Removal of Eye Blink Artifacts From EEG Signals Using Sparsity

S. R. Sreeja, Rajiv Ranjan Sahay, Debasis Samanta, and Pabitra Mitra

***Abstract*—Neural activities recorded using electroencephalography (EEG) are mostly contaminated with eye blink (EB) artifact. This results in undesired activation of brain–computer interface (BCI) systems. Hence, removal of EB artifact is an important issue in EEG signal analysis. Of late, several artifact removal methods have been reported in the literature and they are based on independent component analysis (ICA), thresholding, wavelet transformation, etc. These methods are computationally expensive and result in information loss which makes them unsuitable for online BCI system development. To address the above problems, we have investigated sparsity-based EB artifact removal methods. Two sparsity-based techniques namely morphological component analysis (MCA) and K-SVD-based artifact removal method have been evaluated in our work. MCA-based algorithm exploits the morphological characteristics of EEG and EB using predefined Dirac and discrete cosine transform (DCT) dictionaries. Next, in K-SVD-based algorithm an overcomplete dictionary is learned from the EEG data itself and is designed to model EB characteristics. To substantiate the efficacy of the two algorithms, we have carried out our experiments with both synthetic and real EEG data. We observe that the K-SVD algorithm, which uses a learned dictionary, delivers superior performance for suppressing EB artifacts when compared to MCA technique. Finally, the results of both the techniques are compared with the recent state-of-the-art FORCe method. We demonstrate that the proposed sparsity-based algorithms perform equal to the state-of-the-art technique. It is shown that without using any computationally expensive algorithms, only with the use of over-complete dictionaries the proposed sparsity-based algorithms eliminate EB artifacts accurately from the EEG signals.**

***Index Terms*—Artifact removal, brain computer interface (BCI), dictionary learning, electroencephalography (EEG), K-SVD, morphological component analysis (MCA).**

I. Introduction

THE objective of brain-computer interface systems (BCIs) is to provide human beings with an alternative mode of communication for directly conveying human intentions to a computer without the use of peripheral nerves and muscles of human body [1]. For people with severe physical disabilities, such as limb damage, brainstem stroke, amyotrophic lateral sclerosis (ALS) [2], [3], spinal cord injury [4] or other neuromuscular diseases, BCI is found to be the best mode to communicate [5]. Among the available non-invasive devices, EEG is unique and most often used in BCI research, as it provides high temporal resolution of the measured brain signals, is less bulky and mobile, relatively convenient, affordable and safe for users [3]. EEG measures the synaptic activity as voltage fluctuations along the brain scalp due to the flow of electrical signals between the neurons [5]. Traditionally, EEG has been used for medical BCI applications, but in recent years, there has been a great interest in non-medical BCI applications such as device control, training and education, and gaming/entertainment [3], [6], [7].

However, raw EEG signals are often contaminated by non-cerebral signals such as ocular artifacts, muscle artifacts, cardiac artifacts, power line interference and electrode artifacts [8]. Although, reliable solutions exist for dealing with most EEG artifacts it is really hard for human subjects to avoid ocular artifacts. The levator muscle potential developed as a result of eye blink (EB) is 10 times larger in amplitude than the EEG and is the most dominant over other artifacts. On an average, a human involuntarily blinks once in 5 s and the blinking activity lasts for 100–400 ms [9], [10]. Due to high magnitude of the EB signal and high resistance of the skull and scalp tissues, the EB artifact contaminates majority of the electrode signals [11]. Hence, EB artifact removal is an important problem in EEG signal processing. It is crucial as it can affect the detection and extraction of EEG features which leads to undesired activations of a BCI system besides causing frustration in users.

Although, in the literature there are several algorithms available to remove EB artifacts, they still have major limitations. Some techniques use extra elctrooculography (EOG) channels [12], which may not always be available. Other blind source separation techniques such as principal component analysis (PCA) [13], independent component analysis (ICA) [14] require prior knowledge of artifacts to detect artifactual sources or contaminated components and they are computationally expensive. In thresholding-based ICA technique [15], the threshold value needs to be decided a priori in order to avoid any information loss.

To address the above problems, we need a method or representation which can efficiently and accurately separate EB

Manuscript received May 23, 2017; revised September 9, 2017 and October 24, 2017; accepted October 30, 2017. Date of publication November 13, 2017; date of current version August 31, 2018. *(Corresponding author: Sreeja SR.)*

S. R. Sreeja, D. Samanta, and P. Mitra are with the Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur 701302, India (e-mail: sreejasr@iitkgp.ac.in; dsamanta@sit.iitkgp.ernet.in; pabitra@cse.iitkgp.ernet.in).

R. R. Sahay is with the Department of Electrical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 701302, India (e-mail: rajiv@ee.iitkgp.ernet.in).

Digital Object Identifier 10.1109/JBHI.2017.2771783

and artifact-free EEG signal without any prior information from the user. Recently, sparse representation of natural signals has received great attention and has been used in many applications such as image denoising [16], image compression [17], speech signal compression [18] and speech signal classification [19]. According to compressive sensing theory, many natural signals can be sparsely represented when expressed using a convenient basis [20]. Subsequently, numerical optimization procedures can be used to perfectly reconstruct the original signal from the sparse representation [21]. However, these advantages of sparse representation are rarely studied for the purpose of artifact removal. Hence, in this work, we investigated two different sparsity-based approaches to remove EB artifacts from EEG data.

In the first sparsity-based approach, we propose to use morphological component analysis (MCA) technique, which separates the EEG signals into components that have different morphological characteristics. Each morphological component is sparse in an over-complete dictionary constructed by merging pre-defined DCT and Dirac basis. From the obtained morphological components, the component corresponding to EB is subtracted from the original EEG signals to yield artifact-free EEG. In the second proposed sparsity-based method, K-SVD algorithm is applied wherein a sparsifying over-complete dictionary is learned from the observed EEG signals contaminated with EBs. Then using sparse coefficients and the learned dictionary, EBs are estimated and subtracted from the original recorded EEG signal. To the best of our knowledge, no work has been done to remove EB artifacts from EEG using sparsity and over-complete learned dictionaries.

The proposed sparsity-based methods are designed to remove EB artifacts from the EEG signals accurately without the use of additional EOG recordings. Besides, unlike other existing methods [22], no knowledge of artifacts and no selection of artifactual sources is required. To demonstrate the efficacy of our proposed methods we have conducted several experiments on both synthetic and real EEG data. The results have been evaluated using different statistical performance metrics. Our experiments confirm the capability of the two proposed sparsity-based techniques for separation of EB artifacts from the EEG signal. The results show that, out of the two proposed techniques, K-SVD based artifact removal method outperforms the MCA approach. Finally, the efficacy of these sparsity-based methods has been evaluated and compared with the recent state-of-the-art FORCe (Fully Online and Automated Artifact Removal for Brain-Computer Interfacing) method [15]. Note that we do not use independent component analysis (ICA), wavelet decomposition and higher order statistical properties of EEG signal. Only with a single over-complete dictionary learned from the given EEG data, the proposed K-SVD based method eliminates EB artifacts accurately from the observed EEG signal.

Our paper is organised as follows. In Section II, we present an overview of related research works. In Section III, we provide the mathematical background of proposed techniques along with algorithms for the purpose of EB artifact removal. The experimental results using both simulated and real EEG data are presented in Section IV. Comparison of the proposed approaches with the existing state-of-the-art techniques is given in Section V. In Section VI, a discussion regarding phase delay and computation time of the proposed methods are given. Finally, conclusions and directions for future work are outlined in Section VII.

II. Literature Survey

Investigating EB artifact properties is important for researchers to devise novel techniques for their removal. A variety of methods have been proposed to remove ocular artifacts [12]–[15] [22]–[29]. One of the primary approaches to restrict EB is eye fixation on a stationary target [23]. In this method the user is asked to stare at a stable point for fixed amount of time which is unrealistic and not suitable for BCI applications. Further, this fixation method [23] does not eliminate involuntary eye blinks [14]. Regression based methods [25] play an important role in the removal of ocular artifacts including blinks. Artifacts in EEG can be removed by subtracting the weighted noise (EOG) from signals. However, the main disadvantage of this method is that it requires extra ocular channels to record ocular artifacts [12].

There is also accord among researchers in using techniques such as principal component analysis (PCA) [30], independent component analysis (ICA) [14], wavelet based denoising [24] for ocular artifact correction. Among these algorithms, ICA is a well-known blind source separation (BSS) technique to remove ocular artifacts from EEG signals. ICA based methods assume that the EEG recordings can be considered to be a linear mixture of several components, namely, non-artifactual and artifactual independent components (ICs). Artifactual ICs represent potential of non-cerebral artifacts whereas non-artifactual ICs represent electrical signals originating from brain. After detecting the ocular artifact components and eliminating them manually, the clean EEG can be acquired by remixing remaining components [26]. The most important issue in ICA based artifact removal methods is selecting the correct number of ICs, that is, extracted components which should be deflated [27]. Since these ICs do not unavoidably only contain artifact data, but also contain the fundamental EEG data, removing contaminated ICs will lead to loss of useful data [28]. Hence, some researchers incorporated empirical mode decomposition (EMD) [29], wavelet decomposition [12], wavelet neural network [22] and higher order statistics such as multiscale sample entropy [10] and kurtosis [15] into their ICA-based algorithms to extract specific sources with known behaviours. However, many of these approaches are not suitable for BCI applications since they are computationally expensive.

In recent years, algorithms exploiting sparse representation of signals have witnessed a remarkable growth in variety of applications such as image compression [17], speech signal compression [18] and image denoising [16], but they have been rarely studied for the purpose of artifact removal from EEG signals. For instance, in [31] and [32], pre-defined basis are applied to sparsely represent EEG signals for removing existing artifacts. The work in [33] successfully removes ballistocardiogram (BCG) artifact modeled from EEG-fMRI data taking advantage of sparse representation and learned dictionaries. MCA technique finds use in image-content separation problem [34] where

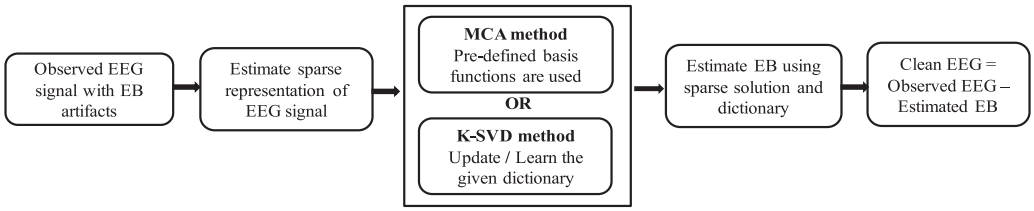

Fig. 1. Framework of the proposed approach.

it is used to separate the texture and the cartoon parts. It has also been applied to signal separation, image separation, biomedical engineering and speech processing [35]. The success of MCA is dependent upon the choice of appropriate dictionaries (basis functions). In the literature, curvelet dictionaries are found appropriate to represent piecewise smooth images [36], wavelet dictionaries can be used to represent isotropic structures [37] and ridgelet dictionaries are highly recommended to represent global lines in an image [38]. If an appropriate dictionary is identified, the use of sparse representation will lead to the desired signal separation [35]. In K-SVD based artifact removal method instead of using pre-chosen set of basis functions, it is proposed to learn the over-complete dictionary from examples [16]. K-SVD algorithm is an iterative method that updates sparse coefficients and dictionary atoms jointly, which results in accelerated convergence [39]. K-SVD algorithm is proven to be simple, flexible and efficient for the image denoising problem [16].

III. Methodology

In this section, each of the two proposed sparsity-based methodologies is described briefly and their mathematical background is given.

A. Proposed EB Removal Techniques

The framework of the proposed approach is given in Fig. 1. The first task in sparsity-based approach is to obtain the l_0-norm sparse representation for the observed EEG signal. With the obtained sparse solution, the degraded EEG signals are modeled using MCA and K-SVD approaches. In MCA method the dictionaries are pre-defined and remain unchanged throughout the process of estimation of sparse coefficients. In K-SVD method the randomly initialized dictionary is updated using the estimated sparse solution. Then with the estimated sparse solution and the updated dictionary matrix, EB artifacts are estimated and subtracted from the degraded EEG signal to obtain clean EEG data. The proposed algorithms are explained briefly in the following sections.

1) Sparse representation of signals: The objective of sparse representation is to compute the sparse coefficients $\boldsymbol{s} \in \mathbb{R}^k$, given the observed signal $\boldsymbol{y} \in \mathbb{R}^m$ and over-complete dictionary matrix $D \in \mathbb{R}^{m \times k}$ (where $k > m$ implies redundancy) [16]. The EEG signal $\boldsymbol{y}$ of length m is divided into n segments $\boldsymbol{y} = \{\boldsymbol{y}_i\}_{i=1}^n$. Each of n segments $\{\boldsymbol{y}_i\}_{i=1}^n$ are represented by $\{\boldsymbol{s}_i\}_{i=1}^n$ sparse coefficient vectors. The dictionary D can be randomly generated or produced from the collection of several known basis functions such as wavelet, curvelet, Fourier, DCT, Dirac bases [40]. The columns of the dictionary matrix D are called atoms and denoted as $\boldsymbol{d}_j, j = 1, 2.., k$. Assuming D as *fixed*, the signal segment $\boldsymbol{y}_i$ can be represented as a sparse linear combination of atoms of dictionary $\{\boldsymbol{d_j}\}_{j=1}^k$, such that $\boldsymbol{y}_i = D\boldsymbol{s}_i$ with good approximation. Infinite number of solutions are possible if D is a full-rank matrix; hence prior constraints must be imposed on the solution [39], i.e.

$$\min_{\boldsymbol{s}_i} \|\boldsymbol{s}_i\|_0 \text{ subject to } \boldsymbol{y}_i = D\boldsymbol{s}_i \tag{1}$$

where $\|.\|_0$ is l_0- norm, counting non-zero entries in $\boldsymbol{s}_i$. Solving equation (1) is both numerically unstable and NP-hard, requiring enumeration of all possible locations of nonzero entries in $\boldsymbol{s}_i$ [21]. Nevertheless, methods exist in the literature for obtaining an approximate solution which is found to be quite accurate if the solution is sparse enough [41], [42]. In this work, orthogonal matching pursuit (OMP) algorithm has been used to find an approximate solution of equation (1) because of its efficiency and simplicity [43].

Algorithm 1: OMP algorithm: s = OMP (y, D, T).

Input: Observed signal $\boldsymbol{y}$, Dictionary D, sparsity threshold T
Output: $\boldsymbol{s}$
Initialize: the residual $\boldsymbol{r}_0 = \boldsymbol{y}$, the index set $I_0 = \varnothing$, $t = 1$
while $t < T$ **do**
  $i_t = \arg\max_{j=1,2,..,k} |\langle \boldsymbol{r}_{t-1}, \boldsymbol{d}_j \rangle|$
  $I_t = I_{t-1} \cup \{i_t\}$
  $\boldsymbol{s}_t = \boldsymbol{d}_{I_t}^{\dagger} \boldsymbol{y}$
  $\boldsymbol{r}_t = \boldsymbol{y} - D\boldsymbol{s}_t$
  $t = t + 1$
end
return $\boldsymbol{s} = \boldsymbol{s}_t$

The steps of OMP [41] which is a greedy technique have been outlined in Algorithm 1. At each iteration, this algorithm attempts to find the atoms, that are maximally correlated with the residual vector $\boldsymbol{r}$. This implies that the selected atom has maximum information and hence maximally minimizes the error in reconstruction. Then OMP appends the index of selected atoms i_t to the active set I_t. In the next step, orthogonalization is carried out, where $\boldsymbol{d}_{I_t}^{\dagger}$ indicates Moore-Penrose pseudo-inverse of $\boldsymbol{d}_{I_t}$ i.e., $\boldsymbol{d}_{I_t}^{\dagger} := (\boldsymbol{d}_{I_t}^T \boldsymbol{d}_{I_t})^{-1} \boldsymbol{d}_{I_t}^T$. More importantly, this algorithm ensures that the residual $\boldsymbol{r}_t$ is always orthogonal to atoms indexed in I_t. Accordingly, the correlation of the active atoms I_t will be zero in the next iteration. Hence, no atom is selected

Algorithm 2: MCA algorithm.

Input: Observed signal $\boldsymbol{y}$, dictionaries $\psi_1, \psi_2, ..\psi_N$, number of iterations N_{iter}, sparsity threshold T
Output: Morphological components $\hat{\boldsymbol{y}}_1, \hat{\boldsymbol{y}}_2, ..\hat{\boldsymbol{y}}_N$
Initialize: $\mathbf{r}_n^0 = \boldsymbol{y}$, $\hat{\boldsymbol{y}}_n^0 = 0$
for $t = 1...N_{iter}$ **do**
  for $n = 1...N$ **do**
    update residual $\mathbf{r}_n^t = \mathbf{r}_n^{t-1}$
    compute sparse coefficient vector, $\mathbf{s}_n^t$, using OMP
    $\mathbf{s}_n^t$ = OMP $(\mathbf{r}_n^t, \psi_n, T)$
    update n^{th} component $\hat{\boldsymbol{y}}_n^t = \psi_n \mathbf{s}_n^t$
    compute residual $\mathbf{r}_n^t = \boldsymbol{y} - \sum_{n=1}^{N} \hat{\boldsymbol{y}}_n^t$
  end
end

twice in OMP. The iteration is repeated until the pre-defined threshold T which controls the sparsity is attained.

2) Morphological component analysis (MCA): This is a method for decomposing a signal into its parts possessing different morphological characteristics [31], [32]. MCA depends on the hypothesis that for every signal to be separated, there is a dictionary of atoms that allows its reconstruction using a sparse representation [35]. With the benefit of sparse representation $\boldsymbol{s}$, it assumes that a signal $\boldsymbol{y}$ can be represented as linear combination of N morphological components, $\boldsymbol{y}_i$:

$$\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{s} = \sum_{i=1}^{N} \psi_i \mathbf{s}_i = \sum_{i=1}^{N} \boldsymbol{y}_i \tag{2}$$

where $\boldsymbol{\Phi} = \{\psi_1, \psi_2, ...\psi_N\}$. Note that $\{\psi_1, \psi_2, ...\psi_N\}$ are the basis matrices or dictionaries and $\boldsymbol{s} = \{\mathbf{s}_1, \mathbf{s}_2, ..\mathbf{s}_N\}$ are the corresponding sparse coefficient vectors. Choosing the correct dictionary is crucial to create a good sparse decomposition. If appropriate dictionaries are used, then the OMP algorithm searching for the sparsest solution can lead to the desired separation of EEG and EB artifacts. For the problem at hand, discrete cosine transform dictionaries are used to represent the background EEG signals [32] and Dirac dictionaries have been used to represent EB artifacts [44].

In this work, the degraded EEG signal $\boldsymbol{y}$ is considered as a linear combination of two morphological components $\hat{\boldsymbol{y}}_1$, $\hat{\boldsymbol{y}}_2$ obtained using DCT and Dirac basis and it is represented as

$$\boldsymbol{y} = \hat{\boldsymbol{y}}_1 + \hat{\boldsymbol{y}}_2 = \psi_{DCT}\mathbf{s}_{DCT} + \psi_D \mathbf{s}_D \tag{3}$$

where $\mathbf{s}_{DCT}$ and $\mathbf{s}_D$ are the sparse coefficient vectors corresponding to the DCT and Dirac basis, ψ_{DCT} and ψ_D, respectively. The steps followed to separate EEG and EB components using MCA technique are given in Algorithm 2. Using the estimated morphological components, the components corresponding to the EB artifact are subtracted from the original degraded EEG to obtain the artifact-free signal.

3) K-SVD based artifact removal method: In the previous section, to find sparse vector $\boldsymbol{s}$, the dictionary ($D = \boldsymbol{\Phi}$) is predefined and assumed fixed. However, such a matrix may not be the optimal over-complete dictionary. Hence, in this work K-SVD algorithm is used to fine-tune or update the dictionary

Algorithm 3: K-SVD algorithm.

Input: Observed signal $\boldsymbol{y}$, initial estimate of dictionary $D = D_0$, sparsity parameter T, K denotes number of atoms in dictionary, N is the number of K-SVD iterations.
Output: $D, \boldsymbol{s}$
for $t = 1...N$ **do**
  $\forall i$ $\mathbf{s_i}$ = OMP $(\boldsymbol{y}_i, D, T)$
  for $l = 1...K$ **do**
    Find the set of signals that use the current atom, $w_l = \{i \mid 1 \le i \le n, s^l \ne 0\}$
    Compute overall representation error E_l $E_l = \boldsymbol{y} - \sum_{j \ne l} \boldsymbol{d}_j \boldsymbol{s}^j$
    Set E_l as matrix by choosing only those columns in w_l and obtain E_l^R
    $[U, S, V^T] = SVD(E_l^R)$
    Update the dictionary column $\boldsymbol{d}_l$ with the first column of U
    $\boldsymbol{s}^l \Leftarrow S_{1,1} * V_1$
  end
end

D, one atom at a time, so as to further reduce the reconstruction error. This algorithm is flexible and works with any pursuit technique. Mathematically, the dictionary learning problem for the given signal $\boldsymbol{y} = \{\boldsymbol{y}_i\}_{i=1}^n$ can be defined as

$$\min_{D, \boldsymbol{s}_i} \sum_{i}^{n} \|\boldsymbol{y}_i - D\boldsymbol{s}_i\|_F^2 \text{ subject to } \|\boldsymbol{s}_i\|_0 \le T \tag{4}$$

where $\|.\|_F$ denotes Frobenius norm. The parameter T controls the number of non-zero entries in $\boldsymbol{s}_i$ and should satisfy $T \ll k$, where k is the number of atoms in the dictionary D i.e., $D = \{\boldsymbol{d}_j\}_{j=1}^k$. K-SVD algorithm consists of two steps. In the first step it finds the sparse representation vector $\boldsymbol{s}_i$ using OMP algorithm. Subsequently, in the second step the dictionary D is updated such that it best represents the observed signal $\boldsymbol{y}_i$ for the estimated coefficient vector $\boldsymbol{s}_i$. For a detailed study of K-SVD algorithm we refer the reader to [39]. This algorithm is called "K-SVD" because it carries out K number of singular value decomposition (SVD) computations to obtain the learned dictionary D [39]. Here, $K = k$, the number of atoms in the dictionary, and each atom in D is updated as outlined in Algorithm 3.

We denote the given raw EEG signal by $\boldsymbol{y}$, the clean EEG signal as $\boldsymbol{c}$ and EB artifact as $\boldsymbol{z}$, then these quantities can be related as

$$\boldsymbol{y} = \boldsymbol{z} + \boldsymbol{c}. \tag{5}$$

The aim of this work is to model the eye blink $\boldsymbol{z}$ using dictionary learning technique so as to separate it from $\boldsymbol{y}$. Initially, this method is applied to single channel EEG data $\boldsymbol{y}$ of length m. For learning the dictionary, we require n training signals as input. Hence the original signal is divided into n smaller segments which can have varying degrees of overlap with each other. In this work, we used fully overlapping segments in order to capture

the behaviour of the entire observed EEG signal. However, our main objective is to remove EB artifact from EEG, hence the original dictionary learning problem is modified as:

$$\min_{D,s_i,z} \lambda \left\|\boldsymbol{y}-\boldsymbol{z}\right\|_2^2 + \sum_i \mu_i \left\|\boldsymbol{s}_i\right\|_0 + \sum_i \left\|D\boldsymbol{s}_i - R_i\boldsymbol{z}\right\|_2^2 . \quad (6)$$

Here, the first term is the Euclidean norm between observed EEG $\boldsymbol{y}$ and eye blink $\boldsymbol{z}$ and λ is a Lagrange multiplier. The other terms are equivalent to the problem of dictionary learning. R_i is a binary matrix which extracts ith segment from $\boldsymbol{z}$ i.e. $\boldsymbol{z}_i = R_i\boldsymbol{z}$. The parameter μ_i controls the level of sparsity. To calculate $\boldsymbol{z}$ we assume D and $\boldsymbol{s}_i$ as fixed in equation (6) and take its derivative with respect to $\boldsymbol{z}$ equating it to zero. The solution is given as

$$\hat{\boldsymbol{z}} = \left(\lambda I + \sum_i R_i^T R_i\right)^{-1} \left(\lambda\boldsymbol{y} + \sum_i R_i^T D\boldsymbol{s}_i\right). \quad (7)$$

In the above equation, I refers to an identity matrix, $\hat{\boldsymbol{z}}$ is the estimated EB artifact of the observed EEG signal. Subtracting it from $\boldsymbol{y}$ gives us the clean EEG $\hat{\boldsymbol{c}}$ i.e. $\hat{\boldsymbol{c}} = \boldsymbol{y} - \hat{\boldsymbol{z}}$. In equation (7) $\hat{\boldsymbol{z}}$ can be calculated sample-wise so that handling huge matrices can be avoided.

B. Performance Metrics

The performance of our proposed artifact removal methods is evaluated using various quantitative metrics such as root mean square error, signal to artifact ratio, correlation coefficient, mutual information and power spectral density.

Root Mean Square Error (RMSE) or Root Mean Square Deviation (RMSD) [22] is the measure of difference between the estimated artifact-free EEG signal ($\hat{\boldsymbol{c}}$) and the data actually observed ($\boldsymbol{y}$). This difference is also called as residuals and RMSE aggregates them into a measure of predictive power. It is computed using the equation:

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}\left(\boldsymbol{y}^t - \hat{\boldsymbol{c}}^t\right)^2}{n}} \quad (8)$$

where n is the length of the corrupted EEG signal $\boldsymbol{y}$.

Signal to Artifact Ratio (SAR) is a method to measure the amount of artifact removed from the observed EEG signal after processing with the proposed methodologies [45]. SAR between the observed EEG signal with artifacts ($\boldsymbol{y}$) and the artifact-free signal ($\hat{\boldsymbol{c}}$) is given as

$$\text{SAR} = 10\log_{10}\frac{std(\boldsymbol{y})}{std(\boldsymbol{y}-\hat{\boldsymbol{c}})} \quad (9)$$

where std refers to standard deviation. Higher value of SAR index indicates superior performance of the algorithm.

Correlation Coefficient (CC) measures the degree of association between the degraded original EEG with EB artifacts ($\boldsymbol{y}$) and the artifact-free EEG ($\hat{\boldsymbol{c}}$). More positive the correlation value, stronger is the correlation [46]. Suppose $C(\boldsymbol{y},\hat{\boldsymbol{c}})$ is the covariance, then CC of $\boldsymbol{y}$ and $\hat{\boldsymbol{c}}$ is given as:

$$CC = \frac{C(\boldsymbol{y},\hat{\boldsymbol{c}})}{\sqrt{C(\boldsymbol{y},\boldsymbol{y}) * C(\hat{\boldsymbol{c}},\hat{\boldsymbol{c}})}}. \quad (10)$$

Mutual information (MI) measures how much relevant information the degraded original EEG with EB artifacts ($\boldsymbol{y}$) shares with the artifact-free EEG ($\hat{\boldsymbol{c}}$) [47]. Mathematically, it is defined as

$$MI = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(\boldsymbol{y},\hat{\boldsymbol{c}})\log\left(\frac{p(\boldsymbol{y},\hat{\boldsymbol{c}})}{p(\boldsymbol{y})p(\hat{\boldsymbol{c}})}\right) d\boldsymbol{y}d\hat{\boldsymbol{c}} \quad (11)$$

where $p(\boldsymbol{y},\hat{\boldsymbol{c}})$ is the joint probability density function and $p(\boldsymbol{y})$ and $p(\hat{\boldsymbol{c}})$ are the marginal probability density functions of $\boldsymbol{y}$ and $\hat{\boldsymbol{c}}$, respectively. The resemblance between the observed EEG signal and the artifact-free EEG signal is more if the magnitude of MI is large [10].

To estimate the similarity in frequency content between the raw EEG signal ($\boldsymbol{y}$) and the artifact-free EEG signal ($\hat{\boldsymbol{c}}$), magnitude squared coherence (MSC) [10] measure is computed. Mathematically, it is defined as follows:

$$MSC_{\boldsymbol{y},\hat{\boldsymbol{c}}}(f) = \frac{PSD_{\boldsymbol{y},\hat{\boldsymbol{c}}}(f)^2}{PSD_{\boldsymbol{y},\boldsymbol{y}}(f)\,PSD_{\hat{\boldsymbol{c}},\hat{\boldsymbol{c}}}(f)} \quad (12)$$

where $PSD_{\boldsymbol{y},\hat{\boldsymbol{c}}}$ denote the cross-spectral density between $\boldsymbol{y}$ and $\hat{\boldsymbol{c}}$, $PSD_{\boldsymbol{y},\boldsymbol{y}}$ and $PSD_{\hat{\boldsymbol{c}},\hat{\boldsymbol{c}}}$ denote the auto-spectral density of $\boldsymbol{y}$ and $\hat{\boldsymbol{c}}$, respectively.

IV. Experiments and Results

We used both synthetic and real EEG data in our experiments. The synthetic data was used to evaluate the performance of the proposed approaches whereas the real EEG dataset was used to validate them.

A. Experiments With Synthetic EEG Data

The design of synthetic EEG data consists of two parts: creation of EEG-like signal and EB-like artifact. Considering the noise-like nature of EEG signal we decided to choose $1/f$ or pink noise to simulate our EEG signal. The EB-like artifact signal is created using several combinations of sinusoids with different amplitudes and frequencies. The created EEG-like signal $\boldsymbol{c}$, EB-like signal $\boldsymbol{z}$ and their summation as noisy EEG signal $\boldsymbol{y}$ are shown in Fig. 2(a), (b) and (c), respectively. As in a real-world EEG signal, pink noise which resembles EEG also has lower amplitude than EB signal.

Initially, we use the MCA technique to obtain the morphological components from the synthetically generated corrupted EEG signal of Fig. 2(c). DCT and Dirac basis are selected as the two dictionaries in the MCA algorithm because the noise-like low amplitude EEG-like signal can be represented by DCT basis [32] and the spike-like components caused by EB-like signal can be sparsely represented by Dirac basis [44]. The MCA algorithm decomposes the generated noisy EEG signal into EB artifact and EEG signal as shown in Fig. 3(b) and (c), respectively. The extracted EB artifact in Fig. 3(b) is similar to the generated EB artifact in Fig. 2(b). The correlation coefficient between the extracted EB artifact and the generated EB artifact is 0.7832. As our aim is to extract artifact-free EEG, the quantitative metrics are measured between the generated pink noise in Fig. 2(a) and extracted EEG signal in Fig. 2(c). The quantitative

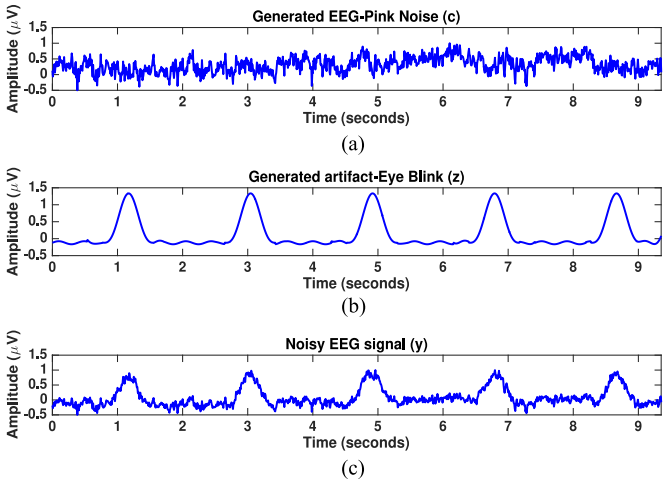

Fig. 2. (a) Simulated EEG-like signal, $\boldsymbol{c}$; (b) EB-like signal, z; (c) their summation, $\boldsymbol{y}$.

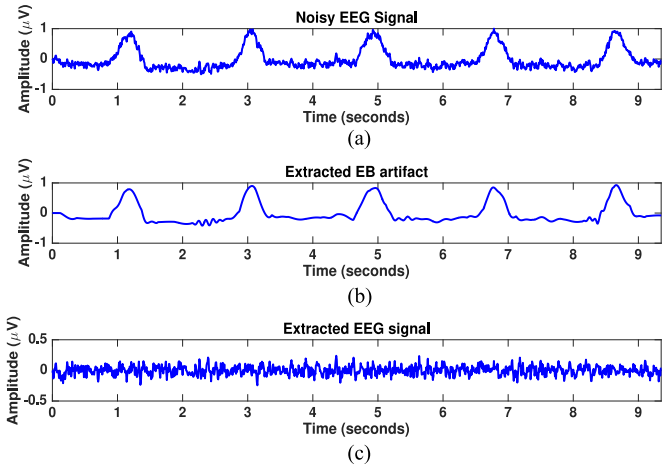

Fig. 3. (a) Generated noisy EEG signal; (b) Extracted EB-like signal; (c) Artifact-free EEG-like signal estimated using MCA method.

TABLE I
COMPARISON OF RMSE, CC, SAR AND MI VALUES OF K-SVD AND MCA METHODS COMPUTED USING SIMULATED EEG (PINK NOISE) LIKE DATA AND EXTRACTED EEG LIKE DATA

Metrics	K-SVD	MCA
Root Mean Square Error (RMSE)	4.6731	6.7164
Correlation Coefficient (CC)	0.7892	0.6831
Signal to Artifact Ratio (SAR)	3.4592	1.7722
Mutual Information (MI)	0.7241	0.6468

metrics such as RMSE, CC, SAR and MI values listed in Table I show good performance of MCA algorithm.

Next, the K-SVD based dictionary learning artifact removal method is applied over the noisy EEG signal $\boldsymbol{y}$ of Fig. 2(c). The signal $\boldsymbol{y}$ of length m is divided into n smaller segments. For example, if a random dictionary of size 100×150 is chosen, then the signal of length 1200 samples is divided into n segments where each segment includes $N(=100)$ samples which depends on the size of the dictionary. That is, if the dictionary size is

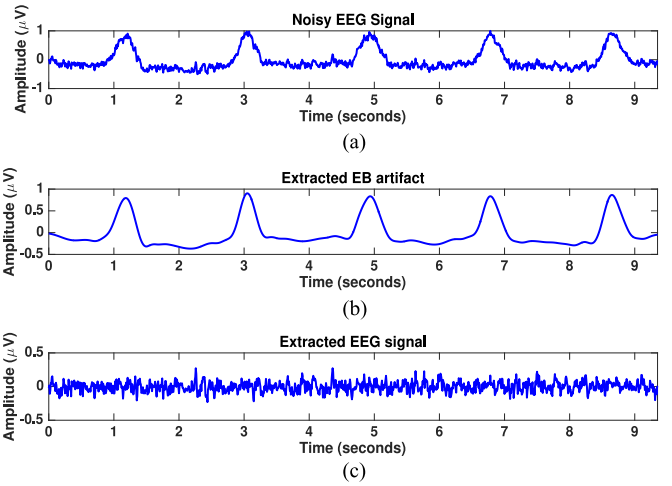

Fig. 4. (a) Generated noisy EEG signal; (b) Extracted EB-like signal; (c) Artifact-free EEG-like signal estimated uisng K-SVD method.

$p \times q$, then the number of samples N in each segment is equal to p. The number of segments n, depends on the degree of overlapping. Considering again the EEG signal, $\boldsymbol{y}$, of length $m = 1200$ samples, if there is no overlap between segments then the value of n is 12. For partial overlap with 50% coverage n is 23 and for fully overlapping segments n is 1101. In this work we consider fully overlapping segments and hence the K-SVD technique outlined in Algorithm 3 is applied to 100 samples of each segment to obtain the sparse coefficient vector of size 150×1 with 5 non-zero coefficients and the learned dictionary $D \in \mathbb{R}^{100\times150}$.

This process is repeated for each segment with full overlap until the entire signal is covered. Then the obtained sparse coefficient vectors and the learned dictionary are used in equation (7) to separate EB-like signal and EEG-like signal which are shown in Fig. 4(b) and (c), respectively. The extracted EB artifact using K-SVD technique is very similar to the generated EB artifact in Fig. 2(b) and the correlation coefficient between them is 0.8214. The quantitative metrics measured between the generated EEG signal in Fig. 2(a) and the extracted EEG like signal in Fig. 4(c) are listed in Table I. From the values in the Table I it is observed that K-SVD based artifact removal method outperforms MCA method since it produces smaller RMSE value and higher CC, SAR and MI values. The magnitude squared coherence (MSC) between the generated noisy EEG signal $\boldsymbol{y}$ and the extracted noise-free EEG signal by MCA and K-SVD method are shown in Fig. 5(a) and (b), respectively. It clearly shows that both the sparsity-based methods have strongest coherence above 10 Hz and therefore, preserve the neural signal frequency after removing the EB-like artifacts.

B. Experiments With Real EEG Data

1) Experimental setup: Seven healthy subjects (four males and three female), between ages of 23 and 30, volunteered for the experiment. Emotiv EPOC+ [48] device with sampling rate of 128 samples per second, with low pass filter cut-off frequency 45 Hz and resolution of 14 bits was used to record data. Users

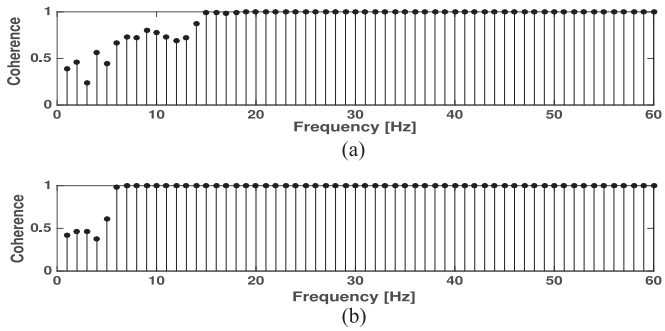

Fig. 5. MSC of the generated noisy EEG signal $\boldsymbol{y}$ and the extracted artifact-free EEG-like signal using (a) MCA method and (b) K-SVD method.

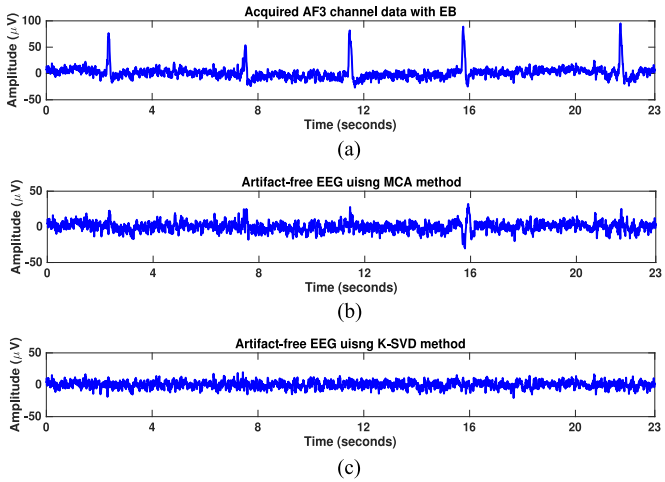

Fig. 6. (a) Acquired AF3 channel data with EB for subject S1_1; (b) Artifact-free EEG estimated using MCA method; (c) Artifact-free EEG estimated using K-SVD method.

were seated in a chair with their arms and legs extended, resting on a desk and resting on a footrest, respectively. We made sure that there was no background noise while the data was recorded. Initially, the user is asked to close his/her eyes for 10 seconds. Then they were asked to open and sit relaxed for one minute and again instructed to close their eyes for 10 seconds. We could see that most of the users blinked their eyes once in 5 seconds or at least eight times in a minute. So, EEG data of total one minute 20 second duration is collected in each trial for each user. The experiment is repeated two times with all users. Hence, a total of 14 datasets were collected and used. Each session lasted about 10–15 minutes, including the time for experimental setup and data recording. The data was captured on all 14 channels which were labelled as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 and channels P3, P4 were used as reference channels. The recorded EEG signals are transferred to the computer using wireless USB connector.

2) Experiments With Single Channel EEG Data: Case 1 (Subject S1 - dataset 1, S1_1): Out of 14 channels acquired using Emotiv EPOC+ device, the data of AF3 channel is severely affected by EB artifacts as it is located over the dorsolateral prefrontal cortex [10]. Fig. 6(a) shows the acquired AF3 EEG signal with EBs for the subject S1_1. Before using the proposed artifact removal techniques, we normalized the data by subtracting the mean value of the signal from each data point. Then the proposed MCA and K-SVD based EB artifact suppression techniques were applied to AF3 channel data. MCA method decomposes the acquired EEG signal in Fig. 6(a) into EB component and artifact-free EEG component. The artifact-free EEG obtained from MCA method is shown in Fig. 6(b). The result shows that MCA method removes EB artifacts from the AF3 channel data to a significant degree. However, one can observe that some EB artifatcs still exist in the estimated artifact-free EEG signal.

TABLE II
COMPARISON OF RMSE, CC, SAR AND MI VALUES OF K-SVD AND MCA METHODS FOR SINGLE CHANNEL DATA FOR THE SUBJECT S1_1

Metrics	K-SVD	MCA
Root Mean Square Error (RMSE)	8.0143	9.3249
Correlation Coefficient (CC)	0.6687	0.5131
Signal to Artifact Ratio (SAR)	1.9946	1.2561
Mutual Information (MI)	0.6984	0.6244

Similar to experiments on synthetic data, the normalised signal of length m samples is divided into n segments. Each segment of $N(=100)$ samples is processed using Algorithm 3 to obtain the sparse coefficients and the learned dictionary. As explained previously in Section IV-A, the choice of value of N depends on size of the dictionary. Finally, we use equation (7) to separate EB contribution from the AF3 channel data. The artifact-free EEG obtained from K-SVD technique is shown in Fig. 6(c). The result shows that K-SVD method successfully removes EB artifacts from single channel EEG data.

By comparing the results of applying the proposed methods on the EEG signal corrupted with EB artifacts, we observe that the K-SVD method removes all the EB artifacts and yields an artifact-free EEG signal. Hence, the learned dictionary of K-SVD algorithm delivers superior performance for removing artifacts than using pre-defined DCT and Dirac dictionaries in the MCA technique. We reach the same conclusion by evaluating quantitative metrics such as RMSE, CC, SAR and MI for both the sparsity-based approaches. The metrics are calculated between the acquired raw AF3 channel signal and the artifact-free signal obtained by applying the proposed sparsity-based approaches. The values obtained are listed in Table II. We observe from Table II that the proposed K-SVD based artifact removal method outperforms MCA method since it produces smaller RMSE value and higher CC, SAR and MI values. The MSC plots obtained for the acquired AF3 channel data with EBs of subject S1_1, and the estimated artifact-free EEG signal using MCA and K-SVD method are shown in Fig. 7(a) and (b), respectively. The coherence values show that the sparsity-based artifact removal algorithms preserve the neural signal frequency after removing the EB artifacts.

Case 2 (Subject S3 - dataset 2, S3_2): Out of seven subjects, subject S3 involuntarily blinked two times at certain instances during experimental trials. We found that this was a mannerism of the particular subject. We observed that our proposed sparsity-based methods remove such blinks accurately from the EEG signals. The acquired AF3 channel data for the subject S3_2 with double EBs is shown in Fig. 8(a) and the observed

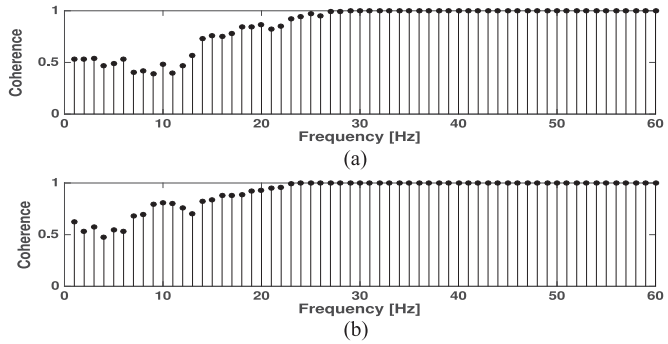

Fig. 7. MSC of the acquired AF3 channel data for the subject S1_1 and the extracted artifact-free EEG signal using (a) MCA method and (b) K-SVD method.

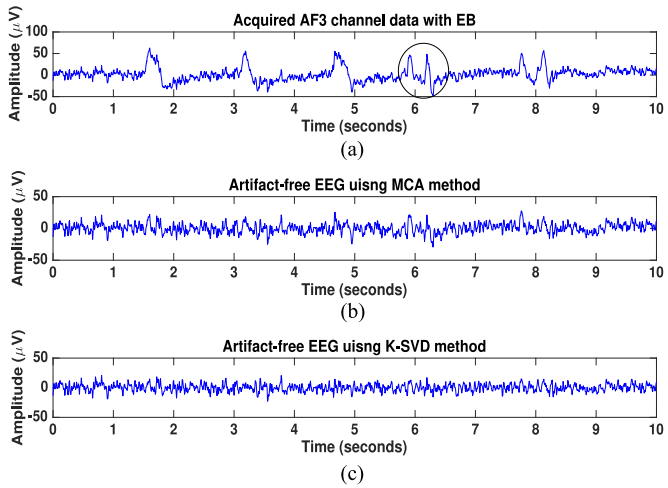

Fig. 8. (a) Acquired AF3 channel data with double EBs for the subject S3_2; (b) Artifact-free EEG estimated using MCA method; (c) Artifact-free EEG estimated using K-SVD method.

TABLE III
COMPARISON OF RMSE, CC, SAR AND MI VALUES OF K-SVD AND MCA METHODS FOR SINGLE CHANNEL DATA FOR THE SUBJECT S3_2

Metrics	K-SVD	MCA
Root Mean Square Error (RMSE)	8.3279	10.5682
Correlation Coefficient (CC)	0.6598	0.5001
Signal to Artifact Ratio (SAR)	1.9346	1.2282
Mutual Information (MI)	0.6874	0.6053

double EB artifacts are shown encircled with an ellipse. Artifact-free EEG signals obtained using MCA and K-SVD methods for the subject S3_2 are shown in Fig. 8(b) and (c), respectively. The quantitative metrics obtained are listed in Table III and it is observed that K-SVD based artifact removal method outperforms MCA method in removing EB artifacts. The MSC obtained for the acquired AF3 channel data with double EBs for the subject S3_2 and the extracted artifact-free EEG signal using MCA and K-SVD method is shown in Fig. 9(a) and (b), respectively.

3) Experiment with 14-channel EEG Data (Subject S4 - dataset 1, S4_1): Next, we use both MCA and K-SVD based artifact removal methods on all the 14 channel EEG data acquired

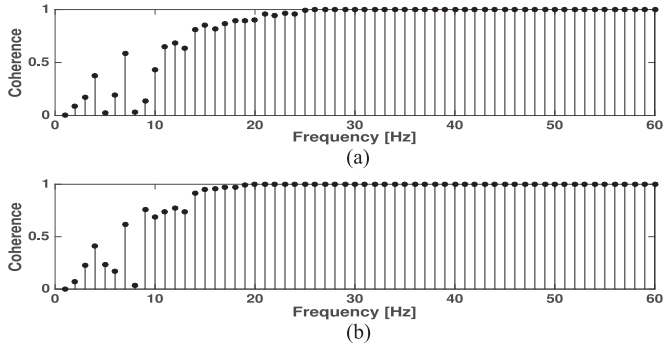

Fig. 9. MSC of the acquired AF3 channel data with double EBs for the subject S3_2 and the extracted artifact-free EEG signal using (a) MCA method and (b) K-SVD method.

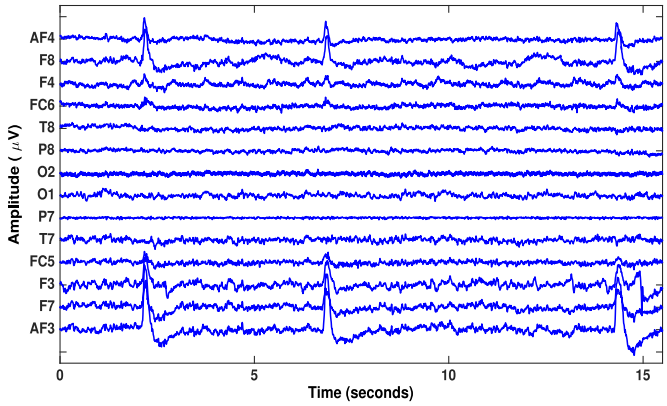

Fig. 10. Observed 14 channel data with EB artifatcs for the subject S4_1.

using Emotiv Epoc+ device. Fig. 10 shows the observed EEG signals of all the 14 channels with EB artifacts for the subject S4_1. First, MCA technique is applied to all the 14 channels by keeping the two pre-defined DCT and Dirac dictionaries unchanged throughout the process. To attain this task, the extracted signals from all the channels are divided into segments of size N samples, and each segment is then used as input to Algorithm 2. The MCA algorithm decomposes the acquired EEG signal into EB artifact and artifact-free EEG signal. We follow a similar procedure for processing 14-channel EEG data using K-SVD method. The only difference here is that instead of pre-defined dictionaries, a single over-complete dictionary is learned. Algorithm 3 and equation (7) are applied to each segment of raw EEG signal to learn a single over-complete dictionary. Fig. 11 shows the artifact-free EEG obtained by applying the proposed K-SVD based technique on the acquired EEG data of all the 14 channels for the subject S4_1 shown in Fig. 10.

The performance of the proposed sparsity-based techniques on all the fourteen datasets is evaluated using metrics such as RMSE, CC, SAR and MI. Each dataset consists of 14 channel EEG data. The performance metrics are applied separately to each channel data of each dataset. The averaged RMSE, CC, SAR and MI values of each channel for all the fourteen datasets is enumerated in Table IV. The total average value of RMSE (7.6801, 8.5167), CC (0.6738, 0.5562), SAR (2.0282, 1.3063) and MI (1.0102, 0.8795) demonstrate that K-SVD based artifact

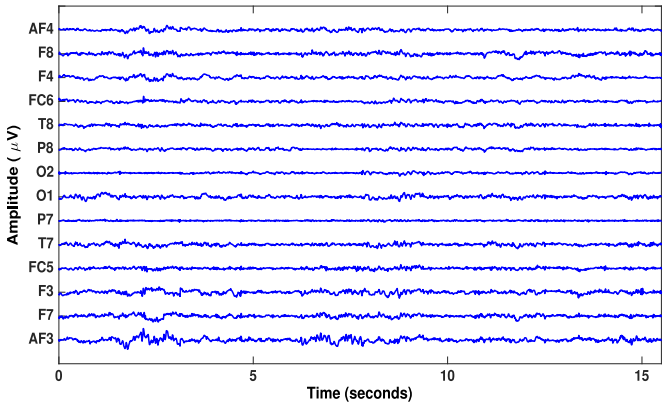

Fig. 11. Artifact-free EEG signals using K-SVD method for the subject S4_1.

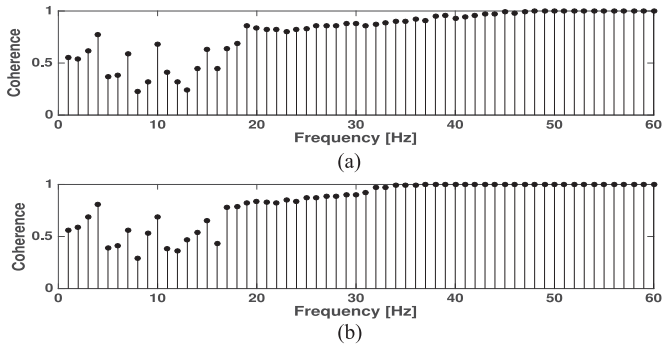

Fig. 12. Average MSC of all the 14 channel data for the subject S4_1 and the extracted artifact-free EEG signal using (a) MCA method and (b) K-SVD method.

removal method outperforms MCA method. The average MSC obtained for all the 14 channels for one of the subjects (S4_1) is shown in Fig. 12(a) and (b), respectively.

V. Comparison With Existing Work

To evaluate the performance of the proposed K-SVD-based artifact removal method, we provide a comparison of statistical metrics with a very recently proposed artifact removal method namely, Fully Online and Automated Artifact Removal for Brain-Computer Interfacing (FORCe) method [15]. FORCe technique is based upon a combination of wavelet decomposition, ICA, hard and soft thresholding. FORCe is mainly intended for the removal of participant generated artifacts and it has been shown to outperform state-of-the-art artifact removal methods such as Lagged Auto-manual Information Clustering (LAMIC) [49] and Fully Automated Statistical Thresholding for EEG artifact Rejection (FASTER) [50].

The average RMSE, CC, SAR and MI values of our proposed K-SVD method and the existing FORCe method for all the fourteen datasets are enumerated in Table V and the average values of each channel for all the fourteen dataset is shown in Fig. 13. From Table V it is shown that the RMSE (7.6801, 7.6702), CC (0.6738, 0.6853), SAR (2.0282, 2.0565) and MI (1.0102, 1.0154) values of K-SVD algorithm and FORCe method, respectively are nearly equivalent. It is therefore demonstrated that the proposed K-SVD based artifact removal technique performs equivalent to FORCe method.

TABLE IV
Comparison of Average RMSE, CC, SAR and MI Values of K-SVD and MCA Based EB Artifact Removal Methods for all the 14 Channel EEG Data of all the Fourteen Datasets

Metrics	Channel location	K-SVD	MCA
Average RMSE	AF3	12.0352	13.4814
	F7	9.5141	10.6113
	F3	7.4345	8.1279
	FC5	11.8764	12.4589
	T7	5.5465	6.0214
	P7	5.2163	5.5242
	O1	5.3690	5.6348
	O2	5.8525	6.0259
	P8	6.3476	6.9884
	T8	7.3451	8.2486
	FC6	4.3430	5.5473
	F4	8.4465	8.9759
	F8	6.5932	8.6424
	AF4	11.6016	12.9456
Total average RMSE		**7.6801**	**8.5167**
Average CC	AF3	0.5048	0.4124
	F7	0.6812	0.5182
	F3	0.6488	0.5876
	FC5	0.7028	0.4849
	T7	0.7998	0.6731
	P7	0.7368	0.6444
	O1	0.7285	0.6744
	O2	0.7960	0.6417
	P8	0.6679	0.5037
	T8	0.6843	0.6284
	FC6	0.7025	0.5638
	F4	0.6215	0.4515
	F8	0.6299	0.5184
	AF4	0.5191	0.4837
Total average CC		**0.6738**	**0.5562**
Average SAR	AF3	1.4129	0.7378
	F7	1.6427	0.8721
	F3	2.1261	0.9666
	FC5	1.3122	0.5107
	T7	3.1374	1.9571
	P7	2.6438	1.5914
	O1	1.5953	1.5356
	O2	2.9674	1.8752
	P8	1.3994	1.1735
	T8	3.5603	2.8642
	FC6	3.2868	2.3035
	F4	1.2353	0.5461
	F8	1.4598	0.9890
	AF4	0.6150	0.3657
Total average SAR		**2.0282**	**1.3063**
Average MI	AF3	0.6963	0.6585
	F7	0.7939	0.7489
	F3	1.1629	1.0580
	FC5	1.1696	1.0468
	T7	1.3416	0.9841
	P7	1.1728	1.0603
	O1	1.0723	0.9419
	O2	1.2517	1.0423
	P8	1.1590	0.8482
	T8	0.9141	0.8705
	FC6	1.0438	0.9619
	F4	1.0576	0.8453
	F8	0.6588	0.6157
	AF4	0.6476	0.6298
Total average MI		**1.0102**	**0.8795**

TABLE V
COMPARISON OF AVERAGE RMSE, CC, SAR, AND MI VALUES OF THE PROPOSED K-SVD METHOD WITH FORCE METHOD FOR ALL THE FOURTEEN DATASET

Metrics	K-SVD	FORCe
Root Mean Square Error (RMSE)	7.6801	7.6702
Correlation Coefficient (CC)	0.6738	0.6853
Signal to Artifact Ratio (SAR)	2.0282	2.0565
Mutual Information (CC)	1.0102	1.0154

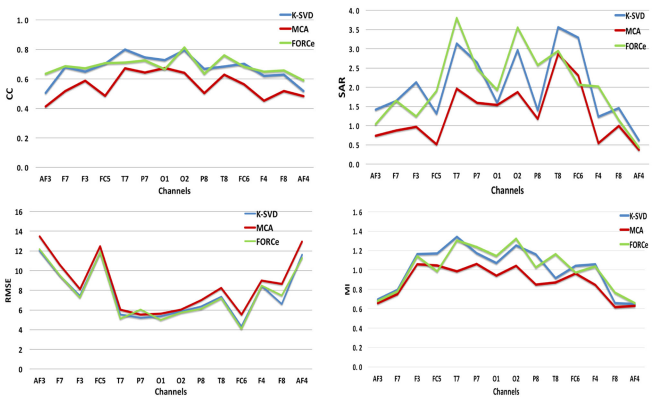

Fig. 13. Average CC, SAR, RMSE and MI values obtained by K-SVD, MCA and FORCe methods for all the 14 channels of all the fourteen datasets.

VI. DISCUSSION

A. Phase Delay

In our proposed sparsity-based methods, we are processing the EEG signals in time-domain without performing any transformations such as Fourier, wavelet, etc. Also, we are not using any filters for processing the signals. The acquired EEG signals are segmented and directly passed through the proposed MCA and K-SVD algorithms to estimate artifact-free EEG signals. Hence, no phase delay involved in the proposed methods.

To validate our argument, we estimated time-delay between the degraded EEG signal and the artifact-free EEG signal. The generalized cross-correlation (GCC) [51] between the acquired raw EEG signal $x_1(t)$ and the artifact-free EEG signal $x_2(t)$ is computed as

$$GCC_{x_1x_2}(\tau) = \int_{-\infty}^{\infty} \psi_g(f) X_1(f) X_2^*(f) e^{j2\pi f\tau} df \qquad (13)$$

where $X_1(f)$ and $X_2(f)$ are the Fourier transform of the signals $x_1(t)$ and $x_2(t)$, respectively, and $\psi_g(f)$ is the weighting function and it is chosen as $\psi_g(f) = |X_1(f)X_2^*(f)|^{-1}$. Fig. 14 shows the GCC of the degraded EEG signal and the estimated artifact-free EEG signal (300 samples) and we observe no time-delay between the signals.

B. Computation Time

The speed of the algorithm is an important issue in real-time BCI applications. In this work, we have compared the execution times of each algorithm. The proposed sparsity-based algorithm and FORCe algorithm have been executed on the same PC with

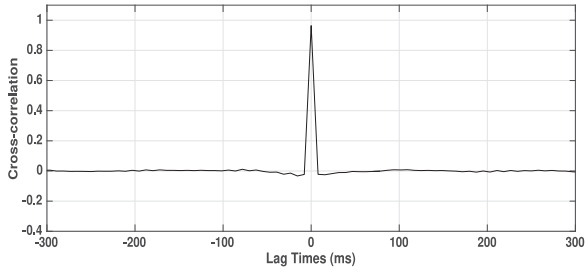

Fig. 14. Generalized cross-correlation (GCC) of the acquired raw EEG signal and the artifact-free EEG signal using K-SVD method for the subject S1_1. The time-delay is 0 samples between the two EEG signals.

3.2 GHz CPU, 4 GB RAM and software (MATLAB R2015b). MCA took an average execution time of 24.3 s, K-SVD requires 31.02 s and FORCe consumes 27.48 s for executing all the 14 channels. MCA method provides lesser execution time since it uses pre-defined dictionaries. In recent literature [52], [53], the original K-SVD algorithm has been modified to perform parallel updation of atoms with less computation time. For the sake of simplicity and in order to provide a benchmark comparison, we used the original K-SVD algorithm with sequential updation of atoms.

VII. CONCLUSION

In this paper, we proposed two sparsity-based approaches, namely, MCA and K-SVD to remove EB artifacts from the raw EEG signal. The advantage of the proposed K-SVD based method is that the learned dictionary adaptively estimates the EB artifact from the given EEG signal. On the other hand, if the pre-defined basis functions are not chosen appropriately in MCA method, the morphological components of the given EEG signal may not be modeled accurately. The experiments on synthetic and real EEG data demonstrate that learning the dictionary from the given data using K-SVD algorithm yields superior performance as compared to using pre-defined dictionaries in the MCA algorithm. The proposed K-SVD algorithm is also compared with the recent state-of-the-art EEG artifact removal method FORCe [15]. Experimental results show that our K-SVD based artifact removal algorithm performs equivalent to the FORCe method. The proposed sparsity-based methods do not require any channel information, parameter tuning such as thresholding or additional equipments/EOG channels to remove EB artifacts from the EEG signal. Although, the algorithms proposed in this paper are intended only for EB artifact correction, they can also be applied to other artifacts in raw EEG data.

REFERENCES

[1] J. Wolpaw and E. W. Wolpaw, *Brain-Computer Interfaces: Principles and Practice*. New York, NY, USA: Oxford Univ. Press, 2012.
[2] G. R. Müller-Putz, R. Scherer, G. Pfurtscheller, and R. Rupp, "EEG-based neuroprosthesis control: A step towards clinical practice," *Neurosci. Lett.*, vol. 382, no. 1, pp. 169–174, 2005.
[3] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw, "A brain-computer interface for long-term independent home use," *Amyotrophic Lateral Sclerosis*, vol. 11, no. 5, pp. 449–455, 2010.
[4] A. Kübler *et al.*, "Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface," *Neurology*, vol. 64, no. 10, pp. 1775–1777, 2005.

[5] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767–791, 2002.

[6] J. Long, Y. Li, H. Wang, T. Yu, J. Pan, and F. Li, "A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 20, no. 5, pp. 720–729, Sep. 2012.

[7] V. Gandhi, G. Prasad, D. Coyle, L. Behera, and T. M. McGinnity, "EEG-based mobile robot control through an adaptive brain-robot interface," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 44, no. 9, pp. 1278–1285, Sep. 2014.

[8] W. O. Tatum, B. A. Dworetzky, and D. L. Schomer, "Artifact and recording concepts in EEG," *J. Clin. Neurophysiol.*, vol. 28, no. 3, pp. 252–263, 2011.

[9] D. Hagemann and E. Naumann, "The effects of ocular artifacts on (lateralized) broadband power in the EEG," *Clin. Neurophysiol.*, vol. 112, no. 2, pp. 215–231, 2001.

[10] R. Mahajan and B. I. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 158–165, Jan. 2015.

[11] K. Nazarpour *et al.*, "Removal of the eye-blink artifacts from EEGs via STF-TS modeling and robust minimum variance beamforming," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 9, pp. 2221–2231, Sep. 2008.

[12] M. T. Akhtar, W. Mitsuhashi, and C. J. James, "Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data," *Signal Process.*, vol. 92, no. 2, pp. 401–416, 2012.

[13] P. Berg and M. Scherg, "Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG," *Clin. Phys. Physiol. Meas.*, vol. 12, no. A, pp. 49–54, 1991.

[14] T.-P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 02, pp. 163–178, 2000.

[15] I. Daly, R. Scherer, M. Billinger, and G. Muller-Putz, "FORCe: Fully online and automated artifact removal for brain-computer interfacing," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 5, pp. 725–736, Sep. 2015.

[16] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[17] O. Bryt and M. Elad, "Compression of facial images using the K-SVD algorithm," *J. Vis. Commun. Image Representation*, vol. 19, no. 4, pp. 270–282, 2008.

[18] A. I. Koutrouvelis, A. Härmä, and A. Mouchtaris, "Compressive sensing in footstep sounds, hand tremors and speech using K-SVD dictionaries," in *Proc. 18th IEEE Int. Conf. Digit. Signal Process.*, 2013, pp. 1–6.

[19] J. F. Gemmeke, T. Virtanen, and A. Hurmalainen, "Exemplar-based sparse representations for noise robust automatic speech recognition," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 7, pp. 2067–2080, Sep. 2011.

[20] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[21] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[22] C. Burger and D. J. van den Heever, "Removal of EOG artefacts by combining wavelet neural network and independent component analysis," *Biomed. Signal Process. Control*, vol. 15, pp. 67–79, 2015.

[23] S. A. Hillyard and R. Galambos, "Eye movement artifact in the CNV," *Electroencephalography Clin. Neurophysiol.*, vol. 28, no. 2, pp. 173–182, 1970.

[24] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, "Automatic identification and removal of ocular artifacts from EEG using wavelet transform," *Meas. Sci. Rev.*, vol. 6, no. 4, pp. 45–57, 2006.

[25] P. He, G. Wilson, C. Russell, and M. Gerschutz, "Removal of ocular artifacts from the EEG: A comparison between time-domain regression method and adaptive filtering method using simulated data," *Med. Biol. Eng. Comput.*, vol. 45, no. 5, pp. 495–503, 2007.

[26] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *Neuroimage*, vol. 34, no. 4, pp. 1443–1449, 2007.

[27] M. Kirkove, C. François, and J. Verly, "Comparative evaluation of existing and new methods for correcting ocular artifacts in electroencephalographic recordings," *Signal Process.*, vol. 98, pp. 102–120, 2014.

[28] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas, "Automatic removal of eye movement and blink artifacts from EEG data using blind component separation," *Psychophysiology*, vol. 41, no. 2, pp. 313–325, 2004.

[29] M. H. Soomro, N. Badruddin, M. Z. Yusoff, and A. S. Malik, "A method for automatic removal of eye blink artifacts from EEG based on EMD-ICA," in *Proc. IEEE 9th Int. Colloq. Signal Process. Appl.*, 2013, pp. 129–134.

[30] T. D. Lagerlund, F. W. Sharbrough, and N. E. Busacker, "Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition," *J. Clin. Neurophysiol.*, vol. 14, no. 1, pp. 73–82, 1997.

[31] J. W. Matiko, S. Beeby, and J. Tudor, "Real time eye blink noise removal from EEG signals using morphological component analysis," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 13–16.

[32] X. Yong, R. K. Ward, and G. E. Birch, "Artifact removal in EEG using morphological component analysis," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 345–348.

[33] V. Abolghasemi and S. Ferdowsi, "EEG–fMRI: Dictionary learning for removal of ballistocardiogram artifact from EEG," *Biomed. Signal Process. Control*, vol. 18, pp. 186–194, 2015.

[34] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmonic Anal.*, vol. 19, no. 3, pp. 340–358, 2005.

[35] J. M. Fadili, J.-L. Starck, M. Elad, and D. Donoho, "MCAlab: Reproducible research in signal and image decomposition and inpainting," *IEEE Comput. Sci. Eng.*, vol. 12, no. 1, pp. 44–63, Jan./Feb. 2010.

[36] E. Candes, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, 2006.

[37] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. New York, NY, USA: Academic, 2008.

[38] E. J. Candès and D. L. Donoho, "Ridgelets: A key to higher-dimensional intermittency?" *Philosophical Trans. Roy. Soc. Lond. A, Math., Phys. Eng. Sci.*, vol. 357, no. 1760, pp. 2495–2509, 1999.

[39] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[40] Y. Li, Z. L. Yu, N. Bi, Y. Xu, Z. Gu, and S.-i. Amari, "Sparse representation for brain signal processing: A tutorial on methods and applications," *IEEE Signal Process. Mag.*, vol. 31, no. 3, pp. 96–106, May 2014.

[41] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[42] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.

[43] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, Jul. 2011.

[44] X. Yong, R. K. Ward, and G. E. Birch, "Generalized morphological component analysis for EEG source separation and artifact removal," in *Proc. 4th Int. IEEE/EMBS Conf. Neural Eng.*, 2009, pp. 343–346.

[45] S. Khatun, R. Mahajan, and B. I. Morshed, "Comparative study of wavelet-based unsupervised ocular artifact removal techniques for single-channel EEG data," *IEEE J. Trans. Eng. Health Med.*, vol. 4, pp. 1–8, 2016.

[46] M. Zima, P. Tichavskỳ, K. Paul, and V. Krajča, "Robust removal of short-duration artifacts in long neonatal EEG recordings using wavelet-enhanced ICA and adaptive combining of tentative reconstructions," *Physiol. Meas.*, vol. 33, no. 8, pp. N39–N49, 2012.

[47] J. W. Kelly, D. P. Siewiorek, A. Smailagic, J. L. Collinger, D. J. Weber, and W. Wang, "Fully automated reduction of ocular artifacts in high-dimensional neural data," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 598–606, Mar. 2011.

[48] Emotiv-EPOC, "Software development kit," 2010. [Online]. Available: http://www.emotiv.com/researchers

[49] I. Daly, N. Nicolaou, S. J. Nasuto, and K. Warwick, "Automated artifact removal from the electroencephalogram: A comparative study," *Clin. EEG Neurosci.*, vol. 44, no. 4, pp. 291–306, 2013.

[50] H. Nolan, R. Whelan, and R. Reilly, "Faster: Fully automated statistical thresholding for EEG artifact rejection," *J. Neurosci. Methods*, vol. 192, no. 1, pp. 152–162, Sep. 2010.

[51] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, Aug. 1976.

[52] M. Sadeghi, M. Babaie-Zadeh, and C. Jutten, "Learning overcomplete dictionaries based on atom-by-atom updating," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 883–891, Feb. 2014.

[53] L. He, T. Miskell, R. Liu, H. Yu, H. Xu, and Y. Luo, "Scalable 2D K-SVD parallel algorithm for dictionary learning on GPUs," in *Proc. ACM Int. Conf. Comput. Frontiers*, 2016, pp. 11–18.