Entropy-based Multichannel Measure of Stationarity for Characterization of Motor Imagery Patterns

J.D. Martínez-Vargas[1] C. Castro-Hoyos[1†] G. Castellanos-Dominguez[1]

Abstract— **We propose a novel approach for measuring the stationarity level of multichannel time-series. This measure is based on stationarity definition over time-varying spectra and aims to quantify the relationship between local (single-channel dynamics) and global (multichannel dynamics) stationarity. With the purpose of separate among several motor/imagery tasks, we asssume that movement imagination implies an increase on the EEG variability, consequently, as discriminant features, we first compute the non-stationary components of input signals, and we further obtain its stationary level throughout the proposed measure. To assess the separability level of the proposed features, we employ the *t*-student test. Obtained results evidence that our measure is able to accurately detect brain areas projected on the scalp where motor tasks are performed.**

I. INTRODUCTION

Brain Computer Interface (BCI) relates communication and control system that creates a non-muscular output channel for the brain [1], [2]. BCI systems are based on the cognitive neuroscience paradigm termed Motor Imagery (MI) that consists on the imagination of a motor action without any actual movement, when patterns of the human sensorimotor functions are characterized [3].

In this regard, the most common employed method for monitoring brain activity is the electroencephalogram (EEG) that is a non-invasive technique with high temporal resolution and low-cost. Nevertheless, brain activity has strong spatio-temporal dynamics (non-stationarity) reflected in EEG dipoles measurements and it poses a challenge for accurate characterization of patterns related with movement tasks. Thus, the sources (cluster of aligned and synchronously activated/deactivated neurons) that produce different types of MI (e.g., right hand versus left hand) can be considered as spatially distinct [4], yielding non-stationary behavior in measured time-series.

One of the approaches proposed to cope with this issue is the piecewise stationary analysis, yet, the problem arises when measuring stationarity over real stochastic data. To this end, the weak sense stationarity is the most common definition that assumes time-invariability of the first and second statistical moments. Based on the weak sense stationarity, separation of stationary and non-stationary sources is carried out from multichannel recordings in [5], [6], while in [7] from a single time-series.

Nonetheless, to improve interpretability of the provided separation, some authors have proposed to include information about meaningful neurophysiological EEG spectral components [8]. Thus, spectro-temporal evolution allows measuring the stationary degree of a given time-series. Particularly, several indexes are addressed in [9] to measure the level of non–stationarity influence of a single channel signal. However, those proposed indexes, based on marginal frequency distributions and computed from a given time–frequency representation, are not bounded, and they rely on the energy distribution over time of a given signal. Furthermore, to deal with multichannel non–stationary signals estimated individual channel indexes might not reflect the actual relationship between local and global stationarity of signals.

To overcome the aforementioned problems, we propose as stationarity index the use of a kernel-based entropy of the marginal frequency distribution. This index measures the degree of certain stationary process according to the dynamic variability of multichannel signals. From obtained results on real EEG data, we show that non-stationary neural activity can be used to differentiate among several motor-imagery tasks.

II. METHODS

A. Signal separation filtering task

Let $\boldsymbol{X}^s \in \mathbb{R}^{N_c \times N_t}$ denote a multichannel stationary time-series, measured by N_c sensors at N_t time samples, that is assumed to be corrupted by an observed non-stationary multichannel signal $\boldsymbol{X}^n \in \mathbb{R}^{N_c \times N_t}$, so that the measured observation of linearly mixing signal is given by $\boldsymbol{X} = \boldsymbol{X}^s + \boldsymbol{X}^n$. The problem of separability is, by definition, to determine conditions on $\boldsymbol{X}^s$ and $\boldsymbol{X}^n$ such that an estimate of the desired signal $\widehat{\boldsymbol{X}}^s$ can be obtained, from filtered $\boldsymbol{X}$, to a given degree of accuracy. Consequently, observed time-series $\boldsymbol{X}$ can be model as a linear superposition of stationary sources $\boldsymbol{S}^s \in \mathbb{R}^{N_s \times N_t}$ and non-stationary sources $\boldsymbol{S}^n \in \mathbb{R}^{N_n \times N_t}$, where N_s and N_n denote the number of stationary and non-stationary sources, respectively, as follows [6]:

$$\boldsymbol{X} = \boldsymbol{A}\boldsymbol{S} = [\boldsymbol{A}^s \quad \boldsymbol{A}^n] \begin{bmatrix} \boldsymbol{S}^s \\ \boldsymbol{S}^n \end{bmatrix} \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{N_c \times N_c}$ is an invertible mixing matrix, whereas $\boldsymbol{A}^s \in \mathbb{R}^{N_c \times N_s}$ and $\boldsymbol{A}^n \in \mathbb{R}^{N_c \times N_n}$ are the stationary and non-stationary subspaces, respectively.

[1]Authors are working with the Universidad Nacional de Colombia, Sede Manizales. Signal Processing and Recognition Group.

† Corresponding author ccastroh at unal.edu.co

*This work is carried out under COLCIENCIAS grants: Programa Nacional de Formacion de Investigadores 2011 and 2013, and *Evaluación asistida de potenciales evocados cognitivos como marcador del transtorno por déficit de atensión e hiperactividad (TDAH).*

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

Therefore, we aim to factorize an observed time-series into both the stationary and non-stationary sources by finding the following inverse mixing matrix:

$$\boldsymbol{A}^{-1} = \boldsymbol{B} = [\boldsymbol{B}^{s\top}\boldsymbol{B}^{n\top}]^{\top} \tag{2a}$$

$$\boldsymbol{S}^{s} = \boldsymbol{B}^{s}\boldsymbol{X}, \quad \boldsymbol{S}^{n} = \boldsymbol{B}^{n}\boldsymbol{X} \tag{2b}$$

Thus, by splitting the N_t time samples into N_e, that is, splitting the time–series $\boldsymbol{X}$ into the set $\left\{\widehat{\boldsymbol{X}}_i : \forall i \in N_e\right\}$ epochs, each one with mean $\boldsymbol{\mu}_i \in \mathbb{R}^{N_c\times 1}$ and covariance matrix $\boldsymbol{\Sigma}_i \in \mathbb{R}^{N_c\times N_c}$, we consider the time series to be stationary in the weak sense *iff* the corresponding values of epoch mean and covariance equal to the average: $\boldsymbol{u}_i = \bar{\boldsymbol{u}}$ and $\boldsymbol{\Sigma_i} = \bar{\boldsymbol{\Sigma}}$;, where $\bar{\boldsymbol{u}} = \boldsymbol{E}\{\boldsymbol{u}_i : \forall N\}$ and $\bar{\boldsymbol{\Sigma}} = \boldsymbol{E}\{\boldsymbol{\Sigma}_i : \forall N\}$ are the average epoch mean and covariance matrix, respectively. This can be achieved by solving the following optimization problem:

$$\min_{\boldsymbol{B}^s} \operatorname{tr}[\boldsymbol{B}^s\boldsymbol{S}\boldsymbol{B}^{s\top}] \quad \text{s.t.} \quad \boldsymbol{B}^s\bar{\boldsymbol{\Sigma}}\boldsymbol{B}^{s\top} = \boldsymbol{I}_{N_c}, \tag{3}$$

where $\boldsymbol{I}_{N_c}$ is an identity matrix and $\boldsymbol{\Xi}$ is given by:

$$\boldsymbol{\Xi} = \boldsymbol{E}\{\boldsymbol{u}_i\boldsymbol{u}_i^{\top} + 2\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_i\} - \bar{\boldsymbol{u}}\bar{\boldsymbol{u}}^{\top} - 2\bar{\boldsymbol{\Sigma}}, \tag{4}$$

that can be represented by the following generalized eigenvalue problem:

$$\boldsymbol{\Xi\Phi} = \lambda\bar{\boldsymbol{\Sigma}}\boldsymbol{\Phi}. \tag{5}$$

Solution is given by a set of $\lambda_j \in \mathbb{R}, \boldsymbol{\phi}_j \in \mathbb{R}^{N_c\times 1} : \forall j \in N_c$ generalized eigenvalues and $\bar{\boldsymbol{\Sigma}}$-orthonormal eigenvectors, where the stationary projection $\boldsymbol{B}^s$ is given by the N_s eigenvectors with smallest eigenvalues, $\boldsymbol{B}^s = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{N_s}]^{\top}$, and the non-stationary projection is the remaining eigenvectors. Further description of this method, called Analytic Stationary Subspace Analysis (ASSA), is found in [5].

B. Multichannel Non–Stationarity Measure using Kernel Based Entropy

Let $\boldsymbol{x}_c \in \mathbb{R}^{N_t}$ be the $c-$th channel of the observed time-series $\boldsymbol{X}$, $\forall c \in N_c$, which is transformed to its time-frequency representation $\boldsymbol{\Omega}^c \in \mathbb{R}^{N_F\times N_t}$, where N_F stands for the number of frequency bins. Thus, we define the kernel-based marginal frequency entropy as [10], [9]:

$$h^c_{\boldsymbol{\Omega}}(f) = \frac{1}{1-\gamma}\log\left(\frac{1}{N_t{}^{\gamma}}\sum_{t=1}^{N_t}\left(\sum_{t'=1}^{N_t}\mathcal{K}_{\sigma}(\omega^c_t(f), \omega^c_{t'}(f))\right)\right)^{\gamma-1} \tag{6}$$

where $\omega^c_t(f) \in \mathbb{R}$ is the value of $\boldsymbol{\Omega}^c$ over the f-th frequency band at time instant t, $\mathcal{K}_{\sigma}(\cdot)$ is a Gaussian kernel function with parameter σ over its argument used for the Parzen non–parametric estimation of the probability density function, and $\gamma \in \mathbb{R}^+$ is the Rényi entropy degree. To normalize the measure within variability ranges of all-channels stationarity, each c channel entropy is normalized as follows:

$$\hat{h}^c_{\boldsymbol{\Omega}}(f) = h^c_{\boldsymbol{\Omega}}(f) / \sum\nolimits_{\forall c} h^c_{\boldsymbol{\Omega}}(f). \tag{7}$$

Therefore, we are measuring uncertainty over each frequency band along the time, thus, the higher the normalized entropy value, the more non-stationary the frequency band f of the channel c with respect to the stationary dynamics of the multichannel signal.

III. Experimental Set-Up

A. Motor movement/imagery (MI) database (Dataset I, BCI competition IV-2008)

To validate the proposed approach measuring the stationary degree of multichannel recordings, we use the well-known EEG motor imagery database publicly available [11]. Data hold recordings from 7 subjects, who were asked to perform two different motor/imagery tasks, at a given cue trigger, selected from three classes: i)left hand, ii) right hand, and iii) foot (side chosen by the subject). Position of the 59 EEG electrodes covers mostly sensorimotor area. Signals are band-pass filtered between the rank from 0.05 to 200 Hz, sampled at 1000 Hz. Preprocessing is carried out by means of a 10-order low-pass Chebyshev II filtering with 50dB stop-band ripple and 49 Hz stop-band frequency and down-sampled at 100 Hz. The database holds 100 runs for each MI class per subject during 4 s duration.

B. Separation between stationary and non-stationary signals

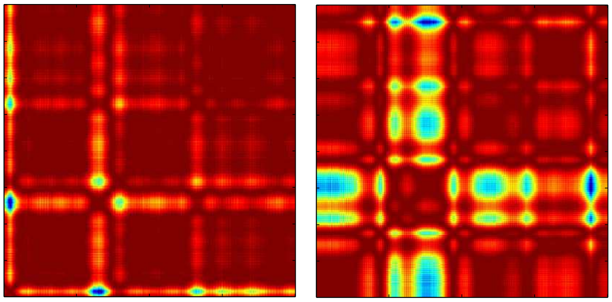
We use the ASSA algorithm to separate stationary and non-stationary components from EEG recordings. Thus, for each EEG trial $\boldsymbol{X_k} \in \mathbb{R}^{59\times 400}, k=1, \ldots, 200$, from Eq. (1), we compute the non-stationary signal by setting as an all-zeros matrix $\boldsymbol{0} \in \mathbb{R}^{59\times N_s}$ the columns belonging to the stationary components in the mixing matrix $\boldsymbol{A}$, and we use a similar procedure to estimate the stationary signal. For this purpose, all ASSA parameters are empirically tuned, namely, we use 40 epochs, while the number of non-stationary sources N_n is obtained as the elbow of the singular values curve $\boldsymbol{\lambda}$ decreasingly ranked, while the number of stationary components is set as the remaining eigenvalues. Therefore, for each trial, we obtain an estimation of its stationary $\boldsymbol{X}^s_k$ and non-stationary $\boldsymbol{X}^n_k$ components.

C. Entropy-based stationary measure

Once we obtain $\boldsymbol{X}^s_k$ and $\boldsymbol{X}^n_k$, we compute the stationary measure of each channel, as shown in Eq. (7). To this end, the Short Time Fourier Transform with 256 frequency bins resolution is employed using a hamming window lasting 63 samples length. The entropy degree is selected as $\alpha = 3$, and the Gaussian kernel parameter is tuned according to each channel variance. Figure 1 shows the Kernel used for computing the entropy in a randomly selected trial of Subject 7 in the C_1 channel, and for a randomly selected frequency band. As expected, the kernel for the stationary signal (Figure 1(a)) has closer values (smooth image) than the kernel for the non-stationary signal (Figure 1(b)).

D. Separability between motor-imagery tasks

Meanly, the different dynamics of motor/imagery tasks are assumed to be highly concentrated within specific frequency bands [2]. Thus, we calculate the stationarity index in the following bands: $\delta(0-4)\,Hz$, $\theta(4-8)\,Hz$, $\alpha(8-15)\,Hz$ and $\beta(15-30)\,Hz$. Besides, it has been shown that non-stationary data is most informative for detecting state changes of the time series [12]. Thus, we assume that any movement response (MI tasks), produced by the activation of sources



(a) Gassian Kernel for stationary signal (b) Gassian Kernel for non-stationary signal

Fig. 1. Influence of non-stationarities for the Gaussian kernel construction.

in a particular region on the brain, should increase the non-stationary dynamics (activation of sources), and thus, source activation leads to higher levels of entropy. Consequently, to separate among MI classes, we use as discriminant features the sum of the entropy-based stationary measure in the channel c of the filtered non-stationary signal $\boldsymbol{X}_k^n$, relating to one of the four estimated EEG frequency bands as follows:

$$\xi^c = \sum_{\forall f \in f_b} \hat{h}_{\boldsymbol{S}}^c(f), \tag{8}$$

where $\hat{h}_{\boldsymbol{S}}^c(f)$ is the stationary measure computed over the $c-$th channel of $\boldsymbol{X}_k^n$ and f_b is one of the four aforementioned frequency ranks. Additionally, to compare the proposed characterization, we use the energy content of original time-series as input feature. Figure 2 shows the energy mapped to the scalp of an EEG recording randomly selected from subject 7 of the original signal and its respective estimated stationary and non-stationary dynamics.

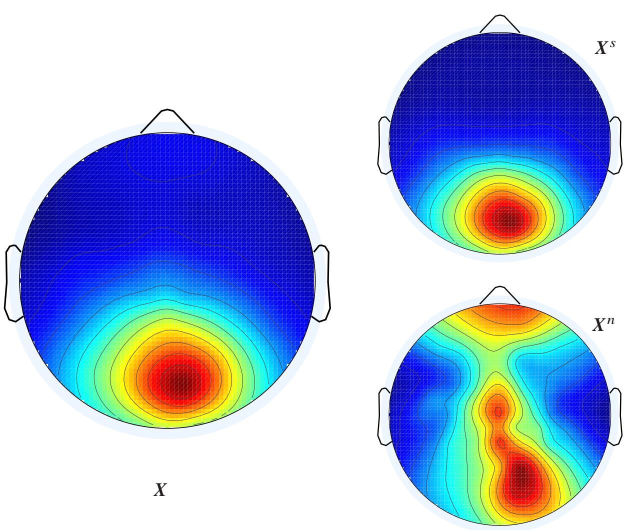
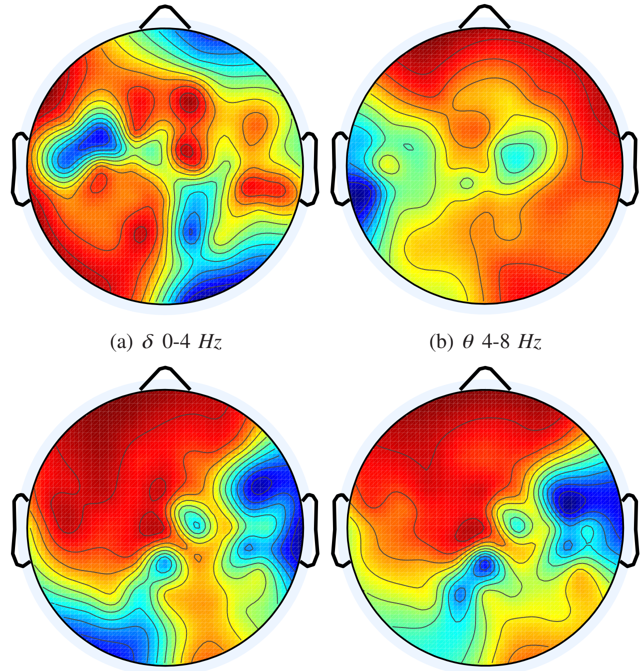


Fig. 2. Example of a single-trial randomly selected EEG recording and its energy topological plot.

Figure 3 shows an example of estimated features for the same subject in Figure 2; as seen, there are different energy values distributed over the left/right sensorimotor areas that should be related to hand movement tasks in the α and β bands.



(a) δ 0-4 Hz (b) θ 4-8 Hz (c) α 8-16 Hz (d) β 16-32 Hz

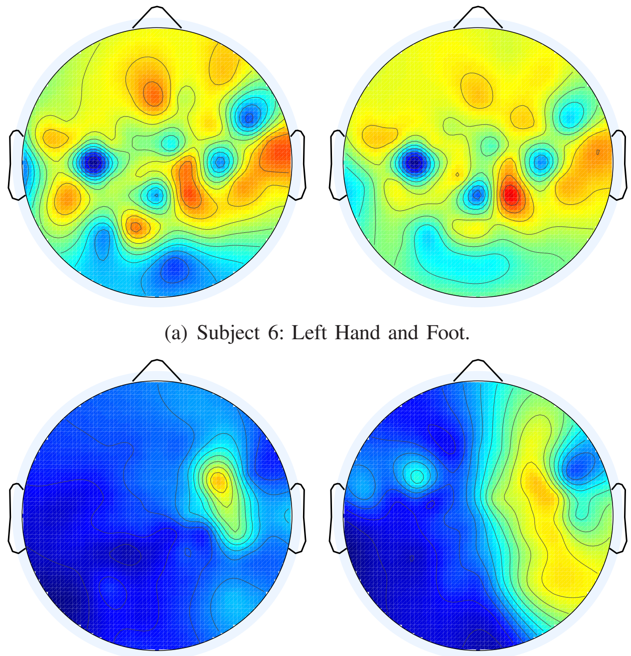
Fig. 3. Stationary measure of estimated non-stationary signal $\boldsymbol{X}^n$ for left hand movement.

Summarizing, we obtain five different feature sets, namely Ξ_δ, Ξ_θ, Ξ_α, Ξ_β and $\Xi_{\boldsymbol{X}}$, each $\in \mathbb{R}^{200\times 59}$ (number of trials $\times$ number of features that, in this case, corresponds with the number of EEG channels). Lastly, to make clear differences among activated brain areas performing each motor/imagery task, we carry out a paired t-student test by using as input each computed feature matrix, with the null hypothesis that there are not differences between MI classes. Figure 4 displays statistical results for most discriminative used features for two different subjects.

IV. Discussion and conclusion remarks

We discussed a novel approach for measuring the stationary level of multichannel recordings. This approach that is based on the uncertainty of time-varying spectra measures the single-channel stationary dynamics accordingly to the global stationarity of the multichannel recording. Experiments that are carried out over a BCI database show that measured variations in the non-stationary dynamics are useful to distinguish between motor/imagery tasks.

The used ASSA method to obtain estimated stationary and non-stationary signal components takes advantage of the weak-sense stationarity definition. Nonetheless, it is important to remark that parameters tuning in the separation method may affect the estimation results, which are not considered in the present work. In this sense, it would also be useful to include other moments as the kurtosis or skewness for which non-Gaussian assumptions are to be imposed.



(a) Subject 6: Left Hand and Foot.

(b) Subject 7: Left Hand and Right Hand.

Fig. 4. Results of the *t*-test for α and β frequency bands (left and right columns respectively) for two different subjects.

Regarding the proposed measure, we initially use the spectrogram to compute the time-varying spectrum for the sake of simplicity. Nevertheless, there are not restrictions about the kind of the involved TFR. In fact, used spectral representation can be modified according to the considered EEG properties. Lastly, as information measure, we use the Renyi's entropy [8]. Figures 1(a) and 1(b) show the used Gaussian kernels to compute the stationarity measure. In these Figures, it is readily noted that non-stationary recording presents less uniform behavior because of the sudden changes in the distances among frequency bands.

Figure 3 shows an example of the estimated features based on the stationary measure over the non-stationary estimation $\boldsymbol{X}_k^n$, for a left-hand movement task of subject 7. Activation in the frontal lobe (the electrodes fixed around the eyes) can be explained as an ocular artifact, that can be evidenced by the observed higher values in the θ and α bands. Additionally, response to the trigger cue can be seen in the scalp visual area. Furthermore, in α and β bands, highly concentrated activity is shown in the sensory motor area (centrally positioned electrodes); this activity should be related to the movement task. To compare this results, Figure 2 shows the signal energy mapped to the scalp. We see that there is no relationship between activated zones and motor-imagery tasks. Nevertheless, it can be seen that most information is found in the non-stationary estimation (lower-right plot). Even in that cases when frontal and visual areas are activated, the sensory motor area gets no response associated with hand movements. Consequently, we can infer that the proposed measure of stationarity can distinguish among different motor-imagery classes.

It is worth noting that to assess the proposed measure as a possible discriminant feature, we use the t-student test score to compare differences between active brain areas during different motor-imagery tasks. According to the results shown in Figure 3, we use α and β bands as discriminant features. Figure 4 shows obtained results. As seen, main differences are found around the leg sensory motor for the Subject 6 for whom we consider the left-hand and foot tasks. In contrast, prominent differences are focused around the hand movement motor area for the Subject 7, when left and right hand movements are considered. Those results display the measure ability for distinguishing among motor-imagery tasks being related to sudden changes of EEG recordings. In any case, we are providing a novel approach for measuring stationarity of multichannel time-series instead of a characterization approach.

As future work, Authors plan to use the obtained measure for training a BCI classification system, and to take the measure of stationarity into an on-line framework.

References

[1] X. Yu, P. Chum, and K.-B. Sim, "Analysis the effect of PCA for feature reduction in non-stationary EEG based motor imagery of BCI system," *Optik - International Journal for Light and Electron Optics*, vol. 125, no. 3, pp. 1498 – 1502, 2014.

[2] T.-H. Nguyen, S.-M. Park, K.-E. Ko, and K.-B. Sim, "Multi-class stationary csp for optimal feature separation of brain source in bci system," in *Control, Automation and Systems (ICCAS), 2012 12th International Conference on*, Oct 2012, pp. 1035–1039.

[3] A. Alvarez-Meza, L. Velasquez-Martinez, and G. Castellanos-Dominguez, "Feature relevance analysis supporting automatic motor imagery discrimination in eeg based bci systems," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, July 2013, pp. 7068–7071.

[4] C. Gouy-Pailler, M. Congedo, C. Brunner, C. Jutten, and G. Pfurtscheller, "Nonstationary brain source separation for multiclass motor imagery," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 2, pp. 469–478, Feb 2010.

[5] S. Hara, Y. Kawahara, T. Washio, P. von Bünau, T. Tokunaga, and K. Yumoto, "Separation of stationary and non-stationary sources with a generalized eigenvalue problem," *Neural Networks*, vol. 33, no. 0, pp. 7 – 20, 2012.

[6] P. von Bünau, F. C. Meinecke, F. C. Király, and K.-R. Müller, "Finding stationary subspaces in multivariate time series," *Phys. Rev. Lett.*, vol. 103, p. 214101, Nov 2009.

[7] J. Martinez-Vargas, D. Cardenas-Pena, and G. Castellanos-Dominguez, "Extraction of stationary components in biosignal discrimination," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, Aug 2012, pp. 1–4.

[8] C. Cao and S. Slobounov, "Application of a novel measure of EEG non-stationarity as shannon- entropy of the peak frequency shifting for detecting residual abnormalities in concussed individuals," *Clinical Neurophysiology*, vol. 122, no. 7, pp. 1314 – 1321, 2011.

[9] S. Tong, Z. Li, Y. Zhu, and N. V. Thakor, "Describing the non-stationarity level of neurological signals based on quantifications of Time-Frequency representation," *IEEE Transactions of Biomedical Engineering*, vol. 54, no. 10, pp. 1780–1785, 2007.

[10] J. Principe, *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*, ser. Information Science and Statistics. Springer, 2010.

[11] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. MÃijller, and G. Curio, "The non-invasive berlin brainâĂŞcomputer interface: Fast acquisition of effective performance in untrained subjects," *NeuroImage*, vol. 37, no. 2, pp. 539 – 550, 2007.

[12] D. Blythe, P. von Bunau, F. Meinecke, and K. Muller, "Feature extraction for change-point detection using stationary subspace analysis," *Neural Networks and Learning Systems, IEEE Transactions on*, vol. 23, no. 4, pp. 631–643, April 2012.