Reference-free Reduction of Ballistocardiogram Artifact from EEG Data Using EMD-PCA

Ehtasham Javed[1], Ibrahima Faye[*, 2], Aamir Saeed Malik[1]
Centre for Intelligent Signal and Imaging Research (CISIR)
[1]Department of Electrical and Electronics Engineering,
[2]Department of Fundamental and Applied Sciences,
University Teknologi PETRONAS
31750, Tronoh, Perak, Malaysia
rajaehti1@gmail.com, ibrahima_faye@petronas.com.my,
aamir_saeed@petronas.com.my
[*]Corresponding author

Jafri Malin Abdullah
Department of Neuroscience,
School of Medical Sciences
Hospital Universiti Sains Malaysia,
16150, Kubang Kerian, Kelantan, Malaysia
deptneurosciencesppspusm@yahoo.com

***Abstract*— Concurrent electroencephalograph (EEG) and functional magnetic resonance image (fMRI) led researchers to acquire neuronal activities in detail over the past few decades. Regardless of the advantages of combining these modalities, artifacts posed a greater challenge to attain good quality data. One such problematic artifact which contaminates EEG recordings is Ballistocardiogram (BCG) artifact. A reference-free composite algorithm which combines empirical mode decomposition (EMD) and principal component analysis (PCA) named as EMD-PCA has been introduced in this study. The results show that the algorithm can efficiently reduce the BCG artifact by preserving original neuronal signals. The proposed algorithm showed improvement in reducing the BCG artifact as well as in the preservation of brain activities, when compared with two renowned existing methods that are average artifact subtraction (AAS) and optimal basis set (OBS).**

***Index Terms*—Simultaneous EEG-fMRI, Ballistocardiogram artifact, Empirical Mode Decomposition; Principal Component Analysis; Event-related potential**

I. Introduction

Understanding functions of the human brain has been an explorative area for researchers over the past few decades. The complexity of its operations has prescribed them to use various methods to discover new insights. One such method was coined in the 1920s known as electroencephalography (EEG), which measures the electric potential in the brain [1]. This was later followed by magnetic resonance imaging (MRI) which offers signals correlated to the hemodynamic response [2]. However, the usage of both the methods independently has limitations; EEG alone provides low resolution in spatial and MRI lacks the ability to provide detailed information in the temporal domain. Hence, the combination of EEG with functional magnetic resonance imaging (fMRI) was utilized to study the human brain in detail. This concurrent acquisition setup provides information up to milliseconds and millimeters in the temporal and spatial domains respectively. Recently, due to the increased amount of information, EEG-fMRI has become popular in the areas of clinical studies; assessment of working memory of brain tumor patients [3, 4], localization of epileptic seizure [5, 6] and effect of aging on brain processes [7].

Combining two modalities (EEG and fMRI) encompass the limitations when used separately but, lead to another challenging aspect, in precise; fMRI environment contaminates EEG data by artifacts known as gradient artifact (GA) and ballistocardiogram (BCG) artifact [8]. These artifacts distort the original neuronal activities and can lead to wrong assessment. Hence, it is required to remove these artifacts before further analysis of EEG data. Among the above artifacts, the gradient magnetic fields used for spatial encryption of fMRI data result in the GA which has the amplitude of 10 to 100 times the amplitude of the EEG data. The pattern of GA is periodic for each channel and it can easily be removed [2]. The artifact which is more complex, dynamic and problematic to remove is the BCG artifact. The BCG artifact varies significantly within magnetic scanners, channels and individuals [9].

The BCG artifact is an artifact, which results from the slight movement of head, contraction and expansion of scalp arteries due to blood flow at each cardiac pulsation inside the magnetic scanner. The onset of the BCG artifact is synchronous to the onset of the QRS pulse i.e. delay between both waveforms is ~210ms and duration of BCG artifact is approximately around 500ms. The maximum amplitude with which it contaminates EEG data is 200μV at 3T and has main effect on the frequency distribution ranging from 0.5-13 Hz [10]. The variation in temporal and spatial domain over the scalp regions and synchronicity of the BCG waveform with every heart-beat can be seen clearly from Fig. 1. The dynamic pattern of the BCG artifact makes it more challenging compared to the GA.

Number of researchers implemented different methods in their studies to reduce BCG artifact. These methods include average artifact subtraction (AAS) [11], adaptive filtering [12], algorithms of independent component analysis (ICA) and principal component analysis (PCA) as blind source separation

978-1-4799-4653-2/14/$31.00 © 2014 IEEE

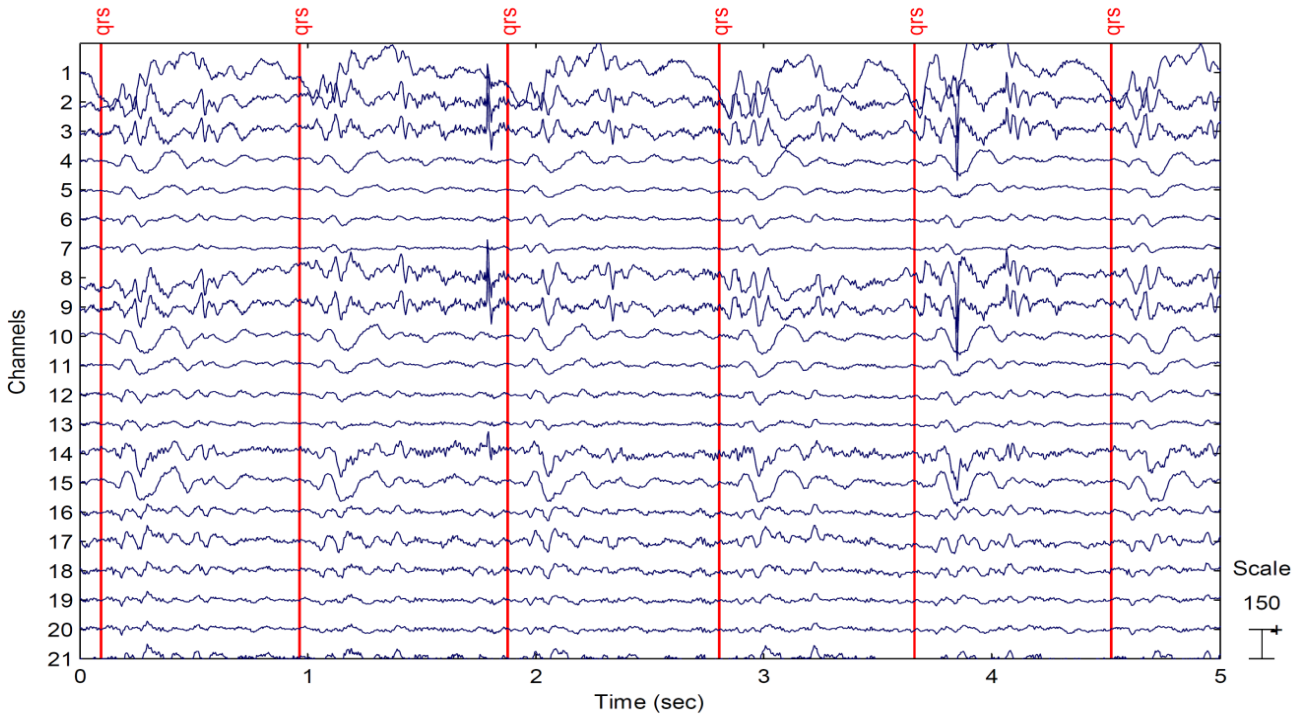


Fig. 1. Synchronous behaviour of BCG artifact with QRS pulse and variations in temporal and spatial domain.

(BSS) methods [13]–[15]. Recently, a short and long term prediction model was used to estimate the BCG artifact [16] along with other methods which heavily rely on reference signal [17], [18]. Even though all approaches claim to reduce the BCG artifact, substantial conflicts exist between them because of residuals of the BCG artifact left behind. The main limitations of above mentioned methods are: dependency over a reference (ECG) signal, assumption of temporal stability in BCG waveforms, selection of components in BSS approaches that belongs to the BCG artifact and absence of parameters to evaluate the quality of EEG data after reduction of the BCG artifact. Moreover, the comparative studies were also unable to find a suitable method which can reduce the BCG artifact from all types of dataset of EEG [10].

A composite algorithm has been proposed in this study to overcome the limitation of dependency over the reference signal, temporal stability assumption of the BCG waveform. The proposed algorithm combines two well-known algorithms named empirical mode decomposition (EMD) and principal component analysis (PCA). The algorithm decomposes the contaminated EEG data into set of components using EMD which are the inputs for the PCA to extract the principal component that belongs to BCG. Overall the algorithm works in temporal domain (channel-wise) as it is also suggested in literature to overcome the spatial influence of the BCG artifact [19]. The capability of the proposed algorithm (EMD-PCA) to reduce the BCG artifact has been evaluated using peak-to-peak value before and after the reduction of the BCG artifact. The quality of the reconstructed signal after reduction of BCG artifact has been assessed by using two parameters; power spectrum and extraction of event-related potential (ERP). In addition, the results of the proposed algorithm has also been compared with average artifact subtraction (AAS) and optimal basis set (OBS) methods proposed in [11], [14] respectively.

II. Data and Methods

A. Measurements

The data used in this study have been recorded using magnetic-compatible EEG cap (128-channel HydroCel Geodesic Sensor Net shown in Fig. 2). A normal subject having no history of neuronal disorder volunteered himself for the acquisition of ERP dataset and signed the consent from prior to the start of the experiment. The two stimulus oddball paradigm has been presented via a magnetic-compatible projector placed behind the magnetic scanner [20]. The target

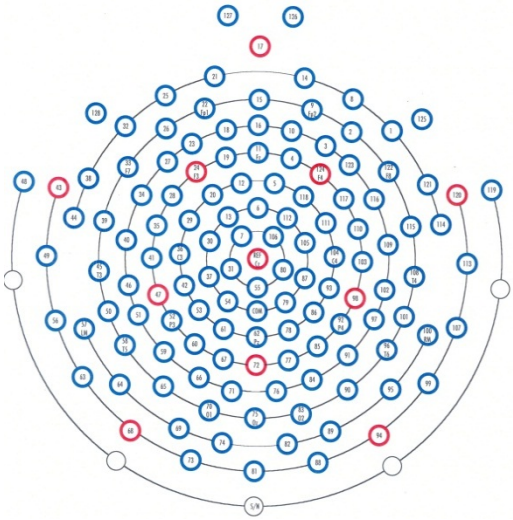


Fig. 2. 128-channel HydroCel Geodesic Sensor Net [24]

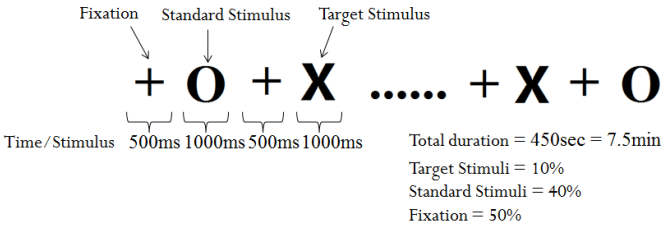


Fig. 3. Two stimulus oddball paradigm used in the study.

stimuli presented were 10% and standard stimuli were 40% of the all experiment contents whereas, fixation symbol which appear between every stimuli irrespective to target or standard were remaining 50%. Figure 3 describes the experiment design used in this study. The subject was instructed to press button when target stimulus appears and do nothing on appearance of standard stimulus. The subject also got familiar with the equipment and all setup in a practiced session before the start of the final acquisition. The fMRI scanner used for simultaneous EEG-fMRI acquisition has the magnetic field of 3.0 T. Moreover, the same experiment has been repeated outside the fMRI scanner and was considered as EEG dataset at 0 T.

B. Data Analysis

The acquired contaminated EEG was pre-processed in which filtering has been performed using Chebyshev bandpass filter. Since the intention was to extract the ERP, the low and high frequency was set to 0.3 and 30 Hz respectively. Both dataset (EEG recorded at 3 T and EEG recorded at 0 T) were filtered using same specifications of bandpass filter. The filtered data were used further to reduce the BCG artifact via EMD-PCA, AAS and OBS. Cleaned data using above mentioned methods were used for analyzing the reduction capabilities and quality of data. The parameters used for evaluation are peak-to-peak value, power spectrum and extraction of ERP. Matlab has been used to process and extract the parameters.

C. Methods

The two well-known methods named as empirical mode decomposition (EMD) [21] and principal component analysis (PCA) [22] has been combined in the proposed algorithm. In the proposed algorithm, EMD has been used to decompose the signals into components known as intrinsic mode functions (IMFs) in order to overcome the limitation of assuming temporal stability of the BCG waveform while using PCA. The method is adaptive to temporal changes in the original and unlike other decomposition methods; it does not require prior information about the sources mixed in the signal [23]. These IMFs are components derived directly from the original data, having different frequencies and variable amplitudes. Secondly, PCA has been performed over the extracted IMFs since it can decompose the correlated set of signals into linearly uncorrelated components named as principal components (PCs). Generally PCs are less than or equal to the number of IMFs and are arranged in descending order of variance. Later on, the PC having more similarity index (PC_{maxSI}) with the contaminated data has been considered as

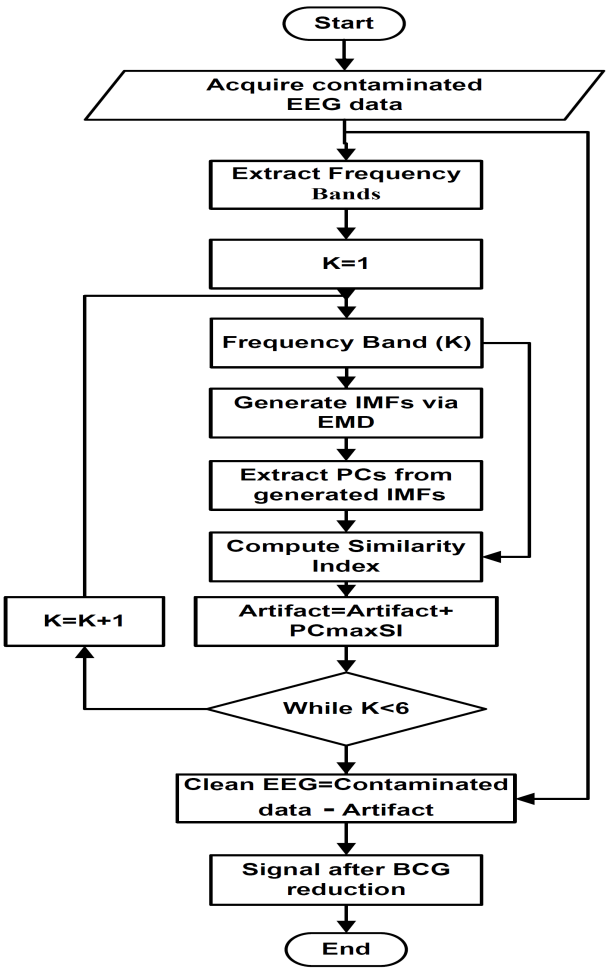


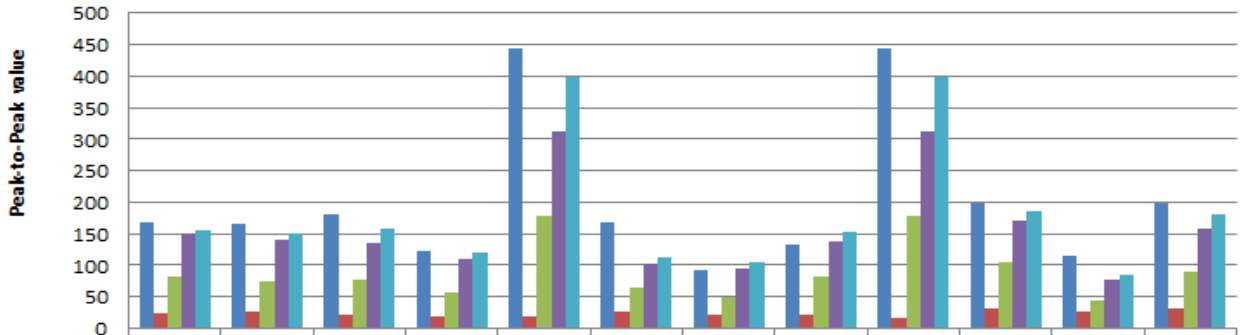
Fig. 4. Flow chart of EMD-PCA algorithm.

BCG artifact and then this component has been subtracted from the contaminated data to remove the BCG artifact. The flow of the proposed algorithm has been presented in Fig. 4.

III. Results and Discussion

The reduction of the BCG artifact using the proposed algorithm has been evaluated via time and frequency domain parameters that are peak-to-peak value and power spectrum. Moreover, event-related potential has also been extracted to assess the capability of the proposed algorithm in preserving the original neuronal signals. The study also provides a comparison of the proposed algorithm with two existing methods named as AAS and OBS. In addition, the EEG data recorded at 0T have been used as a reference data while comparing the results of the above mentioned methods.

First, the EEG data acquired using 128-electrodes for the analysis of peak-to-peak value in time domain were grouped over 11 scalp regions as: Prefrontal (PF): (9, 14, 15, 21, 22), Anterior Frontal (AF): (2, 3, 12, 16, 23, 26, 34), Frontal (F): (1, 4, 11, 14, 21, 15, 19, 24, 27, 32, 33 122, 123, 124), Fronto-Central (FC): (6, 13, 28, 29, 111, 112, 117), Fronto-Temporal (FT): (34, 38, 116, 121), Central (C): (30, 36, 41, 103, 104, 105), Centro-Parietal (CP): (37, 42, 47, 55, 87, 93, 98), Parietal



	PF	AF	F	FC	FT	C	CP	P	T	TP	O	Mean
Contaminated	167.36	164.89	179.62	123.23	443.65	168.04	91.89	133.33	443.65	199.05	114.79	199.54
Outside	25.00	27.36	21.64	19.51	19.41	25.89	21.90	20.72	16.53	31.92	25.89	32.12
EMD-PCA	82.89	74.72	75.90	57.24	178.43	65.61	48.95	81.06	178.43	105.63	43.15	89.76
OBS	149.88	141.38	135.16	109.06	312.32	103.50	95.37	138.84	312.32	171.29	75.94	156.99
AAS	156.73	150.75	158.54	120.80	397.41	112.38	103.94	151.75	397.41	185.53	85.69	181.58

Fig. 5. Region-wise peak-to-peak values of contaminated data (Blue), EEG data at 0T (Red), data corrected via EMD-PCA (Green), data corrected using OBS (Purple) and data corrected by AAS (light blue).

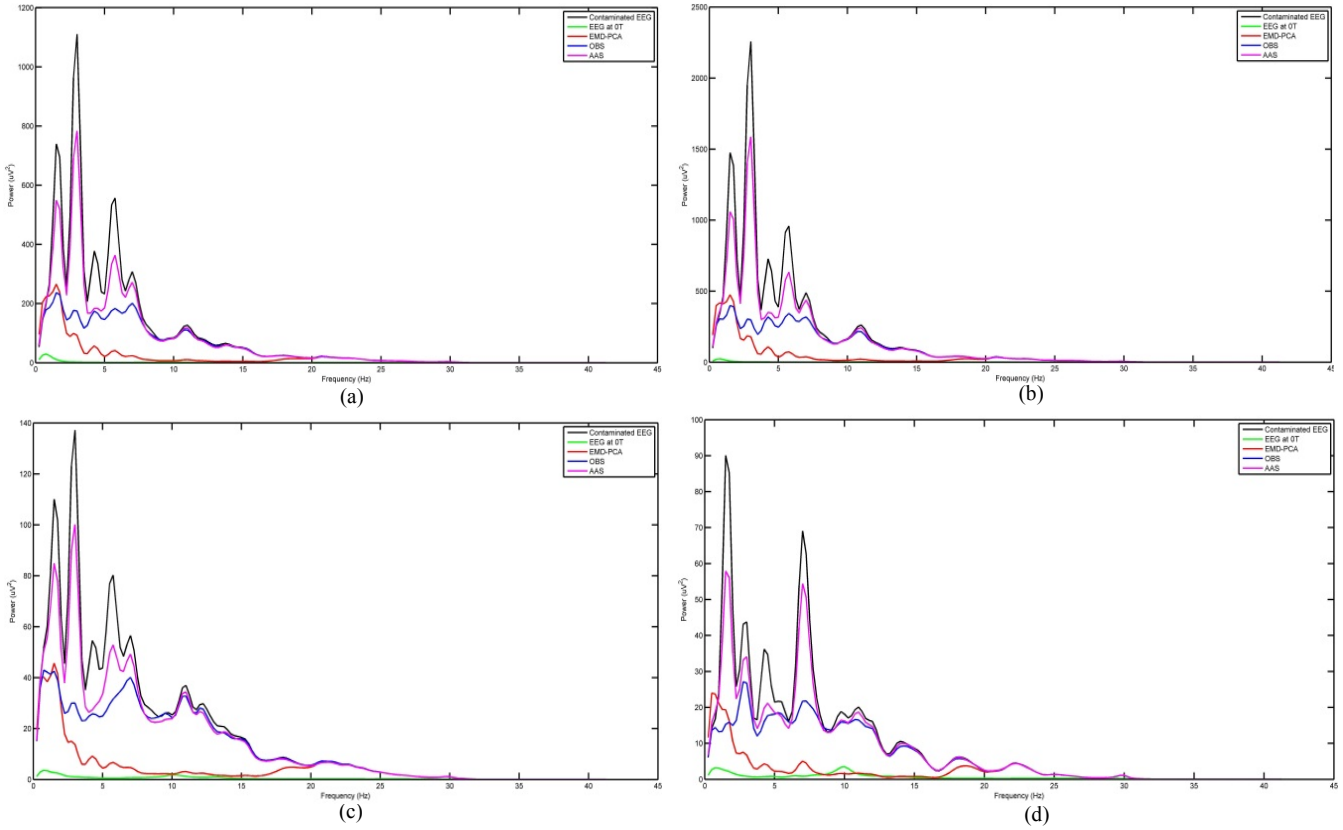


Fig. 6. Power spectrum of Contaminated EEG (Black), EEG at 0T (Green) and data after removal of BCG artifact using EMD-PCA (Red), OBS (Blue) and AAS (Pink) of: (a) Frontal, (b) Temporal, (c) Central and (d) Occipital regions.

(P): (51, 52, 60, 64 85, 92, 95, 97), Temporal (T): (34, 38, 44, 45, 108, 114, 116, 121), Tempo-Parietal (TP): (46, 57, 58, 96, 100, 102) and Occipital (O): (70, 75, 83). The difference between maximum peak and minimum peak in the EEG epochs based on R-R interval of ECG has been calculated and averaged per channel and then averaged over the electrodes per region. Equation 1 describes the calculation of the peak-to-peak value.

$$Peak-to-Peak\ value = \frac{1}{M}\left[\sum_{j=1}^{M}\left(\frac{1}{N}\sum_{i=1}^{N} epoch_i\right)_j\right] \quad (1)$$

where M is the number of electrodes in the region and N is the total number of epochs per channel. The peak-to-peak value of the contaminated EEG data (Contaminated data), EEG data recorded at 0T (Outside data) and EEG data after reducing BCG artifact via the proposed algorithm (EMD-PCA), OBS and AAS has been calculated and presented in Fig. 5. The

results show that all methods can significantly reduce the BCG artifact when compared to the contaminated data. The comparison with the EEG data recorded at 0T shows that both OBS and AAS have residuals of the BCG artifact that still have significant influence over the peak-to-peak value compared to the value of EMD-PCA which is closer to the outside data. The results of the peak-to-peak value describe the improvement in reduction of the BCG artifact via EMD-PCA compared to OBS and AAS. Moreover, it can be noticed that the temporal region has the most influence of the BCG and without an appropriate reduction it may hinder the assessment of brain activities at this lobe.

Secondly, Fast Fourier Transform (FFT) has been used to calculate the power spectrum to analyze the reduction of BCG artifact in frequency domain. The power spectrum of four scalp regions: Frontal, Temporal, Central and Occipital has been calculated using welch window with specifications; window size equals to 500 samples, overlapping region was 50% of window size and 1024 are the number of point to compute FFT. The results presented in Fig. 6 show that the BCG artifact affects the EEG recordings more severely over the frequency range of 0.5 to 15 Hz; the fundamental frequencies of EEG. The variation in powers at different region again depicts the conclusion as above that the BCG effect is more in the temporal region. Whereas, comparing the powers of contaminated data, EEG at 0T, BCG reduced using EMD-PCA, BCG reduced using OBS and AAS show that EMD-PCA has better capability of reduction compared to AAS and OBS as the power spectrum is relatively closer to the power spectrum of EEG at 0T.

The intuition behind choosing peak-to-peak value and power spectrum is to evaluate the capability of the proposed algorithm in reducing the BCG artifact from the contaminated data in respective domains (time and frequency). It can also be seen from Fig. 6 that the power of EEG data contaminated by the BCG artifact has larger values compared to EEG data without contamination. The increase in power is because of larger amplitudes in contaminated data. However, these metrics are not enough to validate the proposed algorithm while looking at the preservation of original neuronal signals. The preservation capability assures that the quality of EEG data is conserved or undistorted while reducing the artifact. Event-related potential (ERP) from the EEG data after reduction of the BCG artifact has been extracted to analyze the preservation capability. The extracted ERP from the dataset, reconstructed using the proposed algorithm, AAS and OBS has been presented in Fig. 7, in which P300 has been analyzed at electrode Pz. From Fig. 7 it can be concluded that ERP can be extracted from all datasets, but the strength and latency of P300 are the parameters that should be looked after for the comparison of the methods. Table 1 shows that the proposed method has improved the SNR and latency of P300 compared to the OBS and AAS.

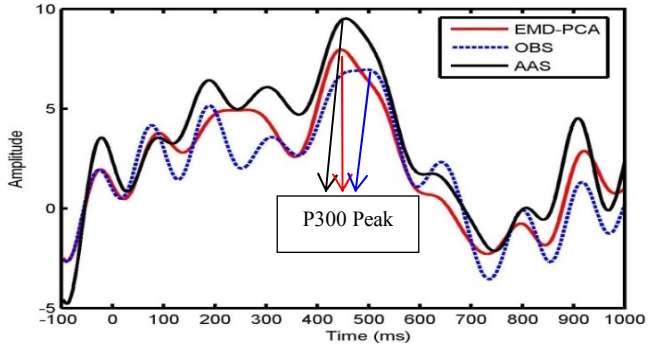


Fig. 7. The extracted ERP from the dataset reconstructed using proposed algorithm (red), AAS (black) and OBS (blue) at PZ electrode.

TABLE I. THE SIGNAL-TO-NOISE RATIO (SNR) AND LATENCY OF P300 EXTRACTED FROM THE DATA RECONSTRUCTED BY EMD-PCA, OBS AND AAS AT PZ ELECTRODE.

Methods	SNR (dB)	Latency (ms)
AAS	3.4939	456
OBS	6.5762	496
EMD-PCA	7.2343	444

IV. CONCLUSION

BCG artifact contaminates EEG data when recorded concurrently with fMRI and hinders in assessing the original neuronal activations. A reference free composite algorithm has been introduced in this study to reduce the contamination of the BCG artifact. The algorithm overcomes the drawback of existing method as they rely on the ECG acquisition to identify the onsets of the BCG waveform which can lead to erroneous identification and loss of neuronal data, if recordings of ECG are also corrupted. Moreover, the proposed algorithm does not require assuming any kind of stability in the BCG waveform which algorithms like AAS and OBS do. The reduction of the BCG artifact using the proposed algorithm has been evaluated in both time and frequency domains. In addition to validate the preservation capability of the proposed algorithm ERP has been extracted from the BCG reduced data. Compared to to AAS and OBS, the proposed algorithm showed better performances. In future works, the algorithm will be tested over large dataset. .

REFERENCES

[1] H. Berger, "On the electroencephalogram of man. Sixth report.," *Electroencephalogr. Clin. Neurophysiol. Suppl-28.*, 1969.

[2] H. Laufs, J. Daunizeau, D. W. Carmichael, and a Kleinschmidt, "Recent advances in recording electrophysiological data simultaneously with magnetic resonance imaging.," *Neuroimage*, vol. 40, no. 2, pp. 515–28, Apr. 2008.

[3] N. G. Müller and R. T. Knight, "The functional neuroanatomy of working memory: contributions of human brain lesion studies.," *Neuroscience*, vol. 139, no. 1, pp. 51–58, 2006.

[4] H. Amin and A. S. Malik, "Human memory retention and recall processes: A review of EEG and fMRI studies," *Neurosciences*, vol. 18, no. 4, pp. 330–344, 2013.

[5] J. Gotman, E. Kobayashi, A. P. Bagshaw, C.-G. Bénar, and F. Dubeau, "Combining EEG and fMRI: a multimodal tool for epilepsy research.," *J. Magn. Reson. Imaging*, vol. 23, no. 6, pp. 906–920, 2006.

[6] R. F. Ahmad, A. S. Malik, N. Kamel, and F. Reza, "A proposed frame work for real time epileptic seizure prediction using scalp EEG," *2013 IEEE Int. Conf. Control Syst. Comput. Eng.*, pp. 284–289, Nov. 2013.

[7] P. M. Rossini, S. Rossi, C. Babiloni, and J. Polich, "Clinical neurophysiology of aging brain: from normal aging to neurodegeneration.," *Prog. Neurobiol.*, vol. 83, no. 6, pp. 375–400, 2007.

[8] W. Nakamura, K. Anami, T. Mori, O. Saitoh, A. Cichocki, and S. Amari, "Removal of ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data using independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1294–308, Jul. 2006.

[9] F. Ghaderi, K. Nazarpour, J. G. McWhirter, and S. Sanei, "Removal of Ballistocardiogram Artifacts Using the Cyclostationary Source Extraction Method," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 11, pp. 2667–2676, Jul. 2010.

[10] F. Grouiller, L. Vercueil, A. Krainik, C. Segebarth, P. Kahane, and O. David, "A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI.," *Neuroimage*, vol. 38, no. 1, pp. 124–37, Oct. 2007.

[11] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux, "Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction.," *Neuroimage*, vol. 8, no. 3, pp. 229–39, Oct. 1998.

[12] G. Bonmassar, "Motion and Ballistocardiogram Artifact Removal for Interleaved Recording of EEG and EPs during MRI," *Neuroimage*, vol. 16, no. 4, pp. 1127–1141, Aug. 2002.

[13] C.-G. Bénar, Y. Aghakhani, Y. Wang, A. Izenberg, A. Al-Asmi, F. Dubeau, and J. Gotman, "Quality of EEG in simultaneous EEG-fMRI for epilepsy," *Clin. Neurophysiol.*, vol. 114, no. 3, pp. 569–580, Mar. 2003.

[14] R. K. Niazy, C. F. Beckmann, G. D. Iannetti, J. M. Brady, and S. M. Smith, "Removal of FMRI environment artifacts from EEG data using optimal basis sets," *Neuroimage*, vol. 28, no. 3, pp. 720–37, Nov. 2005.

[15] T. Rasheed, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "Attenuation of artifacts in EEG signals measured inside an MRI scanner using constrained independent component analysis," *Physiol. Meas.*, vol. 30, no. 4, p. 387, 2009.

[16] S. Ferdowsi, S. Sanei, V. Abolghasemi, J. Nottage, and O. O'Daly, "Removing ballistocardiogram artifact from EEG using short- and long-term linear predictor," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1900–11, Jul. 2013.

[17] Z. Liu, J. a de Zwart, P. van Gelderen, L.-W. Kuo, and J. H. Duyn, "Statistical feature extraction for artifact removal from concurrent fMRI-EEG recordings," *Neuroimage*, vol. 59, no. 3, pp. 2073–87, Feb. 2012.

[18] J. C. De Munck, P. J. van Houdt, S. I. Gonçalves, E. van Wegen, and P. P. W. Ossenblok, "Novel artefact removal algorithms for co-registered EEG/fMRI based on selective averaging and subtraction," *Neuroimage*, vol. 64, pp. 407–15, Jan. 2013.

[19] S. Debener, K. J. Mullinger, R. K. Niazy, and R. W. Bowtell, "Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength," *Int. J. Psychophysiol.*, vol. 67, no. 3, pp. 189–99, Mar. 2008.

[20] A. A. Stevens, P. Skudlarski, J. C. Gatenby, and J. C. Gore, "Event-related fMRI of auditory and visual oddball tasks.," *Magn. Reson. Imaging*, vol. 18, no. 5, pp. 495–502, 2000.

[21] 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," *Proc. R. Soc. A Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[22] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *J. Educ. Psychol.*, vol. 24, no. 6, pp. 417–441, 1933.

[23] T. Wang, M. Zhang, Q. Yu, and H. Zhang, "Comparing the applications of EMD and EEMD on time–frequency analysis of seismic signal," *J. Appl. Geophys.*, vol. 83, pp. 29–34, Aug. 2012.

[24] "HydroCel Geodesic Sensor Net 128 Channel Map." [Online]. Available: ftp.egi.com/pub/documentation/.../hcgsn_128.pdf.