Leveraging Sparsity: A Low-Rank + Sparse Decomposition (LR+SD) Method for Automatic EEG Artifact Removal

J. Gilles[1], T.Meyer[2], and P.K. Douglas[2,3]

[1] Department of Mathematics, University of California, Los Angeles,CA 90095, USA
[2] Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, California, USA 90024
[3] Wellcome Trust Centre for Neuroimaging at UCL, London WC1N 3BG

Abstract. In theory, multimodal EEG-fMRI recordings represent an excellent tool for studying bioelectric-hemodynamic coupling in the human brain without incurring added complexity due to nonstationarity. However, ballistocardiogram (BCG) artifacts as opposed to magnetic gradient noise have made analysis of EEG data collected in the MRI environment very challenging. Conventionally, BCG artifacts have been removed only partially after meticulous user-guided identification of independent components associated with noise. In this paper, we present a novel method for automatically removing BCG artifact from event related EEG data by leveraging sparsity in the time domain. Our method, low rank + sparse decomposition (LR+SD) extends robust PCA and requires tuning of only a single regularization parameter. We apply this method first to simulated data,and then to real simultaneous EEG-fMRI data, collected while subjects viewed photic stimuli. We found that LR+SD improved the signal-to-noise ratio by 34 and 36 percent, as compared to either manual or automatic IC methods respectively. This method appears quantitatively superior to IC methods, and may improve the feasibility of analyzing event related EEG-fMRI data collected concurrently.

Keywords: Concurrent EEG-fMRI, balistocardiogram, artifact removal, robust, PCA, low rank, sparsity, sparse decomposition

1 Introduction

Independently, electroencephalogram (EEG) and functional MRI (fMRI) offer either rich temporal (EEG) or spatial (fMRI) information related to neuronal dynamics in the brain. Ideally, these imaging techniques could be combined in a complementary fashion to harness their respective strengths [3], [16], and potentially improve our ability to localize epileptiform generators [15]. However, analysis of EEG data collected in the MRI environment has proven quite challenging, given a number of artifacts introduced during concurrent recordings.

EEG recordings putatively reflect the superposition of electric dipoles associated with synchronous activity from neural populations measured at the

scalp [2]. When collected inside the MR scanning environment, these signals are corrupted by noise due to the switching of magnetic fields, which creates a prominent gradient artifact. This gradient signal intiially appeared problematic, however, a number of template-subtraction methods have been developed that can effectively remove this large signal [7].

More troublesome to this analysis is the quasi-periodic signal known as the ballistocardiogram (BCG) artifact, which cannot be easily removed with template based methods. The BCG is generated as EEG electrodes move due to pulsatile motion during the cardiac cycle, and presents broadly in the spectral frequencies often analyzed in EEG (0.5-25 Hz) [4]. The presence of these artifacts can dramatically change the spectral properties of the signal, and obscure ability to perform trial-by-trial analyses.

Thus far, a variety of techniques have been tested to remove BCG artifact from these data including template based average artifact subtraction based on cardiac r-wave timing [1], filtering [11], independent component analysis (ICA) [9], optimal basis sets (OBS) [13], clustering [20], and combined methods [4], many of which can be applied using manual or automated algorithms. Although each of these methods have acheived some degree of efficacy, most of these either require collection of additional data (e.g. ECG data) for template characterization purposes, require tedious manual artifact component identification, or require cardiac signal identification within the EEG itself, which is often only intermittently identifiable throughout a recording session.

Here, we introduce a new algorithm for removing artifact from EEG signals that uses low rank + sparse decomposition (LR+SD). To do so, we propose a mathematical model based on a reasonable experimental assumption that artifact components will be mathematically expressed differently than the data themselves. Importantly, this method obviates the need for any reference or template artifact signal. As such, the combined effects of many types of artifacts can be removed in a single decomposition without the need for manual identification of artifact components in the data. We then assess the utility of this new algorithm on simulalted and real data.

2 Methods

2.1 Low Rank + Sparse Decomposition Method

We denote $\{\tilde{f}_i(k)\}_{i=1}^N$ the set of recorded EEG signals. The index i corresponds to the channel index, assuming we have a total of N electrodes distributed over the scalp. Moreover, we assume that each signal is recorded over K samples, i.e. $k \in \{1, \ldots, K\}$. We consider J (unknown) artifacts, and will denote them $\{f_j^A(k)\}_{j=1}^J$ where the index j identifies different artifacts. The goal of the artifact removal procedure is then to retrieve cleaned EEG signals, $\{f_i(k)\}_{i=1}^N$.
In the proposed method we assume the following model: each recorded EEG channel is a linear combination of its cleaned version and the different artifacts.

This model is equivalent to write:

$$\tilde{f}_i(k) = f_i(k) + \sum_{j=1}^{J} a_{ij} f_j^A(k), \tag{1}$$

where the mixing coefficients a_{ij} are unknown. In the following, we use a matrix formalism to model the global processes. To do so, we cast each EEG and artifact channels as columns of $K \times N$ matrices:

$$\underbrace{\begin{pmatrix} | & | & & | \\ \tilde{f}_1 & \tilde{f}_2 & \dots & \tilde{f}_N \\ | & | & & | \end{pmatrix}}_{\tilde{F}} = \underbrace{\begin{pmatrix} | & | & & | \\ f_1 & f_2 & \dots & f_N \\ | & | & & | \end{pmatrix}}_{F} + \underbrace{\begin{pmatrix} | & | & & | \\ \tilde{f}_1^A & \tilde{f}_2^A & \dots & \tilde{f}_N^A \\ | & | & & | \end{pmatrix}}_{\tilde{F}^A}, \tag{2}$$

where $\tilde{f}_i^A(k) = \sum_{j=1}^{J} a_{ij} f_j^A(k)$. The matrix $\tilde{F}$ contains all recorded EEG channels, F the wanted cleaned EEG signals and $\tilde{F}^A$ contains the mixing of all artifacts. The latter can be written as $\tilde{F}^A = \sum_{j=1}^{J} F_j^A$ where

$$F_j^A = \begin{pmatrix} | & | & & | \\ a_{1j}\tilde{f}_j^A & a_{2j}\tilde{f}_j^A & \dots & a_{Nj}\tilde{f}_j^A \\ | & | & & | \end{pmatrix}. \tag{3}$$

The key to our method is to notice that each matrix F_j^A has its columns proportional to the same vector $\tilde{f}_j^A$ implying that $rank(F_j^A) = 1$ and consequently $rank(\tilde{F}^A) \leq J$. Otherwise, the matrix F should contain events resulting from true EEG data. In our case, we focus on event related spectral perturbations (ERSPs), which occur at specific times and affect a limited number of electrodes. Therefore, it is reasonable to assume that F is a sparse matrix. Thus the artifact removal problem is equivalent to performing a low-rank + sparse decomposition of $\tilde{F}$. The resulting sparse component therefore corresponds to the cleaned EEG signals. Such decomposition can be done by solving the following minimization problem:

$$\begin{aligned} (F, F^A) = \arg\min \|F^A\|_* + \lambda \|F\|_1 \\ \text{such that } \tilde{F} = F + F^A, \end{aligned} \tag{4}$$

where $\|.\|_*$ denotes the nuclear norm of a matrix (i.e. the sum of its singular values), $\|.\|_1$ denotes the sum of the absolute value of the matrix entries and λ is a positive parameter allowing us to control the rank of F^A. Such model, also called Robust PCA, was actively studied in the mathematics community [10]. In

our experiments we extend the Lin et al. (2009) algorithm for artifact removal, and select the regularization parameter that maximizes signal to noise ration by running a sweep across all possible ranks. Upon publication, all code developed for this project will be available on the NITRC repository.

2.2 Simulated Dataset

In general, there is no ground truth EEG signal when data are empirical, making it difficult to assess the utility of artifact-removal algorithms. We therefore created a simulated dataset using the free BESA (Brain Electrical Source Analysis) to generate simulated EEG signals generated by three distributed dipole sources corrupted by known artifacts using a spherical head forward model. In order to add realistic noise to the data, we used ECG, EMG, and right and left EOG reference artifact recordings extracted from the free sample of the SHHS Polysomnography Database. These reference artifacts were normalized and added to the pure simulated EEGs using randomized mixing coefficients accordingly to a uniform distribution.

2.3 Empirical Data: Concurrent EEG-fMRI

Twenty healthy individuals (ages 23-30, 12 male) provided written informed consent to participate in this study, which was approved by the UCLA IRB. Concurrent recordings took place while subjects passively viewed 140 Gabor flashes, presented via an MR projector screen with a varied inter stimulus interval of 13.85 +/ -2.8 sec, a task known to generate reproducible occipital ERSPs in the alpha (8-12 Hz) spectral band [8]. EEG were recorded using a 256-channel GES 300 Geodesic Sensor Net (Electrical Geodesics, Inc.) at 500 Hz. MRI clock signals were synced with EEG data collection for subsequent MR artifact removal. Functional scans were acquired using 3-T Siemens Trio MRI Scanner using echo planar imaging gradient-echo sequence with echo time (TE) of 25msec, repetition time of 1s, 6mm slices, 2mm gap, flip angle 90 degrees, with 3mm in-plane resolution, ascending acquisition. EEG data then underwent MR gradient artifact removal by subtracting an exponentially weighted moving average template, according to methods described in [7].

We compare LR+SD to the established InfoMax ICA cleaning method, as implemented in Brain Analyzer v.2.0.2 software (Brain Products) using manual identification of cardiac signal within the EEG followed by the automated solutions procedure for identifying IC components correlated with cardiac signal. For comparison purposes, we also collected single modality EEG data outside the MR environment using the same stimuli and parameters in a copper sheilded room (referred to as "Outside Scanner" in figures).

3 Results

3.1 Simulated Data Results. In the case of simulated data, we know the "true" solution. We adopt the time-frequency representation (TFR) to visualize

results computed via a continuous wavelet transform (CWT) using the Morlet wavelet, to assess the efficiency of the proposed method. Figure 1 shows the TFRs for simulated data arising from three distributed dipole sources. The TFRs corresponding to the pure and artifact signals are depicted in the two upper right plots while the TFR obtained from the raw EEGs (pure EEGs mixed with pure artifacts) is given on the upper left plot. Notice that the time-frequency energy corresponding to pure EEGs is nearly undetectable due to the artifact energy. The signatures of each event are not visible in the raw EEGs' TFRs yet they are clearly visible in the sparse component.In both the single and multiple source experiments the regularization parameter λ had value 5.10^{-3} which resulted in ranks of 4 and 5 for the low-rank artifact components, respectively. The proposed method shows excellent results in separating the artifact parts from the EEG signals of interest.

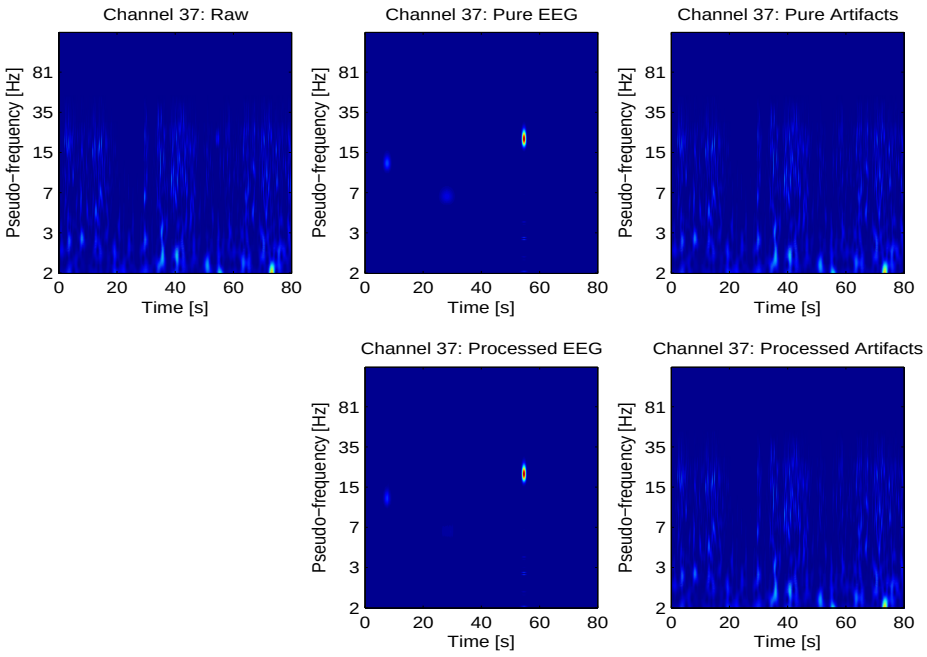

Fig. 1. Simulated Data Results. Three sources measured with an electrode located close to the primary motor cortex (source 3). (Top Panel) CWT of original simulated data consisting of three dipole sources corrupted by ECG, EMG and right and left EOG template artifacts (left), the true EEG data alone (middle), and artifact alone (right). (Bottom Panel) LR+SD result TFRs separating the cleaned EEG (left) from artifact (right). Values are normalized to maximum for each display.

3.2 EEG-fMRI Empirical Data Results. Group level ERSP results for LR+SD artifact removal of our experimental data are summarized and compared

to ICA artifact removal as well as out of scanner data in Figure 2(a-d). Signal-to-noise ratio (SNR) was computed by calculating the ratio of the maximum absolute signal diminution in alpha power from 0 to 500 msec following stimulus presentation to the standard deviation of alpha power from the following 1000msec post stimulus. SNR was 8.5, 11.4, and 15.2 for ICA, LR+SD, and out of scanner data respectively. Figure 2(d) shows group level alpha spectral EEG data projected topographically for pre-stimulus (-250 msec), ERSP (50msec), and post stimulus (500 msec), with timings with respect to the stimulus occurring at time equal to zero. Figure 3 shows single-patient alpha power averaged over all stimuli for a window of 2sec pre stimulus to 8sec post. The raw data is shown in comparison with the sparse component from LR+SD and ICA using an average of the time-frequency intensity over alpha band frequencies. The strength at the specific frequency of 10Hz with bounds of one standard deviation is also shown for each dataset.

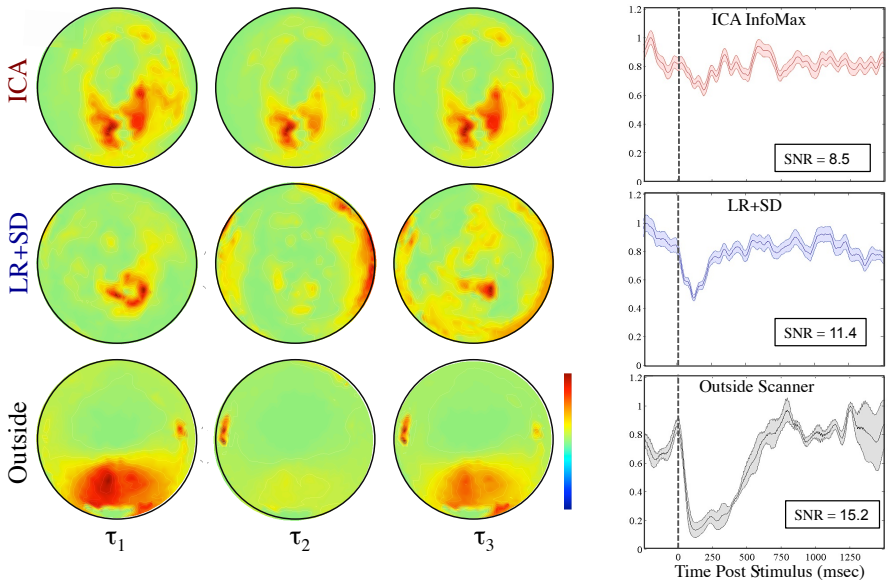

Fig. 2. Group Level Results comparing independent component analysis (ICA) and LR+SD based cleaning to EEG data collected outside of the MR scanner environment. (TOP PANEL) Normalized Alpha Power time courses derived from the ocular EEG channel averaged across all subjects plotted with mean +/- SEM for ICA, LR+SD and Outside Scanner Results. Signal-to-Noise rations are shown in the lower left corner for each result with the stimulus occurring at time equal to zero. (LOWER PANEL) Group level alpha power results projected topographically for 500msec prior to stimulus onset (PRE), 50 msec following stimulus onset(ERSD), and 500msec following stimulus onset (POST).

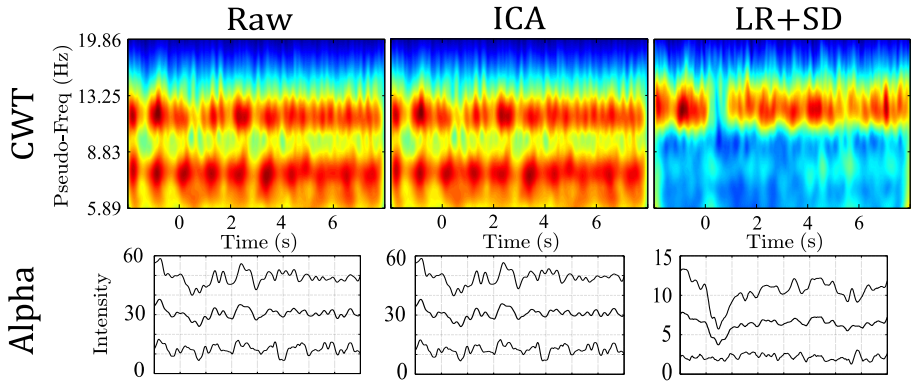

Fig. 3. Single-subject alpha band TFR averaged over stimuli. Columns denote (left) the raw MRI signal, (center) the ICA-cleaned signal (right) the sparse component from LR+SD. For each dataset is shown (top row) TFR over a frequency range about the alpha band and (bottom row) the signal strength at 10Hz with single standard deviation bounds.

4 Discussion

In this paper, we introduce a novel method for removing artifacts from EEG signals, which decomposes data into the sum of two matrices: a sparse matrix representing the cleaned data, and a low-rank matrix, which corresponds to the artifact portion of the data. We applied this algorithm to remove artifact from simulated and empirical data. Overall, LR+SD was quantitatively more effective EEG cleaning that ICA from an SNR perspective, and qualitatively more more robust at recovering the diminution in alpha power topographically. Overall, the LR+SD algorithm is quite similar to the robust PCA algorithm, which assumes that the data events themselves are sparsely represented in the time domain. Artifacts are conversely assumed to be broadly distributed at the channel level across the scalp, which appears to be a relevant assumption given what has been observed about the distribution of EEG artifacts, mentioned above, thus far. Here, we used only the Infomax algorithm for comparison. We did this primarily because previous studies have shown that this algorithm was most effective at BCG removal, however, results using ICA for artifact removal of concurrent EEG-fMRI have varied [6].

After optimization, we found that a rank of 25 was required for the low rank matrix to describe the BCG artifact. Lower ranks were ineffective at isolating the BCG noise, due to the complexity of the BCG signal itself. In the clinical setting, the ECG cardiac signal is often low-pass filtered and gross changes of its signature (e.g. ST segment elevation) are examined. However, lower amplitude changes in higher frequencies of the ECG exist and have been shown to correlate with cardiac ischemia even up to 250Hz [5]. Given this broad spectral signature, it is unsurprising that more that a few sparse components were required to capture

the artifact signals in real data. In summary, LR+SD provides an automatic method for parsing data and artifact into separate groups, with the need to tune only one regularization parameter. Further work may use spatial information from electrode topographies as further constraints.

Acknowledgements The authors want to thank the Keck Foundation for their support.

5 Bibliography

Rabinowitz, P.: PJ. Allen et al.: Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. NeuroImage. 8, 229-239 (1998)
S. Dahne et al.: Integration of Multivariate Data Streams With Bandpower Signals. IEEE Transactions on Multimedia. 15(5), 1001–1013 (2013)
J. Daunizeau and H. Laufs and K.J. Friston: EEGfMRI Information Fusion: Biophysics and Data Analysis: Springer Press: EEG-fMRI. 511–526 (2009)
S. Debener et al.: Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. NeuroImage. 34(2), 587–97 (2007)
P.K. Douglas et al.: Temporal and postural variation of 12-lead high-frequency QRS. J Electrocardiol. 39(3), 259–265 (2006)
T. Eichele et al.: Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. PNAS. 102(49), 17798-803 (2005)
R.I. Goldman et al.: Simultaneous EEG and fMRI of the alpha rhythm. Neuroreport. 13(18), 2487–2492 (2002)
S.A. Huettel et al.: Linking hemodynamic and electrophysiological measures of brain activity. Cerebral Cortex. 14(2), 165–173 (2004)
Tzyy-Ping Jung et al.: Removing Electroencephalographic Artifacts by Blind Source Separation. Psychophysiology. 37, 163–178 (2000)
Z. Lin et al.: The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. UILU-ENG-09-2215. arXiv:1009.5055v2 (2009)
A.J. Masterton et al.: Measurement and reduction of motion and BCG from simultaneous EEG and fMRI recordings. NeuroImage. 37, 202–211 (2007)
K.J. Mullinger et al.: Identifying the sources of the pulse artifact in EEG recordings made inside the MRI Scanner. NeuroImage. 71(1), 75–83 (2013)
R.K. Niazy et al.: Removal of fMRI environment artifacts from EEG data using optimal basis sets. NeuroImage. 28(3), 720–737 (2010)
R.Thornton et al.: Epileptic networks in focal cortical dysplasia revealed using electroencephalography-functional magnetic resonance imaging. NeuroImage. 70(5),822–837 (2011)
P.A. Valdes-Sosa et al.: Model driven EEG/fMRI fusion of brain oscillations. Human Brain Mapping. 30(9),2701–2721 (2009)
Y. Zou et al.: Automatic EEG artifact removal based on ICA and hierarchical clustering. IEEE ICASSP. 649–652 (2012)