

Signal Processing Techniques to Reduce Energy Consumption in EEG Acquisition and Transmission for WBAN

Student Name: Ankita Shukla

IIIT-D MTech-EC-12-092

Indraprastha Institute of Information Technology
New Delhi

Thesis Committee

Prof. Ashwin Srinivasan (Chair)

Dr. S. P. Maitey

Dr. Angshul Majumdar

Dr. A. V. Subramanyam

Dr. Saket Anand

Submitted in partial fulfillment of the requirements
for the Degree of M.Tech. in Electronics and Communication

©2014 Ankita Shukla

All rights reserved

Keywords: Compressed Sensing, EEG, Low Rank, Joint Sparsity, Kronecker Compressed Sensing

Certificate

This is to certify that the thesis titled “**Signal Processing Techniques to Reduce Energy Consumption in EEG acquisition and transmission for WBAN**” submitted by **Ankita Shukla** for the partial fulfillment of the requirements for the degree of *Master of Technology in Electronics & Communication Engineering* is a record of the bonafide work carried out by her under my guidance and supervision in the SALSA group at Indraprastha Institute of Information Technology, Delhi. This work has not been submitted anywhere else for the reward of any other degree.

Dr. Angshul Majumdar
Indraprastha Institute of Information Technology, New Delhi

Abstract

Recently there has been a lot of interest in health tele-monitoring. Some vital health indicators like EEC, ECG, MEG, Pulse Plethysmograph etc. are collected by sensor nodes; the acquired signal is transmitted to a remote healthcare facility for monitoring and analysis. In effect, this comprises a Wireless Body Area Network (WBAN). Such a system allows for health monitoring of the elderly in western countries without the intruding presence of a paramedic. In India, such systems would benefit healthcare penetration into remote rural areas.

In a WBAN, the main challenge is to preserve energy. There are three main energy sinks namely : sensing, processing and transmission. The transmission energy is the highest, therefore the priority is to reduce it. This requires signal compression; however since the computational capacity at the sensor nodes is limited; this precludes application of sophisticated transform coding techniques. Recent studies proposed an alternate Compressed Sensing (CS) based approach wherein the sampled signal is projected onto a lower dimension by a random matrix, thereby effecting compression. The compressed signal is transmitted to the base station, where it is recovered by sophisticated CS recovery algorithms. Since computational power at the base station is not at a premium, this is a plausible model.

In this work, we are particularly interested in EEG. However the contribution of this work extends to any biomedical signal in general. EEG signals are acquired via multiple probes. All the probes monitor the same brain activity; therefore the EEG signals from the multiple probes/channels are correlated. We reviewed prior studies in this area and observed that all these studies concentrated on piecemeal recovery of the compressed EEG signal - they did not exploit the inter-channel correlations. For the first time, we exploited inter-channel correlations for EEG reconstructions. We proposed three methods - group-sparse recovery; Kronecker CS recovery and low-rank matrix recovery. The fundamental idea in all of the three remain the same, these are three different approaches to model the inter-channel correlations. We were able to show that the recovery results improved considerably compared to previous CS based methods.

The second contribution of this work is even more significant. We showed that by exploiting the inter-channel correlations we can actually reduce sensing energy and eliminate the requirement for processing / compression altogether. This simply requires randomly under-sampling the EEG signals. Such an operation reduces sensing cost and as the signal is compressed during acquisition, it eliminates processing costs. We found that our recovery techniques works well in this framework and yields good recovery, i.e. accuracy from random under-sampling is the same as accuracy from full sampling followed by compression.

The final contribution of this work is towards the signal processing community. We have derived algorithms for solving group-sparse recovery problems, low-rank matrix recovery problems and rank-aware sparse / group-sparse recovery problems. These algorithms are derived on the Split Bregman approach. These general purpose algorithms will find other areas of signal processing like multi-spectral imaging, X-Ray CT imaging and Magnetic Resonance Imaging to name a few.

Acknowledgments

This work would not have been possible without the guidance and support of several individuals who in one way or the other assisted in the preparation and completion of this work.

I would like to express my sincere gratitude and respect towards Dr. Angshul Majumdar, for being my supervisor and under whose able and encouraging guidance, I was able to complete my work in a better and exhaustive manner . He has inspired me to do good work, and has supported me all along the way.

I would like to thank Dr. Pankaj Jalote for making IIIT-Delhi such a wonderful place to work.

I owe particular thanks to my parents and siblings who have always encouraged me in my pursuits, without their help all this would have not been possible.

Lastly, I would like to thank my friends, batch mates and PhD students for providing such a good environment for fruitful discussions, many of which helped me immensely.

Ankita Shukla

Contents

1	Introduction	1
1.1	Background	1
1.2	EEG Based Body Area Network	3
1.2.1	Electroencephalogram Signals	3
1.2.2	Challenges and Shortcomings in EEG WBAN's	3
1.3	Aim of the Thesis	4
1.4	Organization of Thesis	5
2	Precursors	6
2.1	Background on Compressed Sensing	6
2.1.1	Sparse Recovery	6
2.1.2	Noisy Compressible Signals	8
2.1.3	Practical Signal Recovery	8
2.1.4	Measurement Basis	10
2.2	Alternating Direction Method of Multipliers	10
2.3	Split Bregman Method	12
3	Literature Survey	15
3.1	Gaussian Compression Basis with Gabor Sparsifying Transform	15
3.2	Non-Uniform Sampling and Reconstruction using Slepian Basis	16
3.3	Comparative study of different Sparsifying Basis	18
3.4	Block Sparse Bayesian learning	19
3.5	Joint reconstruction	19
3.6	Analysis Prior	20
3.7	Other Studies	21
4	Proposed Formulation	23
4.1	Sensing Power Reduction: Sub-Sampling	23
4.2	Kronecker Compressed Sensing Formulation	25
4.3	Low Rank Recovery Formulation	28

4.4	Row Sparse MMV Recovery Formulation	30
4.5	Combined Low Rank and Row Sparse Recovery Formulation	31
4.6	Row Sparse BCS Formulation	32
5	Proposed Algorithm	34
6	Experimental Setup, Evaluation and Results	38
6.1	Dataset	38
6.2	Evaluation Parameter	38
6.3	Results	38
6.3.1	Reconstruction Accuracy	39
6.3.2	Power Analysis	40
7	Conclusion and Future Scope	42
7.1	Conclusion	42
7.2	Future Work	42

List of Figures

1.1	Clinical Setup for EEG monitoring	1
1.2	WBAN setup for EEG monitoring	2
4.1	Data Acquisition and Transmission for Previous method(CS based)	24
4.2	Data Acquisition and Transmission for Proposed Method	24
4.3	Schematic Diagram of Sparsifying operation	26
4.4	Decay of Fourier Coefficients for EEG signal ensemble	27
4.5	Decay of Fourier Coefficients for EEG signal ensemble(magnified)	27
4.6	Similarity of EEG signals from 4 channels	29
4.7	Decay of Singular Values for EEG signal ensemble	29
4.8	Common Sparse Support of Wavelet Coefficients	31

List of Tables

6.1	Comparative Reconstruction Results (NMSE)for various Algorithms	39
-----	---	----

Chapter 1

Introduction

1.1 Background

In the current scenario, with people's busy work and home schedules, finding time to visit a doctor or a nutritionist to get an update on one's health is definitely not something most people can do on a regular basis. Also, an increase in elderly population and chronic diseases patients is observed in last few years. The traditional health care system cannot provide scalability as it relies on physical, one to one relationship between caregiver and patient.

Also, the test equipment and clinical setup involved in monitoring many health signals are time consuming and involves video monitoring along with long cables to amplifier and recording unit making the process cumbersome. Thus, making monitoring an impatient task for old and sick people. An EEG test equipment setup at hospital is shown in Fig. 1.1 [1]. So, there is a need of handy portable device that can monitor the health of a person. This allows one to sit comfortably at his/her home and also reduces the burden on medical staff and hospitals. Thus patients can carry out most of their activities normally and save trips to doctors and hospitals. And once the monitored data is captured, it can be sent to a hospital for analysis.



Figure 1.1: Clinical Setup for EEG monitoring

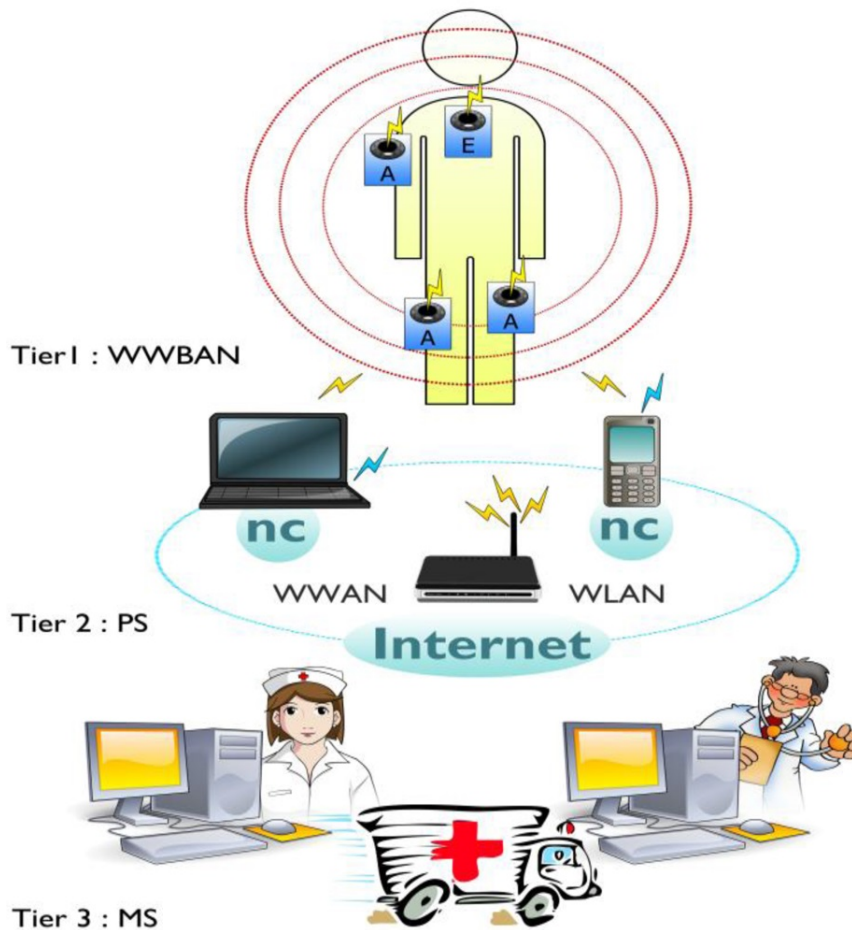


Figure 1.2: WBAN setup for EEG monitoring

With the rapid development in wireless communication, tele-monitoring of health signals by Wireless Body Area Networks(WBANs) is an evolving field for home based e-health where various sensors are located on patient's body [2, 3] as shown in Fig.1.2.

In the figure, first stage named Wearable Wireless Body Area Network (WWBAN) integrates various wireless medical sensor nodes that monitor several health signals such as electroencephalogram, electrocardiograph etc. The second stage includes personal server that runs on a cell phone or a home personal computer. The personal server provides an interface for the sensor and the medical server. It also stores the data collected from medical sensors locally and uploads to the medical server when network connection is available. The last stage, tagged as medical server retrieves data via internet and stores them in medical records. This stage may also include other servers such as informal caregivers and emergency servers.

1.2 EEG Based Body Area Network

1.2.1 Electroencephalogram Signals

In this work, the health signal we are particularly interested in monitoring over WBAN's is electroencephalogram(EEG). EEG signals measure the electrical signals generated within the brain. These signals are captured by placing sensors at various locations on the scalp of the head. EEG signals provide high temporal resolution but have poor spatial resolution. It is an important and commonly used brain imaging technique for diagnosis of various neurological disorders such as epileptic seizures [4, 5]. Monitoring of EEG with WBAN's offers several advantages in detection of seizure. Since seizure has rare occurrence, detection of it requires continuous monitoring for long time durations. This makes the process resource intensive for the hospital. Using WBAN allows the patients to monitor EEG themselves and then consult the doctor when the relevant data is collected.

Along with this, EEG signals are profusely used in many non medical applications such as Brain Computer Interface (BCI) [6, 7] wherein the EEG pattern associated with a particular activity is detected to understand a person's response.

1.2.2 Challenges and Shortcomings in EEG WBAN's

In WBAN's EEG is collected by number of sensors placed on scalp of head and are transmitted via gateway node over a existing wireless transmission channel as shown in Fig.1.2. The main problem in any wireless sensor network (including WBAN) is its limited energy resources at the sensor nodes. So, in order to operate the device for longer duration there is a need to reduce energy consumption as much as possible. The idea of using larger batteries to run for elongated period of time will make the device bulkier and overheating of device might create safety concerns for the patient. But now, if we look at the power consumption profile of the system, which involves capturing, sampling, digitizing and transmitting of samples, we find that some of process have higher power needs than other. On the basis of this, the entire process can be clubbed into three components in their increasing order of power consumption as follows:

- Processing
- Sensing
- Transmission

Energy saving in processing part can be achieved by using algorithms that have low computational complexity and hence are less power hungry. Transmission power can be reduced by reducing the number of samples that needs to be transmitted i.e. by compressing the signal.

So, we need to find a simple system that consumes less power and uses a suitable technique for compression.

Using Transform coding to obtain signal compression is not appropriate for this scenario as encoder (to perform compression) is complex i.e. power hungry whereas decoder (used to extract original signal from compressed signal) is simpler. This is exactly opposite to our situation's requirement where, the encoder needs to be simple (less power hungry) while the decoder being at the base station (hospital) has no premium on computational resources.

Compressed Sensing (CS) based methods are suited for this scenario. CS uses random projections (from higher to lower dimensions) for compression; this operation is computationally cheap (reduces processing power). Since the signal is now compressed, the transmission power is also reduced. However, the decoding is computationally expensive and requires solving a non-linear optimization problem. As mentioned before, this poses no issue since there are powerful computers at the base station.

But, CS is not the best and only solution for power saving. In CS, the amount of power expended in sensing remained unaltered, which contribute significantly to total power consumption. Hence, a better energy efficient WBAN system can be developed if one can reduce the sensing power as well.

1.3 Aim of the Thesis

To make a significant improvement in power saving for WBAN's over existing CS based methods, we present this work. As discussed before, in order to design a comfortable wearable power efficient device, the computation power and circuitry needs to be minimized. As listed above sensing is the second highest power consuming unit, so a lot more power at the encoder can be saved if this can be brought down.

Based on the brief discussion of the problem, the goal of our work can be divided into following sub categories:

1. At first, we review the recent CS based compression and reconstruction techniques for EEG signals.
2. We present a technique to reduce the total energy requirement of the EEG sensor nodes by significantly. Prior techniques could only reduce transmission energy, but our method can reduce acquisition energy as well and eliminates processing energy required for compression totally.

3. For the first time, to the best of our knowledge we exploit inter channel correlation to formulate a theoretically sound framework based on Kronecker Compressed Sensing (KCS) for recovering multiple signals simultaneously acquired via method developed in (2), since existing CS methods fail to do so.
4. Exploiting channel correlation we propose to model signal recovery as a low rank minimization problem for better reconstruction .
5. We also propose to model signal recovery as row sparse problem by exploiting channel correlation in transform domain.
6. We combine the models proposed above in (4) and (5) and present signal recovery as a low rank and row sparse formulation and achieve better reconstruction accuracy over all existing works.
7. Lastly, we modify the existing BCS (Blind Compressed Sensing) framework to exploit channel correlation to achieve improved reconstruction as opposed to prior CS recovery methods.

1.4 Organization of Thesis

This thesis is organized as follows.

Chapter 2 provides a brief understanding of Compressed Sensing and description of concepts for developing the algorithm to solve various proposed signal recovery formulation.

Chapter 3 presents the literary survey of the work done on CS based techniques for acquisition and reconstruction of EEG signals.

Chapter 4 proposes method to reduce sensing power and also proposes formulation for signal recovery at base station that exploits inter channel correlation.

Chapter 5 proposes algorithm developed on the concepts discussed in Chapter 2 to solve the proposed formulation in Chapter 4.

Chapter 6 applies the formulation proposed in Chapter 4 on EEG data to test the performance of the proposed work against existing methods. This chapter also comments on the power saving achieved by proposed work as compared to existing CS based methods.

Finally, Chapter 7 concludes this thesis by summarizing the contribution of the work and research done along with the possible extension of our work that can be explored.

Chapter 2

Precursors

This chapter briefly discusses the key theoretical concepts of Compressed Sensing and provides a basic know how of underlying concepts to understand our proposed algorithm. The two main pillars of our work, Alternating Direction Method of Multipliers (ADMM) algorithm and Split Bregman Algorithm are described in this chapter.

2.1 Background on Compressed Sensing

One of the central and important element in signal processing is the Nyquist or Shannon sampling theory, which states that number of samples needed to reconstruct a signal without error is governed by the smallest time interval or maximum frequency present in the signal [8]. Nearly all signal acquisition protocols are based on the fact that sampling rate must be atleast twice the maximum frequency component to recover the signal back. In the recent years, a new theory of compressed sampling is used which asserts that super resolved signals and images can be recovered with far fewer samples than required by Nyquist sampling [8]. Although, there is large body of literature around the theory of compressed sensing, this section only provides a brief overview and focusses on basics needed for our work.

2.1.1 Sparse Recovery

Compressed Sensing studies the problem of solving an under-determined system of linear equations, when the solution is known to be sparse:

$$y = A_{m \times n} x_{n \times 1} \tag{2.1}$$

In general, such an under-determined system will have an infinite number of solutions. However, when the solution is sparse, the situation is more tractable. We assume that the solution is k -sparse, i.e. it has only k non-zero elements. In such a scenario, the k -sparse vector has at most $2k$ (k positions and corresponding k values) unknowns. When the number of equations (m)

is larger than the number of independent variables ($2k$ in this case), intuition tells us that, we should be able to solve the problem in Eq. 2.1. This arises the question if the problem is solvable or not.

Even though an under-determined system has infinitely many solutions, if the solution is known to be sparse, it is necessarily unique [9]. But, in most cases Eq. 2.1 will have two solutions, both of which are k -sparse or one of which is k -sparse while the other is $k + 1$ -sparse. Thus, when it is known that the solution to Eq. 2.1 is sparse, one may as well search for the sparsest solution.

$$\min_x \|x\|_0 \tag{2.2}$$

subject to $y=Ax$

The l_0 -norm is not strictly a norm, but is a diversity measure - it counts the number of non-zeroes in the solution; therefore minimizing it would yield the sparsest solution. Unfortunately minimizing the l_0 -norm is an NP hard problem [9], and thus is not feasible for any practical large scale system.

There are two ways to address this. The first way is to approximately solve the NP hard problem using greedy algorithms like Orthogonal Matching Pursuit (OMP) [10]. However, they cannot solve ALL k -sparse solutions and have poor recovery guarantees. The other way to address the problem in Eq. 2.2 is to relax the NP hard l_0 -norm by its nearest convex surrogate - the l_1 -norm.

$$\min_x \|x\|_1 \tag{2.3}$$

subject to $y=Ax$

The l_1 -norm (convex) minimization problem can be easily solved via linear programming. It has been shown in [9] and in other works, that when certain conditions are met, the solution from Eq. 2.2 and Eq. 2.3 are the same; i.e. both of them are guaranteed to recover the sparsest solution. The l_1 -norm minimization is also dubbed as the Basis Pursuit (BP); BP can recover ALL k -sparse solutions.

However, nothing comes for free. The minimum number of equations required by l_0 -norm in order to recover the solution is $m=2k + 1$. However, the minimum number of equations required by l_1 -norm is:

$$m = Ck \log n \tag{2.4}$$

The required number of measurements have increased by a factor of $\log(n)$. So, the ease of solvability is achieved at a price of increased measurements (convex l_1 -norm as compared to NP hard l_0 -norm).

2.1.2 Noisy Compressible Signals

Most of the practical signals are approximately sparse or compressible in nature i.e. their coefficients show fast decay behaviour when sorted in decreasing order of magnitude. Also, the signals in most situations are corrupted by noise. This more realistic situation can be modelled as follows:

$$y = Ax + \eta \quad \eta \sim N(0, \sigma^2) \quad (2.5)$$

Here, η denotes noise with normalized distribution. Under such situation, the l_1 minimization problem is modified as below:

$$\begin{aligned} & \min_x \|x\|_1 & (2.6) \\ & \text{subject to } \|y - Ax\|_2 \leq \epsilon, \quad \epsilon = m\sigma^2 \end{aligned}$$

In theory, it has been shown that error between solution to Eq. 2.6, assume it to be x^* and actual solution x^0 is bounded by Eq. 2.7 [11] .

$$\|x^* - x^0\|_2^2 \leq C_1\epsilon + C_2 \frac{\|x^0 - x_s^0\|_1}{s} \quad (2.7)$$

Here, x_s^0 denotes the best s -term approximation of x^0 that indicates the s highest coefficient values in x^0 .

The expression in above equation implies that the error between the actual and reconstructed signal is bounded by the noise variance (ϵ) and the best s -approximation of the compressible signal. The first term, $C_1\epsilon$ arises owing to the noise in the system and the second term is due to the fact that the signal is not exactly sparse, but compressible.

2.1.3 Practical Signal Recovery

So far we have discussed about the signals that are sparse. Unfortunately, practical signals are never sparse in the physical domain. But they do have a sparse representation in a transform domain, e.g. medical images are sparse in wavelet domain, seismic waveforms are sparse in curvelets, speech is sparse in short time fourier transform etc. CS exploits the transform domain sparsity of these signals in order to recover them.

There are two variants for CS based reconstruction: Synthesis prior and Analysis prior. Transforms which are orthogonal or tight frame transform are good to go with any of the variants of CS. These transforms satisfy the following property:

$$\begin{aligned} & \text{Orthogonal} : W^T W = I = W W^T & (2.8) \\ & \text{Tight - frame} : W^T W = I \neq W W^T \end{aligned}$$

Here, W denotes the sparsifying basis. The analysis and the synthesis equations hold for both kinds of transforms, and is given by

$$\begin{aligned} \textit{Analysis} : \alpha &= Wx \\ \textit{Synthesis} : x &= W^T\alpha \end{aligned} \tag{2.9}$$

Here, α is the vector of sparse transform coefficients and W is the sparsifying basis. Now using the synthesis equation one can express Eq.2.5 as follows:

$$y = AW^T\alpha + \eta \tag{2.10}$$

Since, α is the sparse representation of the signal x , the solution to Eq. 2.10 can be obtained by solving l_1 minimization problem.

$$\begin{aligned} \min_{\alpha} \|\alpha\|_1 \\ \text{subject to } \|y - AW^T\alpha\|_2 \leq \epsilon \end{aligned} \tag{2.11}$$

The formulation in Eq. 2.11 is in synthesis prior form. This is the most popular technique for CS recovery. Unfortunately, the synthesis prior form is very restrictive. It can only be formulated for orthogonal and tight-frame transforms.

However, images are modelled as piecewise smooth functions and are sparse under finite differencing, biomedical signals like EEG and ECG are sparse in the Gabor basis, but these transforms (finite differencing, Gabor) are not orthogonal or tight-frame. The synthesis prior formulation precludes employing such powerful transforms. For all linear transforms, the analysis equation holds, i.e. even for Gabor / finite differencing one can express $\alpha = Wx$; but the synthesis equation does not hold for these transforms, i.e. $x \neq W^T\alpha$.

For such transforms, the analysis prior formulation needs to be used:

$$\begin{aligned} \min_x \|Wx\|_1 \\ \text{subject to } \|y - Ax\|_2 \leq \epsilon \end{aligned} \tag{2.12}$$

The analysis prior is more generalized than the synthesis prior and holds for all linear sparsifying transforms. The analysis prior does not enjoy widespread popularity like the synthesis prior; it is only recently that a concerted effort is being expended in understanding this form [12, 13]. It must be noted that the synthesis and the analysis forms are theoretically the same for orthogonal transforms, but are different for tight-frames. It has been empirically verified that the analysis prior with tight-frame transforms yields better results than orthogonal or tight-frame transforms using the synthesis prior [14–16].

2.1.4 Measurement Basis

One of the key component of signal compression in CS is measurement basis. CS relies on two principles: *sparsity* i.e. signal should have sparse representation in some sparsifying basis and *incoherence* which states that the measurement and sparsifying basis are incoherent i.e. the signal which has a sparse representation in a sparsifying basis must spread out in the measurement basis [17]. It has been found that random matrices are maximally incoherent with any fixed basis. The two most commonly used random basis are: i.i.d Gaussian and Binary [17].

Gaussian Matrices

This matrix have entries chosen independently from normal distribution with zero mean and $1/n$ variance. This is the most commonly used sensing or measurement basis. Being a dense matrix i.e. values present at every location, they have large storage and involves lot of computation when a signal is projected onto them.

Binary Matrices

The entries of a random binary matrix takes the value of 1 or 0 with equal probability. These matrices have lesser storage and computation as compared to Gaussian random matrices while performing equally well [18].

2.2 Alternating Direction Method of Multipliers

ADMM algorithm combines the decomposability of dual ascent method with superior convergence property of method of multipliers [19]. Dual ascent method provides decomposability to the problem provided some strict assumptions are considered on the minimization function. Dual ascent solves a large data problem iteratively but provides relatively slower convergence. So, in order to achieve a fast convergence, method of multipliers is used that relaxes the strict assumptions on the function. Method of multipliers is based on augmented Lagrangian method, that adds a additional quadratic penalty term to the Lagrangian of the function. The quadratic nature of penalty function wipes out the decomposability in the method. Considering the positives provided by the two methods, alternating direction Method of multipliers came forth.

ADMM is a variant of augmented Lagrangian method that uses partial updates for dual variables. Before we start with ADMM algorithm lets give a moment to method of multipliers to gain more insight for the situation. Consider a equality constraint minimization problem as given below

$$\min f(x) \tag{2.13}$$

$$\text{subject to } y = Ax$$

The augmented Lagrangian for Eq. 2.13 is given by

$$L_\rho(x, y) = f(x) + y^T(Ax - b) + (\rho/2)\|Ax - b\|_2^2 \quad (2.14)$$

where $\rho > 0$ is the penalty parameter. The method of multipliers is give in following steps:

$$\begin{aligned} x^{k+1} &= \underset{x}{\operatorname{argmin}} L_\rho(x, y^k) \\ y^{k+1} &= y^k + \rho(Ax^{k+1} - b) \end{aligned} \quad (2.15)$$

Method of multiplier provide better convergence under much more relaxed condition i.e. f can be non differentiable as compared to dual ascent method. But, method comes along with the loss of splitting due to quadratic term added in augmented Lagrangian step.

ADMM algorithm provides robustness to method of multipliers while retaining the decomposability of dual ascent method. ADMM algorithm solves the problem of the form given in the Eq. 2.16 with two set of variables x and z .

$$\text{minimize } f(x) + g(z) \quad (2.16)$$

$$\text{subject to } Ax + Bz = c$$

Here, original x is splitted into two parts x and z which provides two separable objective functions $f(x)$ and $g(z)$. ADMM iterates in following steps;

$$\begin{aligned} x^{k+1} &= \underset{x}{\operatorname{argmin}} L_\rho(x, z^k, y^k) \\ z^{k+1} &= \underset{z}{\operatorname{argmin}} L_\rho(x^{k+1}, z, y^k) \\ y^{k+1} &= y^k + \rho(Ax^{k+1} - b) \end{aligned} \quad (2.17)$$

The algorithm is very much similar to dual ascent and method of multiplier with two minimization steps(x and z) and a dual variable update. Its clear from equation that ADMM retains decomposability by alternatively solving for x and z in two separate steps as opposed to method of multiplier, that minimizes both variables simultaneously. This makes $f(x)$ and $g(z)$ separable. The only two assumptions considered by ADMM algorithm are:

1. f and g are closed, proper and convex functions. This implies that x and z update steps are solvable.
2. Excluding the penalty term, Lagrangian of the function posses saddle point.

In context to our work, ADMM in conjunction with Split Bregman is used for problem solving. The next section briefly explains about Split Bregman approach and its connection to ADMM.

2.3 Split Bregman Method

Split Bregman method is derived from Bregman Iterative Algorithms. Bregman iterative algorithms have been used to solve various problems such as image denoising and basis pursuit formulation because of the very nice convergence behaviour it shows. These properties are: monotonic decreases in residual term, convergence of recovered signal to original signal and convergence of bregman distance for noisy data [20]. All Bregman Iterative algorithms are based on the concept of bregman distance.

Bregman distance posses several nice properties, that makes it an effective tool for solving l_1 regularized problems. For a convex function $J : X \rightarrow R$ where $u, v \in X$ and $p \in \partial J$ bregman distance between the points u and v is give by

$$D_j^p(u, v) = J(u) - J(v) - \langle p, u - v \rangle \quad (2.18)$$

Here ∂J denotes sub gradient of J .

Bregman distance is a non-negative quantity but is not a conventional distance metric as it is not symmetric and does not follow triangular inequality. It is just a measure of separation of two points. Consider a constrained optimization problem given below:

$$\begin{aligned} \min_u E(u) & \quad (2.19) \\ \text{subject to } H(u) &= 0 \end{aligned}$$

Where, E and H are convex functions defined on R_n with H being differentiable. The corresponding unconstrained formulation can be written as

$$\min_u E(u) + \lambda H(u) \quad (2.20)$$

Eq. 2.20 is solved iteratively by Bregman iterative algorithm in the following steps:

$$u^{k+1} = \min_u D_E^p(u, u^k) + \lambda H(u) \quad (2.21)$$

$$p^{k+1} = p^k - \nabla H(u^{k+1})$$

Bregman provides excellent convergence results because H monotonically decreases with every iteration and converges to the minimizer of the function. An advancement over Bregman iterations is the linearized Bregman algorithm which makes computation efficient by combining the minimizing of unconstrained problem and Bregman update into a single step; which can be solved exactly. However, linearized Bregman cannot solve problems involving multiple l_1 regularization terms. To handle this issue Split Bregman method [21] was proposed. Consider the objective function given below where $\Phi(u)$ and $H(u)$ are convex and H is differentiable .

$$\min_u \|\Phi(u)\|_1 + H(u) \quad (2.22)$$

The underlying concept of Split Bregman is decomposition of l_1 and l_2 terms such that they form different sub problems which can be solved way easily than the original composite objective function. Rewriting Eq. 2.22 by letting $d=\Phi(u)$ the constrained formulation becomes

$$\begin{aligned} \min_u \quad & \|d\|_1 + H(u) \\ \text{subject to} \quad & d=\Phi(u) \end{aligned} \quad (2.23)$$

The unconstrained equivalent of Eq. 2.23 is obtained by adding a penalization function to the problem as in Eq. 2.24.

$$\min_u \|d\|_1 + H(u) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \quad (2.24)$$

Now comparing this to the form in Eq. 2.19, we consider

$$E(u, d) = \|d\|_1 + H(u)$$

Now, Eq. 2.24 is solved as

$$(u^{k+1}, d^{k+1}) = \min_u D_E^p(u, u^k, d, d^k) + \frac{\lambda}{2}\|d - \Phi(u)\|_2^2 \quad (2.25)$$

$$p_u^{k+1} = p_u^k + \lambda(\nabla\Phi)^T(\Phi u^{k+1} - d^{k+1})$$

$$p_d^{k+1} = p_d^k + \lambda(d^{k+1} - \Phi u^{k+1})$$

The 1st update step can be solved using ADMM (alternating direction method of multipliers) as discussed before by alternatively keeping one variable fixed and optimizing over the other [19]. Simplified form of Eq. 2.25 is given below.

$$u^{k+1} = \min_u H(u) + \frac{\lambda}{2}\|d - \Phi(u) - b^k\|_2^2 \quad (2.26)$$

$$d^{k+1} = \min_d \|d\|_1 + \frac{\lambda}{2}\|d - \Phi(u) - b^k\|_2^2$$

$$b^{k+1} = b^k + (\Phi u^{k+1} - d^{k+1})$$

Since $H(u)$ is smooth and differentiable everywhere, updation for u can be solved analytically. Solution for d is nothing but the solution for synthesis prior formulation and is obtained directly by shrinkage (soft thresholding) operator. Last step is to update the Bregman variable.

One important advantage that Split Bregman provides over other algorithms is that, one can keep lambda (the regularization parameter) constant to a value that achieves fast convergence.

Chapter 3

Literature Survey

A significant amount of work has been done in the field of EEG signal compression. Broadly speaking the work can be clubbed under two heads: lossless compression and lossy compression. One can recover a signal perfectly with lossless compression method, but this is achieved at an expense of greater computational complexity and lower compression ratios. These properties make lossless compression techniques unsuitable for WBAN's application. Whereas, on the other hand lossy compression techniques do not aim for perfect reconstruction and are good to consider if recovery error is within the tolerance limits. They are simpler and provide higher compression ratios as compared to lossless techniques making them apt for our situation. Since our proposed methods falls under lossy compression technique, previous works done with lossy techniques are presented in this section. Lossless compression is out of scope of our work. Most of the lossy compression techniques are developed on Compressed Sensing framework described in section 2.1.

Over the past few years, different CS based techniques for EEG signal compression and reconstruction have been proposed. In this chapter, a review of various techniques is presented. These techniques differ from each other in compression and reconstruction techniques. There is no straightforward way to organize these studies. Therefore, we discuss them in a chronological fashion.

3.1 Gaussian Compression Basis with Gabor Sparsifying Transform

One of the earliest studies in this area is [22], where the EEG signal is fully sampled and is compressed by projection onto an i.i.d Gaussian basis. The i.i.d Gaussian basis is an optimal compression basis for CS recovery [23]. Formally, the compression can be modelled as,

$$y = Gx + \eta \tag{3.1}$$

Here, G is the i.i.d Gaussian basis, y is the compressed signal and x is the EEG signal that has been compressed and needs to be reconstructed at the base station.

The aforementioned work [23] assumed that the EEG signal is sparse in Gabor basis (H); and posed the reconstruction (at the base station) as a synthesis prior problem.

$$\begin{aligned} \min_{\alpha} \|\alpha\|_1 & \tag{3.2} \\ \text{subject to } \|y - GH^T\alpha\|_2 & \leq \epsilon \end{aligned}$$

Once the Gabor coefficients are recovered, the signal is computed using the synthesis equation, i.e. $\hat{x}=H^T\alpha$.

However, this approach is not theoretically correct. The Gabor basis is neither orthogonal nor tight-frame (in general); thus the analysis equation holds but the synthesis equation does not. Hence, it is not correct to express Eq. 3.1 in the following form,

$$y \neq Gx + \eta \tag{3.3}$$

Hence, the synthesis prior form given in Eq. 3.2 is not theoretically correct.

This work proposed Gaussian basis for signal compression. Although mathematically optimal, this is practically infeasible. The Gaussian basis is dense; storing and operating on it is therefore not efficient.

The work in [22] suggests the possibility of jointly reconstructing the EEG signals from all the probes by exploiting the correlation information. Since, EEG signals are collected over multiple channels, these signals are correlated due to monitoring the same brain activity. However, it did not provide any concrete formulation for the same.

A recent study [24] posed reconstruction formulation that is same as in [22]. The only difference being that, prior to compression, the EEG signal is denoised using Independent Component Analysis that resulted in slightly improved reconstruction.

3.2 Non-Uniform Sampling and Reconstruction using Slepian Basis

The work in [25] is theoretical, and claimed nothing to compress the signal. Hence, this has no contribution to compression and transmission of signal. However, it provided a CS based approach for EEG reconstruction.

In this work, the EEG signal was assumed to be sparse in the Slepian Basis [26]. The Slepian basis or the prolate spheroidal wave functions (PSWF) are a set of functions derived by time limiting, low passing, and a second time limiting operation.

Let Q_T denote the time truncation operator, such that $x=Q_Tx$ iff x is time limited within $[-T/2, T/2]$. Similarly, let P_W denote an ideal low-pass filtering operator, such that $x=P_Wx$ iff x is band limited within $[-W, W]$. The operator $Q_T P_W Q_T$ is linear, bounded and self-adjoint. For $n = 1, 2$ we denote with ϕ_n the n_{th} eigenfunction, which is defined as:

$$Q_T P_W Q_T \phi_n = \lambda_n \phi_n \quad (3.4)$$

Here, λ_n denotes the respective associated eigenvalues. The time limited functions $\{\phi_n\}_n$ are the Prolate Spheroidal Wave Functions.

These functions are connected with the sinc functions as eigenfunctions of the integral operator:

$$\phi_n = \frac{1}{\lambda_n} \int_{-T/2}^{T/2} \phi_n(\tau) S(t - \tau) d\tau \quad (3.5)$$

where, $S(t)=sinc(t)$ and eigenvalues λ_n are ordered in decreasing order of magnitudes.

In this work, an unusual problem is studied. It stated that samples are not taken uniformly at $\{nT_s\}$, instead at random times around these values i.e. at $t_0 = 0$ and $t_n=cnT_s + \Delta$, with $c=Nn/M$ and Δ is a random variable uniformly distributed in $[-0.5T_s, 0.5T_s]$.

Using the M-orthogonal expansion, the non uniform samples can be written as:

$$\begin{bmatrix} x(t_0) \\ x(t_1) \\ \vdots \\ x(t_{M-1}) \end{bmatrix} = \begin{bmatrix} \phi_0(t_0) & \phi_1(t_0) & \dots & \phi_{M-1}(t_0) \\ \phi_0(t_1) & \phi_1(t_1) & \dots & \phi_{M-1}(t_1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_0(t_{M-1}) & \phi_1(t_{M-1}) & \dots & \phi_{M-1}(t_{M-1}) \end{bmatrix} \begin{bmatrix} \gamma(t_0) \\ \gamma(t_1) \\ \vdots \\ \gamma(t_{M-1}) \end{bmatrix} \quad (3.6)$$

In the matrix form,

$$x(t_i) = \phi(t_i)\gamma \quad (3.7)$$

Here, ϕ is the Slepian basis corresponding to the random sampling instants. The straightforward solution is obtained via Pseudo-inverse, i.e. coefficient estimation is done using following equation

$$\gamma = [\phi(t_i)]^\dagger x(t_i) \quad (3.8)$$

From which the signal can be reconstructed using Eq. 3.6 i.e.

$$\hat{x} = \Phi\lambda = \Phi[\phi(t_i)]^\dagger x(t_i) \quad (3.9)$$

where,

$$\begin{bmatrix} \phi_0(t_0) & \phi_1(t_0) & \dots & \phi_{M-1}(t_0) \\ \phi_0(t_1) & \phi_1(t_1) & \dots & \phi_{M-1}(t_1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_0(t_{M-1}) & \phi_1(t_{M-1}) & \dots & \phi_{M-1}(t_{M-1}) \end{bmatrix}$$

In this work, an alternate CS based formulation is also proposed as well. Instead of estimating the Slepian coefficients using the pseudo-inverse, they proposed solving it via CS, i.e.

$$\min_{\gamma} \|\gamma\|_1 \text{ subject to } x(t_i) = \Phi(t_i)\gamma \quad (3.10)$$

The reconstruction equation remains the same as Eq. 3.9. This work claimed that the CS reconstruction yields better results than the pseudo-inverse reconstruction.

This is a pedagogic work; at least it is not relevant for our topic of discussion. However, as it is remotely related to the topic we have discussed it in detail over here.

3.3 Comparative study of different Sparsifying Basis

In [27], a comparative study on a variety of sparsifying transforms is performed. The data compression model is the same as 3.1; we repeat it for the sake of convenience

$$y = Gx + \eta \quad (3.11)$$

In [22,24], it is assumed that the EEG signal is sparse in the Gabor basis. In [27], several other sparsifying basis are compared; cubic and linear splines, cubic and linear b-splines, Mexican hat and Gabor. This work suffered from the same problem as [22, 24]. Although, none of the aforementioned transforms are orthogonal or tight-framed, yet in [27] the reconstruction is posed as a synthesis prior problem. We will not discuss again why synthesis prior is theoretically erroneous; the argument for this remains the same as pointed out in section 3.1.

In the previous section, we explained Slepian basis functions, but it is not widely used in signal processing literature. Whereas, transforms used in this work [27] are widely used and does not require any introduction. It was empirically verified in [27] that the Gabor dictionary

yields the best reconstruction among all the sparsifying transforms tested. The i.i.d Gaussian compression basis was used in [27] as well.

3.4 Block Sparse Bayesian learning

In [28] the EEG signal is modelled to be block sparse in wavelet / DCT domain, i.e. the transform domain coefficients are assumed to be divided into certain blocks as defined below:

$$\alpha = \{\underbrace{\alpha_{1,1} \dots \alpha_{1,n}} , \dots , \underbrace{\alpha_{b,1} \dots \alpha_{b,n_b}}\}^T \quad (3.12)$$

Here $\alpha = Wx$ (analysis equation). The vector α is assumed to be divided into b blocks. Block sparsity means that only few blocks are non-zeroes, but in a non-zero block all the elements are non-zeroes.

So far, the previous studies assumed that the EEG signal is sparse in the transform domain (Gabor, spline, etc.). In [28] it is assumed the EEG signal is approximately block sparse in DCT domain. There are well established algorithms for recovering block-sparse signals. However, all these algorithms require the blocks to be specifically delineated. Unfortunately this is not the case for EEG signals; even though such signals are approximately block sparse, the block divisions are not known apriori. In such a situation, standard block sparse recovery algorithms fail.

To recover such block sparse signals where the block information is not available apriori, a Bayesian framework is proposed in [29]; this method is called the Block Sparse Bayesian Learning (BSBL) algorithm. The BSBL method is applied on EEG signals for recovering them from their compressed measurements. This method by far yields the best EEG reconstruction results. It should be kept in mind that the block sparse nature of the EEG signal (in transform domain) is an assumption; there is no physical explanation as to why it should be so.

The other novelty in [28] is the introduction of sparse binary random matrix [30] as compression matrix. Prior studies [23, 24] and [25] used i.i.d Gaussian matrices to compress the EEG signal onto a lower dimension. Binary random matrix contains a fixed number of ones at random position along each row, while the rest of the positions are zeroes. Storing and operating such a matrix is very efficient (linear complexity) and is practical for our problem.

3.5 Joint reconstruction

EEG signals are acquired over multiple probes. The possibility of reconstructing the signal using the multi-channel correlation information was mentioned in [22], but no concrete formulation

was proposed. In [31] a method to jointly reconstruct all the EEG signals from all the channels was proposed. However, this work did not actually exploit the inter-channel correlations.

Extending the compression model in Eq. 2.4 to the multi-channel case:

$$y_c = B_c x_c + \eta, c = 1 \dots C \quad (3.13)$$

where C is the total number of channels/probes and B denotes a sparse random binary matrix.

This work also assumed that a sparse binary random matrix is used for compression. In [32], it was assumed that the EEG signal is sparse in the wavelet basis. Since wavelet basis is orthogonal and the analysis-synthesis equations hold, the synthesis equation can be invoked into Eq. 3.13:

$$y_c = B_c W^T \alpha_c + \eta, c = 1 \dots C \quad (3.14)$$

In matrix vector form, this can be concisely expressed as:

$$y = \Theta \alpha + \eta \quad (3.15)$$

where

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_c \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_c \end{bmatrix} \text{ and } \Theta = \begin{bmatrix} B_1 W^T & \dots & 0 \\ \vdots & & \\ 0 & \dots & B_c W^T \end{bmatrix}$$

Since, each of the α_i s are sparse, the concatenated α is sparse as well. Therefore, it can be reconstructed using l_1 -minimization.

$$\min_{\alpha} \|\alpha\|_1 \text{ subject to } \|y - \Theta \alpha\|_2^2 \leq \epsilon \quad (3.16)$$

This method [32] jointly recovered the transform coefficients of the full EEG signal ensemble, but did not really account for the correlations among the different channels.

3.6 Analysis Prior

In [23, 25] and [28] the Gabor basis is used as the sparsifying basis for reconstructing EEG signals; albeit incorrectly as a synthesis prior recovery. The later works [29, 31], used DCT and wavelets, thus their formulation as a synthesis prior recovery stands justified. However, using DCT or wavelets is not an optimal solution.

EEG signal reconstruction is the first step of information processing pipeline. After reconstruction, the signal is analyzed either manually or automatically. Automatic EEG signal analysis is almost always carried forth by using the Gabor basis. Thus, preserving the high valued Gabor coefficients is of paramount importance for EEG signal analysis.

When wavelets or DCT is used for recovery, CS will only guarantee the preservation of high valued wavelet or DCT coefficients; it cannot guarantee the preservation of Gabor coefficients. To ensure that the high valued Gabor coefficients are preserved, an analysis prior recovery was proposed in [32].

$$\min_x \|Hx\|_1 \text{ subject to } \|Y - Bx\|_2^2 \leq \epsilon \quad (3.17)$$

This is a channel by channel reconstruction technique where B denotes the sparse binary random matrix and H is Gabor transform matrix.

3.7 Other Studies

In this section, we briefly review those studies which do not address the problem at hand but are remotely related to the topic of CS in EEG signal reconstruction.

In [33], the problem is to localize the neuronal activity in the brain. At time k , a small patch of the activated cortex can be represented by an equivalent current dipole with location ($3D$) p_k and momentum ($3D$) $m_k = s_k q_k$ where q_k is the $3D$ orientation vector and s_k is the current amplitude of the dipole. Using this model, the EEG signal acquired by C probes is given by

$$z_k = A_k s_k + \eta \quad (3.18)$$

where $z_k = [z_k^{(1)} \dots z_k^{(C)}]^T$ is the EEG data sampled by the C probes, $s_k = [s_k^{(1)} \dots s_k^{(N_d)}]^T$ is the amplitude of the N_d dipoles at the k^{th} instant and A_k is the $C \times N_d$ gain matrix. The j^{th} column of this matrix is given by $a_k^{(j)} = F(p_k^{(j)}) q_k^{(j)}$, where the lead field $F(p)$ is represented by a $C \times 3$ matrix and is a nonlinear function of the dipole location p .

When represented in the $3D$ Cartesian grid Eq.3.18 is expressed as follows:

$$z_k = \Psi \theta_k + \eta_k \quad (3.19)$$

where $\Psi = [F(r_1) \dots F(r_G)]$ is a $C \times 3G$ matrix, $F(r_i)$ is the $C \times 3$ lead field of the i^{th} grid located at r_i , $\theta_k = [m_k^1 \dots m_k^G]^T$ with $m_k^i = [m_k^{i,x}, m_k^{i,y}, m_k^{i,z}]$ is the moment of a dipole concentrated on the i^{th} grid at time k . Since there are N_d dipoles, only N_d grids have non-zero moments. As a result, at most $3 \times N_d$ elements in θ_k are non-zero, where $N_d \ll G$. Thus, the EEG signal z_k is only $3 \times N_d$ -sparse in the spatial domain.

It is possible to compress the measurements z_k by projecting them with a suitable CS matrix:

$$y_k = Gz_k = G\Psi\theta_k + \eta_k \quad (3.20)$$

Using CS theory, it is possible to recover θ_k and correspondingly z_k . However, the focus of the paper [34] was to track neuronal activity from compressive measurements. For this purpose, they employ a particle filter in the compressed domain to track the evolution of θ_k with time.

Our study is based on the assumption that the EEG signal needs to be efficiently compressed and then transmitted to the base station for reconstruction and further analysis. This is the most general framework for telemonitoring. In [35] it is proposed that certain tasks like automated seizure detection can be performed in the compressed domain itself; there is no need to reconstruct the signal. They claimed that the main advantage of carrying out analysis in the compressed domain is that, there is no need for the computationally expensive reconstruction step; and the analysis can be performed at the nodes itself. They proposed a hardware architecture to detect seizures from EEG signals in the compressed domain. Their proposed technique is more energy efficient than seizure detection techniques in the non-compressed domain.

Chapter 4

Proposed Formulation

Our contribution in this work can be divided into two parts. Firstly, as discussed in the introduction section, sensing process in a WBAN is the second major energy hungry component after transmission. A lot of energy saving can be achieved if this can be brought down in some way. Here, a method is proposed that stands a step ahead of compression provided by existing lower dimension projection in CS that only reduced the transmission power. The proposed method compresses the signal right away during capturing, thus eliminating any further processing as required in CS based compression. So, the proposed technique reduces both transmission as well as sensing power. Secondly, in order to recover the original signal from compressed measurement at base station, several recovery methods are proposed. These methods are developed on the grounds of exploiting channel correlation inherent in EEG data. These proposed recovery formulations over perform the existing CS methods with noticeable margin. This chapter is dedicated to provide an understanding of proposed formulation.

4.1 Sensing Power Reduction: Sub-Sampling

All prior studies aimed at reducing only the transmission costs; they sampled the full signal and compress it by projecting onto a lower dimension by Gaussian / Binary matrices. A model for this has been proposed in the previous studies [24]. A schematic diagram of Compressed Sensing based EEG unit is shown in Fig.4.1.

In these studies, the full signal is sampled (after amplification). After that the signal is compressed by projecting onto a lower dimension; this requires a DSP chip for multiplication which consumes considerable amount of power. This compressed EEG signal is then finally transmitted. Such a compression strategy reduces transmission power but does not curtail the sensing power.

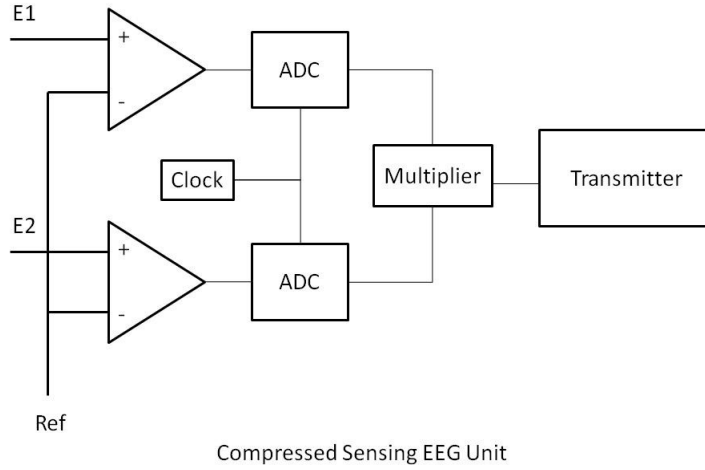


Figure 4.1: Data Acquisition and Transmission for Previous method(CS based)

In this work, we aim to reduce sensing energy as well. The only way to do so is by collecting fewer samples, i.e. by sub-sampling. However, periodic under-sampling is not a viable option for obvious reasons. The other opportunity possible is random under-sampling. So, for the proposed method, data acquisition model is given by:

$$d = R_c x + \eta \tag{4.1}$$

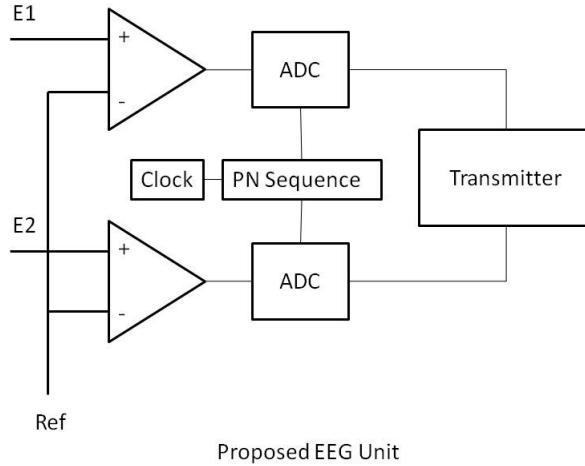


Figure 4.2: Data Acquisition and Transmission for Proposed Method

The schematic for the proposed work is shown in Fig.4.2, here PN sequence generator controls the under-sampling. The clock dictates the sampling rate of the ADC. The proposed modification requires a pseudo-random sequence (PN sequence) generator between clock and ADC (Analog to Digital Converter). The PN sequence determines if the clock samples at a particular instant or not. Effectively this scheme under-samples the signal randomly. In this scenario, the acquired

signal is inherently compressed. Thus, it eliminates energy component needed in processing (for compression) - therefore there is no need of DSP chip. We can express the multi-channel data acquisition problem as follows:

$$y_i = R_i x_i + \eta \quad \forall i \quad (4.2)$$

which can be further arranged as:

$$vec(Y) = R vec(X) + \eta \quad (4.3)$$

where, $Y=[y_1|y_2|\dots|y_c]$, $X=[x_1|x_2|\dots|x_c]$ and

$$R = \begin{bmatrix} R_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & R_2 & \dots & \dots & \dots & 0 \\ 0 & 0 & R_3 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & R_c \end{bmatrix}$$

Here, vec operator has general connotation and c denote the number of channels.

Thus, with under sampling we are able to compress the signal along with reduction in sensing and transmission power as well. The power saving achieved by incorporating this method is discussed later in the work.

4.2 Kronecker Compressed Sensing Formulation

Compressed Sensing assumes that the signal is measured in a basis - in this case the measurement basis is the Dirac; for previous studies the sparse random binary basis was employed. We mentioned before that EEG signals are not sparse but have a sparse representation in a transform domain like wavelet [31] or DCT [28]. Unfortunately, if the signal is measured in the dirac basis, sparsifying transforms like DCT and wavelet are not incoherent with it (Dirac). CS requires that the measurement basis and the sparsifying basis should be maximally incoherent from each other [36]. If the coherence is high (low incoherence), the recovery error is more for a given number of measurements.

Owing to such an incoherence requirement, previous studies were unable to propose a technique that could guarantee reconstruction from under-sampled EEG signals. The Fourier basis is the only viable alternative in this situation since it is maximally incoherent with the Dirac basis. In fact the incoherence between the Fourier and the Dirac basis is theoretically the maximum possible; between any other measurement and sparsifying basis (like Gaussian / Binary measurement basis and DCT / wavelet sparsifying basis), the incoherence is always less than

the Fourier-Dirac pair [36]. Therefore, Fourier basis will yield the best recovery results owing to maximal incoherence with the Dirac sparsifying basis.

A piecemeal reconstruction of individual channels using the sparsity basis would not yield good reconstruction accuracy. Since prior work only looked at reconstructing single EEG channels, they failed to address the problem where the EEG signals were directly under-sampled.

Also, Fourier transform does not sparsify the EEG signals very well; this is evident from Fig.4.4 and Fig 4.5. We can see that, the Fourier coefficients of an EEG signal do not have a very fast decay when transform is applied individually to each channel. This is because, applying Fourier transform to each of the channels separately does not exploit the inter-channel correlations. So, in order to enhance sparsity, we propose to exploit the inter-channel correlation.

Now, what is meant by inter channel correlation is that, since EEG signal is monitoring the same underlying activity, at a particular time instant or a sample the data on all channels are related to each other either by a behaviour of attenuation or phase difference. So, at a particular sampling instant (along the rows of X), owing to the inter-channel correlation, the vector formed by the samples collected at every channel is smooth; i.e. the rows of X have a smooth variation. A smooth signal has compact representation in Fourier domain; therefore a Fourier transform along the row-vectors of X will sparsify it.

So, we propose to sparsify X both along the column (for each channel) and row (across the channels) directions. The process is schematically represented in Fig.4.3

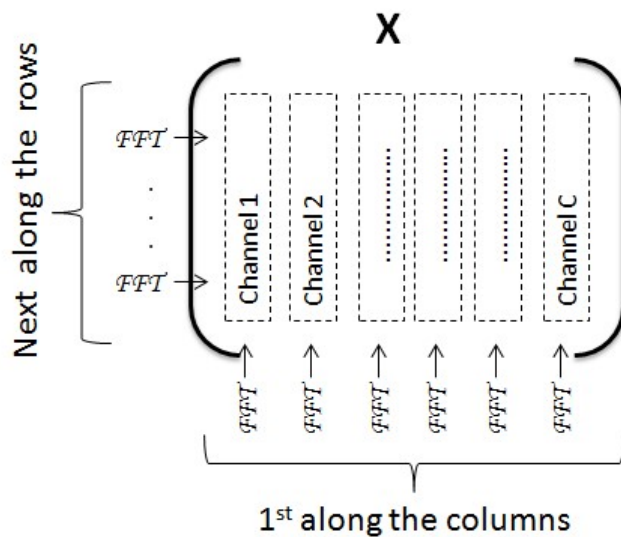


Figure 4.3: Schematic Diagram of Sparsifying operation

Mathematically, the schematic in Fig.4.4, is represented as $\alpha = F_1 X F_2 = F_2^T \otimes F_1 \text{vec}(X)$ making α sparse. Both F_1 and F_2 are 1D Fourier transforms but operate on vectors of different length; this operation is basically the 2D Fourier transform on X . Thus for us, the sparsifying transform is the 2D Fourier transform. The improvement in sparsity can be verified from Fig. 4.4 and 4.5 . It is clear that, the Fourier coefficients after the 2D Fourier transform have a very fast decay as compared to the 1D Fourier transform.

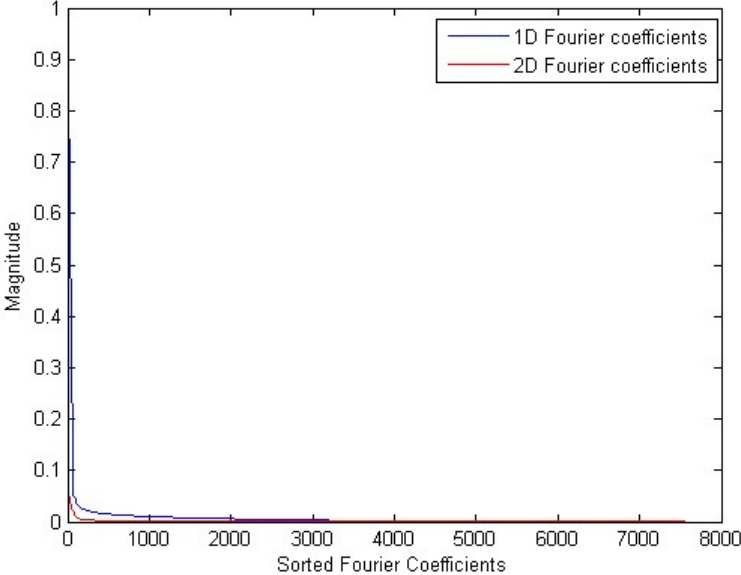


Figure 4.4: Decay of Fourier Coefficients for EEG signal ensemble

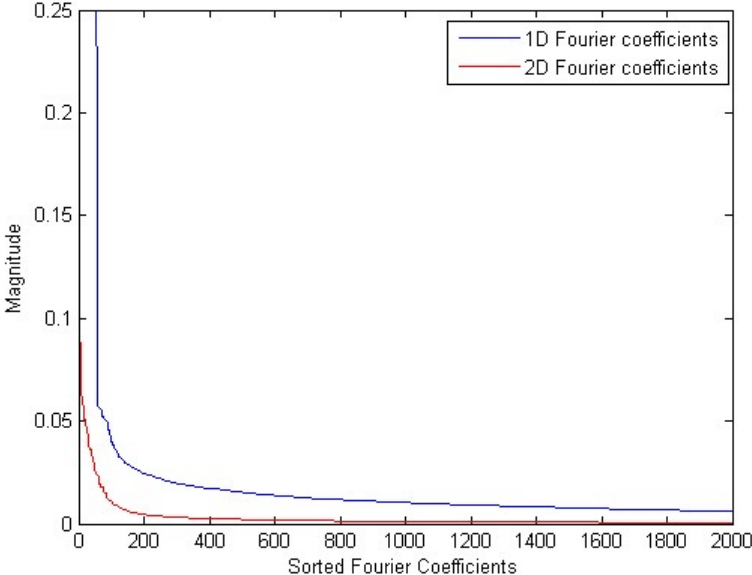


Figure 4.5: Decay of Fourier Coefficients for EEG signal ensemble(magnified)

We could have just used the 1D Fourier transform along the channel for sparsifying the EEG signals. However, as discussed in chapter 3, that for a fixed number of measurements, the sparser the signal, the better is the reconstruction (lower reconstruction error). Due to this reason, we seek maximal possible sparsity.

Following this proposition, we can incorporate the sparsifying operations into the data acquisition model Eq. 4.3 by using the Kronecker product notation as follows:

$$vec(Y) = R(F_2^T \otimes F_1)^T \alpha + \eta \quad (4.4)$$

One can recover the sparse transform coefficients α using l_1 -norm minimization as given in Eq. 3.2 as follows:

$$\begin{aligned} \min_{\alpha} \|\alpha\|_1 & \quad (4.5) \\ \text{subject to } \|y - R(F_2^T \otimes F_1)^T \alpha\|_2 & \leq \epsilon \end{aligned}$$

Thereafter, EEG signals are reconstructed back from recovered Fourier transform coefficients by using the synthesis equation.

4.3 Low Rank Recovery Formulation

Prior studies proposed recovery techniques to reconstruct each EEG signal channel separately. However, in other domains - magnetic resonance imaging [37, 38], hyper-spectral imaging [39], it has been seen that multi channel signal recovery exploiting channel correlation improves reconstruction accuracy. We expect similar improvements for multi-channel EEG signal recovery. Also, the multi channel signal compression can be formulated as

$$y_c = Bx_c + \eta_c, c = 1 \dots C \quad (4.6)$$

Here, C denotes the total number of channels. In a more compact form this can be represented as

$$Y = BX + \eta \quad (4.7)$$

Here, $Y = [y_1 | y_2 | \dots | y_c]$, $X = [x_1 | x_2 | \dots | x_c]$ denotes the column concatenated data of channels.

So, the problem is to recover X from its available measurements Y . Eq.4.7 is an under-determined inverse problem with infinitely many solutions. In order to find a plausible solution, some prior knowledge regarding X is necessary. So, in this section we propose methods that exploit multichannel correlation to recover X .

EEG signals result from brain activity. Different channels monitor the same activity from different locations. Since, the source for all the signals is the same underlying activity, it is

likely that the EEG signals from the different channels are correlated. If this assumption is true, the matrix X as described in Eq. 4.7 is low-rank; therefore one can use low-rank matrix recovery techniques to recover the multi-channel EEG signal ensemble.

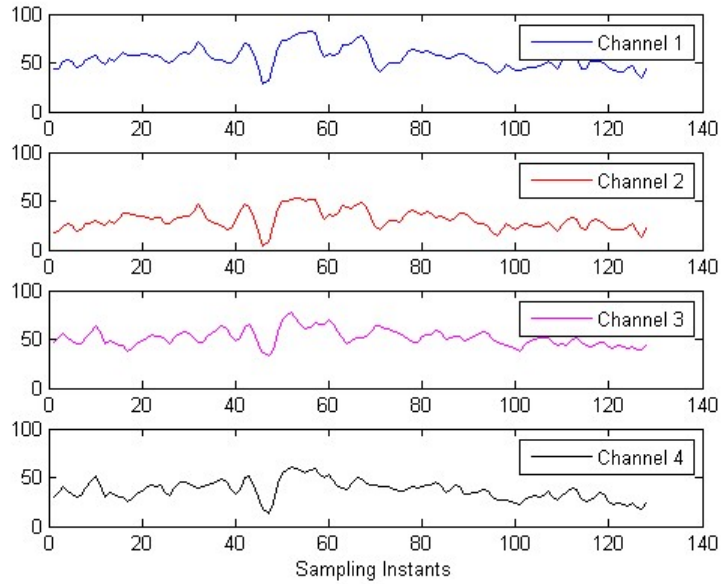


Figure 4.6: Similarity of EEG signals from 4 channels

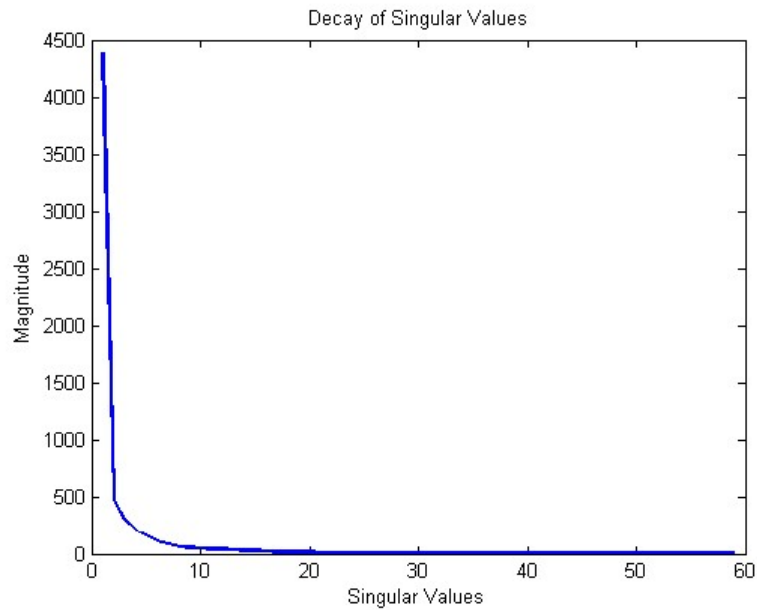


Figure 4.7: Decay of Singular Values for EEG signal ensemble

Our assumption is empirically verified by the plot of EEG signals on different channels and singular value plot respectively in Fig.4.6 and Fig. 4.7.

Fig.4.6 shows EEG signal present on 4 different channels. The similarity among the data over different channels is very clearly visible in the figure.

This implies that the matrix formed by such signals (stacked as columns) will be a low-rank matrix. This fact is further validated by the plot of singular values of the full signal ensemble in Fig.4.7.

The figure clearly indicates that the the singular values decay rapidly; ensuring the approximately low rank nature of the EEG signal ensemble.

Therefore, we can expect to recover X by exploiting its low-rank property. As discussed in the chapter before, the low rank can be approximated by nuclear norm. Thus, reconstruction can be achieved by solving the following nuclear norm regularized least squares minimization problem:

$$\min_X \|Y - BX\|_F^2 + \lambda \|X\|_* \quad (4.8)$$

Here, " F " denotes the Frobenius norm and " $*$ " denotes the nuclear norm. λ is regularization parameter that maintains a trade off between the data constituency and low rank nature of the data.

4.4 Row Sparse MMV Recovery Formulation

As mentioned before, the EEG signals are very similar to each other. The similarity in the time domain leads to a similar sparse signature in the transform domain, i.e. if we take wavelet or Gabor transform of the all the signals from different channels they will have high values at similar positions. This can be verified from Fig.4.8.

Since the EEG signals from different channels are similar, as shown in Fig. 4.6 their wavelet coefficients also have the same sparsity pattern as in Fig.4.8, i.e. the wavelet coefficients are of high values at the same locations. Here, we have used wavelets as an example, the same behaviour holds true for other sparsity promoting transforms. This means that the matrix ΨX will be row-sparse, where Ψ is the sparsifying basis, since all the wavelet transform coefficients have a common sparse support. So, one can solve the row-sparse recovery problem for signal reconstruction.

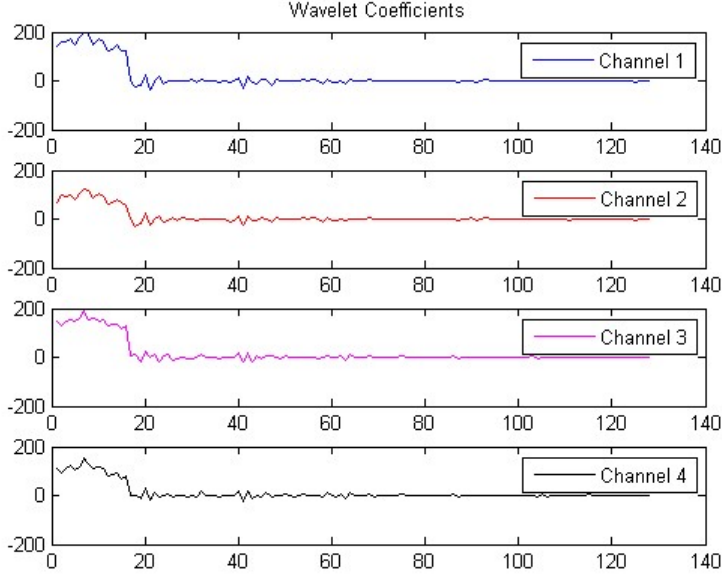


Figure 4.8: Common Sparse Support of Wavelet Coefficients

The two formulations - synthesis prior and analysis prior for the same are given below.

$$\textit{Synthesis Prior} : \min_Z \|Y - B\Psi^T Z\|_F^2 + \lambda \|Z\|_{2,1} \quad (4.9)$$

$$\textit{Analysis Prior} : \min_X \|Y - BX\|_F^2 + \lambda \|\Psi X\|_{2,1} \quad (4.10)$$

The l_{21} norm term promotes sparsity. In order to obtain more insight on this one can refer [40]. Here, λ denotes the regularization parameter and plays the same role as in Eq.4.8.

In this work, we have also derived algorithm to solve the above formulation, and is present in the next chapter.

4.5 Combined Low Rank and Row Sparse Recovery Formulation

In the above two sub sections, we proposed two ways to exploit the inter channel correlation - by formulating the problem as low rank (Eq 4.8) and row-sparse (Eq.4.9and 4.10). So, here arises a question: if these two can be combined to yield better results?. On the basis of prior studies in MRI [41] and hyper-spectral imaging [39], the answer is in affirmative. So here, we propose to combine rank-deficiency with transform domain sparsity in order to improve the recovery

results. Thus, the problem can be formulated as:

$$\textit{Synthesis Prior} : \min_Z \|Y - B\Psi^T Z\|_F^2 + \lambda_1 \|Z\|_{2,1} + \lambda_2 \|Z\|_* \quad (4.11)$$

$$\textit{Analysis Prior} : \min_X \|Y - BX\|_F^2 + \lambda_1 \|\Psi X\|_{2,1} + \lambda_2 \|X\|_* \quad (4.12)$$

4.6 Row Sparse BCS Formulation

We have already discussed row sparse recovery in previous section. In this section, we modify the basic BCS framework to recover row-sparse signal ensembles. This modified framework provides better EEG reconstruction as compared to prior CS recovery methods.

Before we move onto our proposed formulation, a brief overview of BCS is done here. CS assumes sparsity of the signal in a known basis. Blind Compressed Sensing (BCS) [42] framework combines element from both Compressed Sensing and Dictionary Learning (DL). It estimates both the sparsifying dictionary as well as sparse signal coefficients in reconstruction process.

It assumes that the data is sparse in a learned dictionary, i.e. $X=DZ$, where D is the unknown dictionary (to be estimated) and Z is the set of sparse coefficients. Using this formulation, an inverse problem of the form $Y=AX + N$, can be expressed as: $Y=ADZ + N$. In BCS, both D and Z need to be estimated. This is achieved by solving:

$$\min_{D,Z} \|Y - ADZ\|_F^2 + \lambda_1 \|D\|_F^2 + \lambda_2 \|vec(Z)\|_1 \quad (4.13)$$

In [43] BCS was successfully used for recovering dynamic MRI sequences. One must note the subtle different between BCS and other dictionary learning methods. Usually, in dictionary learning, the learning phase is offline i.e. the dictionary is first estimated and then the learned dictionary is employed for signal reconstruction. In BCS, both dictionary learning and signal estimation proceed simultaneously.

But here, we want to exploit inter-channel correlation within the BCS framework for improving signal reconstruction. Similar to above mentioned techniques, there is no algorithm to solve it, we propose a Split Bregman approach to solve the said problem in next section.

We can directly apply the BCS technique in Eq. 4.13 for recovering EEG signals. But BCS only assumes that the signal is sparse; it does not account for row-sparsity. We have already seen similarity in EEG signal over channels in Fig.4.6 pointing channel correlation.

It is already argued before that all the signals are correlated, so they have a common sparse support in the transform domain. Therefore, if we stack the sparse transform coefficients of the multi-channel signals as columns of a matrix, the resulting matrix will have sparse rows. This leads to a row-sparse recovery problem of the following form:

$$\min_{D,Z} \|Y - ADZ\|_F^2 + \lambda_1 \|D\|_F^2 + \lambda_2 \|Z\|_{2,1} \quad (4.14)$$

The original signal can then be recovered, once the sparse coefficients are obtained from above equation.

Chapter 5

Proposed Algorithm

In previous chapter, reconstruction formulation based on exploiting channel correlation are proposed. As stated before, there is no as such algorithm that can solve these optimization problems. In this chapter, we aim to develop an algorithm based on the concepts discussed in the chapter 2 to solve them.

Starting with the first formulation of low rank recovery, there are several algorithms for matrix completion, but for general low-rank matrix recovery there are very few. Also, of all the existing algorithms, we found that the best general purpose algorithm for both matrix recovery and matrix completion is the singular value shrinkage (SVS) [44].

As pointed in the chapter before, the proposed row-sparse recovery has both - synthesis prior and analysis prior formulations. The best algorithm for synthesis prior formulation is perhaps the Spectral Projected Gradient [45]. However, general purpose algorithms for solving both the analysis and the synthesis prior exists [40]; in this work we use these algorithms to solve the proposed formulation after incorporating the concepts discussed in the chapter 2.

In [39] an algorithm is proposed to solve the combined low-rank and row-sparse recovery problem. However, this algorithm is not very efficient. Besides, it can only solve the synthesis prior problem Eq. 4.11. So, there is a need to derive a new efficient algorithm to solve Eq. 4.11 and 4.12. Rewriting these equation in simpler form:

$$\min_X \|Y - BX\|_F^2 + \lambda_1 \|X\|_{2,1} + \lambda_2 \|X\|_* \quad (5.1)$$

$$\min_X \|Y - BX\|_F^2 + \lambda_1 \|\Psi X\|_{2,1} + \lambda_2 \|X\|_* \quad (5.2)$$

In order to solve these above equations, Split Bregman Approach for variable splitting along with ADMM [46] described in Chapter 2 are used. Firstly, algorithm to solve Eq. 5.1 is derived.

The derivation for algorithm to solve Eq.5.2 is similar and is done later in the chapter.

Starting with the derivation for Eq. 5.1, two proxy variables - P and Q for the two penalty functions are introduced respectively. Terms relaxing the equality constraints of each quantity and its proxy are added. Also, in order to enforce equality at convergence, Bregman relaxation variables B1 and B2 are introduced. Thus, the objective function in Eq.5.1 can be re-framed in the form of following equation:

$$\min_{X,P,Q} \|Y - BX\|_F^2 + \lambda_1 \|P\|_{2,1} + \lambda_2 \|Q\|_* + \mu_1 \|P - X - B_1\|_F^2 + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.3)$$

This allows problem in Eq. 5.3 to be splitted into an alternating minimization of the following (easier) sub problems:

$$\min_X \|Y - BX\|_F^2 + \mu_1 \|P - X - B_1\|_F^2 + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.4)$$

$$\min_P \lambda_1 \|P\|_{2,1} + \mu_1 \|P - X - B_1\|_F^2 \quad (5.5)$$

$$\min_Q \lambda_2 \|Q\|_* + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.6)$$

The problem in Eq.5.4 is easy to solve, its a simple least least square problem which can be solved efficiently using any conjugate gradient algorithm by recasting it as:

$$X = \min_X \left\| \begin{pmatrix} Y \\ P - B_1 \\ Q - B_2 \end{pmatrix} - \begin{pmatrix} A \\ \sqrt{\mu_1} I \\ \sqrt{\mu_2} I \end{pmatrix} X \right\|_F^2 \quad (5.7)$$

The sub problem in Eq. 5.5 is a synthesis prior $l_{2,1}$ norm minimization problem. The technique to solve this is available in [40]:

$$P \leftarrow \text{signum}(X + B_1) \max(0, |X + B_1| - \frac{\lambda_1}{\mu_1} \Lambda) \quad (5.8)$$

where, $\Lambda = \text{diag}(\|(X + B_1)^{j \rightarrow}\|_2^{-1}) \text{abs}(X + B_1)$; $(X)^{j \rightarrow}$ denotes the j^{th} row of X.

The sub problem Eq. 5.6 is a nuclear norm minimization problem. The algorithm to solve this was derived in [44] and is given below:

$$USV^T = X + B_1 \quad (5.9)$$

$$\Sigma \leftarrow \text{Soft}(S, \frac{\lambda_2}{\mu_2})$$

$$Q = U \Sigma V^T$$

Soft-thresholding is applied on the singular values of the matrix $X + B_2$; Q is updated by recomposing the matrix using the singular vectors and the thresholded singular values.

The final step in the algorithm is to update the Bregman variables at every iteration:

$$\begin{aligned} B_1 &\leftarrow P - X - B_1 \\ B_2 &\leftarrow Q - X - B_2 \end{aligned}$$

This concludes the derivation for solving Eq. 5.1. Solving Eq. 5.2 is very similar. Applying the Split Bregman technique to this leads to the following formulation:

$$\min_{X,P,Q} \|Y - BX\|_F^2 + \lambda_1 \|P\|_{2,1} + \lambda_2 \|Q\|_* + \mu_1 \|P - \Psi X - B_1\|_F^2 + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.10)$$

This too can be segregated into 3 sub-problems in a similar way as follows:

$$\min_X \|Y - BX\|_F^2 + \mu_1 \|P - \Psi X - B_1\|_F^2 + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.11)$$

$$\min_P \lambda_1 \|P\|_{2,1} + \mu_1 \|P - \Psi X - B_1\|_F^2 \quad (5.12)$$

$$\min_Q \lambda_2 \|Q\|_* + \mu_2 \|Q - X - B_2\|_F^2 \quad (5.13)$$

The solution to Eq.5.11 is easy, since it is a least squares minimization problem. However, the formulation is slightly different from Eq. 5.7.

$$X = \min_X \left\| \begin{pmatrix} Y & A \\ P - B_1 & \sqrt{\mu_1} \Psi \\ Q - B_2 & \sqrt{\mu_2} I \end{pmatrix} \right\|_F^2 \quad (5.14)$$

The sub-problem in Eq.5.12 essentially remains the same, with only a minor variation. The solution to it is as follows:

$$P \leftarrow \text{signum}(\Psi X + B_1) \max(0, |\Psi X + B_1| - \frac{\lambda_1}{\mu_1} \Lambda) \quad (5.15)$$

The sub-problem in Eq.5.13 remains exactly same as in Eq.5.6. The final step to update the Bregman variables is given below:

$$\begin{aligned} B_1 &\leftarrow P - \Psi X - B_1 \\ B_2 &\leftarrow Q - X - B_2 \end{aligned}$$

By introducing the variable splitting, analysis prior problem is converted to a synthesis prior sub-problem, which is easier to solve via iterative soft thresholding.

The proposed split Bregman algorithm requires specifying two internal parameters μ_1 and μ_2 . For our problem we have fixed these values at $\mu_1 = 10^{-3}$ and $\mu_2 = 10^{-4}$. The Bregman variables B_1 and B_2 are initialized to all one's.

Chapter 6

Experimental Setup, Evaluation and Results

6.1 Dataset

The experiments are carried on BCI Competition III data set 1 [47]. EEG is collected over 59 channel and is divided into 2865 epochs. Also, the data on each channel is represented by 128 samples.

6.2 Evaluation Parameter

In order to assess the performance of a method some kind of evaluation criteria is needed. Here, in order to inspect the performance of our proposed algorithms and compare it against existing techniques Normalized Mean Square Error(NMSE) is used. This metric is defined as

$$NMSE = \frac{\|original - reconstructed\|_2}{\|original\|_2}$$

NMSE and its close counterpart SNR are standard metrics used for quantitative reconstruction performance. This is a well accepted measure for comparing reconstruction accuracy in CS recovery problems.

6.3 Results

The proposed work can be broadly classified in two parts: first part is the introduction of sub sampling at the sensor node in WBAN and second is the development of reconstruction algorithm for signal recovery at the base station. So, in order to establish the effectiveness of the work experimental results are present in this section.

6.3.1 Reconstruction Accuracy

Reconstruction accuracy for various existing techniques and proposed methods are given in the following Table 6.1. We present results with both compression where measurement basis is either Gaussian or Binary and with sub sampling i.e. Dirac basis. For all the experiments 50% compression is used. The main fallouts from the table are listed below:

1. Kronecker Compressed Sensing discussed in section 4.2 yields better reconstruction than the best previously known CS techniques that used either DCT or Wavelet as sparsifying basis. KCS and existing CS techniques with either DCT or Wavelet as sparsifying basis are based on l_1 norm formulation.

It is clear from the NMSE values (highlighted in gray color) that the proposed method gives better reconstruction even with partially sampled data as opposed to full data ensemble in exiting CS methods. This better reconstruction also proves the points that using Fourier Kronecker product as sparsifying basis gives sparser representation as opposed to exiting Wavelet or DCT basis when used in conjunction with Dirac basis and hence better reconstruction.

Table 6.1: Comparative Reconstruction Results (NMSE)for various Algorithms

Type	Formulation	Method	NMSE	
			Compression	Sub-Sampling
Existing CS Techniques	Synthesis	l_1 min.with Gabor basis	0.1629	0.3217
		BSBL	0.0229	0.1120
		Joint Reconstruction with Wavelet Basis	0.1667	0.4412
		Joint Reconstruction with DCT Basis	0.1763	0.3967
	Analysis	l_1 min. with Gabor basis	0.0574	0.0741
Proposed Technique		Low Rank	0.0637	0.0911
Proposed Technique	Synthesis	KCS	0.0731	0.0760
		Low Rank and Joint Sparse(Wavelet)	0.1380	0.3638
		Row Sparse BCS	0.0122	0.0379
Proposed Technique	Analysis	Joint Sparse(Gabor)	0.0228	0.0531
		Low Rank and joint sparse (Gabor)	0.0412	0.0375

2. We formulated signal recovery as low rank and row sparse in both analysis and synthesis prior formulation in section 4.5. But, we prefer analysis prior formulation in WBAN's so as to cut down any power needed in processing at the time of compression/under sampling. From the table, results point out that our proposed algorithms both low rank and row sparse formulations outperform the existing CS based methods both for compression as well as for sub- sampled data. Also, modelling signal recovery problem to exploit both low

rank row and sparse nature together using Gabor as sparsifying basis yields even better reconstruction (highlighted in green color).

Similar to the conclusion made in point (1), using Wavelet as sparsifying basis for low rank and joint sparse formulation (highlighted in cyan color) does not perform as good as Gabor due to the incoherence condition with the measurement basis

3. Lastly, we present reconstruction accuracy for row sparse formulation of existing BCS framework (highlighted in pink color) discussed in section 4.6. As can be seen from the table, the results are superior when compared against any other technique. The reason for this lies in the fact that the sparsifying basis best suited to represent signal sparsely is learned in the reconstruction process itself.
4. From the table, it can be observed that the analysis prior formulation yields better results than the synthesis prior formulation; this is not surprising as similar observations have been made in other branches of applied compressed sensing [14–16].

6.3.2 Power Analysis

We have argued in principle how our proposed technique reduces the energy consumption in sensing component of signal and eliminates any energy required for processing component by completely bypassing it. In this section, we analyse the overall energy saving by following a power model proposed in [48]. At first, we discuss power consumption in existing CS based EEG unit. The total power, as explained before is composed of three major components:

$$P_{tot} = P_{sense} + P_{proc} + P_{comm} \quad (6.1)$$

The sensing power is expended under two heads : Amplification (P_{amp}) and Analog-to-digital conversion (P_{ADC}) (see Fig.4.1); therefore one can write $P_{sense}=C(P_{amp} + P_{ADC})$, where C is the total number of channels.

Also, the processing consist of two operations : random number generation (P_{RNG}) and matrix-vector multiplication (P_{mult}); thus, $P_{proc}=P_{RNG}+P_{mult}$. It is difficult to uniquely model the power requirement for communication (transmission) since it is dependent on the communication protocol. However, in general the communication power is expressed as $P_{comm}=CJf_sR$ where J is the transmission power per bit, f_s is the sampling frequency (bits per second) of the ADC and R is the number of bits per sample (resolution).

Both sensing and communication power increases linearly with the number of channels, but processing power does not scale with the number of channels. CS based techniques can only reduce the transmission power by compressing a n dimensional signal to m samples via matrix vector multiplications. So, one can write the total power consumption of a CS based EEG unit

as :

$$P_{tot} = C(P_{amp} + P_{ADC}) + P_{RNG} + P_{mult} + C \frac{m}{n} J f_s R \quad (6.2)$$

Following the same arguments, the proposed unit only consumes (see 4.2)

$$P_{tot} = C(P_{amp} + \frac{m}{n} P_{ADC}) + C \frac{m}{n} J f_s R \quad (6.3)$$

The amplifier being an analog device cannot be switched on and off. Also, we do not need the matrix-vector product for compression; so the corresponding power terms vanish since we compress right when we sample. The transmission power remains the same as that of a CS based unit.

Establishing hard values for power consumption gives a rough idea of the actual power cost and how much we save. For ADC, we get the values from [49] - for a 12 bits per sample (R), 0.5 kSample/sec sampling rate (f_s) as $P_{ADC} = 0.2 \mu W$. The specifications of the power amplifier (for EEG signals) are obtained from [50], 30 dB gain and 30 Hz bandwidth require $P_{amp} = 0.9 \mu W$. The power consumption for the random number generator is $3 \mu W$ [51]. In [48] the power cost of matrix vector multiplication is pegged at $352 \mu W$ assuming that the DSP chip is of TI MSP430 family. The transmission energy required per bit is estimated to be $5 nJ$ per bit [48].

With these figures, we compute the energy requirement for a CS based system using Eq. 6.2.

$$64 \times (0.2 + 0.9) + 3 + 352 + 64 \times 0.2 \times 0.5 \times 5 \times 12 = 809 \mu W$$

Here, it is assumed that the compression ratio is 5:1 (0.2). Now, considering the proposed methodology, the power consumption (for same compression ratio) is give below and is computed using Eq.6.3

$$64 \times (0.2 \times 0.2 + 0.9) + 64 \times 0.2 \times 0.5 \times 5 \times 12 = 444 \mu W$$

This is about half the power consumption required by CS techniques. In other words, with the proposed methodology, the WBAN is expected to last almost twice longer duration. If the sampling ratio is increased to 5:2 (0.4), the power requirement for the CS based method comes out to be 1.19 mW and 0.83 mW for the proposed technique.

Chapter 7

Conclusion and Future Scope

7.1 Conclusion

In this work, we have tackled two important problems that stands in the way of energy efficient transmission of EEG signals on WBAN's. Firstly, to handle the energy constraint due to limited battery life of devices, we proposed a technique to reduce the sensing power along with transmission power. This is achieved by sub sampling the EEG signal in time domain at ADC itself wherein existing CS based methods compress the signal after ADC i.e. after the samples are digitized. Also due to sub sampling in time domain existing CS recovery methods don't fit effectively in the reconstruction scenario. This is because the incoherency constraint of sub sampling basis (Dirac basis in our case) with sparsifying basis is not met as required for CS recovery algorithms.

Secondly, for the very first time, we formulated reconstruction problem such that it exploits inter-channel correlation to recover signal from its sub sampled data received at the base station. We formulated our recovery problem as low rank, joint sparse, combined low rank and joint sparse and achieved better reconstruction than existing CS methods. The existing CS formulation is also modified by incorporating the Kronecker product to exploit channel correlation as discussed in Chapter 4. We also extended the existing BCS framework used in MRI for EEG signals by reformulating it as a row sparse formulation and achieved good reconstruction accuracy.

7.2 Future Work

The work presented in this thesis can have many possible many possible extensions. Here, we point out some of the possible directions where it can be headed. It would be interesting, if the code could be made more efficient by further reducing the computational complexity in the proposed formulation. This will permit the operation of code even on low end devices.

One possible extension of our work can be in classification of EEG data for disease diagnosis. Data classification is used to study sleep stages in order to detect any sleep apnea in a person or to find out if brain is functioning properly or not. This could be achieved by modifying our proposed formulation by incorporating metric learning for data mismatch. Incorporating metric learning will improve the classification and would improve diagnostic accuracy.

Also, we can apply these algorithms to reconstruct other medical signals such as MRI. Other than biomedical signals, we can find application in other areas. One such area is multi channel signal separation wherein one aims to recover the desired data from noisy observations. Multi channel signal separation arises in situations of multi sensor probing systems for underwater acoustics, diagnostics and geological surveys. Similar to EEG signals, these signals also measure the same underlying activity and hence possess correlation in channels.

Other possible application can be in restoration of multichannel color images, which consist of set of image planes acquired by a imaging system that captures the same scene using multiple sensors. This kind of multichannel imaging includes multi spectral satellite images, color images, multiple time frames.

It would be interesting to build hardware to implement our proposed scheme in order to test its performance and achieved reduction in energy consumption in real life working condition.

Finally, it would be a good step if the proposed formulation can detect certain diseases provided some additional health signals data is available.

Bibliography

- [1] Available at http://www.lhspzz.org/gemd_01_img0045.jpg.
- [2] H. Cao, V. Leung, C. Chow, and H. Chan, "Enabling technologies for wireless body area networks: A survey and outlook," *Communications Magazine, IEEE*, vol. 47, no. 12, pp. 84–93, 2009.
- [3] Available at http://www.mdpi.com/sensors/sensors-11-05561/article_deploy/html/images/sensors-11-05561f4-1024.png.
- [4] V. Bajaj and R. B. Pachori, "Classification of seizure and nonseizure eeg signals using empirical mode decomposition," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 6, pp. 1135–1142, 2012.
- [5] A. Temko, C. Nadeu, W. Marnane, G. B. Boylan, and G. Lightbody, "Eeg signal description with spectral-envelope-based speech recognition features for detection of neonatal seizures," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 6, pp. 839–847, 2011.
- [6] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis," *Affective Computing, IEEE Transactions on*, vol. 1, no. 2, pp. 81–97, 2010.
- [7] C. A. Frantzidis, C. Bratsas, M. A. Klados, E. Konstantinidis, C. D. Lithari, A. B. Vivas, C. L. Papadelis, E. Kaldoudi, C. Pappas, and P. D. Bamidis, "On the classification of emotional biosignals evoked while viewing affective pictures: an integrated data-mining-based approach for healthcare applications," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 309–318, 2010.
- [8] E. J. Candès *et al.*, "Compressive sampling," in *Proceedings of the international congress of mathematicians*, vol. 3, pp. 1433–1452, Madrid, Spain, 2006.
- [9] D. L. Donoho, "For most large underdetermined systems of equations, the minimal l1-norm near-solution approximates the sparsest near-solution," tech. rep., Comm. Pure Appl. Math, 2004.

- [10] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [11] E. J. Candes, J. K. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [12] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, “The cosparsity analysis model and algorithms,” *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 30–56, 2013.
- [13] M. Elad, P. Milanfar, and R. Rubinstein, “Analysis versus synthesis in signal priors,” *Inverse problems*, vol. 23, no. 3, p. 947, 2007.
- [14] I. W. Selesnick and M. A. Figueiredo, “Signal restoration with overcomplete wavelet transforms: comparison of analysis and synthesis priors,” in *SPIE Optical Engineering+ Applications*, pp. 74460D–74460D, International Society for Optics and Photonics, 2009.
- [15] A. Majumdar and R. K. Ward, “Under-determined non-cartesian mr reconstruction with non-convex sparsity promoting analysis prior,” in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pp. 513–520, Springer, 2010.
- [16] A. Majumdar and R. K. Ward, “On the choice of compressed sensing priors and sparsifying transforms for mr image reconstruction: An experimental study,” *Signal Processing: Image Communication*, vol. 27, no. 9, pp. 1035–1048, 2012.
- [17] E. J. Candès and M. B. Wakin, “An introduction to compressive sampling,” *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.
- [18] G. Zhang, S. Jiao, X. Xu, and L. Wang, “Compressed sensing and reconstruction with bernoulli matrices,” in *Information and Automation (ICIA), 2010 IEEE International Conference on*, pp. 455–460, IEEE, 2010.
- [19] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [20] H. Nien and J. A. Fessler, “A convergence proof of the split bregman method for regularized least-squares problems,” *arXiv preprint arXiv:1402.4371*, 2014.
- [21] T. Goldstein and S. Osher, “The split bregman method for l1-regularized problems,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.
- [22] S. Aviyente, “Compressed sensing framework for eeg compression,” in *Statistical Signal Processing, 2007. SSP’07. IEEE/SP 14th Workshop on*, pp. 181–184, IEEE, 2007.

- [23] E. J. Candes and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [24] S. Fauvel, A. Agarwal, and R. Ward, “Compressed sensing and energy-aware independent component analysis for compression of eeg signals,” in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pp. 973–977, IEEE, 2013.
- [25] S. Senay, L. F. Chaparro, M. Sun, and R. J. Scwabassi, “Compressive sensing and random filtering of eeg signals using slepian basis,” *Proceedings of the EURASIP EUSIPCO*, vol. 8, 2008.
- [26] D. Slepian and H. O. Pollak, “Prolate spheroidal wave functions, fourier analysis and uncertainty,” *Bell System Technical Journal*, vol. 40, no. 1, pp. 43–63, 1961.
- [27] A. M. Abdulghani, A. J. Casson, and E. Rodriguez-Villegas, “Quantifying the performance of compressive sensing on scalp eeg signals,” in *Applied Sciences in Biomedical and Communication Technologies (ISABEL), 2010 3rd International Symposium on*, pp. 1–5, IEEE, 2010.
- [28] Z. Zhang, T.-P. Jung, S. Makeig, and B. D. Rao, “Compressed sensing of eeg for wireless telemonitoring with low energy consumption and inexpensive hardware,” *Biomedical Engineering, IEEE Transactions on*, vol. 60, no. 1, pp. 221–224, 2013.
- [29] Z. Zhang and B. D. Rao, “Recovery of block sparse signals using the framework of block sparse bayesian learning,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3345–3348, IEEE, 2012.
- [30] A. Gilbert and P. Indyk, “Sparse recovery using sparse matrices,” Institute of Electrical and Electronics Engineers, 2010.
- [31] M. Shoaran, M. Hosseini Kamal, P. Vandergheynst, A. Schmid, and Y. Leblebici, “Compressive multichannel cortical signal recording,” in *38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, no. EPFL-CONF-183014, 2013.
- [32] M. Mohsina and A. Majumdar, “Gabor based analysis prior formulation for eeg signal reconstruction,” *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 951–955, 2013.
- [33] L. Miao, J. J. Zhang, A. Papandreou-Suppappola, and C. Chakrabarti, “Neural activity tracking using spatial compressive particle filtering,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3461–3464, IEEE, 2012.
- [34] A. Majumdar and R. K. Ward, “Calibration-less multi-coil mr image reconstruction,” *Magnetic resonance imaging*, vol. 30, no. 7, pp. 1032–1045, 2012.

- [35] M. Shoaib, N. K. Jha, and N. Verma, “A compressed-domain processor for seizure detection to simultaneously reduce computation and communication energy,” in *Custom Integrated Circuits Conference (CICC), 2012 IEEE*, pp. 1–4, IEEE, 2012.
- [36] E. Candes and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse problems*, vol. 23, no. 3, p. 969, 2007.
- [37] A. Majumdar and R. K. Ward, “Accelerating multi-echo t2 weighted mr imaging: Analysis prior group-sparse optimization,” *Journal of Magnetic Resonance*, vol. 210, no. 1, pp. 90–97, 2011.
- [38] R. Otazo, D. Kim, L. Axel, and D. K. Sodickson, “Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion mri,” *Magnetic Resonance in Medicine*, vol. 64, no. 3, pp. 767–776, 2010.
- [39] M. Golbabaee and P. Vandergheynst, “Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 2741–2744, IEEE, 2012.
- [40] A. Majumdar and R. K. Ward, “Synthesis and analysis prior algorithms for joint-sparse recovery,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3421–3424, IEEE, 2012.
- [41] A. Majumdar and R. Ward, “Rank awareness in group-sparse recovery of multi-echo mr images,” *Sensors*, vol. 13, no. 3, pp. 3902–3921, 2013.
- [42] S. Gleichman and Y. C. Eldar, “Multichannel blind compressed sensing,” in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2010 IEEE*, pp. 129–132, IEEE, 2010.
- [43] S. G. Lingala and M. Jacob, “Blind compressive sensing dynamic mri,” 2013.
- [44] A. Majumdar and R. K. Ward, “Some empirical advances in matrix completion,” *Signal Processing*, vol. 91, no. 5, pp. 1334–1338, 2011.
- [45] E. Van Den Berg and M. P. Friedlander, “Theoretical and empirical results for recovery from multiple measurements,” *Information Theory, IEEE Transactions on*, vol. 56, no. 5, pp. 2516–2527, 2010.
- [46] B. Wohlberg, R. Chartrand, and J. Theiler, “Local principal component pursuit for non-linear datasets,” in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3925–3928, IEEE, 2012.
- [47] Available at <http://www.bbci.de/competition/iii/>.

- [48] A. M. Abdulghani, A. J. Casson, and E. Rodriguez-Villegas, “Quantifying the feasibility of compressive sensing in portable electroencephalography systems,” in *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, pp. 319–328, Springer, 2009.
- [49] N. Verma and A. P. Chandrakasan, “An ultra low energy 12-bit rate-resolution scalable sar adc for wireless sensor nodes,” *Solid-State Circuits, IEEE Journal of*, vol. 42, no. 6, pp. 1196–1205, 2007.
- [50] R. R. Harrison and C. Charles, “A low-power low-noise cmos amplifier for neural recording applications,” *Solid-State Circuits, IEEE Journal of*, vol. 38, no. 6, pp. 958–965, 2003.
- [51] J. Holleman, B. Otis, S. Bridges, A. Mitros, and C. Diorio, “A $2.92\mu\text{w}$ hardware random number generator,” in *Solid-State Circuits Conference, 2006. ESSCIRC 2006. Proceedings of the 32nd European*, pp. 134–137, IEEE, 2006.