorded in the i^{th} electrode on the signal recorded in the j^{th} electrode. Finally, the algorithm infers a set of electrodes $\mathcal{S} \in \{1, \ldots, N\}$ from G, and declares it to be the SOZ.

Before providing a detailed description of the proposed algorithm, in the next section we discuss the DI functional and show how it can be estimated from the ECoG recordings, thus, quantifying the pair-wise causal influence.

III. Quantifying the Pair-wise Causal Influence

A. Directed Information - Definitions and Background

We first recall the definitions of differential entropy and mutual information (MI) [11, Ch. 8]. Consider the RVs $X \in \mathcal{R}_x^d$ and $Y \in \mathcal{R}_y^d$, with marginal PDFs $f_X(x), f_Y(y)$ and joint PDF $f_{X,Y}(x,y)$. The differential entropy of X is defined as $h(X) \triangleq -\mathbb{E}\{\log f_X(x)\}$, where $h(Y)$ is defined similarly. Using these definitions, the MI between X and Y is given by $I(X;Y) = h(Y) - h(Y|X) = h(X) + h(Y) - h(X,Y)$. The mutual information between the *sequences* X_1^N and Y_1^N is similarly given by $I(X_1^N; Y_1^N)$.

[1] In most cases the largest grid was located over the suspected SOZ.

Let $\mathbf{X}$ and $\mathbf{Y}$ be arbitrary discrete-time continuous-amplitude random processes, and let $X^N \in \mathcal{R}^N$ and $Y^N \in \mathcal{R}^N$, be N-length sequences. The DI from X^N to Y^N is defined as [10, *eq.* (1)]:

$$I(X^N \to Y^N) \triangleq \sum_{i=1}^{n} I(X_1^i; Y_i | Y_1^{i-1}). \tag{1}$$

As stated above, DI aims at quantifying the *causal influence* of the sequence X^N on the sequence Y^N. The directed information *rate* between the processes $\mathbf{X}$ and $\mathbf{Y}$ is defined as [10, *eq.* (12)]:

$$I(\mathbf{X} \to \mathbf{Y}) \triangleq \lim_{N\to\infty} \frac{1}{N} I(X^N \to Y^N), \tag{2}$$

provided that this limit exists. We now make the following assumptions regarding the processes $\mathbf{X}$ and $\mathbf{Y}$.

A1) The random processes $\mathbf{X}$ and $\mathbf{Y}$ are assumed to be *stationary, ergodic, and Markovian of order* M in the observed sequences. The stationarity assumption implies that the statistics of the considered random processes is constant throughout the observed sequences. From a practical perspective, stationarity is required to ensure that the causal influence does not change over the observed sequences. Ergodicity is assumed to ensure that the observed sequences truly represent the underlying processes. Finally, the Markovity assumption is common in modeling *real-life systems which have finite memory*, in particular in neuroscience [14], [17], [18]. We formulate the assumption of Markovity of order M in the observed sequences via $f(y_i|Y_1^{i-1}) = f(y_i|Y_{i-M}^{i-1})$ and $f(y_i|Y_1^{i-1}, X_1^i) = f(y_i|Y_{i-M}^{i-1}, X_{i-M}^{i-1}), i > M$. Here, we use the simplifying assumption that the dependence of y_i on past samples of Y_1^{i-1} and X_1^i is of the same order M. Note that in the above Markovity assumption, we implicitly assume that given $(Y_{i-M}^{i-1}, X_{i-M}^{i-1})$, y_i is independent of X_i, which reflects a setting in which X_i and Y_i are simultaneously measured, as in the case of ECoG recording.

A2) The entropy of the first sample y_1 exists: $|H(Y_1)| < \infty$.

A3) The following holds: $|H(Y_{M+1}|Y_1^M, X_1^M)| < \infty$.

Under these assumptions, [14, Lemmas 3.1 and 3.2] imply that $I(\mathbf{X} \to \mathbf{Y})$ exists and is equal to:

$$I(\mathbf{X} \to \mathbf{Y}) = I(X_{i-M}^{i-1}; Y_i | Y_{i-M}^{i-1}), \quad i > M. \tag{3}$$

In view of (3), $I(\mathbf{X} \to \mathbf{Y})$ has the following interpretation: *Given the past of the sequence Y, how much does the past of the sequence X help in predicting the next sample of Y?*

Having defined DI and its interpretation, we now briefly describe our estimator of pair-wise DI. A detailed derivation of the proposed estimator is given in [19].

B. A New DI Estimator

Let $X_1^{N_0}$ and $Y_1^{N_0}$ be two (different) columns in the matrix X, namely, the ECoG recordings after pre-processing. In this subsection we *assume* that $X_1^{N_0}$ and $Y_1^{N_0}$ are stationary, ergodic, and obey the Markov property of order $M < N_0$. In the next section we discuss the relationship between these assumptions and the pre-processing phase. Our estimator extends the MI estimator proposed in [20]. Let $||z||_p, z \in \mathcal{R}^d$, denote the ℓ_p norm of z. We further define $\Gamma(\cdot)$ as Euler's gamma function and $\psi(\cdot)$ as the digamma function [21, Ch. 5]. Finally, we let $c_{d,p}$ denote the volume of the unit ℓ_p-ball in d dimensions given by $c_{d,p} = 2^d \frac{(\Gamma(1+\frac{1}{p}))^d}{\Gamma(1+\frac{1}{p})}$.

To simplify the notation, for $M < i \leq N_0$, we let $X_i^- \triangleq X_{i-M}^{i-1}$ denote the past of X_i, and $Y_i^- \triangleq Y_{i-M}^{i-1}$ denote the past of Y_i. Using this notation, (3) can be written as $I(\mathbf{X} \to \mathbf{Y}) = I(X^-; Y|Y^-)$. For a *fixed* k, let $\rho_{k,i,2}$ denote the $(2M+1)$-dimensional distance between *the tuple* (X_i^-, Y_i^-, Y_i) and its k-NN, calculated using the ℓ_2 norm. Further, for $i > M$, define $n_{Y^-,i,2} \triangleq \sum_{j=M+1, j\neq i}^{N_0} \mathbb{I}(||Y_i^- - Y_j^-||_2 \leq \rho_{k,i,2})$, where $\mathbb{I}(\cdot)$ is the indicator function. Note that $n_{Y^-,i,2}$ is the number of samples, in the (Y^-)-plane, which are within a distance $\rho_{k,i,2}$ from Y_i^-. $n_{(Y^-,Y),i,2}$ and $n_{(Y^-,X^-),i,2}$ are defined similarly. Using the above notations, the DI is estimated as follows:

$$\hat{I}(\mathbf{X}\to\mathbf{Y}) = \psi(k) + \log \frac{c_{M+1,2} \cdot c_{2M,2}}{c_{2M+1,2} \cdot c_{1,2}} + \frac{1}{N_0 - M} \sum_{i=M+1}^{N_0} \big(\log(n_{Y^-,i,2}) - \log(n_{(Y^-,Y),i,2}) - \log(n_{(Y^-,X^-),i,2}) \big). \tag{4}$$

In the next section we provide a detailed description of the proposed algorithm.

IV. Proposed Algorithm for SOZ Localization

Recall that the proposed algorithm consists of three main phases: 1) Pre-processing the data in order to facilitate accurate and computationaly efficient estimation of the pair-wise causal influence; 2) Estimation of the directed graph consisting of the pair-wise causal influences; and 3) Inferring the SOZ from the estimated graph. We now describe each of these steps in more detail.

A. Pre-Processing

The pre-processing consists of 4 steps: First, the common reference is removed from all the recorded signals [22]. Then, each column of the matrix V is filtered using a 60 Hz notch filter to remove line-noise [3]. Next, each column of the matrix V is down-sampled to 100 Hz.[2] Recalling that each block is 10 seconds long, this implies that $N_0 = 1000$. Finally the mean of each column is removed and each column is normalized to have a unit variance.

While removing the common reference, the line noise, and the mean of each recorded signal, as well as normalizing each recorded signal, are straight-forward steps, down sampling requires further discussion. To understand the motivation for down-sampling we first note that ECoG signals are approximately stationary only for few seconds [22]. Therefore, following Assumption **A1**), we use short blocks of 10 seconds to estimate the pair-wise DI. We note here that the works [14], [15] used significantly longer blocks in order to have sufficient

[2]The sampling rate in the iEEG portal is between 500 Hz and 5 KHz.

number of samples for estimating the DI. We further note that the number of samples required for accurate estimation grows exponentially with the Markov order M [23, Sec. V.E]. Clearly, to have an accurate description of the statistical properties of the considered signals, M should capture the memory of the ECoG signals. In [14] this memory was roughly estimated to be on the order of tens of milliseconds, which requires $M > 10$ for $F_{\rm s} = 500$ Hz. Thus, the number of ECoG samples one observes in a 10 seconds block is far from sufficient for accurate estimation.

To tackle this challenge we down-sample the ECoG signals to 100 Hz, thus, the required M is significantly smaller, i.e., smaller than 10. Note that the above down-sampling filters out all data in frequencies larger than 100 Hz. Filtering the high frequencies was also applied in [3], yet, by doing so one ignores high frequency oscillations [24].

B. Estimating the Causal Influence Graph

In recent years several works proposed to analyze ECoG recordings using tools from graph theory, see [25] and references therein. Following this approach we construct a graph G with N nodes, each representing a recording electrode. Let $[\mathsf{X}]_{[:,i]}$ denote the i^{th} column of the matrix X. Then, the weight of the edge between the i^{th} and j^{th} nodes in G is given by $[\mathsf{G}]_{i,j} = \hat{I}([\mathsf{X}]_{[:,i]} \to [\mathsf{X}]_{[:,j]}), [\mathsf{G}]_{i,i} = 0$, where $\hat{I}(\mathbf{X} \to \mathbf{Y})$ is given in (4). For these estimations we used $k = 5$ and $M = 5$, namely, we used 5 nearest-neighbors while the memory length was assumed to be about 50 milliseconds. We observed that the algorithm is fairly robust to minor changes in these parameters. Note that the resulting graph is fully connected and asymmetric (DI is *not* a symmetric functional).

As $[\mathsf{G}]_{i,j}$ is *estimated* from a finite number of samples, it is important to assess the statistical significance of this estimation. A lack of statistical significance may imply that there is no causal influence between the considered signals, and the estimated values is due to either the noise or the estimation error. In this case one may choose to set $[\mathsf{G}]_{i,j} = 0$. This statistical significance can be calculated using the non-parametric bootstrapping procedure described in [26]. However, since this phase of the algorithm already requires a major computational load, applying bootstrapping amounts to multiplying the computational complexity by a factor of at least 20. For this reason, we do not check the statistical significance of the pair-wise estimations, and instead verify the statistical significance of the whole algorithm, as described in the next subsection.

C. Inferring the SOZ

Before discussing how to infer the SOZ, we first note that some data sets may contain multiple seizures, thus, based on the assumption that there is a single focus, we combine all seizures by averaging the estimated graphs, resulting in a new graph $\bar{\mathsf{G}}$.

Recall the observation that focal seizures start in the SOZ and spread to surrounding areas in the brain [6]. Based on this understanding, and since the graph $\bar{\mathsf{G}}$ aims at representing the *causal influences* between the different recordings, the nodes in the SOZ should intuitively have properties reminiscent of *sources*. More precisely, let the net-flow of a node $i \leq N$, be defined as:

$$\mathcal{F}_i \triangleq \sum_{j=1}^{N} ([\bar{\mathsf{G}}]_{i,j} - [\bar{\mathsf{G}}]_{j,i}). \tag{5}$$

Then, intuitively, nodes in the SOZ should have relatively high $\mathcal{F}_i$ compared to other nodes in $\bar{\mathsf{G}}$. Let $\mathcal{S}_0 \in \{1, 2 \ldots, N\}$ be the set of nodes (electrodes) which constitute the top p_0 percentile of $\{\mathcal{F}_i\}_{i=1}^N$. While $\mathcal{S}_0$ is a natural candidate to be the SOZ, we still need to verify that the values associated with these nodes are statistically significant. For this reason we create a null-distribution, for each $\mathcal{F}_i$. Let $N_{\rm S}$ denote the number of seizures in the data-set. We *randomly* choose $N_{\rm s}$ *rest* blocks from the data-set and for these blocks we calculate the net-flow values $\{\tilde{\mathcal{F}}_i\}_{i=1}^N$. We repeat this procedure 200 times and create an empirical PDF *for each one of the* $\tilde{\mathcal{F}}_i$. We define $\mathcal{S}_1 \in \{1, 2 \ldots, N\}$ to be the set of nodes for which $\mathcal{F}_i$ is larger than $(100 - p_1)\%$ of the values in the empirical PDF of $\tilde{\mathcal{F}}_i$. Finally, the algorithm declares $\mathcal{S}$, the SOZ, as: $\mathcal{S} = \mathcal{S}_0 \cap \mathcal{S}_1$.

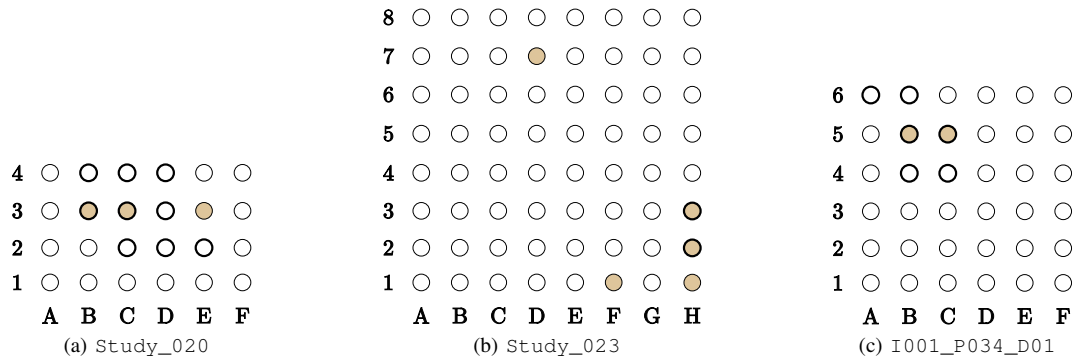
Data-set	Sex	Seizure Onset	Seizure Type	#Seizures	Outcome Class
`Study 020`	M	RF	CPG	8	IV
`Study 023`	M	LO	CP	4	I
`I001_P034_D01`	F	RF	CPG	16	NF

TABLE I: **Patients Information**. RF - Right frontal, LO - Left occipital, CP - complex-partial, CPG - complex-partial with secondary generalization, NF - No follow-up.

V. Results and Discussion

To demonstrate the efficacy of the proposed algorithm we tested it on three data-sets taken from the iEEG portal [16]. The patients' information is detailed in Table I. All the considered patients went through resection. The sampling frequency for the data-sets `Study 020` and `Study 023` is $F_{\rm s} = 500$ Hz, while for `I001_P034_D01` it is $F_{\rm s} = 5$ KHz. For each one of the data-sets we considered the large grid located over the suspected SOZ. The number of electrodes in these grids differs from one data-set to another. The grid in `Study 020` consists of 24 electrodes, the grid in `Study 023` consists of 64 electrodes, and the grid in `Study I001_P034_D01` consists of 36 electrodes. For the seizures start time we used the neurologists' annotations.

The localization results for the data-sets detailed in Table I are presented in Fig. 2. As a ground truth we use the EOIs indicated by the neurologists and detailed in the data-sets reports. In Fig. 2, these electrodes (nodes) are marked by a *bold annulus*, whereas the nodes detected by the proposed algorithm are marked by solid circles. It can be observed that for all three data-sets the algorithm accurately localizes the SOZ. In the SOZ inference phase, see Section IV-C, we used the values $p_0 = 5$ and $p_1 = 10$. We note here that these values control the tradeoff between *false-alarm* and *miss-detection*. This can be clearly observed in Fig. 2b in which node 7D is far from the suspected SOZ. Building upon the assumption of



(a) Study_020 (b) Study_023 (c) I001_P034_D01

Fig. 2: Localization results for the data-sets in Table I. Sub-figure (a) is the RAG grid in data-set Study_020, where node 1F corresponds to electrode RAG1 and node 4F corresponds to electrode RAG19. Sub figure (b) is the LTG grid in data-set Study_023, where node 1A corresponds to electrode LTG1 and node 1H corresponds to electrode LTG8. Sub figure (c) is the GRID grid in data-set I001_P034_D01, where node 1F corresponds to electrode GRID1 and node 1A corresponds to electrode GRID6. Note that even though node 6A–6B are mentioned in the report, their recordings are missing from the data-set.

a single focus, a possible approach to eliminate such false-alarms is to account for the spatial location and choose the nodes that are closely located. In Fig. 2b such an approach leads to the nodes 1H–3H.

We conclude this paper with several future research directions. We first note that due to the decimation applied as part of down-sampling, see Section IV-A, a large number of the samples are not (directly) used in the DI estimation. This can be tackled by processing all the different phases of the decimation output. We further note that currently the algorithm uses a heuristic choice of the Markov order M, see Section IV-B. One of our future research directions is to develop methods for estimating M from the data. Finally, we plan to derive more sophisticated methods to infer the SOZ from G.

REFERENCES

[1] P. van Mierlo, M. Papadopoulou, E. Carrette, P. Boon, S. Vandenberghe, K. Vonck, and D. Marinazzo, "Functional brain connectivity from eeg in epilepsy: Seizure prediction and epileptogenic focus localization," *Progress in Neurobiology*, vol. 121, pp. 19–35, 2014.

[2] S. Pati and A. Cole, "How focal is generalized epilepsy: A distinction with a difference?" *Epilepsy and Behavior*, vol. 34, pp. 127–128, 2014.

[3] A. N. Khambhati, K. A. Davis, B. S. Oommen, S. H. Chen, T. H. Lucas, B. Litt, and D. S. Bassett, "Dynamic network drivers of seizure generation, propagation and termination in human neocortical epilepsy," *PLoS Comput. Biol.*, vol. 11, no. 12, pp. 1–19, 2015.

[4] E. Krook-Magnuson and I. Soltesz, "Beyond the hammer and the scalpel: Selective circuit control for the epilepsies," *Nature Neuroscience*, vol. 18, no. 3, pp. 331–338, 2015.

[5] M. J. M. others, "Responsive cortical stimulation for the treatment of medically intractable partial epilepsy," *Neurology*, vol. 77, no. 13, pp. 1295–1304, 2011.

[6] M. A. Brazier, "Spread of seizure discharges in epilepsy: Anatomical and electrophysiological considerations," *Experimental Neurology,*, pp. 263–272, 1972.

[7] C. J. Quinn, T. P. Coleman, N. Kiyavash, and N. G. Hatsopoulos, "Estimating the directed information to infer causal relationships in ensemble neural spike train recordings," *Journal on Computational Neuroscience*, vol. 30, no. 1, pp. 17–44, 2011.

[8] K. Hlavackova-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya, "Causality detection based on information-theoretic approaches in time series analysis," *Physics Reports*, vol. 441, pp. 1–46, 2007.

[9] K. J. Blinowska, "Review of the methods of determination of directed connectivity from multichannel data," *Medical & Biological Engineering & Computing*, vol. 49, no. 5, pp. 521–529, 2011.

[10] J. Jiao, H. H. Permuter, L. Zhao, Y.-H. Kim, and T. Weissman, "Universal estimation of directed information," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6220–6242, Oct. 2013.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*, 2nd ed. Wiley-Interscience, 2006.

[12] A. Kraskov, H. Stogbauer, and P. Grassberger, "Estimating mutual information," *Physical Review E*, vol. 69, no. 6, 2004.

[13] M. Wibral, R. Vicente, and M. Lindner, "Transfer entropy in neuroscience," *Directed Information Measures in Neuroscience, Understanding Complex Systems*, pp. 3–36, 2014.

[14] R. Malladi, G. Kalamangalam, N. Tandon, and B. Aazhang, "Identifying seizure onset zone from the causal connectivity inferred using directed information," *IEEE Jour. of Sel. Topics in Sig. Proc.*, vol. 10, no. 7, pp. 1267–1283, 2016.

[15] S. Sabesan, L. B. Good, K. S. Tsakalis, A. Spanias, D. M. Treiman, and L. D. Iasemidis, "Information flow and application to epileptogenic focus localization from intracranial eeg," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 17, no. 3, pp. 244–253, 2009.

[16] J. Wagenaar, B. Brinkmann, Z. Ives, G. Worrell, and B. Litt, "A multimodal platform for cloud-based collaborative research," in *Proc. of Int. IEEE EMBS Conf. on Neur. Eng.*, San Diego, CA, USA, Nov. 2013, pp. 1386–1389, see https://www.ieeg.org/.

[17] D. Lederman and J. Tabrikian, "Classification of multichannel eeg patterns using parallel hidden markov models," *Medical & Biological Engineering & Computing*, vol. 50, no. 4, pp. 319–328, Apr. 2012.

[18] T. Wissel, T. Pfeiffer, R. Frysch, R. T. Knight, E. F. Chang, H. Hinrichs, J. W. Rieger, , and G. Rose, "Hidden markov model and support vector machine based decoding of finger movements using electrocorticography," *Journal of Neural Engineering*, vol. 10, no. 5, Oct. 2013.

[19] Y. Murin and A. Goldsmith, "On k-nn estimation of directed information," Tech. Rep., 2016, available at *http://web.stanford.edu/∼moriny*.

[20] W. Gao, S. Oh, and P. Viswanath, "Demystifying fixed k-nearest neighbor information estimators," 2016, available at https://arxiv.org/abs/1604.03006.

[21] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, Eds., *NIST Handbook of Mathematical Functions*, 1st ed. Cambridge University Press, 2010.

[22] M. A. Kramer, U. T. Eden, E. D. Kolaczyk, R. Zepeda, E. N. Eskandar, and S. S. Cash, "Coalescence and fragmentation of cortical networks during focal seizures," *Journal of Neuroscience*, vol. 30, no. 30, pp. 10 076–10 085, 2010.

[23] J. Jiao, K. Venkat, Y. Han, and T. Weissman, "Minimax estimation of functionals of discrete distributions," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2835–2885, May 2015.

[24] M. Zijlmans, P. Jiruska, R. Zelmann, F. S. Leijten, J. G. Jefferys, and J. Gotman, "High-frequency oscillations as a new biomarker in epilepsy," *Annals of Neurology*, vol. 71, no. 2, pp. 169–178, 2012.

[25] S. P. Burns, S. Santanielloa, R. B. Yaffe, C. C. Jouny, N. E. Crone, G. K. Bergey, W. S. Anderson, and S. V. Sarma, "Network dynamics of the brain and influence of the epileptic seizure onset zone," *Proceedings of the National Academy of Sciences*, vol. 111, no. 49, pp. E5321–E5330, 2014.

[26] C. Diks and J. DeGoede, *A general nonparametric bootstrap test for Granger causality*. Institute of Physics Publishing, 2001, in *Global Analysis of Dynamical Systems*.