33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Multiscale Information for Network Characterization in Epilepsy

Catherine Stamoulis, *Member IEEE*, Bernard S. Chang

Abstract— **We have developed a multiscale approach for the estimation of neuronal network coordination in the epileptic brain, from continuous (long-term) non-invasive electroencephalograms (EEG). The proposed approach specifically assesses the effect of large-scale network behavior on local network coordination, at individual dominant frequencies (modes) of the EEG spectrum. For this purpose a set of conditional information parameters is proposed to explicitly quantify the effect of global network correlation in the brain on pairwise (local) mutual information, via conditioning. These parameters are shown to be modulated in a frequency-specific manner at baseline, as well as during seizure evolution.**

I. INTRODUCTION

Mechanisms of modulation of local brain networks at the meso-scale by neural activity at the macro-scale are only partially understood. Improved knowledge of these mechanisms is necessary in order to gain insights into the interaction of neural dynamics at different spatial scales, and may also provide very important insights into the neurophysiological correlates of neurological disorders, such as epilepsy, a common disorder that significantly affects neural dynamics. To date, it remains unclear whether seizure evolution is facilitated by a global, abnormal state of network coordination in the pre-ictal interval, or is triggered by local seizure-related activity originating in the epileptogenic region.

In order to quantify multiscale network interactions from inherently variable electroencephalograms (EEGs), an appropriate, robust to this variability mathematical framework is necessary. Information theory is such a framework, and enables the estimation of local network interactions and their dependence on large-scale dynamic neural states or network coordinations, by conditioning mutual information on spatially global parameters [1][11]. Previous studies have developed and applied information theoretic measures to electrophysiological signals, e.g., [14][10]. A few studies have also been specific to epilepsy, e.g., [9] applied *transfer entropy* for epileptic source localization, [5] proposed *permutation conditional mutual information* as a measure of directional coupling, and [8] estimated modulations of relative entropy in the time-frequency domain during the ictal interval. However, none of the studies addresses the problem of multiscale spatial network interactions, nor do they estimate the effects of global neurodynamic network changes on local networks. Finally, network and/or information theoretic studies involving EEGs typically examine the broadband spectrum of these signals either bounded by the sampling frequency or often at frequencies $\leq$80 Hz. However, it may be more meaningful to examine network coordination measures at the dominant frequencies (or modes) of the EEG. The structural anisotropy of the brain is likely to affect individual frequencies more specifically than entire bands, and thus analysis at the mode level may provide more specific information on directional EEG modulations.

We developed conditional information measures for quantifying the dependence of pairwise (local) network coordination on large-scale and global dynamic synchrony across the brain. These measures were estimated dynamically, for each dominant frequency (mode) of the EEG in the range 1-250 Hz. Modes were estimated using an Empirical Mode Decomposition (EMD) approach, modified to improve the data signal-to-noise ratio (SNR) [4][13]. Mutual information measures were conditioned i) on global or hemisphere-specific mode time-varying cross-correlation and ii) cross-frequency correlations, i.e., mutual information was conditioned on mean global network correlation either at the same frequency or at different modal frequencies, respectively. Cross-frequency coupling in the brain has been previously reported, e.g., modulation of high-frequencies (in the gamma band) by lower frequencies ($\leq$20 Hz) has been reported, even between distant brain regions [2][7][3], as well as during motor behaviors [6]. This analysis was applied to continuous non-invasive recordings from 5 patients with partial focal seizures originating in the temporal lobes.

II. MATERIALS AND METHODS

The proposed methodology consists of: 1) segmentation of continuous non-invasive multi-channel EEGs, collected over a period of multiple days, 2) individual signal decomposition into components with significant amplitude contributions (modes) and distinct characteristic frequencies, 3) estimation of conditional information theoretic parameters, for each segment and dominant component, and thus at each dominant frequency of each EEG signal, to obtain a time-frequency information spatial map of dynamic local network interactions.

A. Data

Continuous (long-term) EEG data were recorded at Beth Israel Deaconess Medical Center, Boston MA, in the Clinical Neurophysiology Laboratory of the Comprehensive Epilepsy Center. All data were part of inpatient, clinical electrophysiology studies for patient evaluation and monitoring, typically spanning several days. Data were recorded with a standard

This work was supported by a Harvard Catalyst grant, and the Harvard Clinical and Translational Science Center (NIH Award #UL1 RR 025758)

C. Stamoulis is with the Departments of Neurology and Radiology and the Clinical Research Program, Children's Hospital Boston and Harvard Medical School, Boston, MA 02115 USA caterina@mit.edu, caterina.stamoulis@childrens.harvard.edu

B.S. Chang is with the Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston MA 02215

TABLE I

PATIENT CLINICAL DEMOGRAPHICS.

Patient #	Age (yrs)	Etiology	Seizure focus	# Recording hrs	# Seizures
1	47	Cryptogenic	L temporal	63.5	3
2	27	Head injury	L temporal	193.5	7
3	24	Cryptogenic	L/R/simult-bilat. temporal	40	4
4	23	Brain malformation	L temporal	48	1
5	27	Cryptogenic	R/L temporal	137.5	6

international 10-20 EEG system and a referential montage, and were sampled at 500 Hz. EEG data are typically contaminated by both noise and artifacts. Power-line noise was attenuated with a stopband filterbank, centered at the 60 Hz harmonics of the noise, in the range 60-250 Hz, with a 1 Hz bandwidth for center frequencies $\leq$ 150 and a 2 Hz bandwidth for center frequencies $>$ 150 Hz. Second order elliptical filters (10 dB attenuation in the stopband, 0.5 dB ripple in the passband) were used. Signals were filtered in both forward and reverse directions to eliminate potential phase distortions due to the non-linear phase of the filter. Artifacts associated with eye blinking and muscle movement were suppressed using a stopband-type matched filter [12]. In addition, continuous recordings typically include segments where patients are disconnected from the recording system. These intervals were removed from the data prior to the analysis. Finally, a board-certified neurologist (B.S.C.) identified all ictal onset and offset times, according to standard clinical methods of visual EEG interpretation. Table I summarizes the patient demographics and data details.

B. Signal Decomposition and Mode Selection

EEG signals are non-stationary, with dynamically varying frequency content and statistics. There are very few strictly data-driven methods for decomposing non-stationary signals, which make no a priori assumptions on the shape of the unknown EEG signal components (referred to as modes). A widely-used method is the *Empirical Mode Decomposition* (EMD) [4], which recursively decomposes a non-stationary signal into modes with significant amplitude contributions. EMD involves fitting polynomials (splines) through local extrema of a waveform to obtain signal envelopes, which are averaged and subtracted from the data to obtain the desired modes. Although a powerful method and currently used in many signal processing applications, EMD does not necessarily estimate an *optimum* set of modes. Thus, additional criteria for mode selection need to be imposed, depending on the application. Here we selected a subset of modes estimated by EMD, based on the following criterion, first developed in an unrelated to EEG study [13]. An EEG signal $x_i(t)$ at electrode i may be decomposed into M modes $d_{i,M}$, which are then superimposed to obtain the estimate $\hat{x}_i(t) = \sum_{m=1}^{M} d_{i,m}(t)$. Note that although a signal may be perfectly reconstructed from a large set of modes, in general only a quasi-optimum subset of modes is selected and $\hat{x}(t)$ is an estimate of the measurement $x(t)$. Thus, for a true source signal $s(t)$, $x_i(t)$ and $\hat{x}_i(t)$ are a measurement and an estimate of this signal, respectively, modulated by the medium in the direction i between source and electrode i, and corrupted by direction-specific noise, i.e.,

$$\begin{aligned} x_i(t) &= s(t) + \varepsilon_i(t) \\ \hat{x}_{i,M}(t) &= \hat{s}_M(t) + w_{i,M}(t) \end{aligned} \tag{1}$$

where $\hat{s}_{i,M}(k)$ is the estimate of the true signal $s(t)$, based on M modes, and $\varepsilon_i \sim N(0, \sigma^2_{\varepsilon_i})$ and $w_{i,M} \sim N(0, \sigma^2_{w_{i,M}})$, with $\sigma^2_{w_{i,M}} < \sigma^2_{\varepsilon_i}$ are the additive Gaussian noise in $x(t)$ and $\hat{x}(t)$, respectively, in the direction of electrode i. In order to optimize a set of estimated modes to represent the true signal rather than noise contributions, we require that the variance of the estimated signal $\hat{s}(t)$ is less than the variance of the noise in the measured signal, i.e., $\sigma^2_{\hat{s}} \leq \sigma^2_{\varepsilon_i}$, When modes are sequentially eliminated from a signal, the noise variance progressively decreases, i.e., $\sigma_{\varepsilon_i} > \sigma_{w_{i,m=1}} > \sigma_{w_{i,m=2}} > \cdots > \sigma_{w_{i,m=M}}$. However, this process may also eliminate significant components, resulting in large signal reconstruction errors. Thus, we need to bound the number of eliminated modes. In a previous study [13], we defined a cost function as:

$$\begin{aligned} R_{m^-}(\hat{s}_{m^-}, s) &= \frac{1}{T}\sum_{k=1}^{T}(x_i(t) - \hat{x}_{i,m^-}(t))^2 \overset{(1)}{=} \\ &= \frac{1}{T}\sum_{t=1}^{T}(s(t) - \hat{s}_{m^-}(t) + \varepsilon_i(t) - w_{i,m^-}(t))^2 = \end{aligned} \tag{2}$$

where $(\cdot)_{m^-}$ refers to a signal in which mode m has been removed. When the reconstructed signal is the optimum estimate of the $s(t)$, then $\hat{s}_M(t) \approx s(t)$, and R_{m^-} becomes

$$\begin{aligned} &R_{m^-,optim}(\hat{s}_{m^-}, s) \approx \frac{1}{T}\sum_{t=1}^{T}(\varepsilon_i(t) - w_{i,m^-}(t))^2 = \\ &\frac{1}{T}\sum_{t=1}^{T}\varepsilon_i^2(t) + \frac{1}{T}\sum_{t=1}^{T}w_{i,m^-}(t)^2 - \frac{2}{T}\sum_{t=1}^{T}\varepsilon_i(t)w_{i,m^-}(t) = \\ &= \sigma^2_{\varepsilon_i} + \sigma^2_{w_{i,m^-}} - 2Cov(\varepsilon_i, w_{i,m^-}) \end{aligned} \tag{3}$$

which ensures the desired $\sigma^2_{\hat{s}} \leq \sigma^2_{\varepsilon_i}$ (assuming w_{i,m^-} and ε_i are positively correlated), with equality when $\varepsilon_i(t)$ or $w_{i,m^-}(t)$ is zero, i.e., a noiseless measurement $x(k) = s(t)$ or noiseless estimate $\hat{x}(t) = \hat{s}_{m^-}(t)$, and asymptotic equality when $T \to \infty$. The proposed approach has the advantage that it does not require detailed estimation of any thresholds for mode selection, although it may result in a quasi-optimum set of modes. Setting the variance of the mode-reduced signal equal to the cost function ensures that mode-reduced EEGs will have higher SNR than the raw signals. Each EEG signal was segmented into 6s intervals, empirically estimated as

an adequate data window that captures the dynamic (and statistical) variation of the EEG. Modes, and subsequently information parameters, were estimated at each interval, and thus at each characteristic (dominant) frequency.

C. Conditional Information Measures

Conditional mutual information of two random variables Y_i and Y_j given a third random variable Z is defined as [1]:

$$I(Y_i,Y_j|Z) = \sum_{y_i,y_j,z} p(y_i,y_j,z) \log \frac{p(Y_i,Y_j|Z)}{p(Y_i|Z)p(Y_j|Z)} \quad (4)$$

where $p(\cdot)$ denotes the probability mass function. We can think of (Y_i,Y_j) as random variables describing pairs of EEGs, measuring signal activity in the region covered by electrodes i and j, and Z as a time-varying process that may influence this pairwise interaction. For example, [14] estimated the information transfer between $Y_i(t)$ and $Y_j(t)$ at a future time interval, i.e., $Y_j(t+\Delta t)$, given the present, i.e., $Z = Y_j(t)$. Here we are interested in the relationship between neural activity at different spatial scales, and thus Z represents a global network coordination 'state' of the brain. Furthermore, I is estimated at individual modal frequencies, i.e., we have $I(Y_{i,m}(f_m,t),Y_{j,m}(f_m,t)|Z(f_l,t))$, where $l = m$ when assessing these effects for the same mode as that of individual EEGs, and $l \neq m$ when assessing cross-frequency effects. We, therefore, defined a set of Zs as follows:

$$Z_1(f_m,t) = \frac{1}{N} \sum_{i,j=1,i\neq j}^{N} C_{Y_i,Y_j}(f_m,t) \quad (5)$$

$$Z_2(f_{l\neq m},t) = \frac{1}{N} \sum_{i,j=1,i\neq j}^{N} C_{Y_i,Y_j}(f_{l\neq m},t) \quad (6)$$

Z_1 is the mean global cross-correlation between all EEGs (excluding autocorrelations), at mode m, Z_2 is the corresponding correlation at mode $l \neq m$, when the mutual information of Y_i and Y_j are calculated at mode m.

III. RESULTS

We first examined multi-modal mutual information (MI) and conditional MI at baseline, separately for each channel. Figure 1 shows these parameters over a baseline period of 100 min, for channel Fp1 and each dominant signal mode averaged over all pairwise estimates of this channel with all others. The top panel shows mean MI, and the bottom panel shows MI conditioned on mean cross-correlation averaged across the entire EEG array.

At baseline, MI between EEGs was highest at the lowest frequencies (≤ 2 Hz), with a quasi-periodic variation from low-to-high information, i.e., pairwise channel coordination, of 2-5 min. Conditioning on the global correlation of the brain, had a negligible effect on this parameter. Also, mutual information decreased with frequency and was lowest for modes with characteristic frequencies > 12 Hz.

We also examined the spatial variation (across channels) of single-mode, mean MI. Figure 2 shows an example of this variation at modal frequencies f=1,5,13 Hz. Although the large-scale EEG correlation over the entire brain did not appear to affect local (pairwise) correlations between channels, as shown in Figure 1, we identified baseline time intervals of temporally locked variation of information across many channels, both at very low frequencies (f=1 Hz) and at mid-frequencies (f>12 Hz). Thus, at specific frequencies and time points, individual networks may locally synchronize, resulting combinatorially in a large-scale coordinated network.

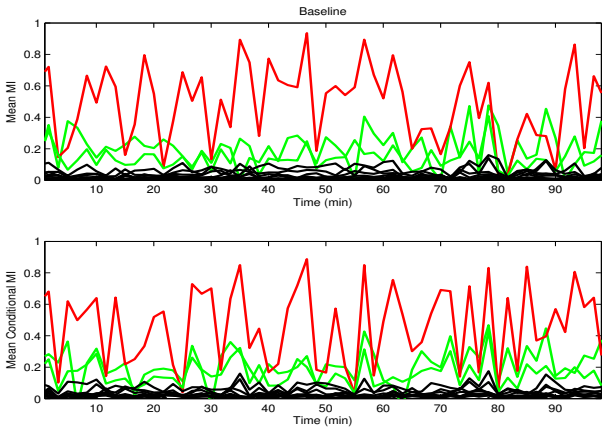

Fig. 1. Mean mutual (top) and conditional mutual information (bottom) between channel Fp1 and all others, at each dominant mode of the baseline EEG: lowest frequency mode (1 Hz) (red), modes in the range $3 < f < 8$ Hz (green), and modes $f > 12$ Hz (black).

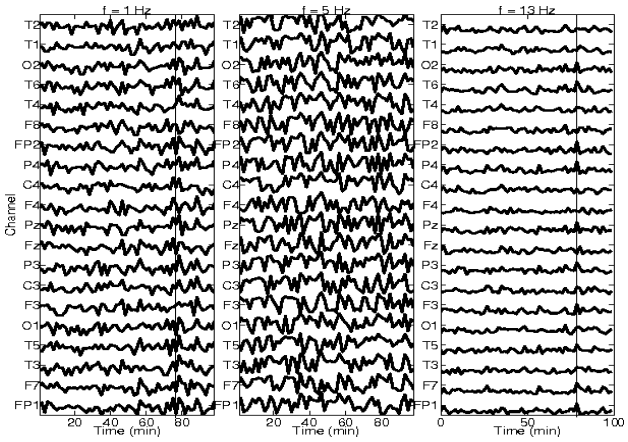

Fig. 2. Spatial variation of mean mutual information, at three dominant EEG frequencies: f=1 Hz (left), f=5 Hz (middle), f=13 Hz (right).

We estimated information parameters for the entire length of the EEGs, and specifically examined their variation prior to during and following seizures. Figure 3 shows an example of the variation of conditional MI at three time intervals: baseline (100 min), 70 min prior to seizure onset, and 42 min that included two seizures and short intervals between (<30 min) and after (<10 min). Very low-frequency (f=1-2 Hz) mean MIT did not vary significantly prior to seizure onset (at least at times $\geq$60 min prior to clinical onset) from baseline, and only slightly increased during ictal and in short inter-seizure intervals (<30 min). In contrast, conditional MI at frequencies $8 < f < 12$ Hz increased significantly during the

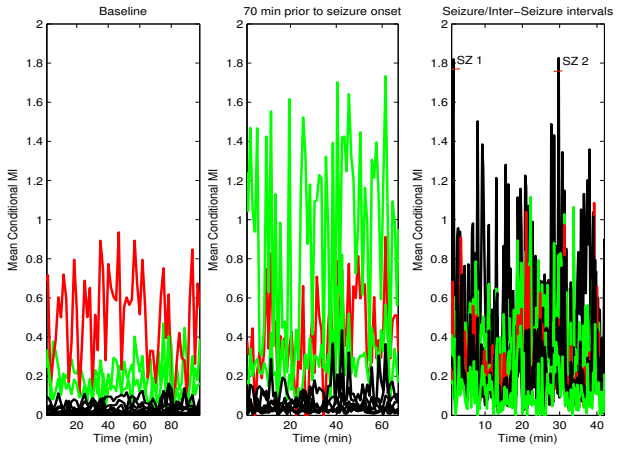

Fig. 3. Conditional mutual information across modes at baseline (left), 70 min prior to any seizure onset and 42 min including 2 ictal and inter/post-seizure intervals. Modes are marked as follows: f=1Hz (red), modes in the range $3 < f < 8$ Hz (green), and modes $f > 12$ Hz (black).

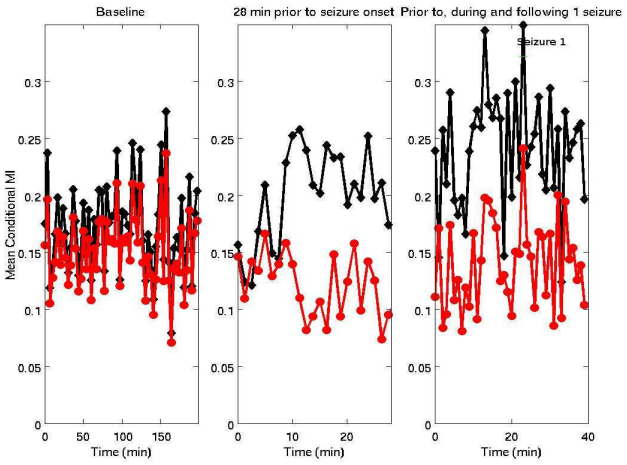

Fig. 4. Mutual information (red) and conditional mutual information (black), averaged over all modes, at baseline (left), 25 min prior to seizure onset and during a 40 min including 1 ictal event (at $\sim$ 22 min from the beginning of the interval) and several min of the post-ictal interval.

pre-ictal interval. Corresponding information at frequencies >50 Hz, increased during ictal periods, as well as between seizures occurring at short intervals from each other, indicating a strong frequency-dependence of these parameters and their potential modulation by seizure evolution. Finally, the periodicity of the dynamics of information measures observed at baseline was undetectable during seizure evolution.

We finally examined the effect of conditioning by cross-frequency correlations. As previously mentioned, conditioning had a negligible effect on MI at baseline. In contrast, approximately 20-45 min prior to seizure onset (depending on the patient), mean conditional MI, cross-frequency conditioned on mean (global) cross-correlation at *low* frequencies ($\leq$4 Hz) was higher then unconditioned information (Figure 4 middle panel). This was also the case during and following seizures, though during the seizure, conditioning on large-scale correlation resulted in relatively smaller increases in pairwise information.

IV. DISCUSSION

We have proposed conditional mutual information parameters to quantify pairwise, local coordination between EEGs at each dominant mode of these signals, and to assess the effect of global neurodynamic synchronization on local brain networks. In a preliminary study, we have applied this analysis to continuous EEGs from 5 patients with multiple seizures. At baseline, local network coordination, appeared to be highest at very low frequencies, and independent of global network correlations. In contrast, both in the pre-ictal (20-45 min) and ictal intervals, very low-frequency, global correlation appeared to facilitate the coordination of local networks, given increased conditional mutual information in these intervals, in comparison with unconditioned information. Finally, during seizure evolution the frequency distribution of these parameters varied significantly, from high information at very low frequencies, to increased information at frequencies $4 < f < 12$ Hz in the pre-ictal interval, to maximum information at frequencies $\geq$50 Hz during the ictal and post-ictal intervals. These are only preliminary, though promising results based on a small number of patients, which indicate frequency-specific and scale-specific network modulations during seizure evolution. Evidently, a larger seizure sample is necessary for validation.

REFERENCES

[1] T. M. Cover and J. A. Thomas (2006), *Elements of information theory*, Second edition, Wiley Interscience.

[2] R.T. Canolty, et al. (2006), High gamma power is phase locked the theta oscillations in human neocortex, *Science*, 313:1626-1628.

[3] F. Darvas et al. (2009), Nonlinear phase-phase cross-frequency coupling mediates communications between distant sites in human neocortex, *J. Neurosci.*. 29(2):426-435, 2009.

[4] N.E. Huang et al. (1998), Empirical Mode Decomposition and Hilbert spectrum for non-linear, non-stationary time series analysis, *Proc. R. Soc. Lond. A*, 454:903-995.

[5] X. Li, G. Ouyang (2010), Estimating coupling direction between neuronal population with permutation conditional mutual information, *Neuroimage*, (in press).

[6] K.J. Miller et al., Spectral changes in cortical surface potentials during motor movement, *J. Neurosci.* 27:2424-2432.

[7] F. Mormann, et al. (2005) Phase/amplitude reset and theta-gamma interaction in the human medial temporal lobe during a continuous word recognition memory task, *Hippocampus*, 15:890-900.

[8] R.Q. Quiroga et al. (2000), Kullback-Leibler and renormalized entropies: applications to electroencephalograms of epilepsy patients, *Phys. Rev. E*, 62(6):8380-8386.

[9] S. Sabesan, et al., (2009), Information flow and application to epileptogenic focus localization from intracranial EEG, *IEEE Trans. Neural Syst. Rehab. Eng.*, 17(3):244-253.

[10] T. Schreiber, (2000), Measuring information transfer, *Phys. Rev. Lett.*, 85(2):461-464.

[11] C. Stamoulis, L. Gruber, B.S. Chang (2011), Differentially modulated preictal brain network dynamics at frequencies >100 Hz, *Epilepsia* (in review)

[12] C. Stamoulis, B.S. Chang (2009), Application of Matched-Filtering to Extract EEG Features and Decouple Signal Contributions from Multiple Seizure Foci in Brain Malformations, *IEEE Proc. 4th Int. Conf. Neural Eng.*, 1:514-517.

[13] C. Stamoulis, R.A. Betensky (2011), A signal processing approach for the detection of copy-number variations in the human genome, *Bioinformatics*, in press.

[14] M. Vejmelka, M. Palus (2008), Inferring the directionality of coupling with conditional mutual information, *Phys. Rev. E*, 77:026214.