Comparison of five directed graph measures for identification of leading interictal epileptic regions

Ladan Amini, Christian Jutten, Sophie Achard, Olivier David, Philippe Kahane, Laurent Vercueil, Lorella Minotti, Gh. Ali Hossein-Zadeh, Hamid Soltanian-Zadeh

▶ **To cite this version:**

Ladan Amini, Christian Jutten, Sophie Achard, Olivier David, Philippe Kahane, et al.. Comparison of five directed graph measures for identification of leading interictal epileptic regions. Physiological Measurement, 2010, 31 (11), pp.1529-46. 10.1088/0967-3334/31/11/009 . inserm-00614037

HAL Id: inserm-00614037

https://inserm.hal.science/inserm-00614037

Submitted on 8 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of five directed graph measures for identification of leading interictal epileptic regions

L. Amini[1,2], C. Jutten[2,3], S. Achard[2], O. David[4,5], P. Kahane[4,6], L. Vercueil[4,6], L. Minotti[4,6], G.A. Hossein-Zadeh[1], H. Soltanian-Zadeh[1,7]

1- Control and Intelligent Processing Center of Excellence (CIPCE), School of ECE, Faculty of Engineering, University of Tehran, Tehran, Iran.

2- Image, Speech, Signal, and Automatic laboratory of Grenoble (GIPSA-LAB), University of Grenoble, Domaine universitaire- BP 46, F-38402 Grenoble Cedex, France.

3- Institut Universitaire de France,103 , bd Saint-Michel 75005 Paris, France.

4- Inserm U836 Grenoble Institute of Neuroscience, Grenoble, France.

5- Neuroradiology Department and MRI Unit, Grenoble University Hospital, Grenoble, France.

6- Neurology Department, Grenoble University Hospital, Grenoble, France.

7- Radiology Image Analysis Laboratory, Henry Ford Health System, Detroit, MI 48202, USA.

E-mail: `hamids@rad.hfh.edu`

Abstract. Directed graphs (digraphs) derived from interictal periods of intracerebral EEG (iEEG) recordings can be used to estimate the leading interictal epileptic regions for presurgery evaluations. For this purpose, quantification of the emittance contribution of each node to the rest of digraph is important. However, the usual digraph measures are not very well suited for this quantification. Here we compare the efficiency of recently introduced local information measure LI and a new measure called total global efficiency with classical measures like global efficiency, local efficiency and node degree. For evaluation, the estimated leading interictal epileptic regions based on five measures are compared with seizure onset zones obtained by visual inspection of epileptologists for five patients. The comparison revealed the superior performance of LI measure. We showed efficiency of different digraph measures for the purpose of source and sink node identification.

Keywords: Directed graph measurements, Epilepsy, Intracerebral EEG recordings, Multiple graph analysis, Mutual information

1. Introduction

The brain functional connectivity is defined as the temporal coupling between the activities recorded from spatially remote brain regions associated with neural units. Effective connectivity which is based on the effect of one neural system over another can be derived from functional connectivity based on the causal information between the interacting regions (Friston 1994, Horwitz 2003, Sporns et al. 2004). The interpretation of the measured functional connectivity can be illustrated by a graph.

Let's assume directed graph (digraph) $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ are the set of vertices (or nodes) and edges (or connections). We describe the digraph $\mathbf{G}$ with an adjacency matrix denoted as $\mathbf{A_G} = [a_{ij}] \in \{0,1\}^{N\times N}$, where N is the number of nodes. In this paper, we focus on the digraphs (oriented graphs) whose adjacency matrix has no symmetric pair of directed edges (each edge is permitted to have a unique direction) and has a diagonal of zeros, i.e. no loops (an edge from a node to itself) are allowed. If there is an edge from node i to node j, then $a_{ij} = 1$ and $a_{ji} = 0$, otherwise if there is no edge, then $a_{ij} = a_{ji} = 0$. For such digraphs, the maximum possible number of edges is equal to $N_c = (N^2 - N)/2$ directed edges ($\sum_{ij} a_{ij} = (N^2 - N)/2$).

The domain of our investigation is to study the brain network related to interictal events of epileptic patients from intracerebral EEG (iEEG) recordings. We try to estimate the leading regions responsible for the generation of interictal epileptiform discharges (IED) which are called leading IED (LIED) regions through graph analysis. The graph studied here is a "differential connectivity graph (DCG)" which was proposed in Amini et al. 2010. Each connection of DCG is a discriminated connection between two different states called "IED" and "non-IED". An IED time interval is a period of iEEG signals including one IED or burst of IEDs, while non-IED time interval is a period excluding any IED or abnormal activities. In DCG it is assumed that the brain connectivity is affected by IED state. Upon this assumption, DCG is designed to identify the connections which change significantly between IED and non-IED states. To be more precise, the connections which their statistical properties change under these two states are aimed to be detected. Therefore the DCG connections are designed to focus on IED events excluding common events which share the same statistical properties under IED and non-IED time intervals like background activity.

In DCG the wavelet cross-correlation was used to measure the functional connectivity and the graph is inferred using the permutation-based multiple testing between the couplings of large number of IED and non-IED time intervals. The causal relationship between participating regions is characterized by the estimation of time delays using the maximum wavelet cross-correlation based on (Amini et al. 2009, Amini et al. 2010). To estimate the leading IED regions from directed DCG (dDCG), a digraph measure is required to determine the source and sink nodes whereas the available classical digraph measures (for review, see Albert et al. 2002, Newman 2003, Boccaletti et al. 2006, Bullmore & Sporns 2009) are not very well suited for this purpose. A source is defined as a node which has high emittance contribution (high positive local information) to the rest of digraph.

Our goal in this paper is to compare the efficiency of usual digraph measures (Watts & Strogatz 1998, Latora & Marchiori 2003, Boccaletti et al. 2006) with two new measures: local information (LI) and total global efficiency. LI measures the amount of information which passes through each node locally (outgoing information minus incoming information) based on lagged mutual information and the digraph

structure. This measure was introduced in Amini et al. 2010 for the purpose of source and sink node definition from digraphs. Positive LI value for a node shows that the node emits information more than it receives which leads to its emittance contribution. The emittance contribution reveals the strength of a source node. Conversely negative LI value shows that the node receives more information than it emits which leads to the strength of a sink node.

Total global efficiency is defined here as an extension to global efficiency which considers both outgoing and incoming paths in order to study the effect of incoming paths, in addition to outgoing ones on the characterization of source and sink nodes.

To evaluate the comparison, each of five measures (three classic ones and two new measures) is used for the characterization of dDCG to estimate the LIED regions for five patients. These patients all had positive (seizure-free) surgical outcome. The leading IED regions based on the different measures are compared with visually inspected seizure onset zones (SOZ), i.e. the region in which the first electrophysiological changes detected at seizure onset, by an epileptologist. Though the SOZ is theoretically the region to be removed in candidates to resective surgery, it is not easy to be defined in many patients partly because of the difficulties regarding recording seizure periods. Thus it is valuable to wonder if one can predict the SOZ by estimating the IED regions. There are several studies wondering if IED regions and specifically leading IED regions can be useful for the resective surgery planning (Alarcon 1996, Le Van Quyen et al. 1998, Hufnagel et al. 2000, Lai et al. 2007, Monto et al. 2007, Ortega et al. 2008, Wilke et al. 2009, Wendling et al. 2009, Amini et al. 2010). Drive-response relationship between IED regions and therefore distinction between source (leading) and sink IED regions may thus be important since it has been shown that sink IED regions may not be necessary to be removed during resective surgery (Alarcon et al. 1997).

Rest of the paper is organized as follows. In Section 2, we describe the definition of digraph measures. Section 3 is devoted to the data protocol, and comparison of estimated leading IED regions based on different measures for the five epileptic patients. A discussion about advantages and disadvantages of TI over other usual measures for the definition of source and sink nodes from the digraphs is brought in Section 4. Concluding remarks are presented in Section 5.

2. Methods

2.1. Background

2.1.1. General definitions for digraphs: To define the usual digraph measures, we start with general definitions for digraph $\mathbf{G}$. A path from node i to node j (or outgoing path of i, or incoming path of j) is an ordered sequence of non-repeated edges and nodes connecting node i to node j (Sporns et al. 2004). Let's assume the shortest path length matrix denoted as $\mathbf{L_G} = [l_{ij}]$, where l_{ij} is the shortest path length between nodes i and j, i.e. the minimum number of edges traversed to get from i to j. If any path does not exist from i to j, then l_{ij} is equal to ∞.

Subgraph $\mathbf{G}_{i^-} = (\mathbf{V}_i, \mathbf{E}_i)$ is the subgraph of the first-order outgoing neighbors of node i excluding node i, i.e. $\{j \neq i \in \mathbf{V} |\ a_{ij} = 1\}$. When node i is included, the subgraph is denoted as $\mathbf{G}_{i^+}$.

For each frequency level, the related digraph is characterized with different digraph measures. In the following, for simplicity the digraph $\mathbf{G}$ and digraph measures

are not represented differently for different frequency levels.

Figure 1 demonstrates a simple example of digraph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ (left) and subgraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$ (right). $\mathbf{V} = \{1, 2, 3, 4, 5, 6, 7\}$, and $\mathbf{E}$ includes 12 directed edges among $N_c = 21$ possible edges. The adjacency ($\mathbf{A_G}$), and shortest path length ($\mathbf{L_G}$) matrices of digraph $\mathbf{G}$ are as follows:

$$\mathbf{A_G} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}, \mathbf{L_G} = \begin{pmatrix} 3 & 1 & 2 & 3 & 4 & 1 & 2 \\ \infty & \infty & 1 & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ 3 & 3 & 1 & 3 & 1 & 2 & 2 \\ 2 & 2 & 2 & 2 & 3 & 1 & 1 \\ 2 & 2 & 2 & 2 & 3 & 3 & 1 \\ 1 & 1 & 1 & 1 & 2 & 2 & 3 \end{pmatrix} \quad (1)$$

The shortest path length matrix ($\mathbf{L}_{\mathbf{G}_{7^-}}$) of subgraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$, $\mathbf{V}_7 = \{1, 2, 3, 4\}$ is as follows:

$$\mathbf{L}_{\mathbf{G}_{7^-}} = \begin{pmatrix} \infty & 1 & 2 & \infty \\ \infty & \infty & 1 & \infty \\ \infty & \infty & \infty & \infty \\ \infty & \infty & 1 & \infty \end{pmatrix} \quad (2)$$

Figure 1. A sample digraph $\mathbf{G}$ (left) and the subgraph of node 7, $\mathbf{G}_{7^-}$ (right).

2.1.2. Indegree and outdegree: The outdegree $K_{out}[i]$ of node i is the number of outgoing edges incident the node (Newman 2003, Boccaletti et al. 2006):

$$K_{out}[i] = \sum_{j \in \mathbf{V}} a_{ij} \quad (3)$$

Accordingly, indegree $K_{in}[i]$ is related to the incoming edges:

$$K_{in}[i] = \sum_{j \in \mathbf{V}} a_{ji} \quad (4)$$

where a_{ij} are the entries of adjacency matrix $\mathbf{A_G}$. The total degree of node i can be defined as the difference between $K_{out}[i]$ and $K_{in}[i]$: $K_{tot}[i] = K_{out}[i] - K_{in}[i]$. K_{tot} is in the range $[-(N-1), N-1]$. The total degree ($\mathbf{K}_{tot} = [K_{tot}[1], \ldots, K_{tot}[7]]^T$) for the digraph $\mathbf{G}$ depicted in figure 1 is equal to the vector $\mathbf{K}_{tot} = [1, -1, -3, 1, 1, -1, 2]^T$.

2.1.3. Global and local efficiencies: If we define the efficiency e_{ij} of the connection between node pair (i, j) equal to $1/l_{ij}$ (Watts & Strogatz 1998, Latora & Marchiori 2003), then the global efficiency of the node $i \in \mathbf{V}$ can be defined as:

$$E_{glob}[i] = \frac{1}{N-1} \sum_{j \in \mathbf{V}, j \neq i} \frac{1}{l_{ij}} \tag{5}$$

and the global efficiency of the whole digraph $\mathbf{G}$ is defined (Latora & Marchiori 2003) as:

$$E_{glob}[\mathbf{G}] = \frac{1}{N} \sum_{i \in \mathbf{V}} E_{glob}[i]. \tag{6}$$

$E_{glob}[i]$ and $E_{glob}[\mathbf{G}]$ are in the range $[0, 1]$. $E_{glob}[i]$ is equal to zero for a node without any outgoing path and equal to one if $l_{ij} = 1, \forall j \neq i, j \in \mathbf{V}$, i.e. we can reach from node i to each node of the digraph with single edge.

The $E_{glob}[\mathbf{G}]$ can also be evaluated for the subgraph $\mathbf{G}_{i^-}$ to measure the local properties. The local efficiency of node i can be defined as the evaluation of (6) for the subgraph $\mathbf{G}_{i^-}$ as follows:

$$E_{loc}[i] = E_{glob}[\mathbf{G}_{i^-}]. \tag{7}$$

$E_{loc}[i]$ shows the efficiency of the connections between the first-order outgoing neighbors of i when i is removed. Equivalently, local efficiency measures the "resilience" of digraph to the damage of node removal, i.e. if we remove a node, how efficient its first-order outgoing neighbors can communicate. $E_{loc}[i]$ is in the range $[0, 1]$.

We calculate global (5)-(6) and local (7) efficiency for the example depicted in figure 1. The global efficiency ($\mathbf{E}_{glob} = [E_{glob}[1], \ldots, E_{glob}[7]]^T$) for the seven nodes of digraph $\mathbf{G}$ and the global efficiency of digraph $\mathbf{G}$ are equal to $\mathbf{E}_{glob} = [0.6, 0.17, 0, 0.61, 0.67, 0.56, 0.83]^T$ and $E_{glob}(\mathbf{G}) = 0.5$, respectively. Global efficiency of node 7 is high since we can reach from this node to the rest of digraph with one or two edges (the seventh row of matrix $\mathbf{L}_{\mathbf{G}}$). The local efficiency ($\mathbf{E}_{loc} = [E_{loc}[1], \ldots, E_{loc}[7]]^T$) for the nodes of digraph $\mathbf{G}$ is equal to $\mathbf{E}_{loc} = [0, 0, 0, 0, 0.5, 0, 0.3]^T$. For calculation of local efficiency of node 7, subgraph $\mathbf{G}_{7^-} = (\mathbf{V}_7, \mathbf{E}_7)$ depicted in figure 1 (right) is used. $E_{loc}[5]$ is higher than $E_{loc}[7]$ because the neighbors of node 7 (nodes 1, 2, 3, and 4) do not communicate as well as neighbors of node 5 (nodes 5, 6, and 7 make a triangle).

2.2. New measures

In this Section, first the evaluation of causal relationships is explained. Next the two new measures local information and total global efficiency are defined for the digraphs.

2.2.1. Causal relationship: Let's assume $\mathbf{x}_i = [x_i[1] \ \ldots \ x_i[T]]^T$ and $\mathbf{x}_j = [x_j[1] \ \ldots \ x_j[T]]^T$ be the time series associated with nodes i and j, where T is the number of samples. We assume the time causality between two signals be the time difference between the events of underlying signals. The time causality can be estimated by different procedures like methods based on phase synchrony (Gotman 1983, Ktonas & Mallart 1991) or based on cross-correlation (Cohn & Leader 1967, Carter 1981, Ianniello 1982, Benesty et al. 2004, Amini et al. 2009, Amini et al. 2010).

Here, we estimate time causality between $\mathbf{x}_i$ and $\mathbf{x}_j$ based on wavelet cross-correlation as it is exploited in Amini et al. 2010.

Let's assume $\mathbf{d}^m_{\mathbf{x}_i}$ and $\mathbf{d}^m_{\mathbf{x}_j}$ be the wavelet (maximal overlap discrete wavelet transform) coefficients of time series $\mathbf{x}_i$ and $\mathbf{x}_j$ for resolution level m, where $m = \{1, \ldots, M\}$. The time causality between wavelet coefficients of $\mathbf{d}^m_{\mathbf{x}_i}$ and $\mathbf{d}^m_{\mathbf{x}_j}$ (associated with temporal time series $\mathbf{x}_i$ and $\mathbf{x}_j$) is estimated as the time lag in which the absolute maximum cross-correlation between underlying wavelet coefficients (Whitcher et al. 2000, Achard et al. 2006) occurs:

$$\widehat{\rho}\left(\mathbf{d}^m_{\mathbf{x}_i}, \mathbf{d}^m_{\mathbf{x}_j}, \tau\right) = \frac{\widehat{cov}\left\{d^m_{\mathbf{x}_i}[k], d^m_{\mathbf{x}_j}[k-\tau]\right\}}{\sqrt{\widehat{var}(d^m_{\mathbf{x}_i}[k])\widehat{var}(d^m_{\mathbf{x}_j}[k-\tau])}} \tag{8}$$

$$\tau^{*m}_{ij} = \arg\max_{\tau}(\left|\widehat{\rho}\left(\mathbf{d}^m_{\mathbf{x}_i}, \mathbf{d}^m_{\mathbf{x}_j}, \tau\right)\right|) \tag{9}$$

where $\widehat{\rho}$, $\widehat{cov}$, and $\widehat{var}$ are the empirical estimations of correlation coefficient, covariance and variance, respectively. The time causality between $\mathbf{x}_i$ and $\mathbf{x}_j$ for the resolution level m is defined through the estimation of time causality between $\mathbf{d}^m_{\mathbf{x}_i}$ and $\mathbf{d}^m_{\mathbf{x}_j}$ denoted as τ^{*m}_{ij}. We assume that the "transfer function" between signals recoded from different brain regions cannot be represented by pure delay, eventually we do not confine the time causality to be the same for all of the frequency bands.

The causal relationship between $\mathbf{d}^m_{\mathbf{x}_i}$ and $\mathbf{d}^m_{\mathbf{x}_j}$ (associated with temporal time series $\mathbf{x}_i$ and $\mathbf{x}_j$) is defined as follows. If there is a connection between node pair (i, j) and τ^{*m}_{ij} is negative then the edge is from i to j. The relationship is reverse if τ^{*m}_{ij} is positive.

Reliability of estimated time lag: The confidence interval of (8) for maximal overlap discrete wavelet transform coefficients of non-stationary fractionally differenced signals (for the definition please refer to Whitcher et al. 2000) was estimated in Whitcher et al. 2000. However, for the non-stationary signals studied here, this property does not hold. Furthermore the estimation of the distribution of maximum wavelet cross-correlation for non-stationary signals is complicated. As such, to test the reliability of time lags calculated in (9) a statistical jackknife method is used. To this end, N_w windows starting at random in the range $[1, T-W]$ are considered where T and W are the length of the selected time series (about one hour) and the length of each window (20 minutes), respectively. The window length is considered large enough to include large number of IED time intervals. The time lag is estimated (9) for each of the randomly chosen windows providing N_w time lag values. These N_w time lag values form a histogram-based probability distribution. For each edge of DCG between node pair (i, j), this histogram is estimated and denoted as $\hat{p}_{\tau^{*m}_{ij}}(u)$. The time lag which has the greatest probability is defined as:

$$\bar{\tau}^{*m}_{ij} = \arg\max_{u}(\hat{p}_{\tau^{*m}_{ij}}(u)) \tag{10}$$

For each node pair (i, j), we test if the N_w time lag values are significantly non-equal to zero. If the time lag values are significantly equal to zero then $\bar{\tau}^{*m}_{ij}$ is set to zero, i.e. the direction of the edge between node pair (i, j) could not be estimated. For the statistical test a bootstrap method with N_b repetitions is used. We compare the $\bar{\tau}^{*m}_{ij}$ values with the τ^{*m}_{ij} estimated (9) between the signals of length T. The percentage

of similar time lags ($\bar{\tau}_{ij}^{*m} \times \tau_{ij}^{*m} > 0$) over number of edges of DCG is in the range [78 95]% for different frequency bands. Since the signal pairs related to the edges of DCG have significant couplings (Amini et al. 2010), (9) provides reliable estimation of the most probable time lag if it is calculated for long enough signals.

2.2.2. LI measure: LI measure (Amini et al. 2010) quantifies the amount of information passing through each node locally, i.e. local information and is defined in terms of lagged mutual information and the structure of the digraph. More precisely, LI of each node depends on the outgoing and incoming edges incident the node and the amount of information which is carried by each of these edges.

Let's assume $\mathbf{d}_{\mathbf{x}_i}^m$ and $\mathbf{d}_{\mathbf{x}_j}^m$ be the observations of two random variables $D_{\mathbf{x}_i}^m$ and $D_{\mathbf{x}_j}^m$ with probability density functions (PDF) $p_i(u)$, $p_j(v)$ and joint PDF $p_{ij}(u,v)$. The mutual information (MI) between these random variables is:

$$MI(D_{\mathbf{x}_i}^m, D_{\mathbf{x}_j}^m) = \int\int dudv p_{i,j}(u,v) \log \frac{p_{i,j}(u,v)}{p_i(u)p_j(v)}. \tag{11}$$

By partitioning the supports of $\mathbf{d}_{\mathbf{x}_i}^m$ and $\mathbf{d}_{\mathbf{x}_j}^m$ into bins of finite size and assuming ergodicity, MI (11) can be approximated as the finite sum:

$$MI(D_{\mathbf{x}_i}^m, D_{\mathbf{x}_j}^m) \approx MI(\mathbf{d}_{\mathbf{x}_i}^m, \mathbf{d}_{\mathbf{x}_j}^m) \equiv \sum_{kn} \hat{p}_{ij}(k,n) log \frac{\hat{p}_{ij}(k,n)}{\hat{p}_i(k)\hat{p}_j(n)} \tag{12}$$

where $\hat{p}_i(k)$, $\hat{p}_j(n)$, and $\hat{p}_{ij}(k,n)$ are the probabilities estimated by relative frequencies.

Let's $\mathbf{d}_{\mathbf{x}_j,\tau_{ij}^{*m}}^m = \left[d_{\mathbf{x}_j}^m[k+1-\tau_{ij}^{*m}] \ \ldots \ d_{\mathbf{x}_j}^m[k+T-\tau_{ij}^{*m}]\right]^T$ be the shifted time series of $\mathbf{d}_{\mathbf{x}_j}^m$ with τ_{ij}^{*m} samples, where τ_{ij}^{*m} (Section 2.2.1) is the time causality between signal pairs $\mathbf{d}_{\mathbf{x}_i}^m$ and $\mathbf{d}_{\mathbf{x}_j}^m$ (associated with $\mathbf{x}_i$ and $\mathbf{x}_j$). The LI of node i is defined (Amini et al. 2010) as:

$$LI[i] = \sum_{\mathbf{V}_{i\to j}} MI(\mathbf{d}_{\mathbf{x}_i}^m, \mathbf{d}_{\mathbf{x}_j,\tau_{ij}^{*m}}^m) - \sum_{\mathbf{V}_{j\to i}} MI(\mathbf{d}_{\mathbf{x}_i}^m, \mathbf{d}_{\mathbf{x}_j,\tau_{ij}^{*m}}^m) \tag{13}$$

where $\mathbf{V}_{i\to j} = \{j \neq i \,|\, a_{ij} = 1\}$, i.e. set of first-order outgoing neighbors of node i.

LI can be in the range $[-\infty, +\infty]$. In theory, LI of the source or sink nodes are positive or negative, respectively and greater positive values demonstrate higher emittance contribution. $LI[i]$ is zero when in and out information flows of node i are equal. If $LI[i]$ is positive or equivalently the node i emits information more than it receives, we can conclude that node i is a source and the value of $LI[i]$ quantifies its emittance contribution. Conversely if node i receives more information than it emits (negative $LI[i]$), then the node can be assumed as a sink.

LI can be interpreted as the total degree of a weighted digraph in which the weight related to the edge from node i to node j is the lagged mutual information (13) between the signal pairs associated with nodes i and j ($MI(\mathbf{d}_{\mathbf{x}_i}^m, \mathbf{d}_{\mathbf{x}_j,\tau_{ij}^{*m}}^m)$). This weight can represent the amount of information carrying by the edge.

2.2.3. Total global efficiency: To consider the incoming paths in global efficiency, we define the total global efficiency in this section. The information about outgoing paths of node i are in the *ith* row of shortest path length matrix ($\mathbf{L_G}$), and the information

about its incoming paths are in the *ith* column of $\mathbf{L_G}$. For each node i of digraph $\mathbf{G}$, we define the total global efficiency as:

$$E_{tglob}[i] = \frac{1}{N-1} \sum_{j \in \mathbf{V}, j \neq i} \left(\frac{1}{l_{ij}} - \frac{1}{l_{ji}} \right) \tag{14}$$

$E_{tglob}[i]$ is in the range $[-1, 1]$. High positive values of $E_{tglob}[i]$ shows that we may reach from node i to the rest of digraph easier than reaching from the rest of digraph to this node. The low negative values show the inverse relationship. The total global efficiency ($\mathbf{E}_{tglob} = [E_{tglob}[1], \ldots, E_{tglob}[7]]^T$) for the nodes of digraph $\mathbf{G}$ depicted in figure 1 is equal to $\mathbf{E}_{tglob} = [0.2, -0.4, -0.75, 0.2, 0.3, 0.05, 0.3]^T$. Nodes 5, and 7 have higher positive values than other nodes. It can be seen in figure 1 that these nodes have more outgoing paths than incoming ones. On the other hand node 3 has very low total global efficiency which shows that we can easily reach from other nodes of the digraph to node 3, while it is impossible to access the rest of digraph from this node.

2.3. Multiple digraph analysis to estimate leading IED regions

So far, we explained different digraph measures. Each of these measures can be used to characterize each of M digraphs related to M different frequency bands (multiple digraphs). To detect the sources using the measure values of M digraphs, a multiple digraph analysis (Amini et al. 2010) is employed. Multiple digraph analysis is based on a multi-objective optimization method (Branke et al. 2008, Deb 1999). Since the preference between different frequency bands is unknown, multi-objective optimization method is suitable to estimate a set of optimum nodes, contrary to single objective optimization methods which suggest a single optimum solution. Here we explain briefly the general ideas of multiple digraph analysis.

For any measure studied, we associate matrix $\mathbf{Q} \in \mathbb{R}^{M \times N} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N]$, so that each column $\mathbf{q}_i = [q^1[i], q^2[i], \ldots, q^M[i]]^T$, $i = \{1, \ldots, N\}$ corresponds to the measure values of node i for M different frequency bands. A multi-objective optimization method (Branke et al. 2008, Deb 1999) is applied on matrix $\mathbf{Q}$ for each measure to select the optimum nodes based on the underlying measure. In other words we apply the multi-objective optimization method as a data mining tool on the M-dimensional measure values to detect the sources. The sources or optimum nodes are a set of nodes which have significantly large measure values in at least one of the frequency bands (detailed definition of optimum nodes can be found in Amini et al. 2010). These optimum nodes provide the estimated LIED regions.

3. Data and results

3.1. Data

The iEEG recordings were obtained during presurgery evaluations from five patients suffering from focal epilepsy (Kahane et al. 2004). They are seizure free after resective surgery. Eleven to fifteen semi-rigid multi-lead intracerebral electrodes with 0.8 mm diameter were bilaterally implanted in suspected seizure origins based on clinical considerations. The multi-lead electrodes (Dixi, Besançon, France) include 5, 10, 15 or 18 leads. Each lead has 2 mm length and the leads are evenly spaced with inter-space of 1.5 mm. The iEEG were recorded with an audio-video-EEG monitoring

system (Micromed, Treviso, Italy) with a maximum of 128 channels and digitized at 512 Hz. The electrode leads were recognized on the patient's implantation scheme, and localized in the Montreal Neurological Institute (MNI) atlas. Bipolar derivations were considered between adjacent leads within each electrode (Nunez & Srinivasan 2006). For simplicity, these adjacent bivariate derivations are represented as e_i instead of $e_{i+1} - e_i$. The 50 Hz is removed by a 5th-order notch Butterworth filter with 3dB lower, and upper band stop frequencies equal to 48 Hz and 52 Hz, respectively.

Table 1. Parameters of jackknife method. N_w: number of windows; W: length of each window in minutes; T: length of selected time series in minutes; N_b: number of bootstrap repetitions; α: the false alarm rate.

parameters	N_w	W (min)	T (min)	N_b	α
values	100	20	55.33	10000	0.05

3.2. Experimental results

The parameter values of jackknife method are given in table 1.

3.2.1. Leading IED regions: For each patient, five dDCGs related to five frequency bands ($M = 5$) from 2-64 Hz (2-4, 4-8, 8-16, 16-32, and 32-64 Hz) are obtained with the method proposed in (Amini et al. 2009, Amini et al. 2010). The nodes and edges of the digraphs are associated with the iEEG bipolar channels and the connections between nodes, respectively. The dDCG for patient 3 in 4-8 Hz is shown in part (a) of figure 2, which includes 29 nodes and 62 edges. This digraph is also depicted (part (b) of figure 2) in the real channel coordinates superimposed on the 3D anatomical mesh (sagittal view)‡.

E_{glob}, E_{tglob}, E_{loc}, K_{tot}, and LI (Section 2) are calculated for each of 25 dDCGs (5 dDCGs for each of 5 patients). The negative E_{tglob}, negative K_{tot}, and negative LI values are set to zero since we are not interested in sink nodes. LI values are normalized to their maximum. Multiple digraph analysis is applied on each of these measures for the five frequency bands (matrix $\mathbf{Q}$) to provide the LIED regions based on the underlying measure.

The estimated LIED regions for five patients are presented in table 2. The visually inspected SOZ (vSOZ) by the epileptologist are also reported for comparison. For each patient, the column labels correspond to either LIED regions estimated by digraph measures or the vSOZ of the patient. The resected regions always include the vSOZ, and all of the patients are seizure free after the resective surgery. Therefore measures which suggest LIED nodes congruent with vSOZ are preferred, i.e. more number of common regions (true positives, or TP), smaller number of uncommon regions (false negatives, or FN), and smaller number of LIED regions which are not included in vSOZ (false positives, or FP). The precision and the sensitivity defined as (#TP)/(#TP + #FP), and (#TP)/(#TP + #FN), respectively are reported in tables 3 and 4. The number of regions is denoted by #. Although both FP and FN are aimed to be minimized, due to the existing trade off between these errors, we prefer

‡ We gratefully acknowledge Julien Bastin, Joseph Fourier University, Grenoble, France for providing the anatomical mesh.

Table 2. Comparison between visually inspected seizure onset zones (vSOZ) and estimated leading IED (LIED) regions for five patients (patients 1-5) based on different measures: E_{glob}: global efficiency; E_{tglob}: total global efficiency; E_{loc}: local efficiency; K_{tot}: total degree; LI: local information. For each patient, the column labels correspond to either LIED regions estimated by digraph measures or the vSOZ of the patient. Each cross shows that the region related to its column is detected by the digraph measure related to its row or by the vSOZ. amyg: amygdala; ant/post/m: anterior/posterior/mesial; CG: cingulate gyrus; entCx: entorhinal cortex; fusi: fusiform gyrus; HC: hippocampus; Ins: insula; midInsG: middle short gyrus of insula; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole.

patient 1	antHC	postHC	amyg	pHcG	mTP			
E_{glob}			×	×	×			
E_{tglob}			×	×	×			
E_{loc} [a]	×	×	×	×	×			
K_{tot}			×	×				
LI	×		×	×				
vSOZ	×	×	×	×	×			
patient 2	antHC	postHC	amyg	pHcG	Ins	fusi		
E_{glob}	×							
E_{tglob}	×			×	×	×		
E_{loc}	×	×		×				
K_{tot}	×			×				
LI	×							
vSOZ	×	×	×	×				
patient 3	antHC	postHC	pHcG	TP	postT4			
E_{glob}	×	×	×					
E_{tglob}	×	×		×	×			
E_{loc}	×	×	×					
K_{tot}	×	×						
LI	×	×						
vSOZ	×	×	×					
patient 4	antHC	postHC	amyg	entCx	mTP	antCG	postT1	T4
E_{glob}	×	×	×	×		×		
E_{tglob}	×	×	×	×	×	×	×	×
E_{loc}	×	×		×				
K_{tot}		×	×	×				
LI	×	×	×	×				
vSOZ	×		×	×	×			
patient 5	midInsG							
E_{glob}	×							
E_{tglob}	×							
E_{loc}	NA							
K_{tot}	×							
LI	×							
vSOZ	×							

[a] See the text.

the measures providing less FP compared to FN. In fact FP indicate normal brain regions that are wrongly detected as SOZ which has the risk of removing normal regions in resective surgery. Conversely, FN show the missed SOZ which has the risk of a second surgery supposing the presurgery evaluations are based on only estimated LIED regions. However, the presurgery evaluations are based on different complementary clinical knowledge including iEEG, fMRI, semiology, etc. Estimated LIED regions are aimed to provide the information extracted from iEEG analysis based on the functional connectivity related to interictal events. Therefore since LIED regions provide complementary information for conducting the resection, we

Table 3. Precision of different digraph measures for five patients. NA: not applicable.

Precision	E_{glob}	E_{loc}	K_{tot}	E_{tglob}	LI
patient 1	1	0.5	1	1	1
patient 2	1	1	1	0.5	1
patient 3	1	1	1	0.5	1
patient 4	0.6	0.67	0.67	0.5	0.75
patient 5	1	NA	1	1	1
mean	0.92	0.8	0.93	0.7	0.95

Table 4. Sensitivity of different digraph measures for five patients. NA: not applicable.

Sensitivity	E_{glob}	E_{loc}	K_{tot}	E_{tglob}	LI
patient 1	0.6	1	0.4	0.6	0.6
patient 2	0.25	0.75	0.5	0.5	0.25
patient 3	1	1	0.67	0.67	0.67
patient 4	0.75	0.5	0.5	1	0.75
patient 5	1	NA	1	1	1
mean	0.72	0.81	0.61	0.75	0.65

prefer to estimate the LIED regions with minimum FP to increase the precision of the estimation in comparison with its sensitivity.

Furthermore, although in this study all patients being seizure free reveals that resected regions (based on vSOZ) included the essential regions generating the seizures, currently there is no way to determine how large was the removed regions during surgery. Therefore it is valuable to study the properties of the uncommon nodes between vSOZ and the set of estimated LIED nodes (FN) based on different measures to analyze their properties more profoundly.

3.2.2. Comparison between LI and classical measures: LI measure is designed for source and sink node identification from digraphs, while E_{glob}, E_{loc}, and K_{tot} are not exactly suited for this purpose. In the following we compare LI measure with these usual measures.

Local efficiency: The estimated LIED regions based on E_{loc} are consistent with vSOZ for patients 2, 3 and 4, while this is not true for patients 1 and 5. For patient 1, the estimated LIED based on E_{loc} includes the vSOZ with the addition of five other regions (FP): {anterior cingulate gyrus, anterior superior temporal gyrus, insular cortex, white matter mesial frontal, frontal operculum}. For avoiding long table, we do not report all of LIED regions detected by E_{loc} for patient 1 in the table 2. The local efficiency measure was not applicable for patient 5 due to great sparseness of related graphs. Here we explain the reason of this sparseness. Patient 5 is a very specific patient, treated for right operculo-insular cavernous malformation after the first resective surgery. After nine years, the patient got reflex seizures (Blauwblomme et al. in press). The iEEG recordings for the second presurgery evaluations include extensive sampling of the insular, opercular and temporal cortices. The SOZ is recognized

visually in middle short gyrus of right insula and this region is removed in the second resective surgery. The patient is seizure free after this surgery. dDCG (Amini et al. 2010) of this patient from iEEG recordings before second surgery include one edge between two adjacent electrode leads in middle short gyrus of right insula in the three frequency bands: 8-16, 16-32 and 32-64 Hz. For such sparse digraphs, E_{loc} is not applicable.

The definition of E_{loc} as the evaluation of the efficiency of the connections between the first-order outgoing neighbors of each node is not well suited for the source and sink node detection. Indeed a node with high E_{loc} may not be a good candidate for a source node. Furthermore, E_{loc} considers neither the incoming paths nor the amount of information of each edge. Finally, this measure may suggest regions that should not be removed during surgery (patient 1), which reduces its precision. Although E_{loc} is highly sensitive, we prefer the other measures with higher precision.

Total degree: For most of the patients (all of the patients except patient 4), all of the LIED regions based on K_{tot} are included in the set of related vSOZ which leads to high mean precision over patients. However this measure provides a smaller number of related vSOZ than LI (less TP) which makes it less powerful. Total degree considers both outgoing and incoming connections, but without weights (the amount of information). Two nodes with the same K_{tot} may have different amounts of local information (table 5, first and second columns), since connections may carry different amounts of information. To benefit from such knowledge (amount of information), LI measure was introduced.

Total global efficiency: The total global efficiency is designed to add the information of incoming paths to the knowledge extracted from outgoing ones. Although it was introduced to refine the global efficiency, the results of this measure are not satisfactory: despite the inclusion of the proposed regions based on this measure in vSOZ, it provides additional regions that do not correspond to the related vSOZ. In other words, although adding the information of incoming paths has increased the sensitivity of total global efficiency in comparison with global efficiency, this information does not increase its precision (tables 3 and 4).

This result shows that considering the incoming paths in addition to the outgoing ones in global efficiency is not sufficient for increasing its precision in source and sink node detection. Consequently, a node with high total global efficiency may not be a good candidate for a source node.

Global efficiency: Global efficiency and LI, both seem to have congruent results with vSOZ (tables 2-4), but LI can provide more information about the emittance contribution of the nodes than global efficiency.

Source and sink node detection: The differences between LI and global efficiency is highlighted in the nodes with both several outgoing and incoming connections and with negative or zero amount of local information, i.e. incoming information is greater or approximately equal to outgoing information. For such nodes, LI values are negative or zero, while E_{glob} can be high if we reach from these nodes to the rest of digraph through the short-length outgoing paths. To show one example of such nodes, we focus on the iEEG analysis of patient 3. In this patient, node 46 located in left parahippocampal gyrus is included in the estimated LIED nodes by E_{glob}, while

it is not estimated by LI. Here we explain the details about this node in the dDCG of different frequency bands. We analyze the role of node 46 based on comparing its different measure values (table 5).

We remind that E_{glob}, E_{loc}, E_{tglob}, K_{tot} ($K_{out} - K_{in}$) and LI are in the range $[0,1]$, $[0,1]$, $[-1,1]$, $[-(N-1), N-1]$, and $[-\infty, \infty]$ respectively. For comparing the values of different measures, LI values are normalized to their absolute maximum (range $[-1,1]$) in table 5. The subgraph of node 46 ($\mathbf{G}_{46^+}$) from dDCG of patient 3 in 4-8 Hz (figure 2) is depicted in figure 3. Node 46 has five outgoing and five incoming connections in this frequency band. The incoming and outgoing information are approximately equal which leads to low LI value. Furthermore we can get to 23 nodes among 29 nodes from node 46 through short-length outgoing paths, which gives relatively high E_{glob} comparing to other nodes of the digraph in this frequency band. Besides from 20 nodes of the digraph, we can also get to node 46 through short-length incoming paths. Global efficiency considers outgoing paths, therefore this measure is blind to the knowledge about the incoming paths and specially to the amount of information related to each connection, while LI value is based on the outgoing minus incoming amount of information.

Node 46 is selected as a source node by the multiple digraph analysis (Amini et al. 2010) based on E_{glob} while the local information of this node (table 5) shows that it may not be considered as a source node. Comparing LI values with other usual measure values reveals that node 46 might be mostly a transit node than a source one. A transit node exerts approximately all of the information receives. However distinguishing between source and transit nodes is challenging and might not be easy to be explored neither by E_{glob} nor by visual inspection.

To end with, E_{glob} is not as precise as LI in source nodes detection since the incoming connections and the amount of information are not considered.

Ranking: Another advantage of LI over E_{glob} is in ranking the estimated LIED nodes in terms of measure values. To explain this issue, we focus on the results of iEEG analysis of patient 2. This patient is chosen since the related LIED nodes based on E_{glob} and LI are the same: nodes 70, 71, and 72 located in left anterior hippocampus (table 2). E_{glob} and LI values of these nodes are reported in table 6 for different frequency bands. E_{glob} and LI values are normalized to their maximum (negative values of LI are set to zero before normalization). To rank the estimated LIED nodes, one way can be to order the ℓ_1 norm of related measure values. ℓ_1 norm is calculated for $\mathbf{q}_i$, $i \in \{\text{estimated LIED nodes}\}$ as: $\sum_{m=1}^{M} q^m[i]$, where $q^m[i]$ are nonnegative real numbers. Normalized ℓ_1 norm values of E_{glob} and LI for LIED nodes related to patient 2 are reported in table 7. The greater the ℓ_1 norm the more reliable the LIED node. As can be seen in table 7, the ranking based on E_{glob} and LI is not the same. Considering the incoming connections and the amount of information lead to smaller LI values for node 70, therefore this decreases the relevance of this node to the emittance contribution. Conversely, E_{glob} of this node is high since the incoming paths and the importance of the edges are not considered which leads to its high relevance.

To easily compare the values of E_{glob} and LI for the estimated LIED nodes (reported in table 6), these measure values are demonstrated in figure 4. The M-dimensional vector $\mathbf{q}_i$ can be demonstrated in a web plot with M rays in two dimensional space. In figure 4, the web plot of $\mathbf{q}_{70}$, $\mathbf{q}_{71}$, and $\mathbf{q}_{72}$ for E_{glob} and LI and for $M = 5$ are depicted. The pentagonals related to nodes 70, 71, and 72 are depicted in solid, dashed and dotted lines, respectively. More expanded pentagonals can refer

to greater values of ℓ_1 norm. Based on E_{glob}, node 70 has the greatest ℓ_1 norm, while node 72 has the greatest ℓ_1 norm based on LI measure. To conclude, LI measure is preferred to be used for ranking the leading IED nodes than E_{glob}. This ranking can be valuable for the presurgery evaluations where there are several LIED nodes located in different brain regions.

Table 5. Comparison of E_{loc}, E_{glob}, E_{tglob}, K_{tot}, and LI values of node 46 located in left parahippocampal gyrus of patient 3. LI values are normalized to their absolute maximum.

	2-4	4-8	8-16	16-32	32-64
E_{loc}	0.02	0	0	0	0
E_{glob}	0.5	0.35	0	0	0.06
E_{tglob}	0.04	-0.05	-0.33	-0.20	-0.12
$K_{out} - K_{in}$	$12 - 12$	$5 - 5$	$0 - 3$	$0 - 1$	$1 - 3$
LI	-0.67	-0.09	-0.08	-0.02	-0.01

Table 6. Comparison between E_{glob} and LI values of the three LIED nodes related to patient 2 in different frequency bands.

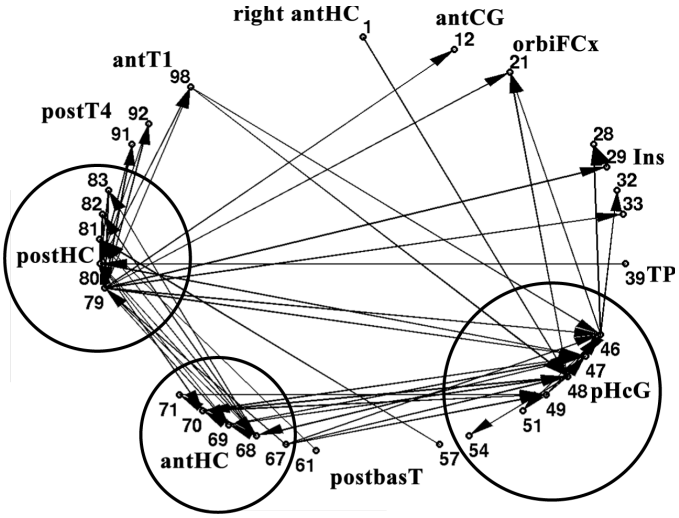
		2-4	4-8	8-16	16-32	32-64
E_{glob}	70	0.68	0.58	0.69	0.67	1
	71	0.85	0.68	0.74	0.69	0.38
	72	0.85	0.72	0.92	0.73	0.25
LI	70	0	0	0	0.08	0.09
	71	1	0.26	0	0.10	0
	72	0.7	0.48	0.27	0.04	0

Table 7. Comparison between ranking estimated LIED nodes based on E_{glob} and LI using ℓ_1 norm for patient 2.

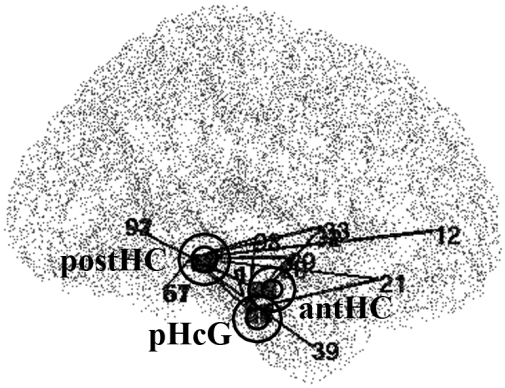
	E_{glob}	ℓ_1 norm	LI	ℓ_1 norm
source1	70	1	72	1
source2	72	0.96	71	0.9
source3	71	0.92	70	0.12

4. Discussion

E_{glob} and K_{tot} provide more comparable results to LI measure among the classical measures studied (E_{glob}, E_{loc}, and K_{tot}) and E_{tglob}. Here we summarize the advantages and disadvantages of LI measure over E_{glob} and K_{tot} measures. We start with the differences in their definitions. LI can be interpreted as the K_{tot} of a weighted digraph in which the weight related to each edge is the amount of information carried by the edge (Section 2.2.2). Accordingly, E_{glob} can be interpreted as the K_{out} of a weighted digraph, except the weight of the edge between nodes i and j is the inverse of the shortest path length between these nodes. The other difference between E_{glob},



(a)



(b)

Figure 2. (a) dDCG for patient 3 in 4-8 Hz, (b) digraph of (a) plotted in real channel coordinates superimposed on the 3D anatomical mesh (sagittal view). The figure of part (b) is added for an anatomical representation to show where the graph and especially the three important regions of the digraph (anterior/posterior hippocampus and parahippocampal gyrus) are located in the brain, while the details of the digraph can be seen in part (a). These three regions are enclosed with circles in parts (a) and (b). The 3D rotatable version of figure of part (b) is available at http://www.gipsa-lab.inpg.fr/∼ladan.amini/mes_images/dDCGPatient3scale4_8Hz.fig. The Matlab software is needed to see this rotatable figure. ant/post/bas/m: anterior/posterior/basal/mesial; CG: cingulate gyrus; Cx: cortex; F: frontal; HC: hippocampus; Ins: insula; orbiF: orbitofrontal; pHcG: parahippocampal gyrus; T: temporal; TP: temporal pole. 12: antCG; 21: orbFCx; 28-33: Ins-mF1; 39: TP; 46-54: antbasT; 57-61: postbasT; 67-71: antHC; 79-83: postHC; 91-92: postT4; 98: antT1. All of bipolar channels are located in left side of the brain except bipolar electrode lead 1 (right antHC).

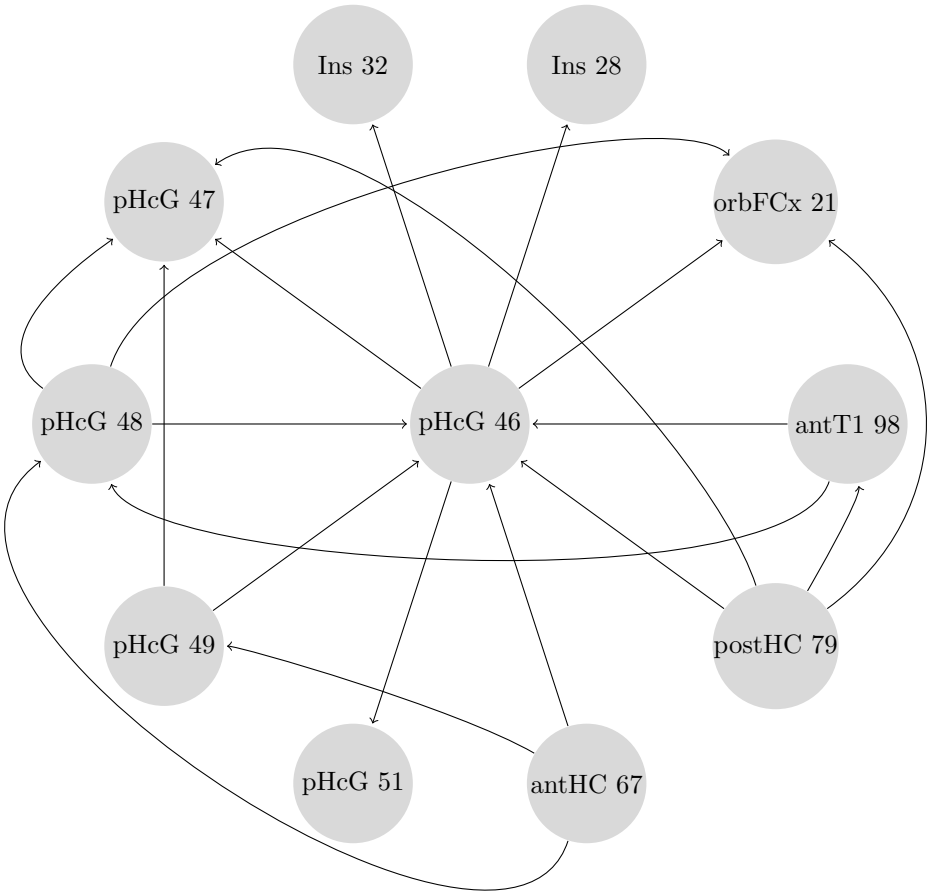
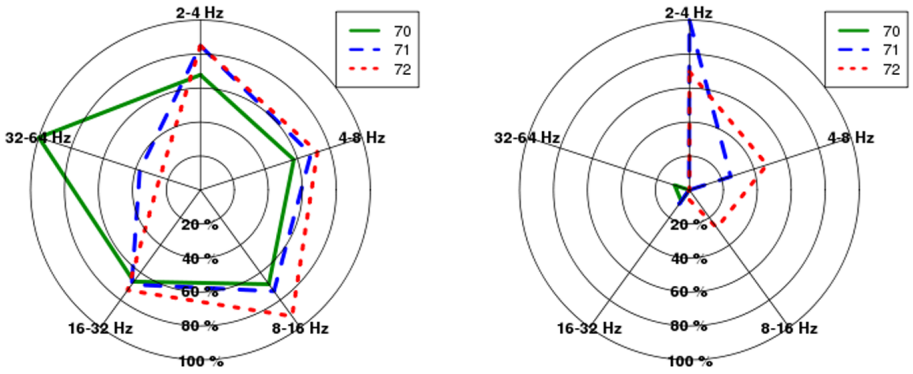


Figure 3. Directed subgraph of node 46 located in parahippocampal gyrus ($\mathbf{G}_{46^+}$) from the dDCG related to patient 3 in 4-8 Hz (figure 2). Each node is represented by its number of bipolar electrode lead and the related brain region. ant/post: anterior/posterior; pHcG: parahippocampal gyrus; Ins: insula; orbiF: orbitofrontal; Cx: cortex; T: temporal; HC: hippocampus.

K_{tot} and LI is that E_{glob} takes into account the higher-order neighborhood (global properties) while K_{tot} and LI consider the first-order neighborhood (local properties). Finally, E_{glob} and K_{tot} are deterministic measures, while LI is a stochastic measure and needs to be estimated correctly. We estimate the accuracy of LI by the estimation of its empirical standard deviation using jackknife resampling method§.

LI measure has several advantages over E_{glob} as follows.

§ For each node $i \in \mathbf{V}$, we calculate the $LI[i]$ (13) for N_w windows. W is the window length and T is the number of samples of underlying signals. The start time of each window is a random number in the range $[1, T - W]$. The standard deviation of $LI[i]$ for node i, is approximated as the standard deviation of N_w recomputed $LI[i]$ values. For more details see Amini et al. 2010.



(a) E_{glob} (b) LI

Figure 4. The demonstration of five dimensional measure vectors $\mathbf{q}_{70}$, $\mathbf{q}_{71}$, and $\mathbf{q}_{72}$ related to patient 2 for (a) E_{glob} and (b) LI in web plot. The five dimensions correspond to five frequency bands from 2 to 64 Hz.

1) LI evaluates each node based on its outgoing and incoming connections and the information which is carried by each of these connections. Therefore, LI uses the structure of the digraph in addition to the information extracted from signal pairs associated with each node pair, while E_{glob} uses only the structure of the digraph. Consequently, the sensitivity of LI is supposed to be less than E_{glob} to the minor changes in the digraph structure.

2) LI measure evaluates the amount of information of both outgoing and incoming connections which increases its efficiency in the definition of source and sink nodes, while E_{glob} considers neither the incoming paths, nor the amount of information related to each edge. A node with high E_{glob} may have negative, zero, or positive LI value because in addition to outgoing connections it may receive several incoming connections as well. Moreover, each outgoing or incoming edge may carry different amounts of information. Hence contrary to E_{glob}, LI measure has the potential to be used for distinguishing between source, transit and sink nodes. To analyze an IED related digraph like dDCG, this information is valuable to characterize the role of different brain regions. Although being all of the patients seizure-free after resective surgery is very valuable, currently there is no way to determine if the removed area is too large. If this is the case, more conservative surgery can be preferred and distinguishing between source, transit and sink regions can be valuable for this purpose.

3) LI is more suited for ranking the LIED nodes than E_{glob}. The node which has greater amount of positive local information is more relevant to a source. Consequently LI which is designed for measuring the emittance contribution is more skilled in ranking the LIED nodes than E_{glob}.

Compared with K_{tot}, LI measure is preferred since LI is a weighted version of K_{tot}, whose weights are the amount of information. Thus, for source and sink definition, LI is more suitable than K_{tot}.

The main disadvantage of LI over E_{glob} and K_{tot} is the computation load. Furthermore, LI is more time consuming since for a proper estimation of mutual

information including the variance, a long period (about one hour with sampling rate equal to 512 Hz) of signal pairs is supposed to be selected. We compare the computation time for E_{glob}, K_{tot} and LI for example for the dDCG related to patient 3 in 4-8 Hz (figure 2). This digraph includes 29 nodes and 62 edges. On a shared 3 GHz, 4 core Xeon 64 bits processor, the computation times for E_{glob} and K_{tot} are equal to 0.6 seconds, while it is 9.7 minutes for LI.

5. Conclusion

We compared different digraph measures with recently introduced LI measure for the source and sink node identification of digraphs. The comparison is evaluated by estimating the leading IED regions from multiple digraphs (digraphs of different frequency bands) related to interictal events for five epileptic patients. The estimated leading IED regions based on different measures are compared with visually inspected SOZ by epileptologist. In this perspective, LI measure is more informative in comparison with usual measures and E_{tglob} for the purpose of source and sink node identification. However, comparing the values of different measures like E_{glob}, K_{out}, K_{in}, and LI provides complementary characterizing information based on different definitions for more profound analysis of the role of each node in the digraph and their specifications.

For source and sink distinction, E_{glob} and K_{tot} outperform the conventional measures and E_{tglob}. E_{tglob} does not increase the precision of E_{glob} in the estimation of LIED regions, which shows that including the incoming paths may not be sufficient without considering the amount of information of each edge. However, LI is preferred in comparison with E_{glob} and K_{tot}, since it is more informative and skilled for the source and sink node detection and ranking the estimated LIED nodes. The power of LI relies on taking into account the amount of information carried by each edge and considering incoming connections in addition to outgoing ones.

6. References

Achard S, Salvador R, Whitcher B, Suckling J & Bullmore E 2006 *The Journal of Neuroscience* **26**, 63–72.

Alarcon G 1996 *Seizure* **5**(1), 7–33.

Alarcon G, Seoane J J G, Binnie C D, Miguel M C M, Juler J, Polkey C E, Elwes R D & Blasco J M O 1997 *Brain* **120 (Pt 12)**, 2259–2282.

Albert R K, Albert R & Barabási A L 2002 *Reviews Of Modern Physics* **74**, 47–97.

Amini L, Jutten C, Achard S, David O, Soltanian-Zadeh H, Hossein-Zadeh G A, Kahane P, Minotti L & Vercueil L 2009 *in* 'IEEE International Workshop, MLSP09, Grenoble, France'.

Amini L, Jutten C, Achard S, David O, Soltanian-Zadeh H, Hossein-Zadeh G A, Kahane P, Minotti L & Vercueil L 2010 *Submitted to IEEE Trans. Biomed. Eng.* .

Benesty J, Chen J & Huang Y 2004 *Speech and Audio Processing, IEEE Transactions on* **12**(5), 509 – 519.

Blauwblomme T, Kahane P, Minotti L, Grouiller F, Krainik A, Vercueil L, Chabardès S, Hoffmann D & David O in press *Journal of Neurology, Neurosurgery & Psychiatry* .

Boccaletti S, Latora V, Moreno Y, Chavez M & Hwang D U 2006 *Physics Reports* **424**(4-5), 175–308.

Branke J, Deb K, Miettinen K & Slowinski R 2008 *Multiobjective Optimization, Interactive And Evolutionary Approaches* Springer.

Bullmore E & Sporns O 2009 *Nature Reviews Neuroscience* . doi: 10.1038/nrn2574.

Carter G 1981 *Acoustics, Speech and Signal Processing, IEEE Transactions on* **29**(3), 463 – 470.

Cohn R & Leader H S 1967 *Electroencephalography and Clinical Neurophysiology* **22**(5), 421 – 428.

Deb K 1999 *Kanpur Genetic Algorithms Lab (KanGal), Technical report 99002* .

Friston K 1994 *Human Brain Mapping* **2**(1-2), 56–78.

Gotman J 1983 *Electroencephalogr Clin Neurophysiol* **56**(5), 501–514.
Horwitz B 2003 *NeuroImage* **19**(2), 466 – 470.
Hufnagel A, Dumpelmann M, Zentner J, Schijns O & Elger C E 2000 *Epilepsia* **41**(4), 467–478.
Ianniello J 1982 *IEEE Trans. Acoust., Speech, Signal Process.* **30**(6), 998 – 1003.
Kahane P, Minotti L, Hoffmann D, Lachaux J & Ryvlin P 2004 *Handbook of Clin. Neurophysiol. Pre-surgical assessment of the epilepsies with clinical neurophysiology and functional neuroimaging (F. Rosenow and H. Lüders, eds.): Elsevier Science* .
Ktonas P & Mallart R 1991 *Electroencephalography and Clinical Neurophysiology* **78**(2), 105 – 110.
Lai Y, van Drongelen W, Hecox K, Frim D, Kohrman M & He B 2007 *Epilepsia* **48**(2), 305–314.
Latora V & Marchiori M 2003 *Eur. Phys. J.* **32**, 249–263.
Le Van Quyen M, Adam C, Baulac M, Martinerie J & Varela F J 1998 *Brain Res* **792**(1), 24–40.
Monto S, Vanhatalo S, Holmes M D & Palva J M 2007 *Cereb Cortex* **17**(6), 1386–1393.
Newman M E J 2003 *SIAM Review* **45**(2), 167–256.
Nunez P L & Srinivasan R 2006 *Electric Fields of the Brain: The Neurophysics of EEG* Oxford University Press, Inc.
Ortega G J, Sola R G & Pastor J 2008 *Neurosci Lett* **447**(2-3), 129–133.
Sporns O, Chialvo D R, Kaiser M & Hilgetag C C 2004 *Trends in Cognitive Science* **8**, 418–425.
Watts D J & Strogatz S H 1998 *letters to nature* **393**, 440–442.
Wendling F, Bartolomei F & Senhadji L 2009 *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **367**, 297–316.
Whitcher B, Guttorp P & Percival D 2000 *J. Geophys. Res. - Atmospheres* **105**, 14941–14962.
Wilke C, van Drongelen W, Kohrman M & He B 2009 *Clin Neurophysiol* **120**(8), 1449–1456.