

Reconstruction of EEG from limited channel acquisition using estimated signal correlation

A.G. Ramakrishnan¹, J.V. Satyanarayana*

Abstract

Nearby scalp channels in multi-channel EEG data exhibit high correlation. A question that naturally arises is whether it is required to record signals from all the electrodes in a group of closely spaced electrodes in a typical measurement setup. One could save on the number of channels that are recorded, if it were possible to reconstruct the omitted channels to the accuracy needed for identifying the relevant information (say, spectral content in the signal), required to carry out a preliminary diagnosis. We address this problem from a compressed sensing perspective and propose a measurement and reconstruction scheme. Working with publicly available EEG database, we have demonstrated that up to 12 channels in the 10-10 system of electrode placement can be estimated within an average error of 2% from recordings of the remaining channels. As a limiting case, all the channels of the 10-10 system can be estimated using recordings on the sparser 10-20 system within an error of less than 20% in each of the significant bands: delta, theta, beta and alpha.

Keywords: Correlated signals, Karhunen-Loeve Transform, Electroencephalography, Motor-Imagery tasks, Compressed Sensing, Convex Optimization, EEG electrode placement

1. Introduction

Nearby scalp channels of multi-channel EEG exhibit high correlation because EEG signals are not produced in the scalp or the neurons (brain) directly under the recording electrodes. Instead, as suggested in [1], they are a consequence of partial synchrony of local field potentials from distinct cortical domains - each domain, in the simplest case, being a patch of cortex of unknown extent. At any electrode, the EEG recording is a weighted linear mixture of underlying cortical source signals. The strong correlations observed between EEG recordings at nearby electrodes can be attributed to the spatial mixing of EEG source signals by volume conduction. Significant research effort has gone into exploring the correlation between EEG recordings at electrodes on different areas of the scalp. Heavy correlation, as high as 0.9, has been reported between anterior-posterior EEG signals in the alpha band in [2, 3, 4]. Very high coherence in the delta band has been reported in [5] between posterior temporal lobe regions. Interhemispheric coherence in the gamma band has been studied in normal adults in [6]. Existence of very high correlation, between temporal regions of the human brain, in the alpha band has been reported in [7].

1.1. Motivation

The primary interest in understanding inter-channel correlation in multi-channel EEG is to identify the scope for information redundancy in a measurement involving the full set of electrodes. This has the following potential applications:

*Corresponding author

Email address: jvsat29@yahoo.co.in (J.V. Satyanarayana)

¹Professor and Chairman, Department of Electrical Engineering, IISc. Bangalore.

1. If the channels are correlated, is it always mandatory to make recordings at all the electrodes, particularly those in close vicinity? For any subject, during the initial training sessions, all the channels are monitored. Once the correlation is learnt, one could avoid sampling some of the electrodes and yet be able to estimate the EEG spectral signature at the locations of the missing electrodes, for the same subject. For instance, using recordings on channels in the international 10-20 system of electrode placement, it is possible to estimate, with fair accuracy, the spectral content of channels in the denser 10-10 system. This could be useful in recording ambulatory EEG [8] which is carried out for an extended period (up to 72 hours) in which the patient can move about freely during the recording and data is stored in a pocket recorder. Thus, EEG recorded on a subject at rest, using a dense set of electrodes, can be used for training. Subsequently when the subject is in motion, all the channels need not be monitored.
2. Small form factor and low power operation are desirable features in a wearable EEG system because of its limited volume and battery capacity [9]. A design that supports simultaneous transmission of 32 channels of broadband data sampled at 30 ks/s is proposed in [10]. Wireless telemonitoring of Fetal ECG and epilepsy EEG based on compressed sensing has been proposed in [11]. Such architectures could benefit considerably from sub-sampling across EEG channels.
3. In sleep studies, it is possible that data is missing on some channels, either due to noise or due to undesirable movement by the subject. In this case, the loss of data, treated as undersampling, can be handled by recovery through compressed sensing. Here there is no intentional sub-sampling.
4. It is possible that deviation between the signal values estimated through compressed sensing and the actual measurements (due to loss of correlation), beyond a threshold, may indicate the onset of seizure in epileptic subjects.
5. In applications such as EEG brain mapping, where we are mainly concerned with the relative signal content in various frequency bands, accuracy of signal reconstruction per se, at such locations can be relaxed.

The focus of this work [[12]] is restricted to providing empirical evidence of obtaining approximate signal reconstruction and a reasonably good estimate of the spectral content in all channels of EEG by recording over limited number of channels. Before exploring the possibility of sub-sampling EEG signals, we present a brief introduction to the major standards of electrode placement.

1.2. Standards for EEG electrode placement

The first internationally accepted standard for electrode placement is the 10/20 system (figure 1a) that describes scalp locations via relative distances over the head surface between cranial landmarks. The primary purpose of this standard is to provide a reproducible method for placing a relatively small number of 21 EEG electrodes for recording. With the advent of multi-channel EEG systems, the need was felt for extending the 10/20 system to higher density electrode settings for use in research and diagnosis. This led to the introduction of the 10/10 system [13] (figure 1b) consisting of 64 electrodes, as a logical extension of the original 10/20 system. While electrodes are placed at distances of ten and twenty percent along certain contours over the scalp in the 10/20 system, they are placed at distances of ten percent along the medial-lateral contours in the 10/10 system. Also, new contours are introduced in between the existing ones. The 10/5 system with even higher electrode density was proposed in [14]. An elaborate description and comparison of all these systems is given in [15].

Consider the scenario where the electroencephalogram of a patient undergoing treatment or a subject voluntarily involved in research, needs to be frequently taken. The first session can constitute the training phase in which data from recordings on the 10-10 system is used to compute the inverse KLT matrix. During the actual recordings that are carried out subsequently, recording is performed at only those electrodes in the 10-10 system that coincide with the positions of the less dense 10-20 electrode placement. Simulation results (subsection 5.2) demonstrate that through compressed sensing recovery, signals at the rest of the electrodes in the 10/10 system can be reconstructed to a fair level of accuracy.

1.3. The Physionet database

We have applied the proposed method to the EEG signal database from Physionet [16]. A detailed description of the database is given in [17]. This data set consists of over 1500 one- and two-minute EEG recordings, obtained from 109 volunteers. The subjects performed different motor/imagery tasks (see Table 1), while 64-channel EEG was recorded at a sampling frequency of 160 Hz using the BCI2000 system [18]. The placement of the electrodes is as per the international 10-10 system (excluding electrodes Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10). Each volunteer performed, in a sequence, a set of 14 tasks: two baseline tasks followed by four different tasks repeated three times. Thus for each subject there are 14 records each containing data recorded from 64 channels. We have assumed that the motor/imagery tasks performed by the subjects (e.g. opening and closing of fists and imagining doing the same) are benign and, on a theoretical basis, unlikely to produce ECG, EOG and EMG artifacts in the recordings. Hence, no explicit steps have been taken to filter out these signals, in case they exist.

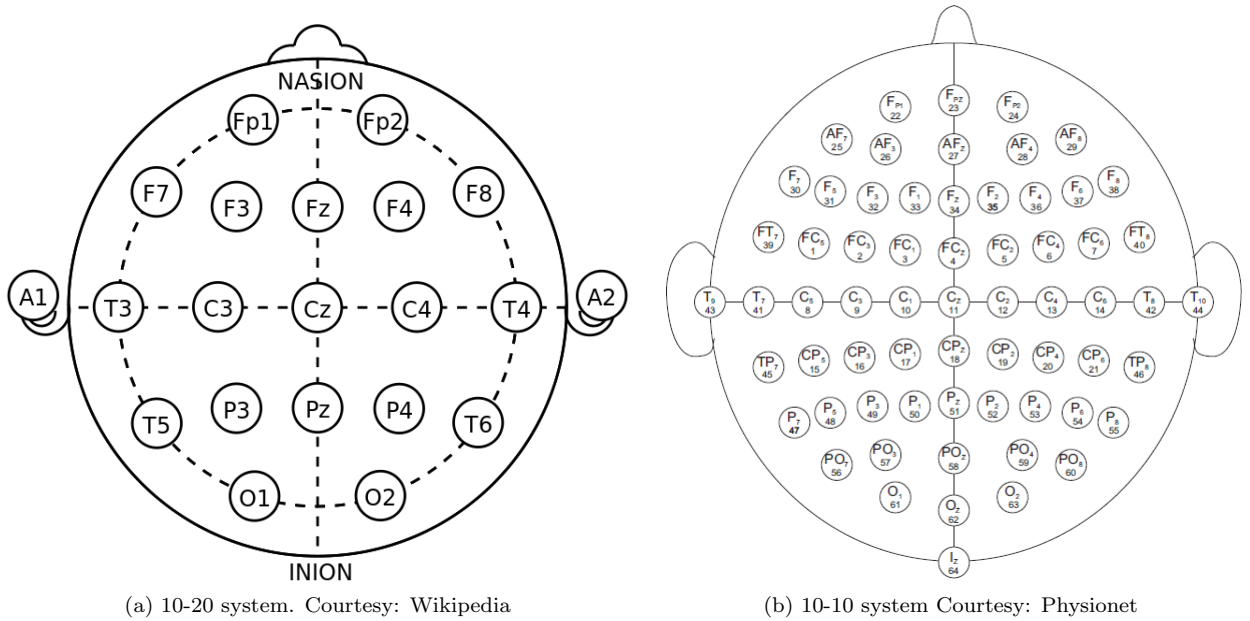


Figure 1: EEG electrode placement systems

2. Sub-sampling and reconstruction

Reduction in the number of EEG recordings involves identification of a suitable subsampling/reconstruction architecture, realization of which based on many different paradigms has been an important research area in itself. Almost all the subsampling schemes proposed for general signals are based on the assumption of signal sparsity in some domain such as time, frequency or space. Many spectral estimation methods have been proposed, where the signal is assumed to be sparse in the frequency domain. These methods [19, 20, 21] are suited to applications like radar, in which targets act as spatially sparse monotonies. Though EEG signals exhibit characteristic spectral peaks for various normal and pathological conditions, they are not in general sparse in the frequency domain. It is at best possible to exploit the intersignal correlation in such cases so that a vector of measurements from a set of electrodes could be transformed into a sparse vector on a basis derived from the inter-channel correlation matrix. It is important to mention here that significant work has been reported previously on the frequency analysis of EEG signals. Time-frequency analysis of EEG data based on adaptive periodogram technique has been proposed in [22]. Identification of the signal

Table 1: Motor/Imagery tasks during which the EEG used for the study has been collected. (see [17])

Record no.	TASK	DESCRIPTION OF THE TASK, DURING WHICH EEG IS RECORDED.
Record 1	Baseline 1	Eyes open DURATION: 1 sec
Record 2	Baseline 2	Eyes closed DURATION: 1 sec
Records 3,7,11	Task 1	A target appears on the left or the right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes. DURATION: 2 sec
Records 4,8,12	Task 2	The stimulus is the same as Task 1. However, in this case the subject only imagines responding to the stimulus the same way as in Task 1 and then relaxes. DURATION: 2 sec
Records 5,9,13	Task 3	A target appears on the top or the bottom of the screen. The subject opens and closes both fists if the target is on top and both feet if the target is on the bottom until the target disappears. Then the subject relaxes. DURATION: 2 sec
Records 6,10,14	Task 4	The stimulus is the same as Task 3. Again the subject only imagines responding to the stimulus the same way as in Task 3 and then relaxes. DURATION: 2 sec

components through decomposition of data into time-frequency-space atoms (based on the Wigner-Ville distribution) using parallel factor analysis has been proposed in [23]. Time-frequency spectral estimation of multichannel EEG has been attempted [24] using smooth, time-frequency localized versions of the Fourier functions. These contributions address the problem of detection and analysis very well.

In [25], the authors have exploited the joint sparsity of EEG signals on the Gabor frame and have achieved a low normalized mean square error. However in this work, the different trials are treated as different electrodes with the assumption that in both cases the same underlying activity is measured. In [26], the authors have suggested a novel approach of structuring individual signals into groups and exploiting the group sparsity by computing the $l_{1,2}$ regularization. However their approach involves the use of the unconventional random sampling based acquisition architecture and does not exploit the joint sparsity of a group of signals. In [27], the authors have demonstrated the use of fast ICA as a preprocessing step before compressed acquisition of EEG signals to achieve a low reconstruction error.

In the approach that we report in this paper, we intend to exploit the inter-channel correlation in EEG and in the process do away with some of the channels altogether during later acquisitions. The focus of our efforts is more on detecting the signal content in various spectral bands using measurements on limited number of channels. Section 4 gives a brief introduction to the well known Karhunen Loeve Transform (KLT), which makes available a sparsifying basis for a set of correlated signals. With such a sparsifying basis, a random subsampling scheme like compressed sensing could be employed for undersampling and reconstruction of the channels. Compressed sensing has made a significant impact in sparse signal processing in the past decade. The next section presents a rudimentary introduction to the area.

3. Compressed sensing paradigm

In classical compression, signals are Nyquist sampled and transformed into a sparse domain, following which only the significant coefficients are retained for transmission or storage. The compressed sensing paradigm [28, 29, 30, 31] provides mechanisms for random sub-sampling of signals that are sparse in an arbitrary transform domain, and subsequent reconstruction of the original signal from the sub-sampled measurement. Under compressed sensing schemes, sampling and compression are combined into a single step, so that only the required smaller number of samples are obtained through non-uniform sampling at

a sub-Nyquist rate. Let $\mathbf{x} \in \mathbb{R}^N$ be a finite length, discrete signal in the time domain which is to be sub-sampled. Assume that \mathbf{x} has a sparse representation in a transform domain, represented by the unitary matrix $\Psi \in \mathbb{R}^{N \times N}$. In other words,

$$\mathbf{x} = \Psi \mathbf{c} \quad (1)$$

where \mathbf{c} is an $N \times 1$ vector that has at most $K < N$ non-zero elements, i.e. it is a K -sparse vector. In practice, for real world signals \mathbf{x} , \mathbf{c} has at most K significant elements and the rest are negligibly small. For signal compression, the negligible coefficients are set to zero and only the significant coefficients are transmitted (stored). A lossy recovery of the original signal is then obtained using (1). Although the signal is efficiently compressed, all the Nyquist samples are required initially. Compressed sensing theory suggests that, $M < N$ linear combinations of the signals be taken using a measurement matrix, $\Phi \in \mathbb{R}^{M \times N}$ we have,

$$\mathbf{f} = \Phi \mathbf{x} = \Phi \Psi \mathbf{c} \quad (2)$$

The next step after taking the compressed measurement is to reconstruct the original signal \mathbf{x} , given the measurement vector \mathbf{f} , the inverse transformation Ψ^{-1} , and the measurement matrix Φ . It is clear that simple linear algebra does not permit us to do so, due to the fact that the set of equations (2) has more number of unknowns than equations.

The earliest reconstruction algorithms, that were proposed to get around this problem, were geometric involving l_1 minimization techniques to find the K -sparse vector, \mathbf{c} from the measurement \mathbf{f} .

$$\begin{aligned} \text{minimize } \|\mathbf{h}\|_1, \quad \text{subject to } \Phi \Psi \mathbf{h} &= \mathbf{f} \\ \hat{\mathbf{c}} &= \mathbf{h} \\ \hat{\mathbf{x}} &= \Psi \hat{\mathbf{c}} \end{aligned} \quad (3)$$

Also known as the basis pursuit [32], this method has been widely used in various applications of compressed sensing. It can be extended to the case of noisy signals by altering the first line in (3) as:

$$\text{minimize } \|\mathbf{h}\|_1, \quad \text{subject to } \|\Phi \Psi \mathbf{h} - \mathbf{f}\|_2 \leq \epsilon \quad (4)$$

where ϵ is a small application specific term bounding the amount of noise in the data.

l_1 minimization technique offers a high reconstruction accuracy. However, its complexity increases significantly with signal dimension. Iterative greedy algorithms have been proposed, which execute faster compromising the reconstruction accuracy. The orthogonal matching pursuit [33], along with its many variants [34, 35], is a greedy, iterative algorithm which finds the support of the sparse vector progressively.

4. Karhunen Loeve Transform

Karhunen Loeve Transform (KLT) [36, 37] is a reversible linear transformation that removes redundancy in signals by decorrelating them. KLT has been extensively used in image compression, wherein the correlation between the pixels is exploited.

Consider the signal matrix, $\mathbf{S} \in \mathbb{R}^{\tau \times N}$, the rows of which are indexed by τ successive time instants and the columns are indexed by N correlated signal sources taken from a training set. The covariance matrix of \mathbf{S} denoted by $\Sigma_{\mathbf{S}}$, is real and symmetric for real valued \mathbf{S} and its eigen vectors, ψ_n , are orthogonal. Consequently, one can construct an orthogonal matrix, $\Psi \equiv [\psi_0, \psi_1, \dots, \psi_{N-1}]$ such that

$$\Sigma_{\mathbf{S}} \Psi = \Psi \Lambda \quad (5)$$

where Λ is a diagonal matrix consisting of the corresponding eigen values. It is to be noted that Ψ is the inverse (or transpose) of the well known KL transform.

Let $\mathbf{x} \in \mathbb{R}^{N \times 1}$ be a test vector comprising samples from N sources. If we assume the inter-signal correlation in the test vector does not vary much from that of the training data, then \mathbf{x} can be represented on the basis Ψ as in (1) and the vector \mathbf{c} can be considered to be sparse. This in turn implies that only a subset

$\Omega \subset \{1 \dots N\}$ (of size $|\Omega| = M$) of the N source signals need to be sampled to form a measurement vector and the remaining sources can be estimated through a suitable reconstruction scheme. Such a subsampling scheme can be represented by a measurement matrix Φ formed by the down-sized identity matrix², $\mathbf{I}_N^{(M)}$ consisting of those rows of an $N \times N$ identity matrix \mathbf{I}_N that are indexed by the set Ω . Thus, in equation (2) we have $\Phi = \mathbf{I}_N^{(M)}$ and an estimate of \mathbf{x} can be obtained by solving (3) where Ψ is the inverse of KLT of \mathbf{S} .

The name KLT has been synonymously used with principal component analysis (PCA) by the signal processing community. In the light of this, it is very pertinent to mention that the method proposed in this paper is not the same as the sparse PCA approach [38, 39] which seeks sparse principal components that span a low dimensional space where the matrix Ψ is found by solving an optimization problem with a sparsity constraint on its entries. Equivalently, one tries to maximize $\psi_n^T \Sigma_S \psi_n$ subject to $\|\psi_n\|_2 = 1$ and also $\|\psi_n\|_0 = K$ where K is the parameter that controls the sparsity.

Our work is similar to regular PCA (or KLT), where the matrix Ψ is formed out of the eigen vectors of the covariance matrix (5). In other words, one tries to maximize $\psi_n^T \Sigma_S \psi_n$ subject to $\|\psi_n\| = 1$. Thus, the sparsity constraint is not imposed on Ψ . Instead, sparsity constraint is applied on the vector \mathbf{c} in the minimization in the equations (3) that also involve the matrix Ψ formed using the standard PCA with the help of plain matrix algebra. Thus, throughout the process of acquisition and reconstruction of the signals, the inverse KLT matrix Ψ , calculated previously from the training data set, remains unaltered.

5. Application to EEG signals

5.1. The experiments

Our experiments are performed on fifty five subjects (about half of the total number in the database) arbitrarily picked up from 1 to 109 in the Physionet database. The experiment comprises the following steps, which are categorized into the training and the testing phases for the sake of clarity:

Training phase

Step 1: For each subject, two of the records are arbitrarily picked up from 3 to 14 (the baseline records are not considered). One of them is used for training and the other is used for testing.

Step 2: The inverse KLT matrix is computed for the training record using all the 64 channels of the 10-10 system (see figure 1b).

Testing phase

Step 3: For each subject, at each successive time instant of the test record, a sub-sampling is carried out. That is, only a subset of the 10-10 system, consisting of nineteen channels that coincide with the locations of electrodes in the 10-20 system (FP1, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, O2) are sampled. The samples from the nineteen channels form the measurement vector, $\mathbf{y}^{(t)}$. The remaining forty five are estimated separately using each of the following reconstruction algorithms:

- basis pursuit (BP)
- orthogonal matching pursuit (OMP)
- compressive sampling matching pursuit (CoSAMP)
- regularized orthogonal matching pursuit (ROMP)

The inverse KLT matrix computed on the training record is used as the measurement matrix in each case.

Step 4: The DFT³ of the original and reconstructed signals is computed for blocks of size 512 samples.

²Informal arguments in the Appendix justify the use of the downsized identity matrix as the measurement matrix.

³To compute the DFT, the entire signal is divided into segments of length 512, with a 50% overlap. To each segment, a Hamming window is applied and the DFT of the windowed segment is calculated. The absolute values of the DFT coefficients are averaged over all the segments. We have chosen 512 as the size of the DFT so as to get a good frequency resolution of 0.3125 Hz in the DFT spectra, given that the sampling frequency is 160 Hz.

Step 5: The original and the reconstructed signals are compared (as explained in the next section) for each of the 45 channels that are compressively reconstructed.

It is to be noted that, for any subject, all the 64 EEG channels are sampled only during initial training. Subsequently, on the same subject, only a subset need to be sampled and the rest can be reconstructed.

5.2. Results

In order to quantify the reconstruction quality, the **normalized mean square error** (NMSE) between the DFT coefficients, \mathbf{z}_{org}^i of the original signal and those, \mathbf{z}_{rec}^i , of the reconstructed signal is computed in each channel for each of the significant frequency bands - delta, theta, alpha and beta (denoted by i). The NMSE for the i^{th} band is defined as

$$NMSE^{(i)} = \frac{\left\| \mathbf{z}_{rec}^{(i)} - \mathbf{z}_{org}^{(i)} \right\|_2^2}{\left\| \mathbf{z}_{org}^{(i)} \right\|_2^2} \quad (6)$$

where i denotes one of the bands: *delta*, *theta*, *alpha* and *beta*. Towards this end, the following values, taken from [40] have been chosen as the lower and upper limits of each band: delta: 0.1-3.5 Hz; theta: 4-7.5 Hz; alpha: 8-13 Hz; beta: 14-30 Hz. Experiments have been carried out on records from 55 subjects in the physionet database to verify proof of performance. For each subject, two records were arbitrarily picked up; one record was used for training and the other for testing.

Sample results (see figure 2) have been shown for Subject 36⁴ in the interval: 60-80 s. The figure gives a time domain comparison of the original and the reconstructed signals for the ‘F2’ (which belongs to the 10-10 system) channel along with a comparison of the corresponding log spectrum. As an additional assessment of the effectiveness of our method, the Fourier spectrum of the reconstructed signal is compared with the spectrum⁵ of the original signal, estimated by linear prediction⁵ (LP) in the bottom sub-plots each of the subfigures: 2a, 2b, 2c, 2d. Comparison with LP spectrum is motivated by [[41]] in which FFT and auto-regressive spectra have been compared for heart rate variability studies. However, it is known that the fidelity of the LP spectrum depends on the model order chosen and will be a poor approximation to the actual spectrum, if the prediction order is low. Hence, a high model order of 50 is chosen. Further, it is to be noted that the LP spectrum is computed from the actual signal itself, which is compared with the Fourier spectrum of the signal reconstructed from the other channels. The plots clearly show that spectra of the compressively reconstructed signals have a reasonably good match with the spectrum of the original signal.

Figure 3 gives the NMSE values, shown as plots, for forty five 10-10 channels, of Subject 36, reconstructed using various recovery algorithms. The NMSE values between the linear prediction spectral coefficients and the DFT spectra of the original signal are also shown (in blue) in the same figure. Table 2 gives the average NMSE across all the reconstructed channels with various recovery methods along with the corresponding values for linear prediction spectrum.

A plot of the NMSE in different EEG bands, averaged over all the reconstructed channels, is shown 55 subjects in figure 4. Please note that while in figure 4, the plots shown are the average NMSE of 45 channels for each of the 55 subjects (x-axis) in different bands, in figure 3, the plots shown are the NMSE of 45 reconstructed channels (x- axis) of the 10-10 system of a single subject. Although the NMSE for a few channels is above 0.2 (mainly in the alpha and beta bands), the average NMSE across all channels for the 55 subjects (figure 4) is close to 0.1 in all the bands.

The rationale behind sampling 19 channels and reconstructing all the 64 channels including the 45 that are not sampled is to be able to reconstruct EEG signals on the 10-10 grid (which uses 64 channels) using measurements on the 10-20 grid (which uses 19-21 channels). However, simulation has been carried out to observe the variation of NMSE as the number of sampled channels is altered from 10 to 63. When the number

⁴Both the subject and the interval have been arbitrarily chosen

⁵Order 50 has been used in the simulations

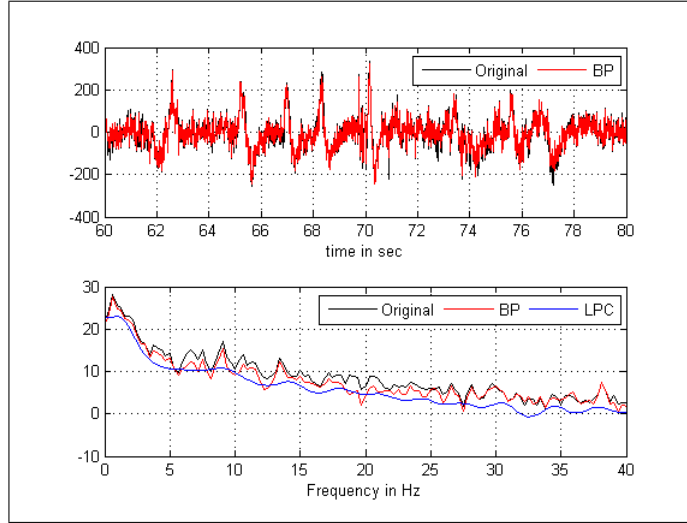
Table 2: Average NMSE for Subject 36 using various reconstruction methods. For comparison, the corresponding figures with LPC spectrum of the original signal are given.

Average NMSE(%)				
	Delta	Theta	Alpha	Beta
BP	3.76	3.26	5.27	4.36
CoSAMP	2.0	1.26	1.87	2.87
OMP	3.35	1.80	2.58	2.53
ROMP	2.52	1.58	3.34	3.86
LPC	7.45	8.26	7.95	8.77

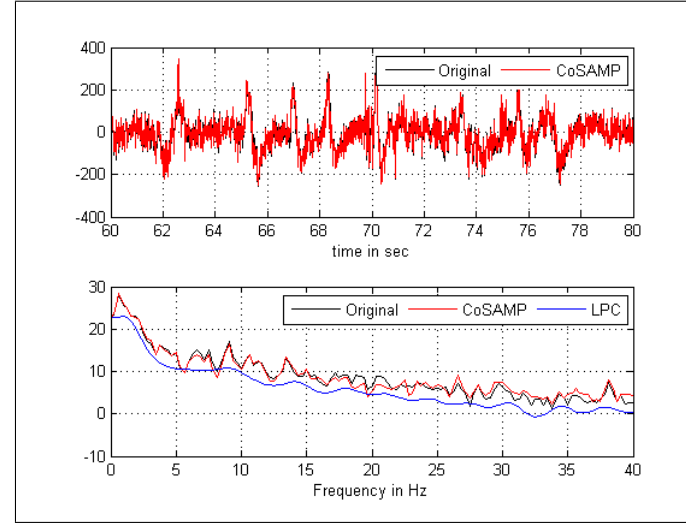
of sampled channels is as low as 10, the number of channels that have to be compressively reconstructed is 54. When the number of sampled channels is 63, only one channel is recovered through compressed sensing. As seen in figure 5 the NMSE steadily decreases as the number of sampled channels increases from 10 to 63, with all the reconstruction methods. The mean NMSE is less than 3% when nearly 50% of the channels are predicted from the remaining 32 channels and it is less than 1% when only 10% of the channels are predicted from the remaining 58 channels.

The simulation has been carried out on 64-bit Matlab operating on a computer with the specifications: Intel core i7 processor @3.4 GHz with 4 GB RAM. The mean computation time per subject for the reconstruction of 64 channels ⁶ of 125s duration EEG is 54.7s for OMP, 63.4s for ROMP, 105.7s for CoSAMP and 2602s for BP.

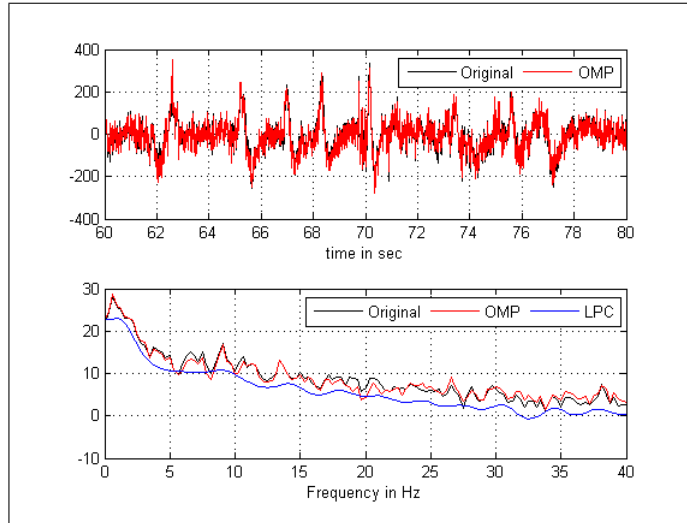
⁶Although 45 of the 10-10 channels have to be estimated, all the 64 channels (including the 19 belonging to the 10-20 system that are directly sampled) are involved in compressed sensing recovery.



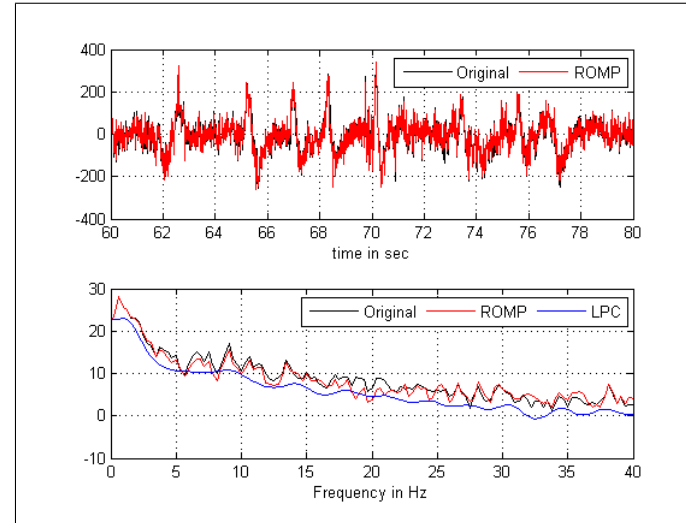
(a) **Basis Pursuit**



(b) **Compressive Sampling Matching Pursuit**

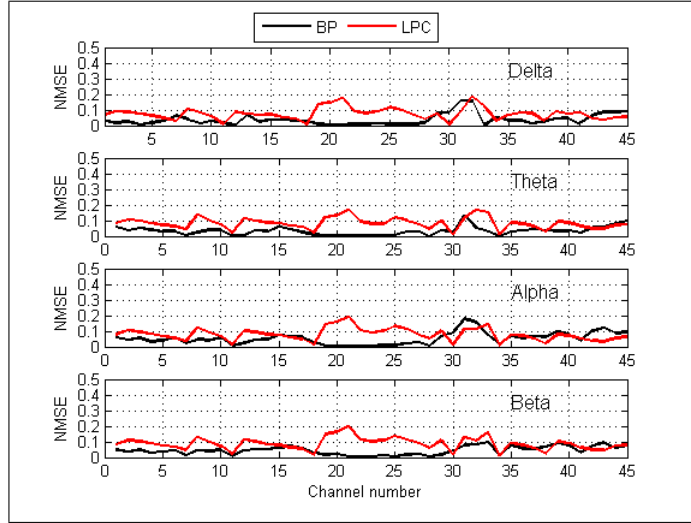


(c) **Orthogonal Matching Pursuit**

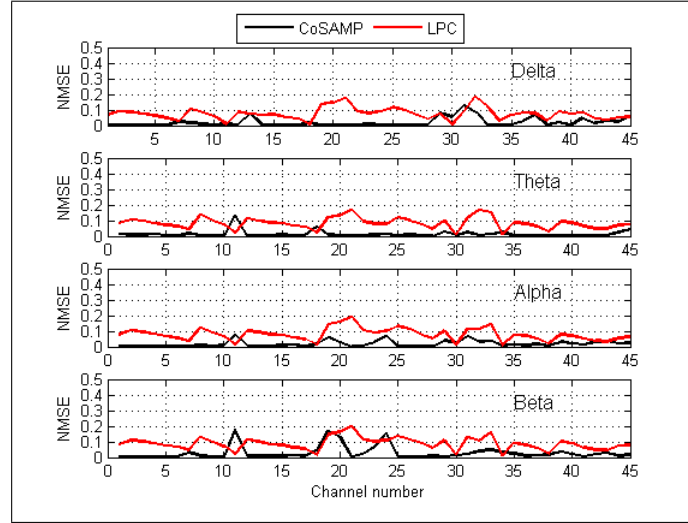


(d) **Regularized Orthogonal Matching Pursuit**

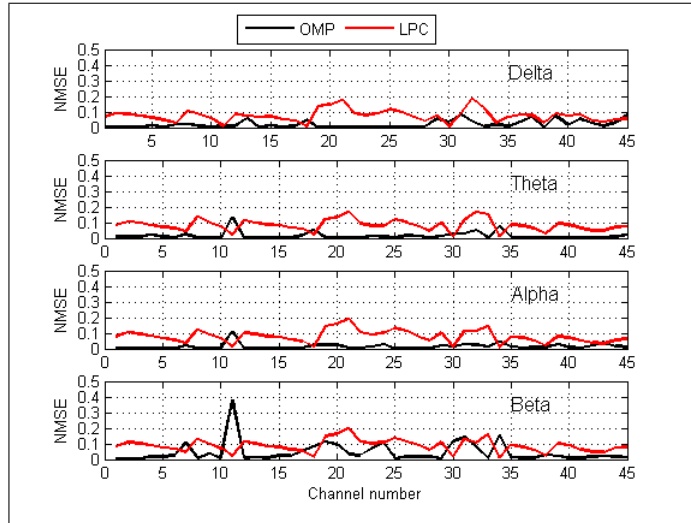
Figure 2: Original (black) and reconstructed (red) signal (upper plot) and log magnitude spectral plots (lower part) of channel F2 (60 - 80 sec) of subject 36, using (a) BP, (b) CoSAMP, (c) OMP and (d) ROMP. The reconstructed signals have a reasonably good fidelity, both in the time and frequency domains. Also, compared is the power spectrum, estimated using a 50-order linear prediction modeling of the original signal.



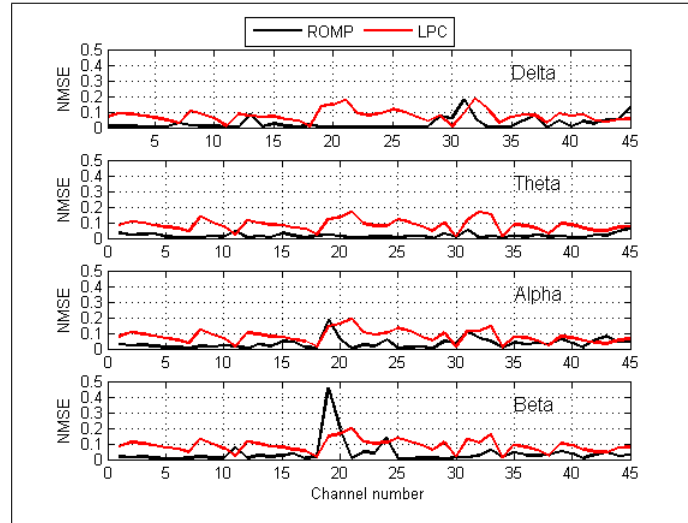
(a) Basis Pursuit



(b) Compressive Sampling Matching Pursuit

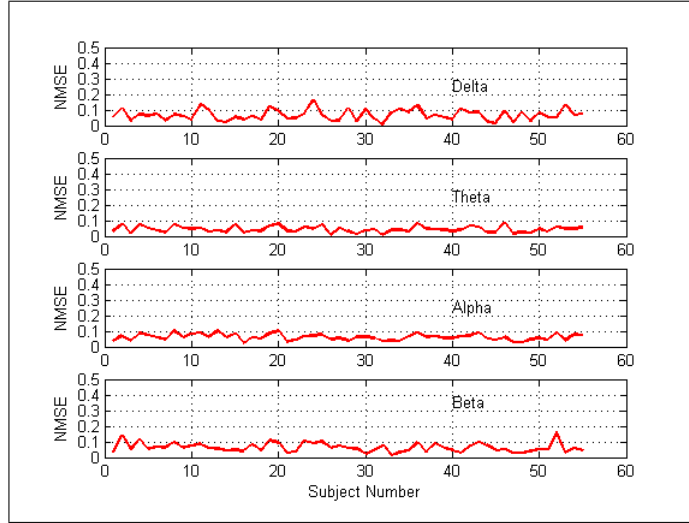


(c) Orthogonal Matching Pursuit

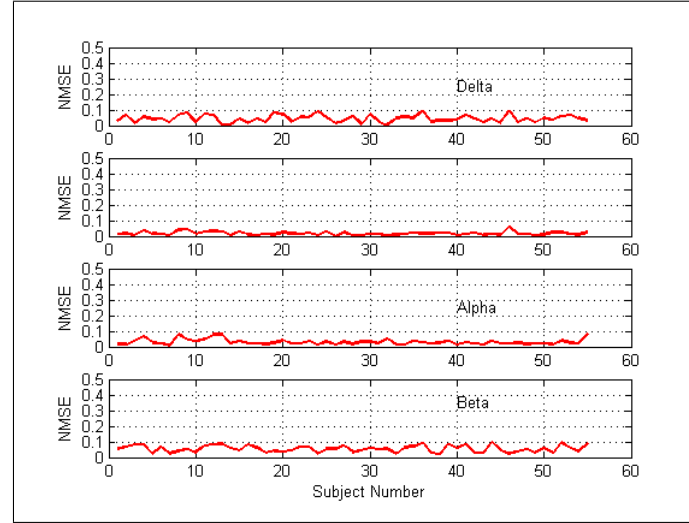


(d) Regularized Orthogonal Matching Pursuit

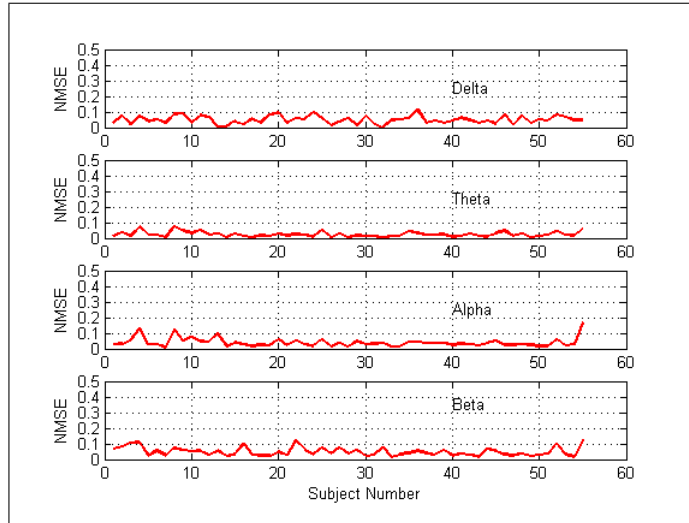
Figure 3: Plot showing NMSE for 45 different “10-10” channels: Fc5, Fc3, Fc1, Fcz, Fc2, Fc4, Fc6, C5, C1, C2, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6, FPZ, AF7, AF3, AFz, AF4, AF8, F5, F1, F2, F7, FT7, FT8, T9, T10 TP7, TP8, P5, P1, P2, P6, PO7, PO3, POz, PO4, PO8, Oz and Iz reconstructed from 19-channel “10-20” data for subject 36, using (a) BP, (b) CoSAMP, (c) OMP and (d) ROMP. Except for a few channels, the NMSE is less than 10%, using any of the four reconstruction methods. Also, compared are the NMSE values for the power spectra estimated using a 50-order linear prediction modeling of the original signal. Since NMSE itself is a measure of the error between DFT coefficients of the original and reconstructed signals, there is no separate plot for the original signal.



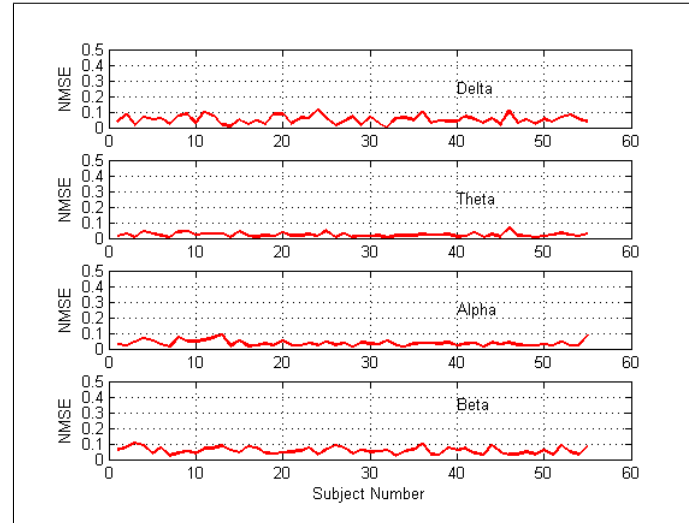
(a) **Basis Pursuit**



(b) **Compressive Sampling Matching Pursuit**



(c) **Orthogonal Matching Pursuit**



(d) **Regularized Orthogonal Matching Pursuit**

Figure 4: NMSE between DFT coefficients of original and reconstructed signals. Mean NMSE values across 45 reconstructed channels for each frequency band are plotted for each of the 55 subjects (x-axis) using (a) BP, (b) CoSAMP, (c) OMP and (d) ROMP

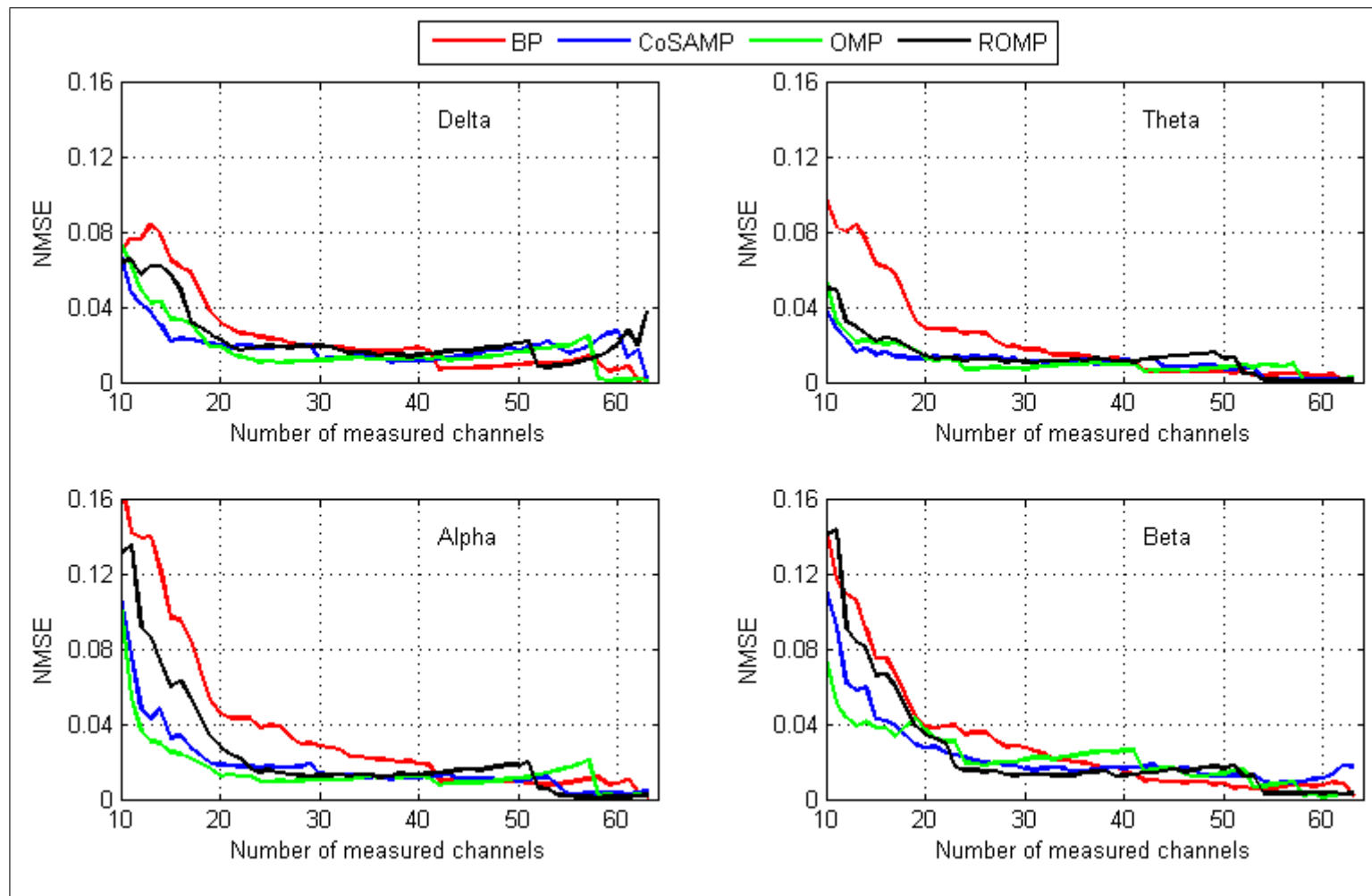


Figure 5: Mean NMSE (Subject 36) between DFT coefficients of original and reconstructed signals vs the number of measured channels for each frequency band using BP, CoSAMP, OMP and ROMP

6. Conclusion and Future work

In this work, we have presented a novel approach of channel subsampling of EEG, by measuring only a subset of electrodes and reconstructing the remainder through compressed sensing. Empirically, we have demonstrated that by recording at only a few locations on the scalp, it is possible to estimate the relative signal content in different frequency bands within an error of 20%. We propose the idea that if it is a matter of only knowing the relative spectral content, measurement of only a few EEG channels suffices, provided the correlation is previously captured in the inverse KLT matrix using data recorded in a training session. The prediction error is expected to be substantially less ($\approx 1\%$), if only a missing channel or two are predicted from otherwise a complete recording. Such a prediction is eminently useful in sleep studies, where occasionally one or two channels may be noisy due to the electrode losing contact with the skin. ~~There are~~ studies of hemispherical dominance [[42]] as a function of forced nostril breathing. We have used NMSE as a measure of fidelity of reconstruction of the missing channels. However, whenever the EEG is to be used for clinical evaluation, it may be desirable to use subjective evaluation of the quality of reconstructed signals by neurologists. Suitable subjective measures may also be defined for this purpose.

As seen from Fig. 5, the prediction error is less than 1%, when only a few channels are predicted from all the remaining channels. At this high level of reconstruction quality, the estimated signals can be directly used for diagnosis. The mean NMSE only increases to 3% when half the channels (32) are predicted from the remaining 32 channels, using any of the matching pursuit based reconstruction algorithms. The mean error being very low, again the predicted signals could be clinically useable. However, when the 10-10 channels are predicted from only the 10-20 data, even though the mean NMSE is still less than 4%, some of the channels have a peak error of up to 20%. One needs to study further to determine if these signals are sufficient for clinical diagnosis. However, for certain research studies such as effect of forced nostril breathing on hemispheric dominance [[42]], one is only looking at the change in the topographic EEG spectral map, and the reconstructed EEG, at 4% mean error, is still useable.

The method proposed is suitable for real time data capture, since the computationally intensive reconstruction process can be done offline. Also it estimates the signal spectrum equally well with different reconstruction algorithms: BP, CoSAMP, OMP and ROMP and better than the linear prediction spectrum.

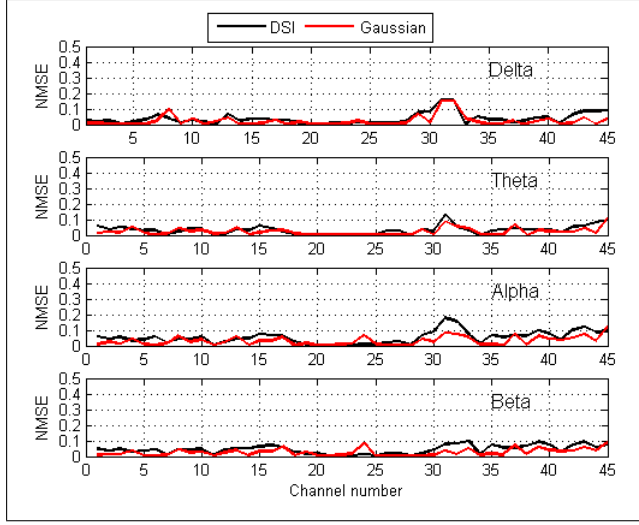
Since the probability of successful reconstruction of the EEG signals using our method will be no better or worse than the guarantees provided by the respective reconstruction algorithms, we have not attempted to present any additional theoretical proof of performance of the proposed method.

One cannot ignore the fact that the inter-channel correlation may be weak and time varying. It is definitely desirable to be able to detect small changes in the correlation pattern, which might have diagnostic significance. However, we are not sure that the proposed method can accomplish this and we shall explore this in future. Also, we intend to explore the possibility of detecting sudden changes such as epileptic seizures manifesting as high reconstruction error in one or more channels, due to lack of correlation.

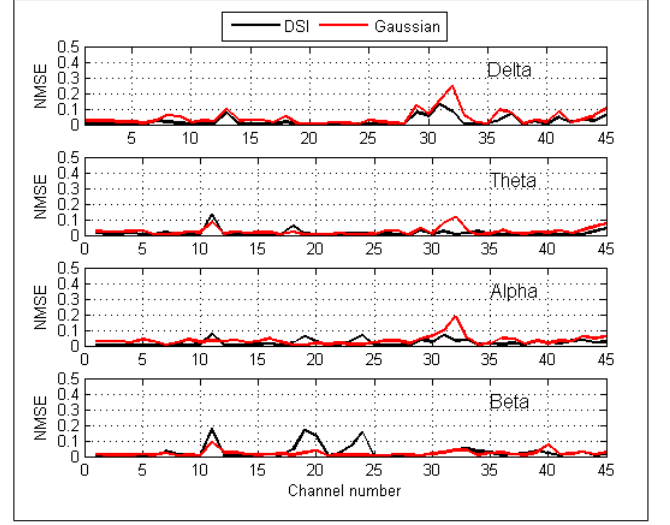
7. Appendix: Justification for using the down-sized identity matrix

One can show that a sufficient condition for stable recovery of a sparse signal is that the matrix product $\Phi\Psi$ in (3) must satisfy the restricted isometry property (RIP) [43, 44]. An alternative approach to stability is to ensure that the measurement matrix Φ is incoherent with the sparsifying basis Ψ in the sense that the vectors $\{\phi_j\}$ cannot sparsely represent the vectors $\{\psi_i\}$ and vice versa [28, 33]. Compressed sensing literature offers several random matrices like Gaussian and Bernoulli which are supported by adequate proof of satisfying RIP, unlike the downsized identity matrix (DSI) that we have used in this work. In the absence of strong theoretical evidence, we make the following informal arguments in favour of the DSI matrix:

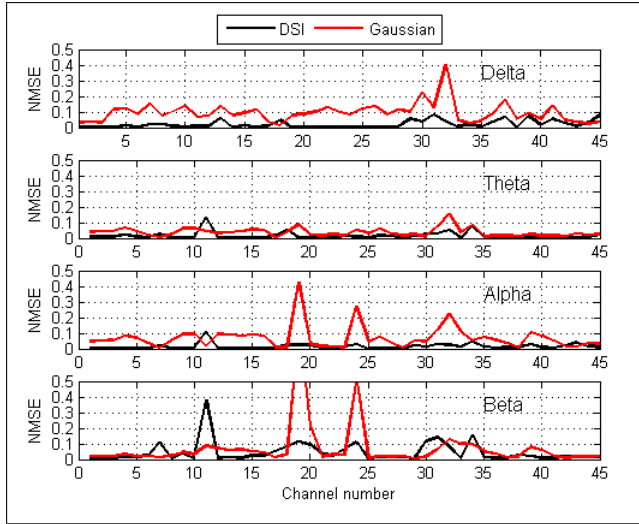
1. As suggested in [45] given an $M \times N$ measurement matrix \mathbf{U}_Ω formed by picking up M rows, indexed by the subset $\Omega \subset \{1 \dots N\}$ of size $|\Omega| = M$, uniformly at random, from an N -by- N orthonormal matrix \mathbf{U} , it is possible to recover a signal $\mathbf{x} \in \mathbb{R}^N$ that is sparse $|\mathbf{x}|_0 \ll N$ from a vector $\mathbf{f} = \mathbf{U}_\Omega \mathbf{x}$. In the scheme we have proposed, \mathbf{U}_Ω , is formed by picking rows from the orthonormal, inverse KLT matrix, Ψ by pre-multiplying with the DSI matrix, Φ .



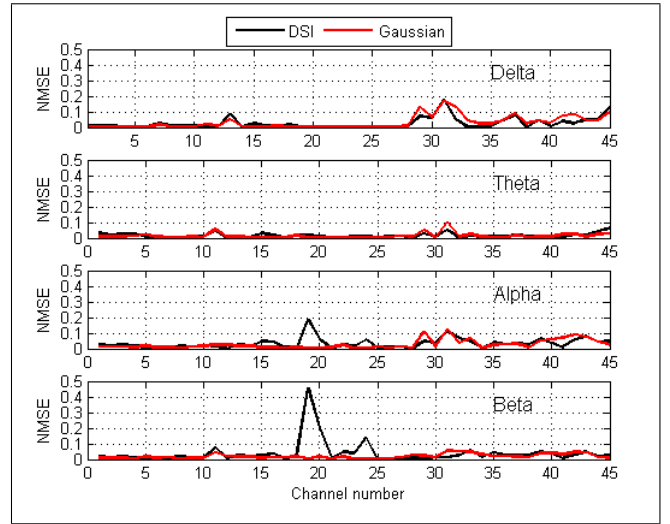
(a) Basis Pursuit



(b) Compressive Sampling Matching Pursuit



(c) Orthogonal Matching Pursuit



(d) Regularized Orthogonal Matching Pursuit

Figure 6: Plot (red color) showing NMSE for 45 different “10-10” channels (as in figure 3) using a Gaussian measurement matrix with different reconstruction methods (a) BP, (b) CoSAMP, (c) OMP and (d) ROMP. For comparison, the corresponding plot (black color) using the down-sized identity matrix (DSI) shown in figure 3 is reproduced.

2. The authors in [45] suggest that although theoretically powerful, the practical relevance of using completely random matrices like the Gaussian and Bernoulli is limited since we do not have the liberty to choose the type of measurements that will be used to acquire the signals. In the same paper, it is also stated that it is quite costly to use random matrices in practical sensing applications due to high computational complexity and memory buffering because of the completely unstructured nature.

For the sake of completeness, we have presented the results of simulation with a Gaussian measurement matrix, in place of the down-sized identity matrix (DSI) with different reconstruction methods in figure 6.

References

- [1] J. Onton, M. Westerfield, J. Townsend, S. Makeig, Imaging human EEG dynamics using independent component analysis, *Neuroscience and Biobehavioral Reviews* 30 (2006) 808–822.
- [2] L. Leocani, T. Locatelli, V. Martinelli, M. Rovaris, M. Falautano, M. Filippi, G. Magnani, G. Comi, Electroencephalographic coherence analysis in multiple sclerosis: correlation with clinical, neuropsychological, and mri findings, *J. Neurol. Neurosurg. Psychiatry* 69 (2000) 192–198.
- [3] I. Manshanden, J. C. D. Munck, N. Simon, F. da Silva, Source localization of MEG sleep spindles and the relation to sources of alpha band rhythms, *Clin. Neurol.* 113 (2002) 1937–1947.
- [4] P. L. Nunez, R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein, P. J. Cadusch, EEG coherence: statistics, reference electrode, volume conduction, laplacians, cortical imaging, and interpretation at multiple scales, *Electroencephalogr. Clin. Neurophysiol.* 103 (1997) 499–515.
- [5] G. Winterer, M. F. Egan, T. Radler, T. Hyde, R. Coppola, D. R. Weinberger, An association between reduced interhemispheric EEG coherence in the temporal lobe genetic risk for schizophrenia, *Schizophr. Res.* 49 (2001) 129–143.
- [6] M. G. Knyazeva, D. C. Kiper, V. Y. Vildavski, P. A. Despland, M. MaederIngvar, G. M. Innocenti, Visual stimulus-dependent changes in interhemispheric EEG coherence in humans, *J. Neurophysiol.* 82 (1999) 3095–3107.
- [7] K. S. Cover, C. J. Stam, B. W. van Dijk, Detection of very high correlation in the alpha band between temporal regions of the human brain using meg, *NeuroImage* 22 (2004) 1432–1437.
- [8] <http://epilepsy.med.nyu.edu/diagnosis-treatment/eeg/ambulatory-eeg#sthash.UmD06W0N.dpbs>.
- [9] T. Roh, K. Song, H. Cho, D. Shin, H. J. Yoo, A wearable neuro-feedback system with EEG-based mental status monitoring and transcranial electrical stimulation, *IEEE Trans. on Biomed. Ckts. and Sys.* 22 (6) (2014) 755–764.
- [10] H. Miranda, V. Gilja, C. A. Chestek, K. V. Shenoy, T. H. Meng, Hermesd: A high-rate long-range wireless transmission system for simultaneous multichannel neural recording applications, *IEEE Trans. on Biomed. Ckts. and Sys.* 4 (3) (2010) 181–191.
- [11] B. Liu, Z. Zhang, G. Xu, H. Fan, Q. Fu, Energy efficient telemonitoring of physiological signals via compressed sensing: A fast algorithm and power consumption evaluation, *Biomedical Signal Processing and Control*, Elsevier 11 (2014) 80–88.
- [12] J. V. Satyanarayana, A. G. Ramakrishnan, Compressed EEG Acquisition with Limited Channels using Estimated Signal Correlation, *arXiv preprint arXiv:1407.1285*.
- [13] G. E. Chatrian, E. Lettich, P. L. Nelson, Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity, *American Journal of EEG technology* 25 (2) (1985) 83–92.
- [14] R. Oostenveld, P. Praamstra, The five percent electrode system for high-resolution EEG and ERP measurements, *Clinical Neurophysiology* 112 (4) (2001) 713–719.
- [15] V. Jurcak, D. Tsuzuki, I. Dan, 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems, *Neuroimage* 49 (3) (2007) 1600–1611.
- [16] A. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. H. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, *circulation* 101(23): e215–e220, *Circulation* 101 (23) (2000 (June 13)) e215–e220.
URL <http://circ.ahajournals.org/cgi/content/full/101/23/e215>
- [17] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, J. R. Wolpaw, BCI2000: A general-purpose brain-computer interface (BCI) system, *IEEE Trans. Biomed. Engg.* 51 (6) (2004) 1034–1043.
- [18] <http://www.bci2000.org/>.
- [19] S. L. Marple, *Digital spectral analysis*, Englewood Cliffs, NJ, Prentice-Hall (1987) 373–378.
- [20] S. M. Kay, S. L. Marple, Spectrum analysis—a modern perspective, *Proc. IEEE*, Reprinted by IEEE Press, (*Modern Spectrum Analysis II*) 69 (11) (1981) 1380–1419.
- [21] S. M. Kay, *Modern spectral estimation: Theory and application*, Englewood Cliffs, NJ, Prentice-Hall,.
- [22] Q. Zhou, Analysis of EEG data using an adaptive periodogram technique, *Intl. Conf. on Biomedical Engineering and Informatics* 2 (2008) 353–357.
- [23] F. Romer, M. Haardt, D. Jannak, P. Husar, Multi-dimensional space-time-frequency component analysis of event related EEG data using closed-form PARAFAC, *Proc. ICASSP* (2009) 349–352.
- [24] S. D. Cranston, H. C. Ombao, R. von Sachs and W. Guo, B. Litt, Time-frequency spectral estimation of multichannel EEG using the auto-slex method, *IEEE Trans. Biomed. Engg.* 49 (9) (2002) 988–996.
- [25] S. Aviyente, Compressed sensing framework for EEG compression, *Proc. fourteenth workshop on statistical signal processing* (2007) 181–184.
- [26] R. D’Angelo, M. Trakimas, S. R. Sonkusale, S. Aeron, Compressed sensing of EEG using a random sampling ADC in 90nm CMOS, *Proc. Intl. Conf. Body Sensor Networks* (2013) 1–5.
- [27] S. Fauvel, A. Agarwal, K. Ward, Compressed sensing and energy-aware independent component analysis for compression of EEG signals, *Proc. ICASSP* (2013) 973–977.
- [28] D. Donoho, Compressed sensing, *IEEE Trans. on Info. Theory* 52 (4) (2006) 1289–1306.
- [29] E. J. Candes, Compressive sampling, *Proc. of the Int. Congress of Mathematics, Madrid, Spain* 3 (2006) 1433–1452.
- [30] Y. Tsaig, D. Donoho, Extensions of compressed sensing, *Signal Proc.*, Elsevier 86 (3) (2006) 549–571.
- [31] E. Candes, T. Tao, Near-optimal signal recovery from random projections: Universal encoding strategies, *IEEE Trans. on Info. Theory* 52 (12) (2006) 5406–5425.
- [32] S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, *SIAM Rev.* 43 (1) (2001) 129–159.
- [33] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. on Info. Theory* 53 (12) (2007) 4655–4666.

- [34] D. L. Donoho, Y. Tsaig, I. Drori, J. Starck, Sparse solution of undetermined linear equations by stagewise orthogonal matching pursuit, Tech. Report, Department of Statistics, Stanford University 2.
- [35] M. F. Duarte, M. B. Wakin, R. G. Baraniuk, Fast reconstruction of piecewise smooth signals from random projections, in: Online Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS), Rennes, France, 2005.
- [36] H. Hotelling, Analysis of a complex of statistical variables into principal components, *J. Educ. Psychol.* 24 (1933) 417–441.
- [37] K. Karhunen, Ueber lineare Methoden in der Wahrscheinlichkeitsrechnung, *Annales Academiae scientiarum Fennicae. Series A. 1, Mathematica-physica*, 1947.
- [38] H. Zou, T. Hastie, R. Tibshirani, Sparse principal component analysis, *Journal of Computational and Graphical Statistics* 15 (2) (2006) 265–286.
- [39] A. d’Aspremont, L. E. Ghaoui, M. I. Jordan, G. R. G. Lanckriet, A direct formulation for sparse pca using semidefinite programming, *SIAM Review* 49 (3) (2007) 434–448.
- [40] E. Niedermeyer, F. da Silva, *Electroencephalography: Basic principles, clinical applications, and related fields*, 5th ed. Lippincott Williams and Wilkins.
- [41] G. Parati, J. P. Saul, M. D. Rienzo, G. Mancia, Spectral Analysis of Blood Pressure and Heart Rate Variability in Evaluating Cardiovascular Regulation A Critical Appraisal, *Hypertension* 25 (6) (1995) 1276–1286.
- [42] A. Stančák, M. Kuna, EEG changes during forced alternate nostril breathing, *International journal of psychophysiology* 18 (1) (1994) 75–79. doi:10.1016/0167-8760(84)90017-5.
- [43] R. Baraniuk, Lecture notes on compressive sensing, *IEEE, Signal Processing Magazine* (2007) 118–124.
- [44] E. Candes, E. J. Romberg, T. Tao, Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory* 52 (2) (2006) 489–509.
- [45] E. Candes, E. J. Romberg, Sparsity and incoherence in compressive sampling, *Inverse problems* 23 (3) (2007) 969–985.