The use of Ensemble Empirical Mode Decomposition with Canonical Correlation Analysis as a Novel Artifact Removal Technique

Kevin T Sweeney, *Student Member, IEEE,* Seán F McLoone, *Senior Member, IEEE,* and Tomás E Ward, *Senior Member, IEEE*

Abstract—Biosignal measurement and processing is increasingly being deployed in ambulatory situations particularly in connected health applications. Such an environment dramatically increases the likelihood of artifacts which can occlude features of interest and reduce the quality of information available in the signal. If multichannel recordings are available for a given signal source then there are currently a considerable range of methods which can suppress or in some cases remove the distorting effect of such artifacts. There are however considerably fewer techniques available if only a single channel measurement is available and yet single channel measurements are important where minimal instrumentation complexity is required. This paper describes a novel artifact removal technique for use in such a context. The technique known as ensemble empirical mode decomposition with canonical correlation analysis (EEMD-CCA) is capable of operating on single channel measurements. The EEMD technique is first used to decompose the single channel signal into a multi-dimensional signal. The CCA technique is then employed to isolate the artifact components from the underlying signal using second order statistics. The new technique is tested against the currently available Wavelet denoising and EEMD-ICA techniques using both electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS) data and is shown to produce significantly improved results.

Index Terms—Ensemble Empirical Mode Decomposition (EEMD), Canonical Correlation Analysis (CCA), Independent Component Analysis (ICA), Wavelet denoising, EEMD-CCA, EEMD-ICA.

I. INTRODUCTION

MEASUREMENT of physiological signals of the human body has long been a crucial step in determining the health of a patient. The current transition from a hospital-centric healthcare system towards ambient healthcare assessment continuously calls for new systems capable of accurately recording physiological signals with minimal instrumentation and low computational complexity. Common reasons for the desired reduction in instrumentation and computational complexity include the aspiration for the systems to be ambulatory and capable of operating for long periods of time without recharging. As the systems are often patient-operated, the smaller the instrumentation complexity (e.g. the number of sensors) the less chance that the subject will fail to operate the device properly. Further, this reduction in complexity typically results in less expensive devices/technologies.

Unfortunately taking the measurement systems out of the clinical environment increases the likelihood that the recorded signal will be contaminated by artifact. Artifacts can be defined as any unwanted signal variations due to any source other than the desired signal source. In terms of physiological signals, these artifacts can come in many forms including instrumentation noise, subject motion or noise from other physiological signals of the body. Artifact removal/reduction is therefore a fundamental area of signal analysis in the ambulatory physiological signal monitoring domain and has been well documented [1]. In this paper we deal with artifacts arising from motion. There currently exist a number of algorithms to remove motion artifacts from physiological signals including Wavelets, independent component analysis (ICA) and adaptive filters [2].

To aid in the reduction of instrumentation complexity many ambulatory systems operate using single channel recordings only (e.g. [3]). Given this single source, it is critically important to extract as much useful information as possible through the suppression or removal of the corrupting artifact. There are a number of artifact removal techniques available to operate on single channel measurements [1] and these techniques differ in properties such as the efficacy and computational cost. Adaptive filtering, for example, has a low computational load in comparison with the other techniques presented in this paper but requires additional channels for reference purposes. Other techniques such as ensemble empirical mode decomposition (EEMD)-ICA [4], used previously to remove ocular and muscle artifacts from EEG data, have higher computational cost but in doing so are more effective in suppressing undesired artifacts.

This paper proposes a new single channel artifact removal technique, namely ensemble empirical mode decomposition with canonical correlation analysis (EEMD-CCA) and compares it to a competing motion artifact removal technique (Wavelet denoising [2]) and to a similar technique, namely EEMD-ICA. Section II gives a brief explanation of the algorithms used in the paper (Wavelets, ICA, CCA and EEMD) as well as the recording methodology for acquisition of the EEG and fNIRS data. Section III provides the results from the artifact removal techniques and finally Section V gives a

K. Sweeney is with the Department of Electronic Engineering, National University of Ireland Maynooth, Co. Kildare, Ireland email: `ksweeney@eeng.nuim.ie`

Dr S. McLoone and Dr T. Ward are with the Department of Electronic Engineering, National University of Ireland Maynooth, Co. Kildare, Ireland email: `sean.mcloone@eeng.nuim.ie` `tomas.ward@eeng.nuim.ie`

brief discussion and conclusion to the paper.

II. METHODS

In this section we describe the various techniques employed in the paper. First we outline the techniques required to implement the existing Wavelet and EEMD-ICA techniques (used for comparison) and then we will describe the proposed new method, EEMD-CCA.

A. Wavelet denoising

The Wavelet transform operates by decomposing a signal into a number of time shifted and scaled versions of a selected mother wavelet. The wavelet expansion of a signal $(x(t))$ can be written as [5]:

$$x(t) = \sum_k c_{Mk}\varphi_{Mk} + \sum_{j=1}^{M}\sum_k d_{jk}\psi_{jk}(t), \qquad (1)$$

where

$$c_{jk} = \int x(t)\varphi_{jk}^*(t)dt \qquad (2)$$

are called the scaling coefficients. The scaling functions φ are defined as:

$$\varphi_{jk}(t) = \frac{1}{\sqrt{2^j}}\varphi\left(\frac{t - k2^j}{2^j}\right). \qquad (3)$$

The wavelet coefficients (d) are given by:

$$d_{jk} = \int x(t)\psi_{jk}^*(t)dt \qquad (4)$$

with the wavelet functions ψ defined as:

$$\psi_{jk}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t - k2^j}{2^j}\right). \qquad (5)$$

The details $(D_j(t))$ and approximations $(A_j(t))$ of the wavelet transform at each level (j) can thus be defined as:

$$A_j(t) = \sum_{k=-\infty}^{+\infty} c_{jk}\varphi_{jk}(t)$$
$$D_j(t) = \sum_{k=-\infty}^{+\infty} d_{jk}\psi_{jk}(t). \qquad (6)$$

The original signal can be reconstructed by adding up all the details calculated and also the approximation for the final level (M), i.e.

$$x(t) = A_M(t) + \sum_{n=1}^{M} D_n(t). \qquad (7)$$

Once the artifact components (i.e. details or approximation) have been identified and removed (Section II-H) the remaining components can be added to reconstruct the cleaned signal $\hat{x}$.

In this paper the Daubechies 5 mother wavelet was implemented [6]. Wavelets are widely used and have previously been employed by a number of researchers for artifact removal including for motion artifact removal from fNIRS by Robertson *et al.* [2] and for ocular artifact removal from EEG by Kumar *et al.* [7].

B. Independent Component Analysis (ICA)

Independent component analysis (ICA) is a blind source separation (BSS) technique in which recorded signals are separated into their independent constituent components or sources [8]. BSS is based on a wide class of unsupervised learning algorithms with the goal of estimating sources and parameters of a mixing system [9]. An important property required to run the ICA technique is that the number of recorded signals is greater or equal to the number of unknown underlying sources. The ICA technique operates under a number of assumptions including linear mixing, square mixing and stationary mixing [10]. Incorporating these assumptions the ICA technique endeavors to determine an un-mixing matrix $\mathbf{W}$ which is used to determine the unknown independent components, $\hat{\mathbf{S}}$:

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{X}, \qquad (8)$$

where $\mathbf{X}$ is a matrix of the recorded multi-channel signals. There are a number of different algorithms currently available to determine this un-mixing matrix including the fast ICA algorithm [11], the Bell-Sejnowski algorithm [12], extended ICA [13] and JADE [14]. All of the ICA algorithms employed to determine the un-mixing matrix $\mathbf{W}$ use higher order statistics (HOS).

The fast ICA algorithm was implemented for this paper due to its shorter computational time compared to the other ICA techniques. Using this algorithm the un-mixing matrix is determined and the underlying source signals $\hat{\mathbf{S}}$ can be estimated. The sources deemed to be artifacts (see Section II-H) can be removed by setting the corresponding columns of the $\hat{\mathbf{S}}$ matrix to zero. When the artifact-suppressed source signals are passed through the inverse of the mixing matrix $\mathbf{W}^{-1}$, the resultant outputs are the original recorded signals minus the artifact components.

C. Ensemble Empirical Mode Decomposition (EEMD)

Empirical mode decomposition (EMD) is a method, first defined in 1998 [15], for nonlinear signal processing and is well suited to non-stationary data. The method decomposes a time series signal into multiple "intrinsic mode functions" (IMFs). The EMD technique differs from other techniques, such as Wavelet analysis, in that the decomposition of the signal is data driven whereas Wavelet analysis relies on the selection of the appropriate wavelet. As the technique is data driven, it is therefore adaptive in nature, making it very flexible.

The IMFs are functions that satisfy two separate conditions: (1) over the full length of the data set the number of maxima and the number of zero crossings must be the same or differ at most by one and (2) at any point over the data set, the mean value of the envelope defined by the maxima and the envelope defined by the minima must be zero [15].

There are a number of steps taken to calculate the IMF of a given time series, $\mathbf{x} \in \mathbb{R}^{\mathbf{N}}$, where N is the number of samples. EMD is implemented using a sifting process that uses only local extrema. Step 1 involves finding all the local maxima and minima over the full length of the time series. Next, the the maxima are connected using a cubic spline creating an upper envelope, with the same process repeated for the local minima. Step 2 then involves calculating the average of the two envelopes $\mathbf{m}$ and subtracting this average from the data signal, resulting in a new signal $\mathbf{h} = \mathbf{r_0} - \mathbf{m}$, where $\mathbf{r_0} = \mathbf{x}$. This signal $\mathbf{h}$ now becomes the new data signal and steps one and two are repeated until $\mathbf{h}$ complies with the properties of IMFs, detailed above. When this occurs the current data set $\mathbf{h}$ becomes the first IMF ($\mathbf{c_1}$). The above steps are then repeated on the residual signal $\mathbf{r_1} = \mathbf{r_0} - \mathbf{c_1}$. The sifting process stops when the residual signal $\mathbf{r_n}$ becomes a monotonic function. Once calculated, the IMFs $\mathbf{c_j}$ have the property that if they are added together then the original data $\mathbf{x}$ (or $\mathbf{r_0}$) is reconstructed, i.e.

$$\mathbf{x} = \sum_{\mathbf{j=1}}^{\mathbf{n}} \mathbf{c_j} + \mathbf{r_n}, \tag{9}$$

where $\mathbf{r_n}$ is the residual of data after n IMFs components are extracted [16].

The EMD algorithm is however very sensitive to noise in the recorded signal. This can lead to complications due to mode mixing. Mode mixing is defined as an IMF that includes oscillations of dramatically disparate scales or a component of similar scale residing in different IMFs [17], and can also be due to the presence of a transient spectral component in the signal. An extension to the EMD algorithm was proposed in [16] which eliminates this mode mixing dilemma. The updated algorithm called Ensemble-EMD (EEMD) uses an average of a number of ensembles of the EMD algorithm as the optimum choice of IMFs. Each run of the EMD algorithm has an independent, identically distributed white noise of the same standard deviation added thus providing a noise-assisted data analysis method.

With the IMFs determined, the artifact components can then be selected, as described in Section II-H, and removed. The remaining IMFs can then either be used to reconstruct the cleaner signal or can be passed as inputs to additional artifact removal techniques to further enhance the signal quality.

D. EEMD-ICA

The use of EEMD in combination with ICA for source separation of single channel measurements was first detailed by Mijović *et al.* in 2010 [4] and was employed for the removal of EMG and ocular artifact from EEG and also ECG from EMG. The EEMD technique can be used to create a multi-channel signal $\mathbf{X}$, comprised of IMFs, from a single channel recording $\mathbf{x}$. This matrix $\mathbf{X}$ can then be employed as the input to the fast ICA algorithm with the aim of estimating the underlying true sources $\hat{\mathbf{S}}$ (see Equation 8). The individual sources determined to be artifacts are selected and the corresponding columns of the matrix $\hat{\mathbf{S}}$ are set to zero. The source matrix is then passed through the inverse of the un-mixing matrix $\mathbf{W}^{-1}$ to return the multi-channel signals $\hat{\mathbf{X}}$ which are now, ideally, free of artifacts. The original single channel signal $\hat{\mathbf{x}}$, now free of artifacts, can be determined by simply adding the IMFs in the matrix $\hat{\mathbf{X}}$.

Mijović *et al.* [4] compared the EEMD-ICA algorithm against two other techniques capable of operating on single channel recordings, namely Single-Channel ICA (SCICA) and Wavelet-ICA (WICA). The algorithms were tested and compared on simulated data and the EEMD-ICA algorithm was then employed on real EEG and EMG data. The results show that the SCICA algorithm has the worst performance when comparing the relative root mean square error (RRMSE). The WICA algorithm, although comparable to the EEMD-ICA technique, showed a slightly weaker performance in the simulations.

E. Canonical Correlation Analysis (CCA)

Canonical correlation analysis (CCA) [18] is another technique which employs the BSS method for separating a number of mixed or contaminated signals. The CCA technique, similar to the ICA method described above, requires that there be a greater or equal number of recorded signals as underlying sources. CCA differs from the ICA in its method of separating the sources. CCA uses second order statistics (SOS) to generate components derived from their uncorrelated nature rather than their independence. Although the decomposition using CCA uses a weaker condition than ICA, it is also less computationally complex due its use of SOS. Further, it is known that if a random vector has a multivariate normal distribution then any two or more of its components that are uncorrelated are also independent [19] and thus CCA can often return the same result as ICA. If this is not the case, then CCA will return components which are uncorrelated but not independent. Additionally, the ICA algorithm does not take temporal correlations into account and thus the data samples can be arranged arbitrarily in time and the ICA method will return the same solution. CCA addresses this point and is capable of finding uncorrelated components that, in addition, have maximum spatial or temporal correlation within each component.

CCA solves the BSS problem by forcing the sources to be maximally autocorrelated and mutually uncorrelated [20]. Given an input signal $\mathbf{x}$, let $\mathbf{y}$ be a linear combination of neighbouring samples (i.e. $y(t) = x(t-1) + x(t+1)$ [21]). Consider the linear combinations of the components in $\mathbf{x}$ and $\mathbf{y}$, called the canonical variates,

$$x = \mathbf{w_x}^T(\mathbf{x} - \hat{\mathbf{x}}) \tag{10}$$
$$y = \mathbf{w_y}^T(\mathbf{y} - \hat{\mathbf{y}}). \tag{11}$$

CCA finds the weight matrices $\mathbf{w_x}$ and $\mathbf{w_y}$ that will maximise the correlation ρ between x and y [21]:

$$\rho = \frac{\mathbf{w_x}^T \mathbf{C}_{xy} \mathbf{w_y}^T}{\sqrt{\mathbf{w_x}^T \mathbf{C}_{xx} \mathbf{w_x} \mathbf{w_y}^T \mathbf{C}_{yy} \mathbf{w_y}}}, \tag{12}$$

where $\mathbf{C}_{xy}$ is the between-sets covariance matrix and $\mathbf{C}_{xx}$ and $\mathbf{C}_{yy}$ are the nonsingular within-set covariance matrices. The

calculation of the maximum of ρ can be found by setting the derivatives of Equation 12 to zero with respect to $\mathbf{w_x}$ and $\mathbf{w_y}$.

$$\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}\mathbf{C}_{yx}^{T}\hat{\mathbf{w}}_{\mathbf{x}} = \rho^2\hat{\mathbf{w}}_{\mathbf{x}} \qquad (13)$$
$$\mathbf{C}_{yy}^{-1}\mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}^{T}\hat{\mathbf{w}}_{\mathbf{y}} = \rho^2\hat{\mathbf{w}}_{\mathbf{y}}.$$

$\mathbf{w_x}$ and $\mathbf{w_y}$ can be determined as the eigenvectors of the matrices $\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}\mathbf{C}_{yy}^{-1}\mathbf{C}_{yx}^{T}$ and $\mathbf{C}_{yy}^{-1}\mathbf{C}_{yx}\mathbf{C}_{xx}^{-1}\mathbf{C}_{xy}^{T}$ respectively and the corresponding eigenvalues ρ^2 are the squared canonical correlations. The first pair of variates are the eigenvectors of $\mathbf{w_x}$ and $\mathbf{w_y}$ that correspond to the largest square correlation coefficient (or eigenvalue) ρ_{max}^2. The following pairs of variates $(\mathbf{w_x}, \mathbf{w_y})_{2...m}$ (with m recording sites) are then the remaining eigenvectors in decreasing order of correlation [22] which are themselves uncorrelated with the previous eigenvectors. The canonical correlation analysis technique therefore creates a weight Matrix $\mathbf{W_x} = [\mathbf{w_{x1}}, \mathbf{w_{x2}}...\mathbf{w_{xm}}]$ that can be used to separate the recorded sources into the self correlated and mutually uncorrelated sources.

Removal of the artifact is then accomplished in a similar manner to the ICA algorithm. in that the columns of $\hat{\mathbf{S}}$, which represent the artifacts in the recording (Section II-H), are set to zero before performing the reconstruction.

CCA has been previously tested against a number of ICA algorithms using multi-channel recordings. Borga and Knutsson [21] compared CCA with Fast ICA for the problem of separating a 5 channel EEG recording. Results illustrated that both methods had qualitatively the same results, but as the CCA algorithm employs SOS, where ICA employs HOS, the CCA method was determined to be more computationally efficient. In 2006 Clercq et al. [20] used CCA to successfully remove muscle artifacts from EEG data. This method was also shown on simulated data to outperform an ICA-based technique (JADE). This result was again documented by Gao et al. [23] in 2010. One possible reason for CCA's improved performance over the ICA method is due to the form of the muscle activity. Muscle artifacts, involve the movement of a group of muscles, which do not have a stereotyped topography [23] and thus the ICA does not function correctly. Similar results have also been demonstrated with other signal modalities. In 2002, fMRI data was extracted using both CCA and ICA [24] and the CCA algorithm was again shown to perform better in terms of computational complexity by at least an order of magnitude while having similar qualitative results. It is also noted that the CCA method always returns the same result when employed with a given data set, this however is not true when using the ICA algorithm.

F. EEMD-CCA

The EEMD-CCA technique operates in a similar manner to the EEMD-ICA technique. The single channel signal $\mathbf{x}$ is again converted into a multi-channel signal $\mathbf{X}$ using the EEMD algorithm. The IMF determined to be artifacts are removed (as described in Section II-C) and then the remaining channels are used in conjunction with the un-mixing matrix $\mathbf{W}$, determined using the CCA algorithm, to identify the underlying source signals $\hat{\mathbf{S}}$. Similar to the EEMD-ICA algorithm, the

sources corresponding to artifacts are set to zero and then the original multi-channel signal is reconstructed, minus the artifact components, using the inverse of the un-mixing matrix $\mathbf{W}^{-1}$ creating the matrix $\hat{\mathbf{X}}$. The original single channel signal without the artifacts $\hat{\mathbf{x}}$ can be determined by adding the new IMFs components in the $\hat{\mathbf{X}}$ matrix.

As stated previously it should be again noted that CCA uses SOS and is therefore computationally more efficient than ICA. It follows that the EEMD-CCA technique will also have a lower computational cost than EEMD-ICA. In Section III a quantitative comparison of the artifact removal capability of all the techniques is provided using SNR and correlation measurements for both EEG and fNIRS data.

G. Data Acquisition

In order to test the validity and efficacy of the proposed EEMD-CCA algorithm, compared to the current EEMD-ICA and Wavelet algorithms, the authors tested both techniques on electroencephalograph (EEG) and functional near-infrared spectroscopy (fNIRS) data. The data for both modalities was recorded independently using the novel recording methodology proposed in [25]. This particular recording methodology is capable of producing two highly correlated signals for each modality. When recording the two channel signal, motion artifact can be induced onto one channel by disturbing the appropriate recording electrode or optode. Therefore, by implementing this methodology, two signals become available: the first signal is free of motion artifact contamination (labeled "ground truth") and the second is intermittently corrupted by artifacts. An example of a recording for both EEG and fNIRS can be seen in Figure 1 and Figure 2 respectively.
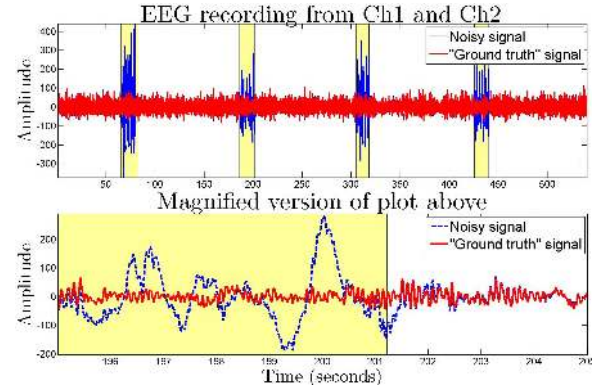

Fig. 1. An example recording of EEG data. Two channels are recorded which are highly correlated during epochs of no motion artifact. Shaded areas show epochs of motion artifact induced by the authors during testing.

During epochs of no movement (as shown by the non-highlighted areas in the figures) the two signals can be observed to have a high correlation coefficient. The average correlation over the clean epochs for all trials was found to be 0.84 for the EEG data and 0.77 for the fNIRS data. However, over the full signal this correlation level drops significantly to 0.40 for the EEG data and 0.58 for the fNIRS data due to the intermittent presence of artifacts. Artifact components are present in 15.7% of the EEG data and 19.8% of the fNIRS data.
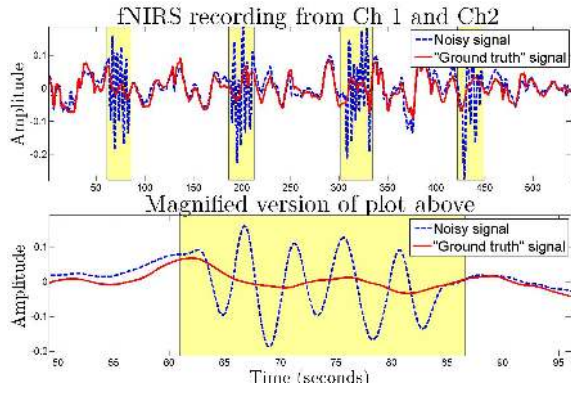
Fig. 2. An example recording of fNIRS data. Two channels are recorded which are highly correlated during epochs of no motion artifact. Shaded areas show epochs of motion artifact induced by the authors during testing.

A total of 24 sessions of EEG data were collected from a cohort of 6 healthy subjects, 3 females and 3 males, with a mean age of 27 years (standard deviation: 4.3 years). The fNIRS database consisted of 16 sessions from 10 healthy subjects. The cohort comprised of 4 males and 6 females with a mean age of 29 years (standard deviation: 5.6 years). These databases for EEG and fNIRS were then used to determine the efficacy of the Wavelet, EEMD-ICA and EEMD-CCA techniques.

H. Removal of Artifact Components

When employing any of the described artifact removal techniques a common issue is determining which components of the decomposed signal are artifacts and thus should be removed. For the methodologies implemented in this paper we utilized the available "ground truth" signal. Any component that, when removed, increased the correlation of the reconstructed cleaned signal with the "ground truth" signal was determined to be artifact and was removed. This ensured that each technique performed optimally during evaluation. Although it not possible to employ this technique for determining artifact components in practical applications, as the required "ground truth" signal will not be available, it effectively decouples the artifact separation process from the component selection process and thus provides the fairest method for comparing the algorithms considered in this paper.

For completeness, a second method was also employed which used the autocorrelation function based automatic artifact component selection technique proposed by Hassan *et al.* [26]. This method assumes that the autocorrelation at lag one of the artifact components differs from that of the true desired signal. Figure 3 and Figure 4 show examples of the autocorrelation coefficients for EEG and fNIRS for an example decomposed trial. For both signal types, a threshold value is required to separate the artifact from the signals of interest. For fNIRS, all components below a threshold of 0.9991 are specified as artifacts whereas for the EEG threshold, all components above a threshold of 0.97 are deemed to be artifacts. These threshold values were determined by employing the "ground truth" reference signal as provided by our methodology for calibration purposes. Using this method,

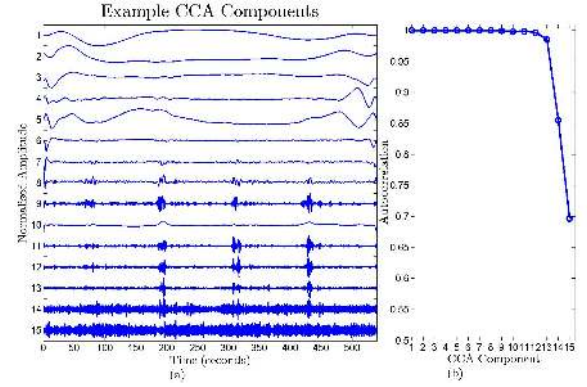all detected artifact components can be automatically removed from the decomposed signals.

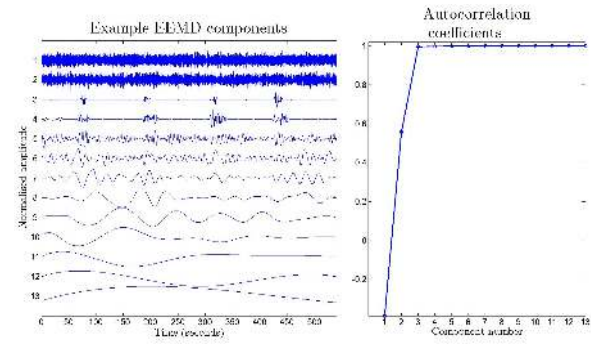Fig. 3. An example output of the CCA algorithm for contaminated EEG data.

Fig. 4. An example output of the EEMD algorithm for contaminated fNIRS data.

III. RESULTS

Using the two available signals for each modality, as described in Section II-G, it is possible to test the efficacy of individual artifact removal techniques. The goal of each artifact removal technique is to return the artifact contaminated signal to its true state. Thus using calculations, such as SNR and correlation, it is possible to quantify the particular technique's ability to remove the artifact. In this study the difference in SNR before and after artifact removal, denoted by Δ SNR, is employed as a performance metric. The Δ SNR was calculated using the following formula:

$$\Delta SNR = 10 \log_{10}\left(\frac{\sigma_x^2}{\sigma_{e_{after}}^2}\right) - 10 \log_{10}\left(\frac{\sigma_x^2}{\sigma_{e_{before}}^2}\right) \quad (14)$$

where σ_x^2 is the variance of the "ground truth" signal and $\sigma_{e_{before}}^2$ and $\sigma_{e_{after}}^2$ are the variances of the error signal before and after employing the artifact removal technique. The error signal is determined as the difference between the noisy signal and the "ground truth" recording , assuming that the motion artifact is additive.

The difference in correlation between signals was used to define a new performance metric, namely the percentage

reduction in artifact λ:

$$\lambda = 100 \left(1 - \frac{R_{clean} - R_{after}}{R_{clean} - R_{before}} \right). \qquad (15)$$

Here R_{before} is the correlation between the "ground truth" and artifact contaminated signal and R_{after} is the correlation with the signal following processing by an artifact removal algorithm. R_{clean} is the correlation over the epochs of known clean data (the non-shaded areas of Figure 1 and 2) and is an estimate of the maximum achievable correlation between the two signals for the experimental setup considered. As such, the denominator of the equation is the improvement in correlation possible if complete artifact removal is achieved while the numerator is the actual correlation improvement obtained for a given artifact removal technique. A high λ therefore equates to good efficacy in artifact removal.

The results obtained using the various artifact removal techniques are presented below. The results presented employ the "ground truth" methodology for artifact component selection as these are deemed the best case results. These results, as well as those obtained when using the autocorrelation based artifact component selection procedure to determine the artifact components, are presented in Table I with their variability presented using the standard deviation metric.

A. EEG

All the algorithms described in Section II were run on the 24 trials of EEG data and the artifact components were removed using both the "ground truth" method and the autocorrelation method as described in Section II-H. Figure 5 displays an example trial of EEG after the removal of the artifact using wavelet analysis, EEMD-ICA and EEMD-CCA with the artifact components selected using the "ground truth" method.

1) Wavelets: When employing wavelet analysis it was calculated that the technique produced an average Δ SNR of 8.08 dB over all trials and there was a 55.3 % reduction in artifact producing a correlation of 0.64 with the "ground truth" signal.

2) EEMD-ICA: Prior to running the technique the ensemble number had to be chosen. The optimum number of ensembles to use was determined by re-running the algorithm while systematically increasing the number of ensembles employed. When the results were observed to plateau the ensemble number was set. When using the EEG data it was found that the performance of the technique became fairly consistent when using 5 or more ensembles. The noise standard deviation was empirically set to 0.1 times [16] the standard deviation of the signal.

The technique produced an average Δ SNR of 8.2 dB and the artifact was reduced by 52.3 % yielding a correlation of 0.63 after employing the EEMD-ICA algorithm.

3) EEMD-CCA: Using the same ensemble number the EEMD-CCA algorithm was determined to generate an average Δ SNR of 8.2 dB over all trials. The correlation also rose to 0.63, representing a 52.2 % reduction in artifact when using this particular technique.

B. fNIRS

The validity of the new technique was also tested using fNIRS data. Again Wavelet denoising, EEMD-ICA and EEMD-CCA were tested on the same sessions and the improvement in signal correlation and SNR with respect to the "ground truth" signal was calculated. Figure 5 shows an example fNIRS trial after use of the three denoising techniques.

1) Wavelets: Over all trials the average Δ SNR was determined to increase by 3.1 dB using wavelet analysis and 43.6 % of the artifact was removed resulting in a correlation of 0.66.

2) EEMD-ICA: For the fNIRS data an ensemble number of 3 for the EEMD algorithm was found to be sufficient to provide accurate results, as increasing the ensemble number did not substantially improve performance. Again, the noise standard deviation was empirically set to 0.1 times the standard deviation of the signal.

EMD-ICA was able to remove some of the contaminating artifacts from the noisy signal as demonstrated by a Δ SNR of 3.4 dB. The algorithm also reduced the artifact by 43.4 % producing a correlation of 0.66 with the "ground truth" signal.

3) EEMD-CCA: Finally the new proposed technique EEMD-CCA was employed to remove the motion artifact from the available fNIRS data. Using this algorithm the percentage artifact reduction was calculated as 49.4 % generating a correlation of 0.68 and the technique also produced an average Δ SNR of 3.5 dB.

IV. DISCUSSION

Table I summarizes the results obtained using the artifact removal techniques described in the paper. Results are also provided for EMD, EMD-ICA and EMD-CCA to demonstrate the requirement for the use of ensembles. For both the EEG and fNIRS modalities the EMD algorithms can be seen to perform substantially inferior to their EEMD counterparts thus validating the use of EEMD over EMD.

When operating on the EEG data the table gives inconclusive results as to the optimum technique. The wavelet algorithm marginally outperforms the EEMD based techniques in terms of correlation improvement ($\approx$ 3 %). The wavelet technique gives the best results for correlation improvement in 18 out of the 24 analyzed trials. However when evaluating the algorithms using the Δ SNR the wavelet function has a lower average level ($\approx$ 0.13 dB) where the EEMD based techniques are seen to have to best results in 19 of the 24 analyzed trials. The EEMD based techniques (EEMD, EEMD-ICA and EEMD-CCA) are seen to have similar results for both λ and Δ SNR demonstrating that the decomposition of the signals into their respective IMF is sufficient to select the artifact components. One possible reason for this is that there is lower spectral overlap between the artifact and the EEG signal (compared to the fNIRS signal) and thus the two decomposition techniques are capable of removing the majority of the artifact.

However, when analyzing the fNIRS data the developed EEMD-CCA technique can be observed to outperform all other tested techniques. As stated previously this may be due to the sources having a multivariate normal distribution. In terms
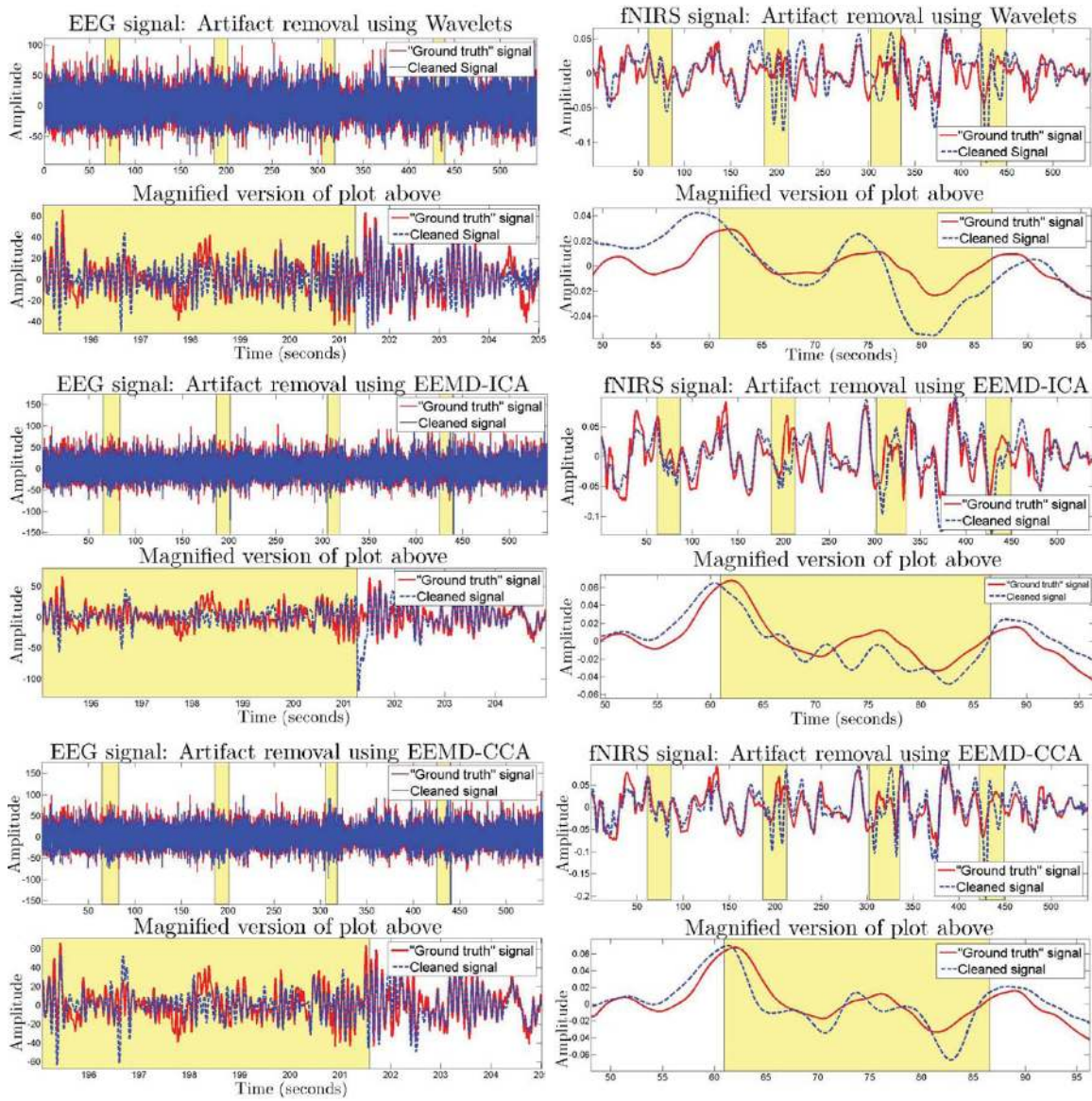
Fig. 5.   Example results of the artifact removal from both EEG and fNIRS using Wavelet denoising, EEMD-ICA and EEMD-CCA.

of correlation improvement (λ) the EEMD-CCA technique was discovered to be $\approx$ 3 % better than the EEMD technique and $\approx$ 6 % better than the EEMD-ICA and wavelet techniques and demonstrated the best results in 12 of the 16 trials. This improvement over the wavelet technique can also be observed in terms of the Δ SNR where EEMD-CCA showed the best results in 11 trials and had on average a 0.56 dB improvement over the wavelet technique. The failure of the wavelet technique to separate the artifact when employed on fNIRS may be due to the higher spectral overlap between the artifacts and the signal of interest. In this case the second artifact removal stage (CCA) succeeds in separating the artifact components which are spread over a number of the wavelet or EEMD components. The reason for the EEMD algorithms improved correlation over EEMD-ICA is the ICA algorithm can often fail to converge [27] and therefore when the signal is reconstructed some data may be lost, reducing the correlation

with the "ground truth" signal.

The computational cost of the EEMD-CCA algorithm was also assessed in order to determine what, if any, additional computational costs were incurred in return for the improved artefact suppression performance of the CCA extension to EEMD. It can be seen from Figure 6 (showing fNIRS data as an example) that the computational time of CCA is considerably less than that of EEMD (using three ensembles) and thus does not add any significant computational complexity to the system. The computational cost of the ICA and Wavelet algorithm is also plotted for comparison. The computation time for the ICA algorithm is significantly greater as it uses higher order statistics whereas the CCA algorithm uses second order statistics to determine the underlying sources. Similar results can be seen for the EEG data. For a number of ambulatory systems, direct feedback is often employed and thus ensuring a low computational load can often be of importance to ensure

TABLE I

FOR BOTH FNIRS AND EEG THE PERCENTAGE REDUCTION IN ARTIFACT (λ) (EQUATION 15) AND THE Δ SNR ARE PRESENTED WITH CORRESPONDING STANDARD DEVIATIONS SHOWN IN BRACKETS. RESULTS ARE SHOWN FOR BOTH ARTIFACT COMPONENT SELECTION USING THE "GROUND TRUTH" SIGNAL (GT) AND USING THE AUTOCORRELATION FUNCTION (RXX) AS DESCRIBED IN SECTION II-H.

			Wavelets	EMD	EEMD	EMD-ICA	EMD-CCA	EEMD-ICA	EEMD-CCA
fNIRS	λ	GT	43.6 % (41.3)	18.9 % (18.6)	46.2 % (45.9)	17.9 % (17.8)	23.3 % (23.1)	43.4 % (43.2)	49.4 % (49.1)
		Rxx	38.2 % (35.3)	13.2 % (12.8)	42.2 % (41.9)	14.9 % (14.8)	17.3 % (17.1)	39.7 % (39.6)	46.4 % (46.3)
	ΔSNR	GT	3.05 dB (2.15)	2.01 dB (1.41)	3.41 dB (2.07)	2.37 dB (1.45)	2.27 dB (1.58)	3.60 dB (1.91)	3.61 dB (2.02)
		Rxx	2.88 dB (2.01)	1.84 dB (1.42)	3.21 dB (1.85)	2.12 dB (1.55)	1.98 dB (1.59)	3.42 dB (1.78)	3.44 dB (1.88)
EEG	λ	GT	55.3 % (35.4)	43.2 % (31.2)	52.2 % (36.3)	44.1 % (30.8)	43.4 % (31.3)	52.3 % (36.2)	52.2 % (36.4)
		Rxx	51.2 % (36.7)	38.7 % (31.8)	48.5 % (34.2)	40.0 % (27.6)	39.6 % (30.8)	48.3 % (37.2)	48.5 % (35.2)
	ΔSNR	GT	8.08 dB (4.01)	7.28 dB (3.67)	8.21 dB (3.82)	7.47 dB (3.53)	7.32 dB (3.67)	8.22 dB (3.81)	8.21 dB (3.82)
		Rxx	7.81 dB (4.28)	7.01 dB (3.35)	7.88 dB (3.46)	7.22 dB (3.70)	6.98 dB (3.42)	8.02 dB (3.75)	8.04 dB (3.72)

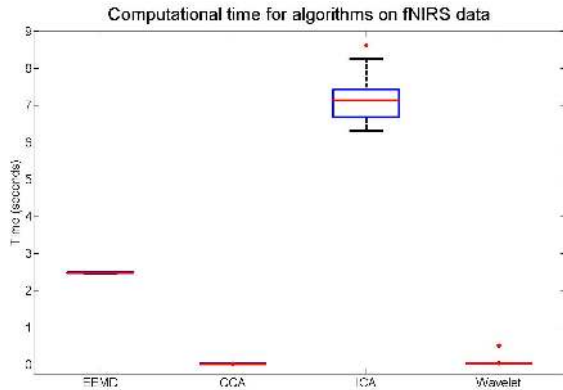fast clean information is available.

Fig. 6. Average computational time, over 100 runs, of the different algorithms on the fNIRS data

It is also possible to see that the automatic artifact selection using the autocorrelation function as described in Section II-H performed quite well in comparison to the "ground truth" results with the EEMD-CCA algorithm again outperforming all of the other techniques for fNIRS and EEG apart from wavelets when analyzing the EEG data using correlation. The results are however consistently lower, as expected due to the "ground truth" referencing allowing for the best possible results. Future work into the optimal artifact selection criterion may increase the accuracy of these results which will further aid in the transition of health monitoring systems from the hospital centric domain to the ambulatory environment.

V. CONCLUSIONS

A novel artifact removal technique has been proposed in this paper, namely ensemble empirical mode decomposition with canonical correlation analysis (EEMD-CCA). The new technique is capable of operating on single channel measurements due to the inclusion of the EEMD algorithm. The efficacy of the algorithm was compared using both EEG and fNIRS data to Wavelet denoising [2] EMD [15], EEMD [16] and the similar EEMD-ICA algorithm [4].

The novel EEMD-CCA technique was shown to provide the leading performance in artifact removal when analyzing the fNIRS data. This was true when using either the "ground truth" signal or the autocorrelation metric for automatic artifact component selection (Section II-H). This performance improvement was shown numerically (Table I) using metrics such as the percentage improvement in correlation (λ) and the change in signal-to-noise ratio (Δ SNR). When analyzing the EEG data the basic EEMD and Wavelet denoising techniques performed best, with the addition of the ICA or CCA algorithm adding little improvement. However, it should also be noted that the EEMD-CCA technique consistently performs as well, if not better than the EEMD technique alone over both the EEG and the fNIRS data. As the computational time of the CCA algorithm is also very low in comparison to both the EEMD and the ICA algorithm, its addition to the EEMD technique has no disadvantage. Therefore by employing the 2 stage artifact removal process of EEMD and CCA, the results will consistently be as good (within 0.01 %) if not better than those achievable using EEMD alone or in conjunction with ICA for both EEG and fNIRS.

REFERENCES

[1] K. Sweeney, T. Ward, and S. McLoone, "Artifact Removal in Physiological Signals -Practices and Possibilities," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 3, pp. 488 –500, may 2012.

[2] F. C. Robertson, T. S. Douglas, and E. M. Meintjes, "Motion Artifact Removal for Functional Near Infrared Spectroscopy: A Comparison of Methods," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 6, pp. 1377 –1387, June 2010.

[3] S. D. Ridwan, R. Thompson, B. T. Jap, S. Lal, and P. Fischer, "Single Channel Wireless EEG: Proposed Application in Train Drivers," *Broadband Communications, Information Technology & Biomedical Applications, 2008 Third International Conference on*, pp. 58–63, Nov. 2008.

[4] B. Mijović, M. De Vos, I. Gligorijevic and, J. Taelman, and S. Van Huffel, "Source Separation From Single-Channel Recordings by Combining Empirical-Mode Decomposition and Independent Component Analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 9, pp. 2188 –2196, sep. 2010.

[5] B. Azzerboni, M. Carpentieri, F. La Foresta, and F. Morabito, "Neural-ICA and wavelet transform for artifacts removal in surface EMG," in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, vol. 4, 25-29 2004, pp. 3223 –3228 vol.4.

[6] I. Daubechies, "Ten lectures on wavelets," *CBM-SNSF Regional Conference Series in Applied Mathematics*, 1992. [Online]. Available: http://ci.nii.ac.jp/naid/10000015709/en/

[7] P. S. Kumar, R. Arumuganathan, K. Sivakumar, and C. Vimal, "Removal of Ocular Artifacts in the EEG through Wavelet Transform without using an EOG Reference Channel," *Int. J. Open Problems Compt. Math.*, vol. 1, 2008.

[8] A. Hyvrinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, S. Haykin, Ed. Wiley-Interscience, June 2001.

[9] S. Choi, A. Cichocki, H. min Park, and S. young Lee, "REVIEW Blind Source Separation and Independent Component Analysis: A Review," *Neural Information Processing - Letters and Reviews*, vol. 6, no. 1, pp. 1–57, 2004.

[10] C. J. James and C. W. Hesse, "Independent component analysis for biomedical signals," *Physiological Measurement*, vol. 26, pp. 15–39, 2005.

[11] A. Hyvrinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.

[12] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.

[13] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources," *Neural Computation*, vol. 11, no. 2, pp. 417–441, 1999.

[14] B. A. Draper, K. Baek, M. S. Bartlett, and J. R. Beveridge, "Recognizing faces with PCA and ICA," in *Computer vision and image understanding, special issue on face recognition*, 2003, pp. 115–137.

[15] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, March 1998.

[16] Z. Wu and N. E. Huang, "Ensemble Empirical Mode Decomposition: a Noise-Assisted Data Analysis Method." *Advances in Adaptive Data Analysis*, vol. 1, no. 1, pp. 1–41, 2009.

[17] Z. Shen, Q. Wang, Y. Shen, J. Jin, and Y. Lin, "Accent extraction of emotional speech based on modified ensemble empirical mode decomposition," in *Instrumentation and Measurement Technology Conference (I2MTC), 2010 IEEE*, may 2010, pp. 600 –604.

[18] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3-4, pp. 321–377, 1936.

[19] M. Borga and C. M. Borga, "Learning multidimensional signal processing," Ph.D. dissertation, 1998.

[20] W. D. Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 12, pp. 2583 –2587, nov. 2006.

[21] M. Borga and H. Knutsson, "A canonical correlation approach to blind source separation," Linkoping University, Sweeden, Tech. Rep. LiU-IMT-EX-0062, June 2001.

[22] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. Wiley-Interscience, 2003.

[23] Gao., Junfeng., Zheng., Chongxun., Wang., and Pei., "Online Removal of Muscle Artifact from Electroencephalogram Signals Based on Canonical Correlation Analysis," *Clinical EEG and Neuroscience*, vol. 41, no. 1, p. 53, 2010.

[24] M. Borga, O. Friman, P. Lundberg, and H. Knutsson, "A Canonical Correlation Approach to Exploratory Data Analysis in fMRI," in *Proceedings of the ISMRM Annual Meeting*, Honolulu, Hawaii, May 2002.

[25] K. Sweeney, H. Ayaz, T. Ward, M. Izzetoglu, S. McLoone, and B. Onaral, "A Methodology for Validating Artifact Removal Techniques for Physiological Signals," *Information Technology in Biomedicine, IEEE Transactions on*, vol. PP, no. 99, p. 1, 2012.

[26] M. Hassan, S. Boudaoud, J. Terrien, B. Karlsson, and C. Marque, "Combination of Canonical Correlation Analysis and Empirical Mode Decomposition Applied to Denoising the Labor Electrohysterogram," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 9, pp. 2441 –2447, sept. 2011.

[27] Z. Wang, "Fixed-point algorithms for constrained ICA and their applications in fMRI data analysis," *Magnetic Resonance Imaging*, vol. 29, no. 9, pp. 1288 – 1303, 2011.