Efficient Design of Embedded Data Acquisition Systems based on Smart Sampling

A Thesis

Submitted for the Degree

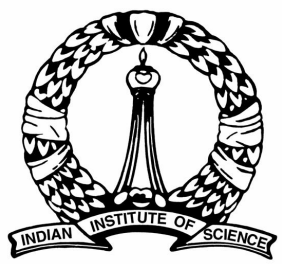
of

Doctor of Philosophy

in the Faculty of Engineering

by

J V Satyanarayana



Department of Electrical Engineering
Indian Institute of Science
Bangalore - 5600012

DECEMBER, 2014

.....The understanding of the enormity of the universe triggers the realization of the insignificance of my existence. In the vast expanse of the creation around me, my individual achievements, my aspirations, my knowledge, my ego and my assets bear no consequence. Yet, I must not give up my endeavor, in making an infinitesimal contribution to knowledge in this world. A world where nothing really changes, only our understanding of it grows in meaningful increments which, at a microscopic level, are built out of such infinitesimal contributions.....

Abstract

Data acquisition from multiple analog channels is one of the important functions in many embedded devices used in avionics, medical electronics, consumer appliances, automotive and industrial control, robotics and space applications. It is desirable to engineer these systems with the objectives of compactness, less power consumption, lower heat dissipation and reduced cost. The goal of this research is to suggest designs that exploit a priori knowledge of the input signals in order to achieve these objectives. In particular, sparsity is a commonly observed property in signals that offers opportunity to perform sub-Nyquist sampling, thereby reducing the number of analogue-to-digital conversions. Compressed sensing provides a mechanism for sub-sampling and reconstruction of sparse signals.

In this research, new architectures are proposed for the real-time, compressed acquisition of streaming signals, in which sampling is performed on a collection of signals in a multiplexed fashion. It is demonstrated that by doing so, it is possible to efficiently utilize all the available sampling cycles of the analogue-to-digital converters (ADCs), facilitating the simultaneous acquisition of multiple signals using fewer ADCs. It is shown how the proposed architecture can be realized using commonly available electronic components. Simulations on signals having Fourier sparsity exhibit that a set of signals is fairly well reconstructed even when the signals are sampled at sub-Nyquist rates by lesser number of ADCs. The proposed method is modified to accommodate more general signals in the case of which spectral leakage, due to occurrence of non-integral number of cycles in the reconstruction window, violates the sparsity assumption. Results of simulation demonstrate that when the primary objective of an application is to only detect the constituent frequencies in the signals, as against exact reconstruction, it can be achieved surprisingly well even in the presence of severe noise (SNR of the order of 5 dB) and considerable undersampling. This has been applied to the detection of the carrier frequency that varies randomly around a central frequency in a noisy FM (frequency-modulated) signal.

Information redundancy, on account of inter-signal correlation, gives scope

for compressed acquisition of a set of signals that may not be individually sparse. In this work, a scheme is proposed in which the correlation structure in a set of signals is progressively learnt within a small fraction of the duration of acquisition, because of which only a few ADCs prove to be adequate for capturing the signals. This also has important practical implications in smart acquisition of electro-encephalogram (EEG). Signals from the different channels of EEG possess significant correlation. Employing signals taken from the *Physionet* database, the correlation structure of nearby EEG electrodes was captured. Subsequent to this training phase, the acquired knowledge has been used on test signals taken from the same database. Results show that the spectral characteristics of signals at all the electrodes are detected with reasonably good accuracy. An average error below 10% has been achieved between the original and reconstructed signals with respect to the estimation of the relative power in various EEG spectral bands: delta, theta, alpha and below 15% in the beta band. It was also possible to demonstrate that the relative spectral power of the channels in the 10-10 system of electrode placement can be estimated, with an average error less than 8% (below 3% in delta band) using recordings on the sparser 10-20 system.

Reduction in the number of ADCs undoubtedly reduces the volume of electronics in embedded designs. It is also possible to downsize other components, for example, the anti-aliasing filter, if as many number of ADCs as the number of signals is used. This thesis proposes a design, wherein a set of signals are collectively sampled on a finer sampling grid using ADCs that are driven by phase-shifted clocks. In this manner, each signal is sampled at an effective rate that is a multiple of the actual rate at which the ADCs operate. Consequently, it is possible to have a transition between the pass band and the stop band that is not too steep, thereby reducing the order of the anti-aliasing filter from 30 to 8 as demonstrated by simulation results. The usefulness of this scheme has been demonstrated in the acquisition of voltages proportional to the deflection of the control surfaces in an aerospace vehicle.

The idle sampling cycles of an ADC that performs compressive sub-sampling of a sparse signal, can be used to acquire the residue left after a coarse low-resolution sample is taken in the preceding cycle, like in a pipelined ADC. Using a general purpose, low resolution ADC, a DAC of the same resolution and a summer, one can acquire a sparse signal with double the resolution of the ADC, without having to use a dedicated pipelined ADC. Results of the work done as part of this research show that the signal-to-quantization ratio

(SQNR) in the reconstructed signal is doubled using such a scheme. It has also been demonstrated how this idea can be applied to achieve higher dynamic range in the acquisition of electro-cardiogram (ECG) signals.

Finally, it is possible to combine more than one of the proposed schemes, to handle acquisition of diverse signals with different kinds of sparsity. The implementation of the proposed schemes in such an integrated design can share common hardware components so as to achieve a compact design.

Acknowledgements

To begin with I would like to profusely thank my research supervisor, Prof. A.G.Ramakrishnan who has literally taught me how to swim in the sea of doctoral research. At every step he has given me guidance on where I should tread next and yet made me feel that it was me who led the journey. He has shown immense patience when I faltered and deftly steered my rocking boat in the right direction. He pulled me out of disappointment umpteen times and gently regulated my over-confidence after success. I acknowledge not only the technical guidance that I have received from him but also advice on many aspects of life, in general.

I sincerely thank Dr. D.R.Jahagirdar, Scientist, Research Center Imarat, DRDO, Hyderabad who has kindly agreed to be my co-supervisor and helped me a lot in providing the right directions during the initial phase of my research. I express my heart-felt gratitude to Prof. P.S. Sastry, Prof. K.R. Ramakrishnan and Prof. T.V.Sreenivas for the wonderful courses they have taught me and to Prof. K. Rajgopal, Dr. Venu Madhav Govindu and Dr. S. Chandrasekhar for the valuable tips they have given me time and again. I sincerely thank Shri B.H.V.S. Narayanamurthy, Director and Shri G. Venkat Reddy, Associate Director of Directorate of Real Time Embedded Computers, RCI for giving me support and encouragement at every stage of my research.

I am highly grateful to all the staff and students of MILE laboratory at Department of Electrical Engineering, IISc who have helped me in every possible way and always treated me as a close friend. I acknowledge the full cooperation extended to me by the staff in the Electrical Engineering office.

I am deeply indebted to my father who, since my childhood, has instilled in me the belief that I can achieve anything with a strong will and dedicated effort and to my mother who has prayed for me innumerable times. No words can express the support given to me by my wife Sandhya, who, on many occasions, when I thought I had reached a dead end, did not allow me to give up. I fall short of words when I attempt to thank the contribution of my daughter Shreya who has let me borrow from the time we spend together, to carry out my research.

Contents

List of Figures **xix**

List of Tables **xxv**

1 Introduction **1**

1.1 Requirements of Embedded Systems 2
1.2 Data Converters in System-On-Chips 4
1.3 Information Redundancy in Signals 5
1.4 Scope for Improvised Embedded Designs 5
1.5 Sampling and Reconstruction of Sparse Signals 7
1.5.1 Sparse Reconstruction Schemes 8
1.6 Outline of Thesis 14

2 Compressed Sensing **17**

2.1 Introduction 17
2.2 Sensing the Signal 21
2.3 Reconstructing the Signal 22
2.4 Stability of Reconstruction 25
2.4.1 Restricted Isometry Property 25
2.4.2 Mutual Incoherence 26
2.4.3 Choosing the Right Measurement Matrix 28
2.5 Robust Compressed Sensing 30
2.6 Greedy Reconstruction Algorithms 31
2.6.1 Orthogonal Matching Pursuit 31

2.6.2 Other Greedy Algorithms 33
2.7 Other Recovery Algorithms 36
2.8 The Compressed Sensing 'Tuple' 37
2.9 Choice of Reconstruction Algorithm 38
2.10 Areas of Research in Compressed Sensing 38
2.11 Goal of this research 40

3 Compressed Acquisition of Multiple Sparse Signals 43
3.1 Introduction 43
3.2 Survey of related literature 44
3.3 Signal Model 45
3.4 Sampling and Reconstruction of the Signal 46
3.4.1 Simple Measurement Matrix 48
3.4.2 The CS Tuple 49
3.4.3 Overlapped Reconstruction Segments 49
3.5 MOSAICS: Multiplexed Optimal Signal Acquisition Involving Compressed Sensing 51
3.5.1 System Input 52
3.5.2 Steps in Execution 53
3.5.3 Derived Parameters 55
3.5.4 Effective Sampling Rate and Utilization Factor . . 58
3.5.5 Simulations and Results 59
3.5.5.1 Test Set 1 59
3.5.5.2 Performance of MOSAICS with other reconstruction algorithms 61
3.5.5.3 Test Set 2 66
3.5.6 Hardware Architecture for Realization of MOSAICS 67
3.5.7 Concluding Remarks on MOSAICS 69
3.5.8 Limitation of MOSAICS 70
3.6 Multiplexed Signal Acquisition for General Sparse Signals . 72
3.6.1 Introduction 72

3.6.2 Signal Reconstruction Based on MUSIC 74
3.6.2.1 Initialization 74
3.6.2.2 Intermediate Signal Estimate 74
3.6.2.3 The *MUSIC* algorithm, $\mathbb{M}(\mathbf{x}, K)$ 75
3.6.2.4 Estimation of DTFT Coefficients 75
3.6.2.5 Estimate the Signal 76
3.6.3 Modification in MOSAICS 76
3.6.4 Simulation and Results 77
3.6.5 Concluding Remarks on MOSAICS with MUSIC . . 78
3.7 The Frequency Detection Problem 79
3.7.1 Introduction 79
3.7.2 Simulation and Results 79
3.7.3 Detection of FM carrier frequency 81
3.8 Conclusion 83

4 Compressed Acquisition of Correlated Signals 87
4.1 Introduction 87
4.2 Joint Sparsity 89
4.2.1 JSM 1: Sparse Component + Innovations 89
4.2.2 JSM 2:Common Sparse Component 90
4.2.3 JSM 3: Non-sparse Common Component + Sparse Innovations 91
4.2.4 JSM 4: Non-sparse Common Component + Non-Sparse Innovations 92
4.3 Sparsity Inducing Basis 92
4.3.1 The CS-tuple 94
4.4 ARCS:Acquisition and Reconstruction of Correlated Signals 94
4.4.1 Objective of ARCS 95
4.4.2 Input 95
4.4.3 Initialization 95
4.4.4 Signal Acquisition and Learning 95

4.5 Simulation and Results 97
4.5.1 Test Signals 97
4.5.2 Simulation 98
4.5.3 Robustness of ARCS 99
4.6 Conclusion 100
4.7 Compressed EEG Acquisition using Estimated Channel Correlation 102
4.7.1 Electroencephalography 102
4.7.2 Standards for EEG electrode placement 103
4.7.3 Background work on compressed sensing of EEG signals 104
4.7.4 Inter-channel correlation in EEG 106
4.7.5 Compressed Sensing of EEG Signals 107
4.7.6 Limitations 109
4.7.7 A note on sparse PCA 109
4.7.8 The Physionet database 110
4.7.9 The experiments 112
4.7.9.1 The Training Phase 112
4.7.9.2 Testing phase 113
4.7.9.3 Results 114
4.7.9.4 Comparison of performance using other reconstruction algorithms 114
4.7.10 Reconstruction of signals at 10-10 locations from measurements at 10-20 locations 116
4.7.11 Conclusion 116

5 Other Compact Embedded Designs for Sparse Signals 125
5.1 Design with Low-Order Anti-Aliasing Filters 125
5.1.1 The Filtering Problem 127
5.1.2 Compressed Acquisition and Reconstruction 127
5.1.2.1 The CS Tuple 129

5.1.2.2 Phase Shifted Sampling 129
5.1.3 Simulation and Results 131
5.1.4 Performance of the proposed scheme with increase in the number of signals 133
5.1.5 Application to Real World Signals 133
5.1.6 Conclusion 135
5.2 Design for Better Resolution 135
5.2.1 Simulation and Results 139
5.2.2 Application to Fetal ECG Acquisition 141
5.2.3 Conclusion 144

6 Conclusion of Thesis **145**
6.1 Integration of proposed methods 148
6.2 Summary of findings in the thesis 150

Appendices **153**

A Embedded hardware design **153**

B Algorithm ARCS **154**

C Justification for using the downsized identity measurement matrix **157**
C.1 Empirical comparison of down-sized identity matrix and gaussian matrix 159
C.2 EEG reconstruction with Gaussian measurement matrix . 161

Bibliography **163**

List of Publications

1. J.V.Satyanarayana and A. G. Ramakrishnan,"MOSAICS: Multiplexed Optimal Signal Acquisition Involving Compressed Sensing". Proc. of the International Conference on Signal Processing and Communications, IISc Bangalore, July 2010.

2. J.V.Satyanarayana and A. G. Ramakrishnan, "Frequency Detection from Multiplexed Compressed Sensing of Noisy Signals". Proc. of the Seventeenth National Conference on Communications, pages 1-5, IISc Bangalore,Jan. 2011.

3. J.V.Satyanarayana and A. G. Ramakrishnan, "Multiplexed Compressed Sensing for General Frequency Sparse Signals". Proc. of the International Conference on Communications and Signal Processing ICCSP, pages 423-427, NIT Kozhikode, Feb. 2011.

4. J.V.Satyanarayana and A. G. Ramakrishnan, "Compressed Acquisition of Correlated Signals". Proc. of the Eighteenth National Conference on Communications,51, IIT Kharagpur, Feb. 2012.

5. J.V.Satyanarayana and A. G. Ramakrishnan,"Low Order Anti-aliasing Filters for Sparse Signals in Embedded Applications", Sadhana: Academy Proceedings in Engineering Sciences, *Springer*, vol. 38, no. 3, pp. 397-405, June 2013.

List of Figures

3.1 Overlapped reconstruction segments 50
3.2 Multiplexed Signal Acquisition Involving Compressed Sensing 53
3.3 Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **basis pursuit** 63
3.4 Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **OMP** 64
3.5 Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **ROMP** 64
3.6 Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **CoSAMP** . . 65
3.7 PSNR histograms for **signal 4** with four reconstruction algorithms 65
3.8 Reconstructed signal (red) over the original signal (black) for four channels of decompressed data obtained from compressed Nyquist samples 66
3.9 PSD of Filtered MOSAICS output during 0–60 μs 68
3.10 Proposed hardware architecture for realization of MOSAICS 69
3.11 Reconstructed signal (red) over the original signal (black) for four channels. The PSNR for the reconstructed signals are: signal 1:10.0 dB, signal 2: 15.74 dB, signal 3: 28.57 dB, signal 4: 26.36 dB 71

3.12 Original signals (black) of test set III and a snapshot of the reconstruction (red) using MOSAICS with MUSIC. The PSNR values for the reconstructed signals are: for signal 1:16.30 dB, for signal 2: 17.23 dB, for signal 3: 15.63 dB, for signal 4: 19.42 dB 77
3.13 Snapshot of signal 2 of Table 3.9 at an SNR of 5 dB during 0–1 sec 81
3.14 Detection of carrier frequency in FM signals with noise at 5 dB SNR. The carrier frequency changes by 100 kHz on either side of the central frequency once in every 100 ms . 82
3.15 First 1000 samples of channel 1 at 5 dB SNR noise 83

4.1 Reconstructed signal (shown in red color) is plotted against the original signal for four channels, 1, 4, 6 and 8 during the interval 35-40sec. The corresponding PSNR values in dB are: 25.6, 23.3, 22.6 and 24.0. 98
4.2 Under the ARCS scheme 10 correlated signals are acquired using 5 ADCs. The curve shows how the correlation structure is incrementally learnt by the system. The y-axis gives the total number of signals learnt at any time. The colored squares represent each of the signals. At 0 s all the signals are strangers; at around 6 s signals 1, 2, 3, 4, 5 are introduced; at around 12 s signal 6 becomes a familiar signal and the total number of familiar signals is 6. This continues until around 41 s when signal 10 becomes familiar. 100
4.3 EEG electrode placement systems 103
4.4 Reconstructed (red) vs original (black) signals for subject 104, record 13 using Basis Pursuit recovery algorithm . . . 118
4.5 Comparison of FSM values of the 10 recovered channels for subjects 1 and 8 120

4.6 Comparison of FSM values of the 10 recovered channels for subjects 41 and 61 121
4.7 Comparison of FSM values of the 10 recovered channels for subjects 77 and 104 122
4.8 Comparison of performance with various reconstruction methods averaged over several subjects in different EEG bands 123
4.9 Comparison (with the original) of FSM, in different bands, of the 10-10 system EEG channels reconstructed through compressed sensing using only recordings done on the 10-20 system. Recovery algorithm used: Basis Pursuit 123
4.10 Estimation of 10-10 channels from 10-20 recordings - reconstructed (red) vs original (black) signals for subject 64, record 12 using Basis Pursuit recovery algorithm 124

5.1 AA filters at the front-end of a data acquisition system . . 126
5.2 Compressed sensing architecture for acquiring signals at a higher sampling rate than the specified sampling rate of the ADC while using low order AA filter 128
5.3 Magnitude response of FIR filter of order 8 131
5.4 Reconstructed signal (red color) vs original signal (black color) for two signals. The deviations are 20.8 dB and 21.5 dB 132
5.5 Reconstructed signal (red color) vs original signal (black color) for four channels of voltage proportional to deflection of control surfaces. **DCT was used as the sparsity matrix**. Additive noise at 20 dB SNR was added to the original signal. 134

5.6 Reconstructed signal (red color) vs original signal (black color) for four channels of voltage proportional to deflection of control surfaces. **DFT was used as the sparsity matrix**. Additive noise at 20 dB SNR was added to the original signal. 134
5.7 Pipeline ADC architecture 138
5.8 Modified pipeline ADC architecture with a single ADC and DAC 138
5.9 a)100 Samples from a Nyquist sampled 4-bit ADC and b) 100 Samples from a 8-bit compressively sampled ADC. In both case the reconstructed signal is shown in red and the original signal in black 140
5.10 a)1000 samples of fetal ECG sensed at the fetal head using a Nyquist sampled 8-bit ADC. b) 1000 samples of fetal ECG reconstructed from a 16-bit compressively sampled ADC. In both cases the reconstructed signal is shown in red and the original signal in black. Voltage range is -3276.8 μV to +3276.8 μV. 142
5.11 a)1000 samples of fetal ECG sensed at the maternal abdomen using a Nyquist sampled 8-bit ADC b) 1000 samples of fetal ECG reconstructed from a 16-bit compressively sampled ADC. In both cases the reconstructed signal is shown in red and the original signal in black. Voltage range is -3276.8 μV to +3276.8 μV. 143
5.12 Comparison of SQNR between 8-bit Nyquist and 16-bit CS acquisition at different input dynamic ranges for fetal ECG captured directly at the fetal head 143

C.1 Probabilities of 56 down-sized identity matrices $p_{\mathbf{\Phi}_{(i)}}$ having a lower mutual coherence with a given inverse KLT sparsity matrix, $\mathbf{\Psi}^{(8)}$ as compared to random Gaussian matrices of dimension 3×8 160

C.2 Comparison of FSM values between reconstructed and original signals for Subject 1, Record 7 using down-sized identity matrices, indicated in the legend as DIM and the gaussian measurement matrix. With both the measurement matrices the CoSAMP recovery algorithm has been used. 161

List of Tables

3.1 FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET I 59
3.2 INPUTS AND DERIVED PARAMETERS FOR TEST SET I 60
3.3 PERFORMANCE OF MOSAICS WITH DIFFERENT RECONSTRUCTION ALGORITHMS AS PSNR VALUES (in dB).THE LAST COLUMN GIVES THE PSNR FOR DECOMPRESSED DATA RECOVERED FROM NYQUIST SAMPLED COMPRESSED DATA. THE LAST ROW GIVES THE EXECUTION TIME FOR A SIMULATED ACQUISITION TIME OF 80 ms 63
3.4 COMPARISON OF MEAN AND STANDARD DEVIATION OF PSNR FROM 100 TRIALS FOR VARIOUS RECONSTRUCTION ALGORITHMS 66
3.5 FREQUENCY CHARACTERISTICS OF SIGNALS IN TEST SET II 67
3.6 MOSAICS OUTPUT (FILTERED) FOR TEST SET II- HIGHEST FREQUENCY COMPONENT 68
3.7 FREQUENCY CHARACTERISTICS OF SAMPLE SIGNALS WITH NON-INTEGRAL NUMBER OF CYCLES IN THE RECONSTRUCTION WINDOW 70
3.8 FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET III 76

3.9 FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET IV AND DETECTED FREQUENCIES, AT DIFFERENT SNR 80

4.1 PSNR VALUES (IN dB) OF THE RECONSTRUCTED SIGNAL IF THE COMMON COMPONENT IN THE ORIGINAL SIGNAL IS REINITIALIZED 100
4.2 MOTOR/IMAGERY TASKS IN THE PHYSIONET DATABASE DURING WHICH THE EEG USED FOR THE STUDY HAS BEEN COLLECTED.(see [1]) 111
4.3 COMPARISON OF FSM IN DIFFERENT BANDS OF THE ORIGINAL AND RECONSTRUCTED EEG FOR DIFFERENT CHANNELS OF SUBJECT 61, RECORD 10 USING DIFFERENT CS RECONSTRUCTION ALGORITHMS . 119

5.1 FREQUENCY CHARACTERISTICS OF TEST SIGNALS 131
5.2 PSNR OF THE SIGNALS OF CONTROL SURFACE FEEDBACKS AT DIFFERENT NOISE LEVELS 135
5.3 SQNR VALUES (dB) AFTER n-BIT NYQUIST SAMPLING AND 2n-BIT COMPRESSIVE SAMPLING 140

6.1 CHARACTERISTICS OF SIGNALS TO BE ACQUIRED UNDER A COMPRESSED SENSING SETUP 148

A.1 DESIGN CONSIDERATIONS FOR EFFICIENT DESIGN OF EMBEDDED SYSTEMS 153

Acronyms

AA	Anti Aliasing.
AB	Acquisition Block.
ADC	Analogue-to-Digital Converter.
ARCS	Acquisition and Reconstruction of Correlated Signals.
BP	Basis Pursuit.
BSS	Blind Source Separation.
COTS	Commercially-Off-The-Shelf.
CS	Compressed Sensing.
DAC	Digital to Analog Converter.
DCM	Digital Clock Manager.
DCS	Distributed Compressed Sensing.
DCT	Discrete Cosine Transform.
DFT	Discrete Fourier Transform.
DMA	Direct Memory Access.
DTFT	Discrete Time Fourier Transform.
DWT	Discrete Wavelet Transform.
EEG	Electro-EncephaloGraph.

FIR	Finite Impulse Response.
FPGA	Field Programmable Gate Array.
FRI	Finite Rate of Innovation.
IF	Intermediate Frequency.
iid	Independently and Identically Distributed.
JSM	Joint Sparsity Model.
KLT	Karhunen Loueve Transform.
LSB	Least Significant Bit.
MMV	Multiple Measurement Vector.
MOSAICS	Multiplexed Optimal Signal Acquisition Involving Compressed Sensing.
MP	Matching Pursuit.
MSL	Multi-Source Location.
MUSIC	Multiple Signal Classification.
OMP	Orthogonal Matching Pursuit.
PCB	Printed Circuit Board.
PSD	Power Spectral Density.
PSNR	Peak Signal to Noise Ratio.
PSS	Piecewise Stationary and Sparse.
RAM	Random Access Memory.
RB	Reconstruction Block.

RIP	Restricted Isometry Property.
RS	Reconstruction Segment.
SCA	Sparse Component Analysis.
SMNR	Signal to Measurement Noise Ratio.
SMV	Single Measurement Vector.
SNR	Signal to Noise Ratio.
SOC	System on Chip.
SQNR	Signal to Quantization Noise Ratio.
SS	Stationary and Sparse.
UF	Utilization Factor.
UUP	Uniform Uncertainty Principle.

List of Symbols

F_{ADC}	Optimal sampling rate of ADC.
F_{MOS}	MOSAICS operating frequency.
K	Number of non-zero elements in a sparse vector.
L	Minimum assured duration of stationary and sparse behavior in a PSS signal.
M	Number of measurements on a sparse vector.
N	Number of elements in sparse vector.
Φ	Measurement matrix.
Ψ	Sparsity matrix.
$\Sigma_{\mathbf{X}}$	Covariance matrix of the signal matrix X.
Θ	Sensing matrix.
Ξ	Signal recovery algorithm.
β	Number of analog to digital conversions in each acquisition cycle in MOSAICS.
η	Number of cycles of a constituent sinusoid within a reconstruction segment.
$\|.\|_0$	l_0-norm of a vector.
$\|.\|_1$	l_1-norm of a vector.

$\|.\|_2$	l_2-norm of a vector.
$\mathbb{R}$	Real numbers.
$\boldsymbol{F}_N^{-1}$	Inverse DFT matrix of order N.
$\mathbf{c}$	Coefficient vector.
$\mathbf{f}$	Measured vector.
$\mathbf{I}_N^{(M)}$	Matrix obtained by randomly picking M out of N rows of an Identity matrix of order N.

Chapter 1

Introduction

Over the past several decades, embedded systems have tremendously evolved with respect to complexity, performance, power consumption and functionality. An embedded device invariably cannot operate in isolation. It has to be aware of the environment in which it operates and be responsive to stimuli. Awareness is primarily realized through a congregation of sensors which interface to the system via a multitude of analog channels. Data acquisition through multiple analog channels is one of the most important functions and is accomplished with analog-to-digital converters. With increasing functionality and number of data acquisition channels, the complexity of the on-board electronics has enormously escalated. While consumer products, common household appliances, biomedical and industrial data acquisition systems need to address this matter, the issue is of greater concern in embedded systems that are employed in aviation, military and space applications. Such systems have stringent budgets for available space, power consumption and dissipation, and in some cases, cost also. In addition, such systems being safety-critical in nature, need to have a very high order of reliability which becomes suspect with increasing number of components and the associated interconnect. Invariably, these systems are built around ruggedized embedded devices performing multiple computational and I/O tasks under harsh environmental conditions. Each unit of embedded hardware is usually a component of a bigger sub-system

and has to perform its function, in real time synchronization with other subsystems and devices with which it has to communicate through several analog and digital interfaces. The incorporation of many IC chips along with the associated passive components makes typical avionics embedded designs highly complex, thereby, in many cases, not being able to meet the requirements of compactness, low power consumption and cost.

A smarter design comprising lesser number of components would go well with the increasing demand for miniaturized avionics. The signals acquired and processed by most of the embedded avionics are well characterized on account of pre-flight simulations and a wealth of data acquired through test flights. A plausible and pertinent question is: can the a priori information about the signal characteristics be utilized to reduce the complexity of embedded designs ? For instance, if the signals have some kind of information redundancy, can we reduce the number of analog-to-digital converters without significantly compromising on the performance ? This research work is a modest attempt to answer such questions and to propose a few innovative design alternatives that exploit information redundancy.

1.1 Requirements of Embedded Systems

Design of embedded systems is different from that of conventional desktop systems in many ways. Embedded devices invariably operate in close coordination with a number of other subsystems, all of which function synergistically to realize a common goal. The following considerations bear significance, specifically in the context of embedded system design:

i) **Small size and low weight**: These complementary requirements are important particularly for small, hand-held systems and those that require portability and mobility, typical examples of which are mobile phones, medical sensors, air-borne systems and robotics, to name a few.

ii) **Low power consumption**: Excessive power requirement demands a built-in, high capacity power source which again contributes to size and weight.

iii) **Harsh operating environment**: In air-borne, space and process control applications, the embedded device has to function in unfriendly conditions: heat and thermal shocks, mechanical shocks and vibrations, electromagnetic interference, unfavorable acoustic and climatic conditions and corrosion. To handle these unavoidable stimuli, the devices have to be designed for very high reliability. Designs with reduced hardware complexity due to lesser number of components, could possibly push the burden of reliability on software. However, in such designs, the probability of failure of the system due to malfunction of components and the associated interconnect reduces. Software anomalies can be removed through simulations and software testing methodologies. Repeatability in behavior can be reasonably expected, but the same is not true for hardware.

iv) **Reliability**: Reliable and predictable operation is essential in air-borne, space and defense applications from the point of human safety. Unlike in traditional desktop systems, failure could be catastrophic.

v) **Deterministic response**: Embedded systems usually operate in real time since they have to function within a bigger system. Due to this, an embedded system needs to respond to an external stimulus within a specified and deterministic time period.

vi) **Cost**: Embedded devices are usually produced in bulk volumes to support multiple deployable systems. Hence, a small reduction in the cost of individual unit has significant financial repercussions.

Table A.1 in Appendix A lists some of the techniques that are typically considered for designing embedded hardware for reliability, compactness,

reduced power consumption and heat dissipation. The focus of this work is on reduction of components, that is common to all the design features.

1.2 Data Converters in System-On-Chips

Development of System-On-Chip (SOC) solutions, which comprise dozens of functional blocks on a single die, has been a significant trend in the embedded world in the last two decades [2]. Examples of applications include consumer appliances like cellular phones, DVD players, set-top boxes and multi-media players to name a few. The largest component of such SOCs is the digital portion which houses multi-core gigahertz processors, multiple megabytes of memory, various media access controllers and dedicated digital signal processors (DSPs) built with close to half a billion transistors. On account of this, assuming that digital is almost available for free, the trend is to keep the analog circuitry to the minimum, pushing as much as possible to the digital domain. However, the data converter continues to remain as the minimum required analog function that has to be available in an SOC, since the analog signal of the outside world has to be in any case brought into the digital world before the available processing power can be made use of. While architectural choices in the design of Analogue-to-Digital Converters (ADC) aid in reducing the analog portion of the SOC, there also has been a trend towards minimizing the associated analog circuitry for any given architecture. For example, if the effective sampling rate of the signal is increased, then the complexity of the anti-aliasing filter before the ADC could be reduced. Reconstruction algorithms in the digital section could overcome any degradation in data conversion process due to the abridged analog circuitry. Such digitally-aided analog design techniques allow for better analog performance, compactness, reduced power requirement and cost.

1.3 Information Redundancy in Signals

Quite often, the signals input to an embedded system have redundant information. For instance, if the signal exhibits sparsity in a transformation domain, then the insignificant coefficients in the transformed vector do not contribute to additional knowledge of the system. Sparsity leads to dimensionality reduction. In classical data compression schemes, such coefficients are omitted before storage or transmission. Commonly used are Fourier and Cosine transforms. The Discrete Wavelet Transform (DWT) has been used for multi-resolution analysis serving as a sparsity inducing transformation for non-stationary signals. The list of signal transforms, that have been introduced by the signal processing community, to handle specific signals in different applications is long. Information redundancy is also exhibited by a set of correlated signals, since knowledge of a few signals in the group, at any instant of time, can be used to predict the rest, given the correlation structure of the signals.

Transform coding schemes like JPEG 2000 typically work by acquiring the full signal, computing the complete set of transform coefficients and encoding the largest coefficients while discarding the others. This process of massive data acquisition followed by compression is wasteful. In this context, a fundamental question can be raised, "Since most signals are compressible, why spend so much effort acquiring all the data when we know that most of it will be discarded ? Would it not be possible to acquire the data in already compressed form so that one does not need to throw away anything?" In other words, does sparsity have bearings on the data acquisition process itself ?

1.4 Scope for Improvised Embedded Designs

Consider a set of n input signals for data acquisition in an embedded system, each having a Nyquist sampling rate of $f_{NYQ}^{(i)}, i = 1...n$. The total

number of degrees of freedom of the system during a time period T of data acquisition is $N = T\sum_{i=1}^{n} f_{NYQ}^{(i)}$. Let the system have information redundancy due to sparsity on some transformation basis or because of inter-signal correlation such that there exist only $M < N$ degrees of freedom. This implies that theoretically there need to be only M analog to digital data conversions required to reconstruct each signal in the digital world of the embedded system. In practice, more analog to digital data conversions are required in order to have a reconstruction probability close to one. In any case, the number of data conversions required is certainly less than N. In the light of this observation, it is reasonable to raise the following questions with respect to realizing efficient embedded hardware for data acquisition of sparse signals:

- Can the input signals in the system be captured with lesser number of analog to digital converters than that required in a Nyquist setup ?
- For a given number of signals and an equal number of analog to digital converters, can we do the acquisition at an effective sampling rate that is higher than Nyquist rate ?
- Assuming that the two questions above can be answered in the affirmative, can we conceptualize designs with less number of components and consequently less floor area requirement on the printed circuit board (PCB)?

One can leverage upon the fact that design of the analog portion in SOCs is usually not general purpose (involves reuse of existing blocks) and has to be tailored to specific system at hand. Thus characteristics of the specific signals that are input to the system, like sparsity can be exploited to evolve efficient designs.

The next section gives a brief description of various sparse signal processing applications with specific focus on data acquisition.

1.5 Sampling and Reconstruction of Sparse Signals

Data acquisition, which is the focus of this research work, is not the only problem that has been addressed in the context of sparse signals. It constitutes only a small space in the vast arena of sparse signal processing that includes such diverse problems as source localization and spectral estimation, to name a few. It is meaningful to briefly explore such problems and their solutions in order to identify the coordinates of this research in this immense space and possibly borrow a few techniques. Before a survey of this field is presented, it is prudent that the criteria which categorize the various applications be understood. The following aspects are associated with any sparse signal processing problem:

i) **Sparsity domain**: This is the domain where the signal exhibits sparsity. For example, in the case of a signal that has only a few frequency components, this is the frequency domain. For a smooth, blurred image this could be the wavelet or Discrete Cosine Transform (DCT) domain.

ii) **Information domain**: The domain in which the signal is sampled is usually different from the sparsity domain and is known as the information domain. In case of a frequency sparse signal, this is simply the time domain signal.

iii) **Type of sampling** in the information domain: The pattern of sampling could be periodic or random.

iv) **Nature of sparsity**: The sparsity could be one of the following:

- band-pass: all the non-zero components in the sparse domain are within a small interval at consecutive locations, e.g. radar applications
- multi-band: the non-zero components are bundled into a number

of distinct small intervals, e.g. the sum of a few narrow-band transmissions, each modulated by distinct high frequency carriers.

- random: the non-zero components can be located anywhere in the sparse domain, e.g. Communication signals, such as transmissions with a frequency hopping modulation scheme that switches a sinusoidal carrier among many frequency channels according to a predefined (often pseudorandom) sequence.

v) **Reconstruction algorithm**: The nature of the algorithm used to reconstruct the original signal from the sub-sampled information domain signal is an important factor as this determines the required computation power. The algorithm could be based on low pass filtering, iterative methods with and without adaptive thresholding, interpolation, filter banks, basis pursuit, matching pursuit, annihilating filter etc. In order to relate the findings of this research to the broader framework of sparse signal acquisition and reconstruction schemes, under which it falls, a survey has been done and is presented in the next subsection. The objective of the survey is to find the sparse signal reconstruction algorithm that is best suited for this research.

1.5.1 Sparse Reconstruction Schemes

While it is a simple matter to sub-sample a signal, reconstructing the original signal from the limited number of samples is the real challenge. Several algorithms have been reported each with its own merits and demerits and suitability to specific applications. A few of these algorithms are described in what follows:

i) **Iterative methods with a priori knowledge of sparsity locations** - When the locations of the non-zero coefficients in the sparsity domain are known, then the number of samples in the information domain that are required to reconstruct the signal should be at least

equal to the number of such coefficients. However, depending upon the nature of the sparsity and the type of sampling, the reconstruction may be unstable[1] and the actual number of samples required for exact reconstruction may be higher. Iterative methods involve alternate projections between the sparsity domain and the information domain. The initial input to the algorithm is the sub-sampled vector. In each iteration, the estimate of the information domain signal is transformed to the sparsity domain; passed through a mask or filter that is localized at the sparsity location and then transformed back to the information domain to get a residue vector. The residue is used to update the signal estimate. If the sparsity is band-pass in nature, then a single iteration is usually sufficient. In the case of random sparsity, more number of iterations is required. The iterations can be accelerated using Chebyshev and conjugate gradient methods [3]. Iterative methods are quite robust against quantization and additive noise and it can be proved that they approach the pseudo-inverse (least squares) in the case of noisy signals.

ii) **Iterative methods for unknown sparsity locations** - When locations of the non-zero coefficients are unknown, one needs to evaluate the number of sparse coefficients (or non-zero samples), the sparsity locations, and the values of the non-zero coefficients. By means of alternate projections between information and sparsity domains and simultaneous adaptive lowering or rising of a threshold in the sparsity domain, the sparse coefficients are gradually picked up after several iterations. The method given in [4] is an example of an iterative method with adaptive threshold. Other iterative methods like Spline interpolation [5], nonlinear/time varying methods [6], Lagrange interpolation [7] and Error Locator Polynomial (ELP) [8] work quite well in the ab-

[1] low probability of accurate reconstruction

sence of additive noise, but may not be robust in the presence of noise.

iii) **Compressed Sensing** (CS) - In compressed sensing, a weighted linear combination of samples, also called compressive measurements, is taken in a basis different from the basis in which the signal is known to be sparse (sparsity domain). It has been proven that even a small number of these compressive measurements contain useful information. Reconstructing the original signal from the linear combination involves solving an under-determined set of equations since the number of compressive measurements taken is smaller than the number of unknown non-zero coefficients in the sparse domain. However, the sparsity assumption constrains the solution set and it is possible to find a solution using a plethora of algorithms proposed in a huge volume of literature spanning the last two decades. Compressed sensing methods do not depend upon any sparsity pattern like band pass. There is no prior knowledge of the sparsity locations; even the exact number of non-zero locations, though within a known upper bound, is not known. The basic tenets of compressed sensing form the subject of the next chapter and therefore, no references are cited here.

iv) **Sampling with finite rate of innovation** - Parametric signals, such as streams of short pulses, appear in many applications including bioimaging, radar, and spread-spectrum communication. The recently developed Finite Rate of Innovation (FRI) framework [9], has paved the way to low rate sampling of such signals, by exploiting the fact that only a small number of parameters which is the number of degrees of freedom or innovation per unit of time, are needed to fully describe them. An elegant and powerful result is that, in many cases, certain types of FRI signals can be reconstructed without error from

samples taken at the rate of innovation [10; 11]. The advantage of this result is self-evident: FRI signals need not be bandlimited, and even if they are, the Nyquist frequency can be much higher than the rate of innovation. Thus, by using FRI techniques, the sampling rate required for perfect reconstruction can be lowered substantially.

v) **Spectral estimation** - Parametric spectral estimation methods like Prony's method [12], Pisarenko Harmonic Decomposition (PHD)[13] and Multiple Signal Classification (MUSIC)[14] have been adapted to give efficient solutions in the case of signals that are sparse in the frequency domain. In such applications, the objective is to estimate the spectral signature of the signal from the sub-sampled measurements, instead of the signal itself.

vi) **Sparse Array Processing** - Three types of array processing problems have been explored by researchers:

- **Estimation of Multi-Source Location (MSL) and Direction of Arrival (DOA)** - In MSL and DOA estimation [15; 16; 17] a sparse (passive or active) array of sensors is used to locate the sources of narrow-band signals. Some applications may assume far-field sources (e.g. radar signal processing), where the array is only capable of DOA estimation, while other applications (e.g. biomedical imaging systems) assume near-field sources, where the array is capable of locating the sources of radiation. The common temporal frequency of the source signals is known. Simultaneous spatial sampling of the signal exhibits a phase change from sensor to sensor, thereby obtaining discrete samples of a complex exponential in which the frequency gets translated into direction of the signal source. This resembles the spectral estimation problem with

the difference that sampling of the array elements is not limited in time. In fact, in array processing, an additional degree of freedom (the number of elements) is present; thus, array processing is more general than spectral estimation. For both MSL and DOA, the angle of arrival (azimuth and elevation) should be estimated; while for MSL, an extra parameter of range is also needed.

- **Sparse array beam forming and design** - In certain applications like radar, sonar, ultrasound imaging and seismology, the challenge is the combinatorial problem of finding the best sparse layout of beam forming elements in one and two dimensions [17; 18]. Linear programming, genetic algorithms and simulated annealing techniques have been used to solve the associated optimization problem.
- **Sparse sensor networks** - Wireless sensor networks typically consist of a large number of sensor nodes, spatially distributed over a region of interest, that observe some physical environment like acoustic, seismic, and thermal fields with applications in a wide range of areas such as health care, geographical monitoring, homeland security, and hazard detection. In general, there are three main tasks that should be implemented efficiently in a wireless sensor network: sensing, communication, and processing. The main challenge in design of practical sensor networks is to find an efficient way of jointly performing these tasks, while using the minimum amount of system resources. In general, sparsity can arise in a sensor network from two main perspectives: 1) Sparsity due to non-uniform spatial distribution of nodes that can be exploited to reduce the amount of sensing, processing, and/or communication [19] and 2) Sparsity of the field to be estimated due to correlation between the data at different nodes [20; 21].

vii) **Sparse Component Analysis** - Recovery of the original source sig-

nals from their mixtures, without having a priori information about the sources and the way they are mixed, is called Blind Source Separation (BSS). BSS algorithms are based on the assumption that the sources may be uncorrelated, statistically independent without any mutual information, or are disjoint in some space. Based on their disjoint characteristics in a suitable domain, in which they are sparse, the signal mixtures can be decomposed with Sparse Component Analysis. SCA algorithms [22; 23] assume that the sources are sparse on an overcomplete dictionary of basis functions. The source separation is performed in two different stages. First, the problem is treated as a clustering problem to extract the unknown mixing matrix. Next the $l_1 - norm$ of the source is minimized subject to the constraint that the mixtures are formed from the sources and the estimated mixing matrix.

viii) **Sparse Dictionary Representation** (SDR) - Closely related to SCA is the sparse dictionary representation problem of finding out a basis or frame in which all the signals in a particular class are sparse [24].

ix) **Multipath Channel Estimation** - In wireless systems, the transmitted signal bounces off different objects and arrives at the receiver from multiple paths. This phenomenon causes the received signal to be a mixture of reflected and scattered versions of the transmitted signal. The mobility of the transmitter, receiver, and scattering objects results in rapid changes in the channel response, and thus the channel estimation process becomes more complicated. Due to the sparse distribution of scattering objects, a multipath channel is sparse in the time domain. By taking sparsity into consideration, channel estimation can be simplified and/or made more accurate.

This completes a very brief survey of some of the major interesting

problems related to sparse signals that are being investigated by the signal processing community. An excellent and exhaustive exposition to sparse signal processing problems has been given by Marvasti et. al. in [25].

1.6 Outline of Thesis

In this research effort, the focus is mainly on streaming data acquisition of multiple sparse signals. Guided by the literature survey, in this work the compressed sensing paradigm is chosen as the core engine for acquisition and reconstruction of signals due to the following reasons:

- Of all the various sparse signal processing schemes, the one that has been most widely used for data acquisition is compressed sensing.
- There is extensive literature in support of compressed sensing with report of proof of performance in a wide variety of applications.
- Well tested and proven, open source toolboxes are available for CS recovery algorithms that can be readily used for simulations.

The next chapter exclusively deals with an introduction to compressed sensing and the associated issues. If it is possible to reconstruct a sparse signal, using sub-Nyquist number of samples, can the idle sampling cycles of an analog-to-digital converter be used to capture many such sparse signals in a multiplexed fashion, such that a single ADC acquires and reconstructs several sparse signals simultaneously ? This question is answered in chapter 3, which presents an efficient data acquisition architecture that samples and reconstructs multiple sparse signals. An improvisation of the proposed scheme for general signals with arbitrary frequencies is also explained in the chapter. The chapter concludes with a suggestion of how the method can be adapted to the problem of detecting sinusoids buried in heavy noise.

Do individual signals always have to be sparse or can we exploit the correlation between multiple non-sparse signals, with the objective of again reducing the number of data converters ? In chapter 4 is proposed an algorithm that gradually learns the correlation between multiple signals that are not necessarily sparse. The chapter also proposes a scheme for exploiting inter-signal correlation to perform compressed acquisition of EEG signals. Instead of reducing the number of data converters, can we use as many of them as there are sparse signals and try to achieve an effective sampling rate for each signal that is higher than the specified sampling rate of each ADC, thereby being able to relax the specifications of the front-end anti-aliasing filter ? This question is probed in chapter 5 and yet another compact data acquisition scheme is proposed. Proof of performance of the ideas put forth in the thesis has been demonstrated through simulations using synthetic data and in some cases with real world signals. The thesis is concluded in chapter 6, along with the presentation of a scheme in which a combination of various methods is used in an integrated fashion to acquire multiple signals with different sparsity properties.

Chapter 2

Compressed Sensing

2.1 Introduction

Conventional approaches to the acquisition of signals or images obey the well known Nyquist/Shannon theorem, which states that the sampling rate must be at least twice the maximum frequency present in the signal. This principle underlies nearly all signal acquisition at the front-end of most applications like consumer audio and visual electronics, medical imaging devices, radio receivers and radar. Where there is an upper limit on the possible sampling rate, an anti-aliasing filter is used to filter out the signal frequencies that are more than half the sampling frequency, assuming that the region of interest in the signal lies in the lower frequencies. While it is true that Nyquist rate sampling is able to completely describe a signal, situations arise where 'twice the maximum frequency in the signal' is so high that it is beyond the sampling capability of conventional analog to digital converters. If such a signal is heavily sparse in the frequency domain is it advisable to sample the signal at a very high rate just because there exists a narrow band of high frequency in the signal ? Sparsity expresses the idea that the information rate of a continuous time signal may be much smaller than suggested by its bandwidth, or that a discrete-time signal depends on a number of degrees of freedom which is comparably much smaller than its (finite) length.

Transform coding methods like JPEG2000 rely on the fact that many signals can be sparsely represented in a fixed basis (e.g. Fourier basis in the case of frequency sparsity). This implies that only a small number of adaptively chosen transform coefficients rather than all the signal samples need to be stored or transmitted. Typically the full signal is acquired from which the complete set of transform coefficients is computed and the largest coefficients are encoded while discarding the rest. Compression after acquisition of huge amount of data is a wasteful exercise. This throws up a basic question: Is it worth acquiring so much data when only a small fraction of it is retained ? Is there a way of acquiring the data in already compressed form, so that there is no need to discard anything ? Is it possible that even the data acquisition process can leverage upon the signal sparsity ? "Compressive sampling", also known as "compressed sensing", shows that it is indeed possible to capture analog signals directly in a compressed digital form. Using a simple and efficient mechanism of signal acquisition it is possible to reconstruct the signal, with the help of computational power, from an incomplete set of measurements obtained at a low sampling rate.

Before giving a formal introduction to compressed sensing, it is necessary to understand the concept of sparsity and compressibility in signals. Consider a vector, $\mathbf{x} \in \mathbb{R}^N$ which can be expanded in an orthonormal basis represented by the $N \times N$ matrix $\mathbf{\Psi} = [\psi_1 \, \psi_2 \, ... \psi_N]$ as:

$$\mathbf{x} = \sum_{i=1}^{N} c_i \psi_i \tag{2.1}$$

where $c_i = \langle \mathbf{x}, \psi_i \rangle \, , i = 1...N$ are the coefficients of the signal on the orthonormal basis. Equivalently, $\mathbf{x}$ can be expressed as

$$\mathbf{x} = \mathbf{\Psi c} \tag{2.2}$$

where $\mathbf{c}$ is a column vector of the coefficients of size $N \times 1$. Hereafter, the vector $\mathbf{c}$ shall be referred to as the **coefficient vector** and the matrix $\mathbf{\Psi}$

as the **sparsity matrix**

Sparsity - When a signal is said to be sparse in any basis like the one above, then most of the elements in the coefficient vector $\mathbf{c}$ are zero. In other words, the number of non-zero elements in vector $\mathbf{c}$ (also called the l_0-norm) is small compared to N.

$$K = \|\mathbf{c}\|_0 << N \tag{2.3}$$

Such signals are called $K-$ **sparse signals**. K in general, will be used to denote the number of non-zero elements in a sparse vector of N elements.
Compressibility - In cases, where the l_0-norm of the coefficient vector is not significantly smaller than N, the signal, although not sparse, can still be called compressible, if the ordered set of coefficients decay exponentially (or $\mathbf{x}$ belongs to a weak-l_p ball[1] of radius R). This can be mathematically expressed as,

$$|c_1| \geq |c_2| \ldots \geq |c_i| \ldots \geq |c_N| \text{ , then} \tag{2.4}$$

$$|c_i| \leq Ri^{-1/p},\ 1 \leq i \leq N \tag{2.5}$$

where p is the **decay constant**[26]. The smaller is p, the faster is the decay.
The K-term linear combination of elements which best approximates $\mathbf{x}$ in an l_2 sense is obtained by keeping only the K largest terms in the expansion (2.1)-

$$\mathbf{x}^{(K)} = \sum_{i=1}^{K} c_i \psi_i \tag{2.6}$$

If the coefficients c_i obey (2.5), then the error between $\mathbf{x}$ and $\mathbf{x}^{(K)}$ also obeys a power law:

$$\left\|\mathbf{x} - \mathbf{x}^{(K)}\right\|_2 \leq C_2 R K^{1/2-1/p} \tag{2.7}$$

[1]For a real number $p \geq 1$, the l_p-norm of a vector $\mathbf{x}$ is defined as $\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_N|^p)^{1/p}$

in the l_2-norm and

$$\left\|\mathbf{x} - \mathbf{x}^{(K)}\right\|_1 \leq C_1 R K^{1-1/p} \tag{2.8}$$

in the l_1-norm[26] for some positive constants C_1 and C_2 .

Sparsity and compressibility have clear implications. When a signal is sparse or compressible then the zero or small coefficients can be discarded without perceptible loss of information. For example, if $\mathbf{c}^{(K)}$ is the coefficient vector containing only K significant coefficients, the remaining elements being trivially zero in case of sparse signals and forced to zero in the case of compressible signals, then the corresponding signal vector is

$$\mathbf{x}^{(K)} = \mathbf{\Psi}\mathbf{c}^{(K)} \tag{2.9}$$

Since $\mathbf{\Psi}$ is an orthonormal basis, we have

$$\left\|\mathbf{x} - \mathbf{x}^{(K)}\right\|_2 = \left\|\mathbf{c} - \mathbf{c}^{(K)}\right\|_2 \tag{2.10}$$

Thus, if $\mathbf{x}$ is sparse or compressible, then $\mathbf{x}$ is well approximated by $\mathbf{x}^{(K)}$ and the error, $\left\|\mathbf{x} - \mathbf{x}^{(K)}\right\|_2$ obeys the power law in (2.7). This is the principle behind most lossy encoders like JPEG2000 in which $\mathbf{c}$ is computed from $\mathbf{x}$ and the K most significant coefficients of $\mathbf{c}$ are encoded before being stored or transmitted. When the signal has to be recovered, the full length sparse vector $\mathbf{c}$ is constructed using the decoded coefficients. The lossy approximation of the original signal is then recovered using (2.2). Many natural signals have concise representations when expressed in a convenient basis. For example, although nearly all pixels in a gray scale image have non-zero values, the wavelet coefficients offer a concise summary: most wavelet coefficients are small, and the relatively few large coefficients capture most of the information about the object.

The sample-then-compress framework in general, is an inefficient scheme: A potentially large number of samples N have to be taken, even if the ul-

timate desired number K is small. Along with the K large coefficients, their locations also have to be encoded by the encoder. Compressed sensing offers a different data acquisition and reconstruction method in which a compressed representation of the signal is directly obtained without requiring to take the N samples.

2.2 Sensing the Signal

Sensing a signal is the mechanism by which information about a signal $\mathbf{x} \in \mathbb{R}^N$ is obtained through linear functionals, ϕ_m as,

$$f_m = \langle \mathbf{x}, \phi_m \rangle \;\; m = 1, ..., M \tag{2.11}$$

or putting it more concisely,

$$\mathbf{f} = \mathbf{\Phi x} \tag{2.12}$$

where $\mathbf{\Phi} = [\phi_1 \, \phi_2 ... \, \phi_M]^T$ is an $M \times N$ matrix ($M \leq N$), which hereafter shall be referred to as the **measurement matrix**. The vector $\mathbf{f}$ shall be called the **measured vector**. That is, we simply correlate the signal we wish to acquire with the waveforms, ϕ_m. The measured vector depends on the sensing waveforms[27]:

- **Dirac deltas** - If the sensing waveforms are Dirac delta functions (spikes), then $\mathbf{f}$ is a vector of sampled values of $\mathbf{x}$ in the time or space domain. As an example of a 2D signal, if the sensing waveforms are indicator functions of pixels, then $\mathbf{f}$ is the image data typically collected by sensors in a digital camera.
- **Sinusoids** - If the sensing waveforms are sinusoids, then $\mathbf{f}$ is a vector of Fourier coefficients; this corresponds to the sensing modality in use in Magnetic Resonance imaging (MRI).

- **Random** - Random measurement matrices with i.i.d.[1] Gaussian entries are suited for compressed sensing of general signals. This is the subject of discussion of a later section.

The design of efficient measurement matrices, tailored to specific applications, is in itself an area of research in compressed sensing.

2.3 Reconstructing the Signal

When $M = N$, the sensing is Nyquist, and the recovery of $\mathbf{x}$ given $\mathbf{f}$ and $\mathbf{\Phi}$, involves just obtaining the solution of the system of equations in (2.12). Compressed sensing deals with undersampled cases in which $M < N$. The solution to (2.12) is ill-posed - there are infinitely many solutions. What is required is a constraint which originates from some a priori knowledge about the signal. Sparsity is such a priori information in the signals dealt with in this research.
Let $\mathbf{x}$ be a signal that has a $K-$ sparse representation on an orthonormal basis $\mathbf{\Psi}$ as in (2.2). Substituting (2.2) in (2.12) we have

$$\mathbf{f} = \mathbf{\Phi}\mathbf{\Psi}\mathbf{c} \tag{2.13}$$

$$\mathbf{f} = \mathbf{\Theta}\mathbf{c} \tag{2.14}$$

$$\text{where, } \mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi} \tag{2.15}$$

In subsequent discussions the matrix $\mathbf{\Theta}$ in (2.14) shall be called the **sensing matrix**. Applications, in which the signal exhibits sparsity in the information domain itself (for example, if the vector $\mathbf{x}$ is itself sparse) are special cases of (2.13) where $\mathbf{\Psi}$ is simply the identity matrix. Since the measurement process is linear and defined in terms of the matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$, solving for $\mathbf{c}$, given $\mathbf{f}$ in (2.14) is just a linear algebra problem, and with

[1] independent and identically distributed

$M < N$, there are fewer equations than unknowns, making the solution ill-posed, in general. If sparsity of $\mathbf{c}$ is the a priori information available, it is meaningful to obtain the sparsest solution of (2.14) by looking for a solution vector $\mathbf{c}$ with the minimal l_p-norm[1].

Minimization of l_2-norm ($p = 2$) - This is nothing but the least squares method, the classical approach of solving inverse problems.

$$\hat{\mathbf{c}} = argmin \left\| \mathbf{c}^{'} \right\|_2 \text{such that } \mathbf{\Theta c}^{'} = \mathbf{f} \tag{2.16}$$

Substituting $\hat{\mathbf{c}}$ in (2.2), an estimate $\hat{\mathbf{x}}$ of the original signal can be obtained. An even more convenient and equivalent solution involves pseudoinverse computation:

$$\hat{\mathbf{c}} = \mathbf{\Theta}^T (\mathbf{\Theta \Theta}^T)^{-1} \mathbf{f} \tag{2.17}$$

Although l_2 minimization is the very fast method, it is incorrect and returns a non-sparse $\hat{\mathbf{c}}$ with plenty of ringing.

Minimization of l_0-norm ($p = 0$) - l_0-norm reflects sparsity in the best possible way.

$$\hat{\mathbf{c}} = argmin \left\| \mathbf{c}^{'} \right\|_0 \text{such that } \mathbf{\Theta c}^{'} = \mathbf{f} \tag{2.18}$$

It can be shown [28; 29] that with just $M \geq K+1$ i.i.d. Gaussian measurements, this optimization will return a $K - sparse$ signal with probability one. Although l_0 minimization guarantees the most accurate results, it cannot be practically used as it is extremely slow. Solution of (2.18) is both numerically unstable [28] and an NP-complete problem that requires an exhaustive enumeration of all $\binom{N}{K}$ possible combinations for the locations of the nonzero entries in $\mathbf{c}$.

Minimization of l_1-norm ($p = 1$) - The compressed sensing paradigm offers a surprise in the form of l_1-norm minimization as a compromise between the fast and inaccurate l_2-norm based solution and the accurate and slow l_0-norm based one. It has been proved [30; 31] that with $M = CKlog(N/K)$

[1]The l_p-norm of a vector $\mathbf{v} \in \mathbb{R}^N$ is defined as $\|\mathbf{v}\|_p \triangleq \left(\sum_{j=1}^{N} |v_j|^p \right)^{1/p}$

independently and identically distributed (i.i.d.) Gaussian measurements, we can exactly reconstruct $K-$ sparse vectors and closely approximate compressible vectors stably with high probability via the l_1 minimization.

$$\hat{\mathbf{c}} = argmin \left\| \mathbf{c}^{'} \right\|_1 \text{ such that } \mathbf{\Theta c}^{'} = \mathbf{f} \tag{2.19}$$

This is a convex optimization problem that conveniently reduces to a linear program known as *Basis Pursuit* (BP) [32], whose computational complexity is about $O(N^3)$.

While introducing the concept of compressed sensing so far, expansion of the sparse signal has been restricted to orthonormal bases. It is important to note that this restriction is not mandatory [26] and the theory and practice of compressed sensing accommodates other types of expansions also. For example, the signal might be the coefficients of a digital image in a tight-frame of curvelets [33].

It is also pertinent here to make a comment on compressed sensing for analog signals. Fourier sparsity in the context of an analog signal implies that the signal can be represented using just K out of N elements of the continuous Fourier sinusoids. However, to facilitate simulations in a digital computer, one is compelled to make use of a discrete sparsity matrix. In support of this argument, the following from [28] is reproduced verbatim,

"*While we have focused on discrete-time signals* $\mathbf{x}$*, compressive sensing also applies to analog signals* $\mathbf{x}(t)$ *that can be represented sparsely using just* K *out of* N *possible elements from some continuous basis or dictionary* $\{\Psi_i(t)\}_{i=1}^{N}$*. While each* $\Psi_i(t)$ *may have large bandwidth (and hence a high Nyquist rate), the signal* $\mathbf{x}(t)$ *has only* K *degrees of freedom, and we can apply the above theory to measure it at a rate below Nyquist.*"

2.4 Stability of Reconstruction

Signals of practical interest may, in general, need not necessarily have a support of relatively small size either in space or in a transform domain. It may however, be possible that the support is only concentrated near a sparse set. Another model that is widely used in signal processing is that of signals in which the coefficients decay rapidly (compressible signals introduced in Chapter 1), typically following a power law. Examples of such signals are smooth signals, piecewise signals and images with bounded variations [34]. In addition, due to finite precision of sensing devices, the measured samples in any practical application will invariably be corrupted by at least a small amount of noise. In the presence of noise or absence of heavy sparsity, what is required is that the signals are reconstructed to the best possible approximation within a precision. In other words, the reconstruction should be stable - small perturbations in the signal caused by noise result in small distortions in the output solution. Clearly, it is not possible to reconstruct the signal if it is distorted during the measurement process itself and information is lost. To ensure that this is not the case, the measurement matrix $\mathbf{\Phi}$ and equivalently, the sensing matrix $\mathbf{\Theta}$ must satisfy certain conditions. To probe into this aspect a key notion, that has proved to be very useful in the study of the robustness of CS, is introduced. This is the so-called **restricted isometry property** (RIP) [35].

2.4.1 Restricted Isometry Property

Definition:. *For each integer $K = 1, 2, ...,$ the isometry constant δ_K of a matrix $\mathbf{\Theta}$ is defined as the smallest number such that*

$$(1 - \delta_K)\|\boldsymbol{c}\|_{l_2}^2 \leq \|\mathbf{\Theta}\boldsymbol{c}\|_{l_2}^2 \leq (1 + \delta_K)\|\boldsymbol{c}\|_{l_2}^2 \tag{2.20}$$

holds for all $K-$ sparse vectors, $\boldsymbol{c}$.

A matrix $\mathbf{\Theta}$ is said to obey RIP of order K if δ_K is small compared to one in which case $\mathbf{\Theta}$ approximately preserves the Euclidean length of $K-$ sparse signals, which in turn implies that $K-$sparse vectors cannot be in the null space of $\mathbf{\Theta}$, thereby making it possible to recover the sparse vector. Equivalently, for a matrix that has the restricted isometry property, every set of columns of cardinality less than K, is approximately orthogonal. Thus, an important requirement for stable reconstruction of the coefficient vector $\mathbf{c}$ and consequently the signal vector, $\mathbf{x}$ is that the sensing matrix, $\mathbf{\Theta}$ must obey RIP. The measured vector $\mathbf{f}$ in (2.14) is just a linear combination of the K columns of $\mathbf{\Theta}$ whose corresponding $c_i \neq 0$. Hence, if we knew a priori which K entries were nonzero, then we could form an $M \times K$ system of linear equations to solve for these nonzero entries, where now the number of equations M equals or exceeds the number of unknowns K. A necessary and sufficient condition to ensure that this $M \times K$ system is well-conditioned and hence sports a stable inverse is that for any vector $\mathbf{v}$ sharing the same K nonzero entries as $\mathbf{c}$, (2.20) is satisfied for a small δ_K. Of course, in practice, the locations of the K nonzero entries in $\mathbf{c}$ are not known. Interestingly, one can show [28] that a sufficient condition for a stable inverse for both $K-$ sparse and compressible signals is that $\mathbf{\Theta}$ must satisfy (2.20) for an arbitrary $3K-$ sparse vector $\mathbf{v}$.

2.4.2 Mutual Incoherence

Definition:. *The mutual coherence between two matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ is defined [26] as*

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \sqrt{N}.max_{k,j} \left|\langle \phi_k, \psi_j \rangle\right| \tag{2.21}$$

where ϕ_k are the rows of $\mathbf{\Phi}$ and ψ_j are the columns of $\mathbf{\Psi}$ and N is the number of basis vectors in $\mathbf{\Psi}$.

To put it simply, the mutual coherence measures the largest correlation between any two elements of $\mathbf{\Phi}$ and $\mathbf{\Psi}$. If $\mathbf{\Phi}$ and $\mathbf{\Psi}$ contain correlated ele-

ments, the coherence is large. Otherwise, it is small. Compressed sensing is mainly concerned with such sensing matrices $\mathbf{\Theta}$ that are constructed from pairs of the measurement and sparsity matrices that have low mutual coherence. Lower the mutual coherence, lesser the number of measurements required for stable reconstruction. For example [27], in classical sampling scheme in time and space, $\mathbf{\Phi}$ is the canonical spike basis $\phi_k = \delta(t-k)$ and $\mathbf{\Psi}$ is the Fourier basis, $\psi_i = (1/\sqrt{N})e^{\frac{j2\pi in}{N}}$ and $\mu(\mathbf{\Phi}, \mathbf{\Psi}) = 1$.

Another simple way to measure the coherence between $\mathbf{\Phi}$ and $\mathbf{\Psi}$ is to look at the columns of $\mathbf{\Theta}$, instead. As $\mathbf{\Theta} = \mathbf{\Phi\Psi}$, the mutual coherence can be defined as the maximum absolute value and normalized inner product between all columns [36] in $\mathbf{\Theta}$ which can be expressed as follows:

$$\mu(\mathbf{\Phi}, \mathbf{\Psi}) = \mu(\mathbf{\Theta}) = \underset{i \neq j, 1 \leq i,j \leq N}{max} \left\{ \frac{|\theta_i^T \theta_j|}{\|\theta_i\| \, . \, \|\theta_j\|} \right\} \tag{2.22}$$

Mutual coherence can also be computed [36] from the Gram matrix $\mathbf{G} = \bar{\mathbf{\Theta}}^T \bar{\mathbf{\Theta}}$ where $\bar{\mathbf{\Theta}}$ is the column-normalized version of $\mathbf{\Theta}$. In this case, $\mu(\mathbf{\Theta})$ is the maximum absolute off-diagonal element of $\mathbf{G}$.

$$\mu(\mathbf{\Theta}) = \underset{i \neq j, 1 \leq i,j \leq N}{max} |g_{ij}| \tag{2.23}$$

In some cases the average of the absolute value of the off-diagonal elements is also used[36].

$$\mu_{av}(\mathbf{\Theta}) = \frac{\sum_{i \neq j} |g_{ij}|}{N(N-1)} \tag{2.24}$$

Suppose that the following inequality holds,

$$K = \|\mathbf{c}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\mathbf{\Theta})}\right) \tag{2.25}$$

then the sparsest possible solution is guaranteed to be obtained[36] for the equations (2.13), (2.14) and (2.15).

Incoherence extends the duality between time and frequency and ex-

presses the idea that objects having a sparse representation in the sparsity domain ($\mathbf{\Psi}$) must be spread out in the information domain ($\mathbf{\Phi}$) in which they are acquired, just as a Dirac or a spike in the time domain is spread out in the frequency domain. In compressed sensing parlance, this notion is called the **Uniform Uncertainty Principle** (UUP). Put differently, incoherence says that unlike the signal of interest, the measurement waveforms $\{\phi_k\}$ cannot be sparsely represented by the vectors $\{\psi_j\}$ (and vice versa) which is another way of saying that there is very little mutual coherence between $\mathbf{\Phi}$ and $\mathbf{\Psi}$.

A natural question is how well one can recover a signal that is just nearly sparse. For an arbitrary vector $\mathbf{x}$ in $\mathbb{R}^N$, let $\mathbf{x}_K$ denote its best $K-$sparse approximation; that is, $\mathbf{x}_K$ is the approximation obtained by applying the inverse transform on $\mathbf{c}_K$ which is a vector formed by keeping the K largest entries of $\mathbf{c}$, the coefficient vector (in the sparsity domain) and setting the others to zero. It turns out [26] that if the sensing matrix obeys the uniform uncertainty principle at level K, then the recovery error is not much worse than $\|\mathbf{x} - \mathbf{x}_K\|_{l_2}$. In other words, the reconstruction is nearly as good as if one had full and perfect knowledge about the signal, and would extract the K most significant elements of the sparse signal.

2.4.3 Choosing the Right Measurement Matrix

Given a sparsifying basis $\mathbf{\Psi}$, is it possible to construct a measurement matrix $\mathbf{\Phi}$ such that $\mathbf{\Theta} = \mathbf{\Phi\Psi}$ has the RIP ? Unfortunately, even simple verification of RIP for a given $\mathbf{\Theta}$ is combinatorially complex. This involves verification of (2.20) for each of the $\binom{N}{K}$ possible combinations of K nonzero entries in the length$-N$ vector $\mathbf{c}$ [28]. In compressed sensing, this issue is avoided by choosing a **random matrix** for $\mathbf{\Phi}$. The restricted isometry property holds for sensing matrices $\mathbf{\Theta} = \mathbf{\Phi\Psi}$, where $\mathbf{\Psi}$ is an arbitrary orthonormal basis and $\mathbf{\Phi}$ is an $M \times N$ measurement matrix, satisfying RIP, that has entries drawn randomly from a suitable distribution. Thus,

random measurements $\mathbf{\Phi}$ are universal [37]in the sense that $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ has the RIP with high probability for every possible $\mathbf{\Psi}$. The sparsity basis need not even be known when designing the measurement system. One needs to confirm the RIP of $\mathbf{\Phi}$ and RIP of $\mathbf{\Theta}$ follows. Several random matrices that have been explored as candidate measurement matrices by researchers are presented here.

i) **Gaussian matrix**-Among the matrices that satisfy the RIP condition (2.20) are Gaussian random matrices consisting of elements drawn as i.i.d. random variables from a zero-mean, $1/N$-variance Gaussian density (white noise) [30; 31]. If $\mathbf{\Phi}$ is an M by N Gaussian random matrix where,

$$M \geq CKlog(N/K) \tag{2.26}$$

and C is a constant, then $\mathbf{\Phi}$ will obey the RIP with a high probability [38]. The proof of this result, using known concentration results about the singular values of Gaussian matrices, is involved and [39; 40] can be referred to for the same. If $\mathbf{\Phi}$ is a Gaussian random matrix with the number of rows satisfying RIP, then $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$, regardless of the choice of (orthonormal) sparsifying basis matrix, is also a Gaussian random matrix with the same number of rows and thus it satisfies RIP.

ii) **Binary matrix**-If the entries of the $M \times N$ measurement matrix, $\mathbf{\Phi}$ are independently drawn from the symmetric Bernoulli distribution, $P(\mathbf{\Phi}_{mn} = \pm\frac{1}{\sqrt{M}}) = 1/2$, then $\mathbf{\Phi}$ will satisfy the RIP.

iii) **Fourier measurements**- The partial Fourier matrix obtained by selecting M rows uniformly at random from the full Fourier matrix of order N and then re-normalizing the columns so that they are unit-normed is used as the measurement matrix $\mathbf{\Phi}$. Candes and Tao have showed in [41] that this construction of $\mathbf{\Phi}$ obeys the UUP.

iv) **Incoherent measurements**- A more general case of Fourier measurements is the measurement matrix $\mathbf{\Phi}$ obtained by selecting K rows

uniformly at random from an $N \times N$ orthonormal matrix U and re-normalizing the columns so that they are unit-normed. The arguments used in [41], to prove that the UUP holds for incomplete Fourier matrices, extend to this more general situation.

2.5 Robust Compressed Sensing

It is important to closely examine the notion of sparsity that has been discussed until this point, in the context of real world signals -

i) First, signals of practical interest possess only approximate sparsity. Very few signals are exactly sparse. Accurate reconstruction of such signals from highly undersampled measurements is an issue.

ii) Second, signals will invariably have measurement noise due to limited precision of the sensors. It is therefore imperative that CS be robust vis a vis such non-idealities.

In the presence of such non-idealities, the CS acquisition and reconstruction procedure must be robust. A small deviation from ideal behavior must not cause a drastic variation in the reconstruction. Fortunately, the recovery procedure may be adapted to be surprisingly stable and robust vis a vis arbitrary perturbations. The measurement process (2.13) is remodeled as follows:

$$\mathbf{f} = \mathbf{\Phi\Psi c} + \mathbf{e} \tag{2.27}$$

where $\mathbf{e}$ is a stochastic or deterministic error term with bounded energy $\|\mathbf{e}\|_{l_2} \leq \epsilon$. The reconstruction program is accordingly altered as

$$\hat{\mathbf{c}} = argmin \left\| \mathbf{c}^{'} \right\|_1 \text{such that } \left\| \mathbf{\Theta c}^{'} - \mathbf{f} \right\|_{l_2} \leq \epsilon \tag{2.28}$$

2.6 Greedy Reconstruction Algorithms

The l_1-minimization algorithm,solved with Basis Pursuit, has strong and rigorous theoretical proofs in support of exact reconstruction under quite general circumstances. However, it has been realized over the years by the compressed sensing research community that Basis Pursuit is much too slow for practical large-scale applications. Quite a few heuristic approaches based on greedy algorithms [1] have been proposed, that are many times faster, albeit without any theoretical guarantee of exact reconstruction. Perhaps the Orthogonal Matching Pursuit (OMP) algorithm proposed by Tropp and Gilbert [42] can be considered as the genesis of several of these greedy approximation methods. A brief outline of OMP is given in the next subsection.

2.6.1 Orthogonal Matching Pursuit

Consider the measured vector $\mathbf{f}$ introduced in (2.14). The vector, $\mathbf{f}$ is just a linear combination of K columns of the matrix $\mathbf{\Theta}$, given that $\mathbf{c}$ is a K-sparse vector. In other words, $\mathbf{f}$ has a K-term representation over the dictionary $\mathbf{\Theta}$. To recover the sparse vector $\mathbf{c}$, which in turn would give the actual signal $\mathbf{x}$ from (2.2), it is required to determine which columns, θ_j of $\mathbf{\Theta}$ participate in $\mathbf{f}$. OMP (see Algorithm 1) picks the columns iteratively, in a greedy fashion. The vector $\mathbf{f}$, that is input to the algorithm, is obtained by sensing the signal $\mathbf{x}$ using the measurement matrix $\mathbf{\Phi}$(see equation 2.12). At each iteration, a column of $\mathbf{\Theta}$ is chosen that is most strongly correlated with the remaining part of the residual. The contribution to $\mathbf{f}$, due to the chosen column, is subtracted from $\mathbf{f}$ and the residual is input to the next iteration.

In this manner, it is expected, without any theoretical guarantee, that after K iterations the algorithm would have identified the correct set of

[1] A greedy algorithm is an algorithm that follows the problem solving heuristic of making the locally optimal choice at each stage with the hope of finding a global optimum.

Algorithm 1 Orthogonal Matching Pursuit

Input: The $M \times N$ sensing matrix $\mathbf{\Theta}$, the $N \times N$ sparsity matrix $\mathbf{\Psi}$, the measured vector $\mathbf{f}$ and K the required number of iterations, the algorithm must execute

Output: An estimate $\hat{\mathbf{x}}$ of the signal vector

Procedure:

1. Initialize:
   the residual, $\mathbf{r}^{(0)} \leftarrow \mathbf{f}$
   the set of indices, $\Lambda^{(0)} \leftarrow \emptyset$
   matrix of atoms, $\mathbf{\Theta}^{(0)} \leftarrow \emptyset$
   iteration-count, $i \leftarrow 1$
2. At every iteration i, find the index $\lambda^{(i)}$ that solves the optimization problem,
   $\lambda^{(i)} \leftarrow argmax_{j=1\ldots N} \left|\left\langle \mathbf{r}^{(i-1)}, \theta_j \right\rangle\right|$ If the maximum occurs for multiple indices, break the tie deterministically.
3. Augment the index set $\Lambda^{(i)} \leftarrow \Lambda^{(i-1)} \cup \lambda^{(i)}$ and the matrix of chosen atoms, $\mathbf{\Theta}^{(i)} \leftarrow \left[\mathbf{\Theta}^{(i-1)} \;\; \theta_j\right]$
4. Solve a least-squares problem to obtain a new estimate of the coefficients:
   $\hat{\mathbf{c}}^{(i)} \leftarrow argmin_{\mathbf{c}'} \left\|\mathbf{\Theta}^{(i-1)}\, \mathbf{c}' - \mathbf{f}\right\|_2$
5. Update the residual:
   $r^{(i)} \leftarrow \mathbf{f} - \mathbf{\Theta}^{(i-1)}\, \hat{\mathbf{c}}^{(i)}$
6. **if** $i < K$ **then**
7. $\quad i \leftarrow i + 1$
8. **else**
9. $\quad$ Form the length$-N$, sparse coefficient vector, $\mathbf{c}$ populated with the elements of $\hat{\mathbf{c}}^{(i)}$ at the appropriate locations
10. $\quad \hat{\mathbf{x}} \leftarrow \mathbf{\Psi c}$
11. $\quad$ quit
12. **end if**

columns. It is important to note in the algorithm listing that $r^{(i)}$ is always orthogonal to the columns of $\mathbf{\Theta}^{(i)}$. The computational complexity of the algorithm is dominated by step 2 whose total cost is $O(KMN)$. The least squares problem in step 4 at iteration i can be solved with marginal cost of $O(iM)$. In comparison the basis pursuit algorithm (see section 2.3), using a dense unstructured sensing matrix, can be solved in time $O(M^2N^{3/2})$ time [43]. Thus in cases where N is much larger than K or M, OMP has clear advantage over BP in terms of speed of computation.

2.6.2 Other Greedy Algorithms

For many applications, OMP can outperform convex optimization methods. For large problems, in which the number of non-zero elements is of the order of several thousands or more, the computational requirements and storage demands of currently available implementations of OMP can easily become too large, and faster alternatives are required. A number of recovery algorithms that are based on OMP and offer some kind of performance enhancement, have been proposed. Significant amongst these are given below:

i) **Stage-wise Orthogonal Matching Pursuit** - In the Stage-wise Orthogonal Matching Pursuit (StOMP) algorithm proposed by Donoho et. al. [44] many coefficients can enter the model at each stage while only one enters per stage in OMP; and StOMP takes a fixed number of stages while OMP can take many. StOMP runs much faster than competing proposals for sparse solutions, such as l_1-minimization and OMP, and so is attractive for solving large-scale problems.

ii) **Gradient Pursuit** - Blumensath and Davies [45] have proposed directional optimization schemes based on - gradient, conjugate gradient and approximation to conjugate gradient. While conjugate gradient solves the Orthogonal Matching Pursuit (OMP) algorithm exactly, the

evaluation of this direction has the same computational complexity as previous implementations of OMP. The gradient as well as the approximate conjugate gradient is much easier to calculate, with the gradient being available in Matching Pursuit (MP) for free.

iii) **Regularized Orthogonal Matching Pursuit (ROMP)** - Needell and Vershynin tried to bridge the two major algorithmic approaches to sparse signal recovery from an incomplete set of linear measurements: l_1- minimization methods and iterative methods (Matching Pursuits) via a simple regularized version of the Orthogonal Matching Pursuit [46]. ROMP has advantages of both approaches: the speed and transparency of OMP and the strong uniform guarantees of the l_1-minimization and reconstructs a sparse signal in a number of iterations linear in the sparsity (in practice even logarithmic), and the reconstruction is exact provided the linear measurements satisfy the Uniform Uncertainty Principle.

iv) **Compressive Sampling Matching Pursuit (CoSAMP)** - CoSAMP is at heart a greedy pursuit that incorporates ideas from the combinatorial algorithms to guarantee speed and to provide rigorous error bounds [47].

v) **Tree Matching Pursuit** - An algorithm to recover piecewise smooth signals that are sparse and have a distinct connected tree structure in the wavelet domain has been proposed by Duarte, Wakin and Baraniuk [48]. The Tree Matching Pursuit (TMP) algorithm significantly reduces the search space of the traditional Matching Pursuit greedy algorithm, resulting in a substantial decrease in computational complexity for recovering piecewise smooth signals.

vi) **Chaining Pursuit**- Given the original signal $\mathbf{f}$ is well-approximated by a vector with K non-zero entries (spikes), the goal of the Chaining Pursuit algorithm, proposed by Gilbert et. al. [49], is to use a

sketch of the signal to obtain a signal approximation with no more than K spikes. To do this, the algorithm first finds an intermediate approximation $\mathbf{g}$ with possibly more than K spikes, then, in the so called *pruning step*, returns $\mathbf{g}_K$, the restriction of $\mathbf{g}$ to the K positions that maximize the coefficient magnitudes of $\mathbf{g}$. In each pass, the algorithm identifies the locations of a constant fraction of the remaining spikes and estimates their magnitudes. Then it encodes these spikes and subtracts them from the sketch to obtain an implicit sketch of the residual signal. These steps are repeated until the number of spikes is reduced to zero. After $O(log\, m)$ passes, the residual has no significant entries remaining. The run time of Chaining Pursuit, namely, $O(K\, log^2(K) log^2(N))$, for an $N-$ length signal of sparsity level K is sub-linear in N.

vii) **Subspace Pursuit**-In the Subspace Pursuit algorithm proposed by Wei Dai [50], a set of K (for a K-sparse signal) codewords of highest reliability that span the code space are first selected. If the distance of the received vector to this space is deemed large, the algorithm incrementally removes and adds new basis vectors according to their reliability values, until a sufficiently close candidate code word is identified. The algorithm has two important characteristics: low computational complexity, comparable to that of orthogonal matching pursuit techniques, and reconstruction accuracy of the same order as that of l_1 optimization methods.

viii) **Simultaneous Orthogonal Matching Pursuit** - Parallel sparse approximation of several input signals that are only weakly correlated has been proposed by Gilbert and Strauss in Simultaneous Orthogonal Matching Pursuit (SOMP) [51].

2.7 Other Recovery Algorithms

In this section, a brief description of a few algorithms, that have drawn the attention of CS research community, is presented for the sake of completion.

i) **Sudocodes**- The method based on sudocodes proposed by Sriram Sarvotham [52] involves non-adaptive construction of a sparse measurement matrix, comprising only the values 0 and 1. Only $O(Klog(N))$ measurements are constructed by summing subsets of the coefficient values of the sparse vector, like in group testing. The reconstruction process receives a stream of measurements and the corresponding rows of the measurement matrix. It has a low worst-case computational complexity of $O(Klog(K)log(N))$.

ii) **Bayesian Compressed Sensing** - Considerable literature [53] has been published in the area of Bayesian compressed sensing [54] that can be considered as a shift in paradigm from the classical compressed sensing. Algorithms based on Bayesian CS, start with a prior belief on the sparsity of the signal in a suitable basis and the objective is to provide a posterior belief (density function) for the values of the elements of the sparse vector. Rather than providing a point (single) estimate of the sparse coefficients, a full posterior density function is provided, which yields error bars on the estimated vector; these error bars may be used to give a sense of confidence in the approximation, and they may also be used to guide the optimal design of additional CS measurements, implemented with the goal of reducing the reconstruction uncertainty. In addition, the Bayesian framework provides an estimate for the posterior density function of additive noise encountered when implementing the compressive measurements.

iii) **Distributed Compressed Sensing** - Considerable work has been done in the area of distributed compressed sensing based on joint

sparsity of multiple-signal ensemble. Literature in this area is cited in a subsequent chapter in which work done, as part of this thesis, on compressed sensing of several correlated signals is described.

iv) **l_p-norm minimization** - Chartrand has proved that instead of minimizing the l_1-norm, if the l_p-norm, where $p < 1$, is minimized then exact reconstruction can be obtained with lesser number of measurements [55].

2.8 The Compressed Sensing 'Tuple'

To apply compressed sensing techniques to various real world problems, it is worthwhile to introduce the '*CS-tuple*' which consists of the following:

i) $\mathbf{\Psi}$ - A **sparsity inducing basis** in which the signal to be acquired has a sparse representation. If the signal is sparse in the information domain itself, then $\mathbf{\Psi} = \mathbf{I}$, the identity matrix.

ii) $\mathbf{\Phi}$ - A suitable **measurement matrix** that is mutually incoherent with $\mathbf{\Psi}$.

iii) Ξ - A suitable **signal recovery algorithm** like Basis Pursuit or Orthogonal Matching Pursuit.

In order to address any sparse recovery problem through the compressed sensing approach, it is required to carefully understand the signal characteristics, the desired accuracy of reconstruction, the expected benefits out of using a CS approach as against a traditional Nyquist scheme and the available hardware to realize the acquisition setup. Based on such an analysis, one has to prudently choose the elements of the tuple, which almost always will be application specific. The schemes that have been proposed in this research are general and therefore, independent of the reconstruction algorithm used. In other words, any reconstruction method that is

not too slow can be used to recover the original signal from the measurements taken. Thus, before proposing an embedded design for compressed acquisition, the '*CS-tuple*' is formed on the basis of which the system is architected.

2.9 Choice of Reconstruction Algorithm

Performance of most of the signal reconstruction schemes that have been proposed in this work are independent of the signal recovery algorithms. In all the simulations, the basis pursuit tool is used to recover the signal, since the primary interest is on the reconstruction accuracy and execution time is not a concern on the computational platform employed for simulations. However, in some cases the proof of performance of the proposed schemes is demonstrated through simulations with other signal recovery tools - OMP, ROMP and CoSAMP and comparison of reconstruction accuracy and execution time is given in tabular form.

2.10 Areas of Research in Compressed Sensing

Research in compressed sensing has been multi-faceted and a number of open issues have been parallelly addressed by the CS community. A bird's eye view of the previous as well as ongoing effort is given in what follows:

i) **Design of measurement matrices** - Focused effort has gone into the design of more effective measurement matrices that are able to capture the signal information with lesser number of samples. In particular, deterministic construction of the matrix [56; 57; 58; 59], wherein the entries are not entirely drawn at random, has been studied in detail to suit specific applications. An open problem that has been investigated is the perturbations/transformations that are admissible on the measurement matrices.

ii) **Development of recovery algorithms** - Computation time and accuracy of reconstruction have been the two major driving factors in the development of newer recovery algorithms. At one end of the spectrum, there exist the convex optimization algorithms that offer a high level of accuracy, but tend to become unmanageably slow as the signal dimension increases. At the other extreme are the greedy methods like OMP which are significantly faster but do not guarantee any bounded reconstruction error. Reconstruction algorithms that have been proposed by various people have tried to find a trade-off - design of more accurate greedy reconstruction methods or alternatively, faster algorithms based on convex minimization. As long as the field of compressed sensing is active, design of newer reconstruction algorithms will continue to happen.

iii) **Restricted Isometry Property of the sensing matrix** - Establishing RIP of candidate sensing/measurement matrices through non-combinatorial complexity algorithms is an active area of research.

iv) **Sparsity behavior of signals** - Study of the existence and type of sparsity of practical, real world signals is a field of interest which opens up the possibility of employing compressed sensing techniques to sub-Nyquist acquisition of such signals.

v) **Efficient sampling architecture**- Design of efficient and practically realizable architecture for compressed acquisition of sparse signals is a problem that demands serious attention for development of commercially viable systems using compressed sensing methods.

vi) **Streaming data acquisition and signal recovery** - Classical compressed sensing deals with applications in which the sub-Nyquist measured vector is available and one has to reconstruct the original signal vector. Recently, significant work [60; 61] has been done towards

acquisition of real-time streaming data and simultaneous signal reconstruction.

vii) **Applications** - Compressed sensing approaches have been applied in numerous areas of engineering - compressive imaging, analog-to-information conversion, geophysical data analysis, radar, hyper-spectral imaging, medical imaging, communications, astronomy, remote sensing, acoustics, audio and speech processing, robotics, control, content-based retrieval and optics to name a few.

Exhaustive information consisting of a large collection of research papers, tutorials, talks, software and toolboxes, and conference/workshop alerts is available in [53].

2.11 Goal of this research

With the necessary background on sparse signal acquisition reviewed, it is pertinent at this juncture to re-emphasize the goal of this research with minimum ambiguity.

What this research is about

- This research effort is primarily focused on the design of compact embedded data acquisition systems which exploit the sparsity characteristic of signals, captured in real time.
- The emphasis is more on how to engineer efficient and cost-effective, front-end analog-to-digital conversion electronics for space and power constrained embedded designs using available off-the shelf components. The research that has been carried out is 'applied' in nature, dealing with a practical problem to give meaningful solutions to the same.

What this research is not about

- As part of this research effort, there is no motive to probe into the intricacies of compressed sensing and reconstruction and the associated algorithms, which is a vast area in itself and is supported by a lot of literature. The designs proposed in this work employ compressed sensing acquisition and reconstruction simply as a tool. The success or failure of the designs are dependent on the accuracy and robustness of the compressed sensing methodology which itself is supported by strong theoretical proofs of performance, given that the signals are sparse. In all the simulations carried out in this research, it is ensured that the test signals satisfy the sparsity constraint imposed by compressed sensing. Therefore, it is felt that there is no necessity to give separate theoretical proofs of reconstruction performance of the methods proposed in the thesis.

- The ideas proposed in this work have more to do with the signal acquisition phase rather than signal processing. They demonstrate how a set of sparse signals can be acquired by a system using lesser number of analog to digital conversions. How the signals are processed, once they enter the digital world, does not fall in the gamut of this research.

- Through simulations, it is intended to demonstrate in this work, how with simple modifications in the data acquisition architecture, it is possible to capture multiple sparse and streaming signals using lesser number of hardware components. It is submitted here that the effort is restricted to providing empirical evidence and actual hardware design based on CAD tools has not been carried out. However, it is anticipated that the implementation of the proposed schemes would be reasonably straightforward and given sufficient proof of performance through simulations, the schemes should work fairly well when implemented in hardware.

To the knowledge of the author, there is no instance of research in the open literature on the application of compressed sensing of sparse signals from the perspective of designing compact embedded data acquisition architectures. Literature on compressed sensing applications gives numerous cases of subsampling sparse signals in various engineering domains. There is not much evidence of smart utilization of ADCs by multiplexing them between different signal sources that generate streaming data. On account of this, it was not possible to compare the results with previous work on the same problem. Consequently, it was not feasible to place a consolidated survey of literature, on existing approaches to the problem, at the beginning of the thesis. However, any (even partially) borrowed facts, figures, information or ideas used in the work have been duly cited throughout the body of the thesis. Thus, citations to the relevant literature have been given at the beginning of each chapter.

Chapter 3

Compressed Acquisition of Multiple Sparse Signals

3.1 Introduction

Embedded systems in practical applications typically perform real time data acquisition of more than one signal. Biomedical signals, seismic signals, flight parameters in aerospace systems, distributed sensor data are examples in which the signals that have to be acquired occur in multitude. Quite often such signals can be grouped based on similarity of characteristics. Each group, for instance, may comprise signals that exhibit sparsity in a particular basis. Assume that the ADC assigned to each signal in every group, is able to sample the signal at its Nyquist rate. Given that the signals are sparse, it is possible to perform the sub-Nyquist acquisition of each signal, in which case the ADC responsible for sampling the signal would operate at much lower than its maximum possible sampling rate. In other words, not all the sampling cycles of the ADC are utilized. Is it possible to utilize the idle cycles of the ADC to sample other signals in the group such that a single ADC is used to acquire more than one signal in a compressed sensing setup ? The work done to explore this possibility, as part of this research, is the content of this chapter. Reducing the number of ADCs at the front end of data acquisition systems will be the first step towards achieving the desirable features in embedded architectures that

were highlighted in Chapter 1.

3.2 Survey of related literature

Efficient sampling of sparse bandlimited wideband signals has been treated at length in [62]. The approach proposed in this paper employs a random demodulator for sampling, which in turn consists of a random number generator, mixer and an accumulator. In [63], the authors have presented the design of a digital compressed sensing architecture employing a mixer, integrator and a random matrix generator. In [64] the authors propose a system, named the modulated wideband converter, which first multiplies the analog signal by a bank of periodic waveforms. The product is then lowpass filtered and sampled uniformly at a low rate, which is orders of magnitude smaller than Nyquist. In [65] the authors propose an analog-to-information converter that uses modulation, filtering, and sampling to produce a low-rate set of digital measurements.

In contrast to the methods proposed in these cited papers, the data acquisition schemes presented in this thesis and this chapter, in particular, operate simultaneously on multiple signals. The primary focus of the research reported in this thesis is to propose data acquisition architectures with simple off-the-shelf components and employ a straightforward sampling scheme, wherein a few random samples are utilized to reconstruct the input signals without involving any elaborate projection matrix and the associated multiplication with the signal vector. In this context, a remark made by the authors in [66] is reproduced verbatim:
"To date, implementing random projections in an analog-to-digital data acquisition device is not straightforward. The most common architectures use an analog high-rate front end implementing the random projections, followed by a low-rate precision analog-to-digital converter decoupled from the analog projection system. Due to the limitations of analog hardware,

the resulting devices are severely limited in the types of projections they can implement. Some systems are based on randomly modulating or randomly sampling the input. Systems that randomly filter and subsequently subsample the input require a significant number of precision analog multipliers or a high-order switched capacitor filter"

3.3 Signal Model

Consider a stationary signal $\mathbf{x}$, that is composed of only K sinusoids where K is small compared to the length of the signal. Such signals are known as *frequency-sparse*. In practice, many real world signals are only piecewise-stationary. Therefore, let $\mathbf{x}(t)$ be a continuous-time, real signal, made up of segments, $\chi_s(t)$ such that

$$\mathbf{x}(t) = \chi_s(t) \text{ for } t_{s-1} \leq t < t_s, \ \forall s = 1, 2, ...\infty \tag{3.1}$$

where each $\chi_s(t)$ is stationary and also sparse on the Fourier basis. Further, let each $\chi_s(t)$ be nearly bandlimited to $[0\ F]$. Clearly, the Nyquist rate of $\mathbf{x}(t)$ is $F_{NYQ} = 2F$. Let any $\chi_s(t)$ have at most K frequency components, ω_k such that $0 \leq \omega_k \leq 2\pi F$. K is much smaller than the number of Nyquist samples in the signal during the interval $l_s = [t_{s-1}\ t_s]$.

$$K \ll l_s F_{NYQ} \tag{3.2}$$

It is important to mention here, that a priori knowledge of K is helpful in estimating the minimum number of measurements (given by 2.26 for Gaussian measurements) required to reconstruct the signal.

Signals like $\mathbf{x}(t)$ shall hereafter be called **piecewise stationary and sparse** (PSS) signals. Thus, each PSS signal is a concatenation of a series of **stationary and sparse** (SS) segments, each of which consists of a different set of frequencies in $[0\ F]$. Real world non-stationary and non-sparse signals can be approximated to PSS signals if each of the intervals l_s

is small compared to the duration of the signal under observation. For example, speech signal, which is non-stationary, is routinely approximated as piecewise stationary in linear prediction analysis as well as in spectrogram based analysis[67].

Although the signal model presented above considers Fourier sparsity, the compressed sensing schemes proposed in this thesis can be used on signals with sparsity on any basis.

3.4 Sampling and Reconstruction of the Signal

Let $\zeta \in \mathbb{R}^N$ be a vector obtained by sampling the signal, within an SS segment $\chi_s(t)$, during the time interval $[\tau_1, \tau_2]$ where $t_{s-1} \leq \tau_1, \tau_2 < t_s$ for some s, at the Nyquist time instants, $\tau_1, \tau_1 + T_{NYQ}, ...\tau_1 + (N-1)T_{NYQ}$ where $T_{NYQ} = 1/F_{NYQ}$. Hereafter, such a vector shall be called a **reconstruction segment** (RS)[1]. Each SS segment will consist of several RSs which may overlap. Instead of taking all the Nyquist samples within an RS, if only $M < N$ samples are taken at random, the measured vector, $\mathbf{f}$ (2.12) can be written as

$$\mathbf{f} = \mathbf{\Phi}\zeta \tag{3.3}$$

where the $M \times N$ measurement matrix $\mathbf{\Phi}$ is constructed by randomly choosing M rows from the $N \times N$ identity matrix $\mathbf{I}_N$. A later sub-section explains why M must have a minimum value. Since each RS, ζ is of finite length, the frequency domain tool of choice for analysis and CS recovery is the Discrete Fourier Transform (DFT)[2].

This process is very similar to the **Windowed Fourier Transform (WFT)** or the **Short Time Fourier Transform (STFT)** which is used to acheive limited time-frequency localization. In order to improve the performance

[1] An implicit assumption made here is that an RS consists of an integral number of periods of each frequency component in the signal. The validity of this assumption is questioned in sub-section 3.5.8 and the means of addressing the issue for general signals is addressed in section 3.6

[2] The DFT of a length-N signal $\mathbf{x}$ is defined as $X[m] = \sum_{n=1}^{N} x[n] e^{-j2\pi mn/N}, 1 \leq m \leq N$, with the inverse transformation defined as $\mathbf{x}[n] = \frac{1}{N} \sum_{l=1}^{N} x[n] e^{j2\pi mn/N}, 1 \leq n \leq N$.

of the STFT, a tapered window (e.g. Kaiser, as used on speech signals to get the spectrogram) is made use of (see sub-section 3.5.2). The factors affecting the choice of the STFT window size are discussed in sub-section 3.5.3.

Given that $\mathbf{\Psi} = \boldsymbol{F}^{-1}$, the inverse DFT matrix is the sparsity basis, equations (2.13-2.15) can then be recast as

$$\mathbf{f} = \mathbf{\Phi}\boldsymbol{F}^{-1}\mathbf{c} \tag{3.4}$$

$$\mathbf{f} = \mathbf{\Theta}\mathbf{c} \tag{3.5}$$

$$\text{where, } \mathbf{\Theta} = \mathbf{\Phi}\boldsymbol{F}^{-1} \tag{3.6}$$

and $\mathbf{c}$ is the sparse vector of DFT coefficients. In order to account for measurement and other noise, (3.4) is modified as in (2.27)

$$\mathbf{f} = \mathbf{\Phi}\boldsymbol{F}^{-1}\mathbf{c} + \mathbf{e} \tag{3.7}$$

where, as before $\mathbf{e}$ is a stochastic or deterministic error term with bounded energy $\|\mathbf{e}\|_{l_2} \leq \epsilon$ and accordingly the solution is obtained in the lines of (2.28)

$$\hat{\mathbf{c}} = argmin\left\|\mathbf{c}^{'}\right\|_1 \text{such that } \left\|\mathbf{\Theta}\mathbf{c}^{'} - \mathbf{f}\right\|_{l_2} \leq \epsilon \tag{3.8}$$

$$\text{and finally, } \zeta = \boldsymbol{F}^{-1}\hat{\mathbf{c}} \tag{3.9}$$

where $\mathbf{\Theta}$ is as given in (3.6). To facilitate a practical implementation of (3.8), the following definition for any general vector is introduced.

Definition:. *The root mean square value of a vector, $\boldsymbol{v} \in \mathbb{R}^D$ is defined as*

$$rms(\boldsymbol{v}) = \sqrt{\frac{v_1^2 + v_2^2 + ... + v_D^2}{D}} \tag{3.10}$$

where $v_1, v_2, ...$ are the elements of $\boldsymbol{v}$

Let the signal to measurement noise ratio be given as

$$SMNR = 20log_{10}(\frac{rms(\mathbf{f})}{\epsilon_{rms}^{max}}) \tag{3.11}$$

where ϵ_{rms}^{max} is the upper limit of the *rms* noise. Then for a specified SMNR value,

$$\epsilon_{rms}^{max} = rms(\boldsymbol{f})/10^{(\frac{SMNR}{20})} \tag{3.12}$$

Instead of (3.8), we can use the alternate constraint,

$$\frac{1}{\sqrt{M}}\left\|\boldsymbol{\Theta}\mathbf{c}^{'} - \mathbf{f}\right\|_2 \leq \epsilon_{rms}^{max} \tag{3.13}$$

where M is the number of elements in the vector, $\left(\boldsymbol{\Theta}\mathbf{c}^{'} - \mathbf{f}\right)$.
Or equivalently from the definition of *rms* above,

$$\hat{\mathbf{c}} = argmin\left\|\mathbf{c}^{'}\right\|_1 \text{such that } rms(\boldsymbol{\Theta}\mathbf{c}^{'} - \mathbf{f}) \leq \epsilon_{rms}^{max} \tag{3.14}$$

3.4.1 Simple Measurement Matrix

It is worth noting here that no specific measurement matrix has been used. The measured vector is obtained by just randomly picking some of the Nyquist samples. This kind of a sampling is represented in the reconstruction equations by a matrix that shall be informally called the 'downsized identity matrix'. This is very unlike using matrices with entries randomly drawn from a distribution like Gaussian or Bernoulli (see sub-section 2.4.3) which requires an elaborate setup for capturing the information domain signal as projections on the measurement matrix.

The author of this thesis is aware that the downsized identity matrix is not supported by a theoretical justification of satisfying RIP or of mutual incoherence with the sparsity matrices employed in the various schemes proposed. In spite of this, it has been used in this work owing to its practical realizability - undersampling just reduces to omitting, uniformly at random, the Nyquist samples in a signal. Verification of proof of per-

formance using this matrix is left to observing the results of simulation with synthetic test data as well as real world signals. It has already been pointed out in sub-section 2.4.3 that verification of RIP is combinatorially complex and has only been done for a few random matrices like Gaussian or Bernoulli. Appendix C lists out arguments, supported by citations, against the practical implementation of fully random measurement matrix constructions like Gaussian or Bernoulli matrices. In addition, it is empirically shown in the Appendix, through a large number of repeated trials, that the probabilty of the downsized identity matrix having a lower mutual coherence with a fixed sparsity matrix is comparable with that of a random matrix of the same dimension[1].

3.4.2 The CS Tuple

Before proceeding further, the elements of the CS tuple that was introduced in section 2.8 must be identified to put the problem formulation in place. Clearly,

- $\boldsymbol{\Psi}$ is the inverse DFT matrix $\boldsymbol{F}_N^{-1}$ of order N
- $\boldsymbol{\Phi}$ is a matrix obtained by randomly choosing M out of the N rows of an identity matrix, $\mathbf{I}_N$ of size N.
- Ξ is the noise constrained l_1 minimization (3.8)

3.4.3 Overlapped Reconstruction Segments

The signals under consideration in this work are input to the system as real-time streaming data that is piece-wise stationary and comprises a series of sparse and stationary segments. It is discernible that around the transition between consecutive SS segments, the stationarity assumption

[1]If the outcome of a random undersampling turns out such that the samples are uniformly spaced, then aliasing can occur. However, it is assumed that the probability of such an event is very small when the number of samples in the reconstruction segment is not a small number.

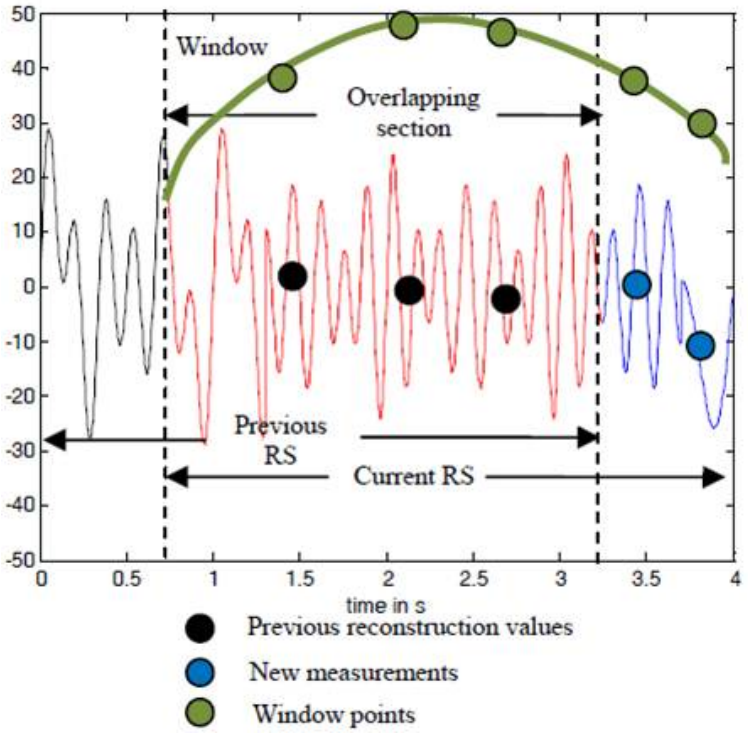


Figure 3.1: Overlapped reconstruction segments

is lost. In other words, a reconstruction segment in which a boundary between two SS segments lies will not necessarily exhibit Fourier sparsity. As a consequence, the CS based reconstruction of such an RS will fail. As a remedy to this problem, reconstruction segments are made to overlap (figure 3.1). Thus, each stationary and sparse segment is acquired and reconstructed as a series of overlapping reconstruction segments. Every new RS will have an overlap (specified as a percentage p of the size N of the RS) with the previous RS that has already been reconstructed. In the region of overlap, $num_O = pN$ samples obtained from reconstruction of the previous RS are utilized as they are. In the remaining non-overlapping portion, num_{NO} sub-Nyquist samples are obtained by actual analog to digital conversion. Clearly, $num_O + num_{NO} \geq M$. Discussion on the appropriate value of p and num_{NO} is deferred to a later subsection. CS reconstruction is then done on the vector comprising the overlapping and non-overlapping samples to recover all the Nyquist samples in the RS. For all the RSs that

lie fully inside an SS segment, the coefficient vector, $\hat{\mathbf{c}}$ in (3.14) will be sparse. At the transition to a new SS segment, the RS (say the j^{th} one) will span two SS segments. As a result, the reconstruction will fail and the coefficient vector will be non-sparse and in general, be different from that of the preceding, $(j-1)^{th}$ RS. Therefore,

$$\|\hat{\mathbf{c}}_j - \hat{\mathbf{c}}_{j-1}\|_2 > \delta \tag{3.15}$$

where δ is a small threshold[1]. Using (3.15), the boundary between two RSs can be detected. Once a transition is found, the RS that immediately follows will not overlap with its predecessor, which is the one that had fallen on a boundary and therefore suffered from a failed reconstruction. Without an overlap, the RS then uses all the samples obtained from actual analog to digital conversion. All the RSs that follow will continue to have overlap.

By ensuring that the RSs overlap, the portion of the reconstructed signal that is erroneous due to occurrence of boundaries is reduced to $(1 - \frac{p}{100})N$. The parameter p, an input to the system, decides the accuracy of reconstruction.

In the next section, an architecture for carrying out multiplexed acquisition of multiple sparse signals is presented.

3.5 MOSAICS: Multiplexed Optimal Signal Acquisition Involving Compressed Sensing

If only sub-Nyquist number num_{NO} of non-overlapping samples are taken, then the remaining Nyquist sampling instants of the ADC can be utilized to capture other signals with similar characteristics. This is the central idea of the procedure for multiplexed sensing of more than one signal. In this section, the operating principle of MOSAICS is explained.

[1]chosen specifically for each application

3.5.1 System Input

Let **S** represent a set of γ signals, $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ...\mathbf{x}^{(\gamma)}$ of the model described in section 3.3 to be input to the system. Let the signals be indexed by i. The following a priori information about each signal $\mathbf{x}^{(i)}$ is assumed to be available:

- Bandwidth $F^{(i)}$ and consequently the Nyquist rate, $F_{NYQ}^{(i)} = 2F^{(i)}$. This decides the operating frequency F_{MOS}, more about which is given in subsection 3.5.3.

- Fourier sparsity, $K^{(i)}$, i.e. the maximum number of non-zero frequency tones in an SS segment of the signal. Only an estimate of the maximum number of frequency tones and not the exact number need to be known [1].

- Minimum duration of an SS segment, $L^{(i)}$ where L for a PSS signal is defined as $L = inf\left\{l_s \in \mathbb{R} : l_s = t_s - t_{s-1}, \forall s = 1, 2, ...\infty\right\}$ where t_{s-1} and t_s are the starting and end points of a SS segment. Thus, $L^{(i)}$ is the minimum duration in which the signal can be assumed stationary. This a priori knowledge aids in deciding the size of a reconstruction segment as explained in sub-section 3.5.3 and the segmentation is not totally blind. However, only an estimate of the minimum assured duration of stationarity, (out of all SS segments) and not the exact value, is required.

- The number of samples in the reconstruction segment, $N^{(i)}$ and the number of measurements, $M^{(i)}$ in each reconstruction segment. These are derived parameters. Explanation of how these are derived is deferred to the next sub-section as this requires a description of the steps in the acquisition and reconstruction.

[1] For any signal i, the sparsity may vary from one SS segment to the other and $K^{(i)}$ can be taken to be equal to an estimate of the maximum value out of all such segments

- Maximum allowed duration $\mu^{(i)}$ of erroneous reconstruction at the transitions between SS segments. This decides the overlap percentage (subsection 3.5.3).
- signal to measurement noise ratio of the ADC that is used

While the list appears to be formidable, for many real world signals, this a priori information is easily available.

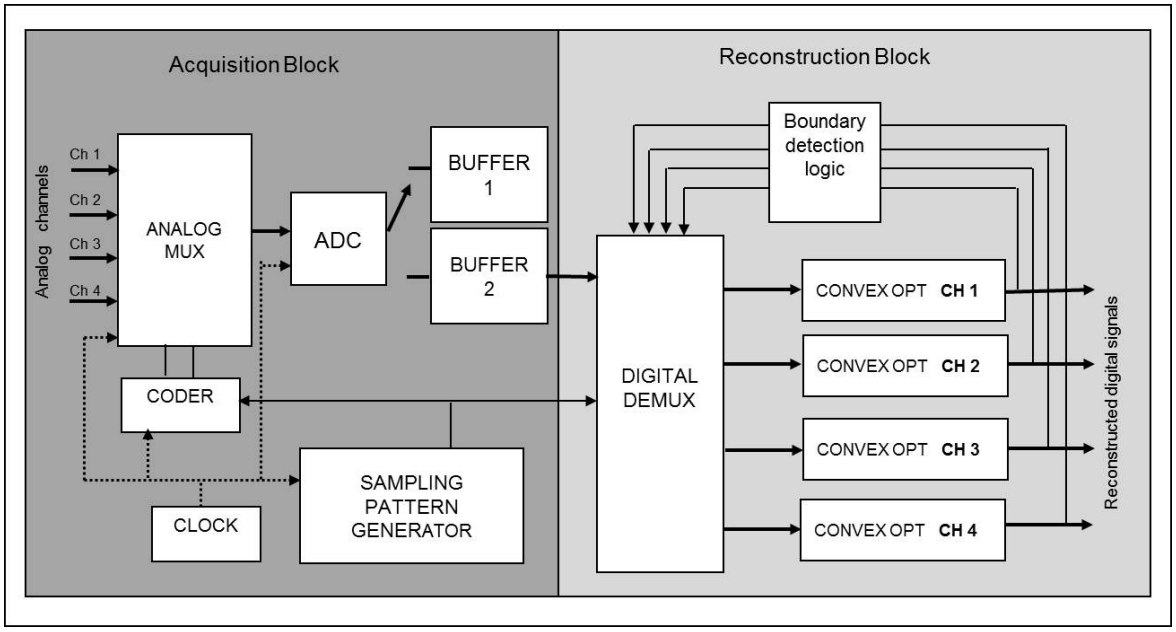


Figure 3.2: Multiplexed Signal Acquisition Involving Compressed Sensing

3.5.2 Steps in Execution

As shown in the figure 3.2, the system consists of two major blocks - the *acquisition block* (AB) and the *reconstruction block* (RB) that operate simultaneously. The AB contains a single ADC that samples data at the MOSAICS operating frequency, F_{MOS} Hz. The system operates in a series of contiguous acquisition cycles, with β analog to digital conversions taking place in each cycle. At the start of each acquisition cycle, the sampling pattern generator releases a sequence, of length β, consisting of random integers between 1 and γ. At each tick of the clock operating at F_{MOS} Hz,

the sampling pattern generator outputs the next element in the sampling pattern which, after being coded as binary, is fed to the analog multiplexer selection lines. The multiplexer, also driven by the clock, selects the chosen channel, which is then sampled by the ADC. This ensures that the sampling of the channels is randomly interspersed throughout the acquisition which is required for successful reconstruction. These samples, in each signal, belong respectively to the non-overlapping portion of the reconstruction segment. Thus, while the overlapping samples from the previous reconstruction segment are always available, num_{NO} of the non-overlapping samples are acquired through a series of contiguous acquisition cycles. Though the ADC operates at a uniform sampling rate, the sampling of the individual channels is non-uniform.

The ADC sample is deposited in one of the two buffers, which is designated as the acquisition buffer for that particular acquisition cycle. Each buffer can hold β samples. At the end of each acquisition cycle, there is a software controlled buffer swap - the acquisition buffer becomes the reconstruction buffer and vice-versa. During any acquisition cycle, the RB operates simultaneously on the data available in the reconstruction buffer, that was acquired in the previous acquisition cycle. The sampling pattern of the previous acquisition cycle is also made available to the digital demultiplexer in RB[1] which segregates the samples in the reconstruction buffer into the separate channels in the current cycle. At the end of every acquisition cycle, a check is carried out for each signal, if the corresponding measured vector $\mathbf{f}^{(i)}, i = 1...\gamma$, for its current reconstruction window, consisting of $M^{(i)}$ elements has been formed. In addition, a window pointer is maintained, so that each of the $M^{(i)}$ elements in the reconstruction vector of the signal i can be multiplied by the appropriate element of a smoothening window, like Kaiser, of size $N^{(i)}$. For all those signals, whose measured vector is ready, the vectors $\mathbf{f}^{(i)}$ are released to the corresponding convex

[1]There is no additional overhead in sending the sampling pattern to the RB as the implementation can be on a single processor as suggested in sub-section 3.5.6

optimization[1] blocks (see equations (3.14) and (3.9)). De-windowing is done on the reconstructed signal vector $\zeta^{(i)} \in \mathbb{R}^{N^{(i)}}$, by dividing the elements of $\zeta^{(i)}$ by the corresponding window elements. Depending upon the value of $N^{(i)}$, the formation of measured vector, for different signals, may consume different number of acquisition cycles. Consequently, the formation of the reconstruction window for the different signals may not, in general, be synchronized. Once a reconstruction window is completed, the boundary detection logic (3.15) is applied for the channel to check if the just completed reconstruction window had overlapped with a boundary between two SS segments, in which case, the decisions described in section 3.4.3 are taken and conveyed to the digital demultiplexer block.

3.5.3 Derived Parameters

MOSAICS operating frequency, F_{MOS}

The MOSAICS operating frequency must be greater than or equal to the maximum of all the Nyquist rates of the individual signals.

$$F_{MOS} \geq 2(max\left\{F^{(i)}, i = 1...\gamma\right\}) \tag{3.16}$$

Each tick at the MOSAICS operating frequency, F_{MOS} represents a Nyquist instant in the reconstruction segment of any signal.

Number of samples in reconstruction segment, $N^{(i)}$

It is convenient if the number of samples, $N^{(i)}$ in the reconstruction segment is a multiple of the number β of sampling instants in an acquisition cycle so that the formation of the measured vector $\mathbf{f}^{(i)}$ always coincides with the end of an acquisition cycle.

$$N^{(i)} = \alpha^{(i)}\beta \tag{3.17}$$

[1]To carry out the convex optimization, the *cvx* toolbox is used. This toolbox can be downloaded from *www.cvxr.com/cvx*

where $\alpha^{(i)}$ is an integer. Larger the size, $N^{(i)}$ of the reconstruction segment, better will be the frequency localization achieved by the STFT and smaller will be the number of times the computationally expensive convex optimization is done during the acquisition of any of the signals. However, it has an upper bound. The window should be narrow enough to make sure that the portion of the signal falling within it is stationary. In other words, it should not be so large that the boundary between two SS segments falls on it which will lead to failure of the sparsity assumption and consequently the reconstruction itself. Given the minimum duration, $L^{(i)}$ of an SS segment as input, the following must be satisfied.

$$\frac{N^{(i)}}{F_{MOS}} \leq L^{(i)} \tag{3.18}$$

Eliminating $\alpha^{(i)}$ from (3.17) and (3.18) and introducing the *floor* function, integral values of $N^{(i)}$ are given by

$$N^{(i)} = \mathit{floor}\left(\frac{L^{(i)}F_{MOS}}{\beta}\right)\beta \tag{3.19}$$

where $floor(\rho)$ is the largest integer that is smaller than ρ.

Percentage of overlap between consecutive RSs

The percentage of samples in an RS (except the one which just follows a boundary between two SS segments), that overlap with the previous RS is given by

$$p^{(i)} = (N^{(i)} - \mu^{(i)}F_{MOS})/N^{(i)} \tag{3.20}$$

Number of measurements in reconstruction segment, $M^{(i)}$

From[1] (2.26)

$$M^{(i)} \geq ceil\left(2CK^{(i)}log(N^{(i)}/2K^{(i)})\right) \tag{3.21}$$

[1]In the absence of a lower bound on the number of mesurements needed in the case of the simple measurement matrix used, the corresponding one for the Gaussian matrix has been used. To handle this issue, a sufficiently large value of 5 is used for the constant C [38] in the simulations

The factor of 2 along with $K^{(i)}$ takes care of the mirror image in the DFT spectrum for real signals. It is to be noted here that since a reconstruction segment is formed with an overlap with the previous RS, the actual number of samples in the RS is much more than $M^{(i)}$ (which only indicates the minimum required number). This only improves the reconstruction quality.

Number of conversions in each acquisition cycle, β

Consider the case of a particular acquisition cycle, at the end of which the reconstruction window and therefore, the measured vectors of all the γ signals are formed. The reconstruction block has to perform the convex minimization on all the channels in the immediately next acquisition cycle. Let the total time taken by the reconstruction block for all the convex optimizations be T^{RB}_{worst}. This is the worst execution time of the reconstruction block which has to be considered when designing the system. The time taken by the reconstruction block is what is available to the acquisition block for sampling the data that streams through the different analog channels. To cater for the worst case, the duration of the acquisition cycle can be fixed at T^{RB}_{worst} which in turn implies that

$$\beta = F_{MOS} T^{RB}_{worst} \tag{3.22}$$

T^{RB}_{worst} depends upon a number of factors:

- Number of signals, γ
- Maximum reconstruction segment size in all signals, $N^{(i)}$
- Computational capability of the processor performing the convex optimizations

A large value of T^{RB}_{worst}, due to non-availability of computational resources increases the length of the acquisition cycle which is not of very serious concern, since the additional time in the acquisition cycle is not idle time

and is utilized in taking new samples. Understandably, the time lag between acquisition and reconstruction is increased which could be accepted if the requirements of the application permit so.

3.5.4 Effective Sampling Rate and Utilization Factor

Since γ signals are being acquired and reconstructed simultaneously, each at the rate of F_{MOS} samples per second, the effective sampling rate of a MOSAICS unit is

$$F_{MOS}^{(eff)} = \gamma F_{MOS} \tag{3.23}$$

In a non-multiplexed Nyquist setup, each signal would require a separate ADC operating at the respective Nyquist rate. The rate of analog to digital conversions would be

$$F_{NYQ} = 2\sum_{i=1}^{\gamma} F^{(i)} \tag{3.24}$$

The benefit under MOSAICS can be quantitatively expressed as the **Utilization Factor** (UF) defined as

$$UF = \frac{F_{NYQ}}{F_{MOS}^{(eff)}} \times 100 \tag{3.25}$$

From (3.16) and (3.24)

$$UF = \frac{\sum_{i=1}^{\gamma} F^{(i)}}{\gamma(max\left\{F^{(i)}\right\})} \times 100 \tag{3.26}$$

If the utilization factor is close to 100, then performance of MOSAICS is optimal. Clearly, this will be true if $F^{(i)}$ have closeby values, that is all the signals are similar to each other. In case this is not so, it is possible to divide the signals into sets based on their bandwidths and allocate a separate MOSAICS unit to each set. Thus a complete data acquisition system can have more than one MOSAICS blocks to cater to varied signals.

Table 3.1: FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET I

SIGNAL 1		SIGNAL 2		SIGNAL 3		SIGNAL 4	
Time(ms)	Freq.(KHz)	Time(ms)	Freq.(KHz)	Time(ms)	Freq.(KHz)	Time(ms)	Freq.(KHz)
0–23.5	7.2, 8.4, 3.8	0–26.1	4.5, 5.1, 6.3	0–18.5	5.4, 8.1, 9.3	0–32.5	3.6, 6.0, 2.4
23.5-45.3	1.8, 2.3, 7.4	26.1–39.6	1.2, 3.5, 8.3	18.5-41.8	1.8, 4.2, 7.5	32.5-47.8	5.9, 7.1, 8.0
45.3-61.2	5.9, 8.1	39.6-53.2	6.4, 9.7	41.8-58.2	8.6, 9.3	47.8-65.6	2.3, 1.7, 3.2
61.2–78.2	3.9, 7.5, 8.2	53.2-76.5	5.3, 8.8, 9.1	58.2-75.2	2.5, 7.8, 9.1	65.6-78.7	3.5, 7.2
$\geq$78.2	5.6, 6.1, 9.3	$\geq$76.5	1.6, 3.3, 7.5	$\geq$75.2	3.3, 6.6	$\geq$78.7	1.8, 3.1, 4.3

3.5.5 Simulations and Results

The performance of MOSAICS has been evaluated on two different sets of synthetic test data. Details of the data sets, the assumptions, the plots and observations are given in the next two sub-sections. The specifications of the workstation on which the simulations have been carried out are as follows:

- Processor: Intel, Xeon @2.67 GHz (2 processors)
- RAM: 4 GB (3 GB usable)
- OS: 32-bit Windows 7
- MATLAB: ver 7.9

All the simulations carried out in this research have been done on the same computational platform.

3.5.5.1 Test Set 1

The time-frequency characteristics of four continuous time PSS signals are shown in Table 3.1. Measurement noise is simulated by adding white Gaussian noise at the SNR of 20 dB to each signal. [The MATLAB function *awgn* is used for this purpose.] The values taken by various input and derived parameters are given in Table 3.2(see footnote) [1].

[1]NOTE: continued on the next page

Table 3.2: INPUTS AND DERIVED PARAMETERS FOR TEST SET I

Input parameters				
Number of signals, $\gamma = 4$				
	Signal 1	Signal 2	Signal 3	Signal 4
Bandwidth (KHz)	9.3	9.7	9.3	8
F_{NYQ} in KHz	72.6			
Number of tones	3	3	3	3
$L^{(i)}$ in ms	10.0	10.0	10.0	10.0
$\mu^{(i)}$ in ms	1.0	1.0	1.0	1.0
Derived parameters				
$N^{(i)}$	200	200	200	200
num_O	180	180	180	180
$M^{(i)}$	46	46	46	46
F_{MOS} in KHz	20			
β	20			
Effective sampling rate in KHz	80			
Utilization factor	90.75			

A snapshot of the reconstructed signal, during the time interval 35 - 50

1. Although F_{MOS} from (3.16) is 18.4 KHz, the value of 20 KHz is chosen since this can be derived from the system clock. This does not affect the reconstruction and only decreases the utilization factor.

2. The effective sampling rate is computed considering a Nyquist rate of 72.6 KHz. However, in practice even with a classical Nyquist setup it is not possible to make use of data converters that can sample at fractional rates. For this example each of the ADCs must have a sampling rate of 20 KHz and F_{NYQ} will thus be 80 KHz in which case the utilization factor will be close to 100 percent.

3. The minimum number of samples required for reconstruction for each signal is 46. Whenever there is overlap with the previous reconstruction segment, 180 samples are already available. The remaining 20 non-overlapping samples are acquired through a series of contiguous acquisition cycles, since in each such cycle only a fraction of the sampling instants are available for any particular signal. For instance if all the four signals have to be reconstructed to the same accuracy, then within each acquisition cycle of 20 sampling instants, randomly chosen 5 instants are available for each signal. Immediately after a boundary between two SS segments is detected, there can be no overlap. In this case, a minimum of 46 samples have to be acquired through direct sampling and will require more number of acquisition cycles than in the overlapping case.

4. The value of $L^{(i)}$ is chosen as 10 ms for all the signals, although the corresponding values are different for each signal and are more than 10 ms, e.g. in case of signal 1 it is 15.9 ms (the SS segment 45.3-61.2 ms). This has been done for the sake of simplicity and does not affect the reconstruction since taking a smaller value is always a more conservative policy because of which the size $N^{(i)}$ of the reconstruction segment is reduced (see equation 3.19). This will only help in further ensuring that a PSS boundary is not missed.

5. Based on the assumption of a value of 1 ms for T^{RB}_{worst}, β gets a value of 20 (see equation 3.22)

ms, is shown in figure 3.3. The plot shows the close match between the original (black color) and the reconstructed signal (red color). Transition from one SS segment to another (see 3.1) is manifested as an erroneous reconstruction, for example around 45 ms in signal 1, 40 ms in signal 2, 42 ms in signal 3 and 48 ms in signal 4.

The erroneous reconstruction is confined to the non-overlapping portion of the reconstruction segment that falls on the boundary between two SS segments which for this test set is equal to 1 ms (this is equal to the corresponding $\mu^{(i)}$), a figure that could be tolerated. A quantitative measure for the deviation between the original signal and the reconstruction is measured by the Peak Signal to Noise Ratio defined as follows:

$$PSNR(orig, rec) = 20\,log\left[\frac{max\,(orig)}{rms\,(rec - orig)}\right] \tag{3.27}$$

3.5.5.2 Performance of MOSAICS with other reconstruction algorithms

The performance of MOSAICS is verified with three other compressed sensing recovery algorithms - OMP, ROMP and CoSAMP through simulations on the same set of signals. Figures 3.4, 3.5 and 3.6 give snapshots of the reconstructed signals for these simulations (during the same time interval of 35-50 ms as in figure 3.3). Table 3.3 compares the performance of the different reconstruction algorithms with respect to the deviation from the original signal and the execution time. The specifications of the computational platform have already been given at the beginning of sub-section 3.5.5.

The reconstruction accuracy obtained using these algorithms can be compared with the accuracy of reconstruction of the signal from classical Nyquist compressed data. The signal in each reconstruction window is transformed into the Fourier domain. Only as many as coefficients in the transformed vector as the assumed sparsity in the signal are retained and the rest made zero and the original signal is recovered from this vector

by applying the inverse transform. The original and the reconstructed signals are shown in figure 3.8. The corresponding PSNR values are listed in the last column of Table 3.3. It is to be noted from the table that the performances of BP and CoSAMP are not much behind that obtained in classical compression.

In order to verify the consistency of reconstruction, MOSAICS was run for 100 trials with each of the recovery algorithms, using the same set of signals in each trial. Table 3.4 gives the mean and standard deviation (separated by comma in each cell of the table) of the PSNR values achieved in the 100 trials with each reconstruction algorithm for each of the signals. Figure 3.7 gives the histogram plots of the PSNR values of the reconstructed signal for signal 4. The histograms for the BP and CoSAMP algorithms have a higher peak and are less spread out. Histograms of OMP and ROMP have slightly lower peaks and are more spread out indicating lower reconstruction consistency.

From these results, the following important observations can be made:

1. Independence from recovery algorithm - The basic functionality of MOSAICS is demonstrated with all the identified recovery algorithms.

2. Accuracy with basis pursuit algorithm - Of all the recovery algorithms, the best reconstruction accuracy is obtained with the BP and CoSAMP algorithms. However, the performance of the other recovery algorithms are comparable.

3. Execution time - There is a drastic reduction in execution time with OMP, ROMP and CoSAMP which is ideally suited for real-time implementation in embedded hardware. Of these algorithms, CoSAMP guarantees higher consistency of reconstruction.

Table 3.3: PERFORMANCE OF MOSAICS WITH DIFFERENT RECONSTRUCTION ALGORITHMS AS PSNR VALUES (in dB).THE LAST COLUMN GIVES THE PSNR FOR DECOMPRESSED DATA RECOVERED FROM NYQUIST SAMPLED COMPRESSED DATA. THE LAST ROW GIVES THE EXECUTION TIME FOR A SIMULATED ACQUISITION TIME OF 80 ms

Signal	**BP**	**OMP**	**ROMP**	**CoSAMP**	**Decompr.**
Signal 1	21.77	11.99	16.81	18.24	24.64
Signal 2	19.39	16.20	14.25	18.83	24.24
Signal 3	19.60	22.70	15.90	19.08	25.90
Signal 4	18.35	12.38	15.00	22.76	25.73
Execution time	253.6 sec	1.37 sec	1.79 sec	1.80 sec	-

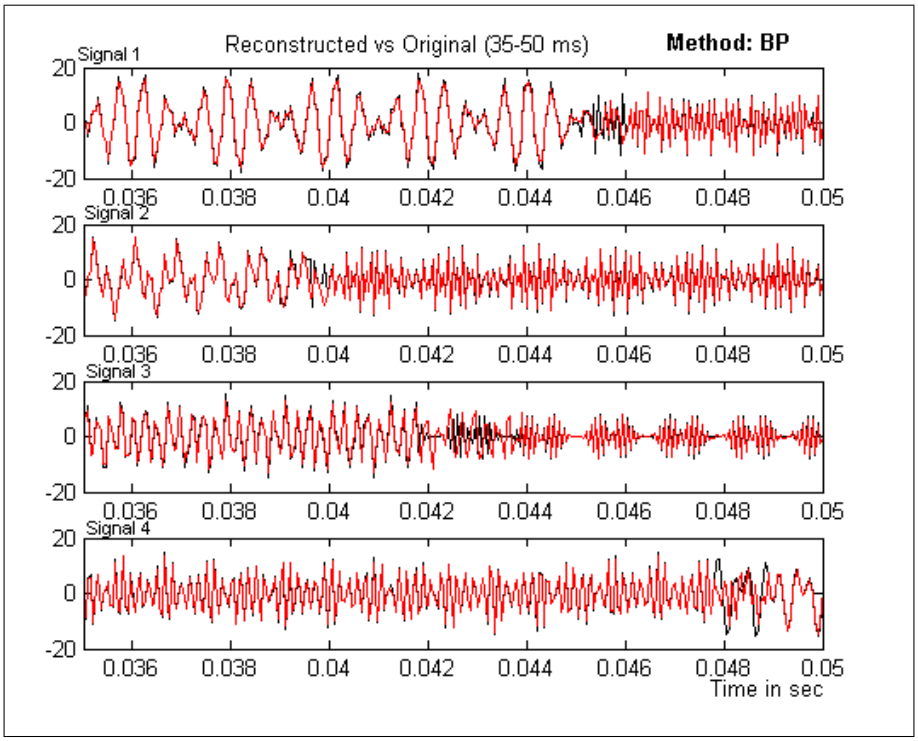


Figure 3.3: Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **basis pursuit**

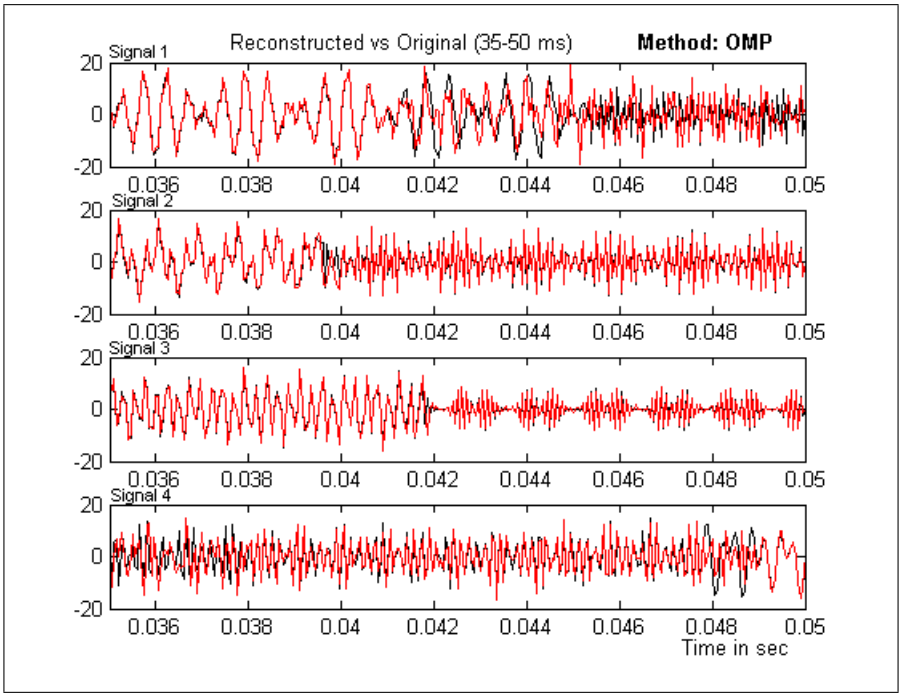


Figure 3.4: Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **OMP**

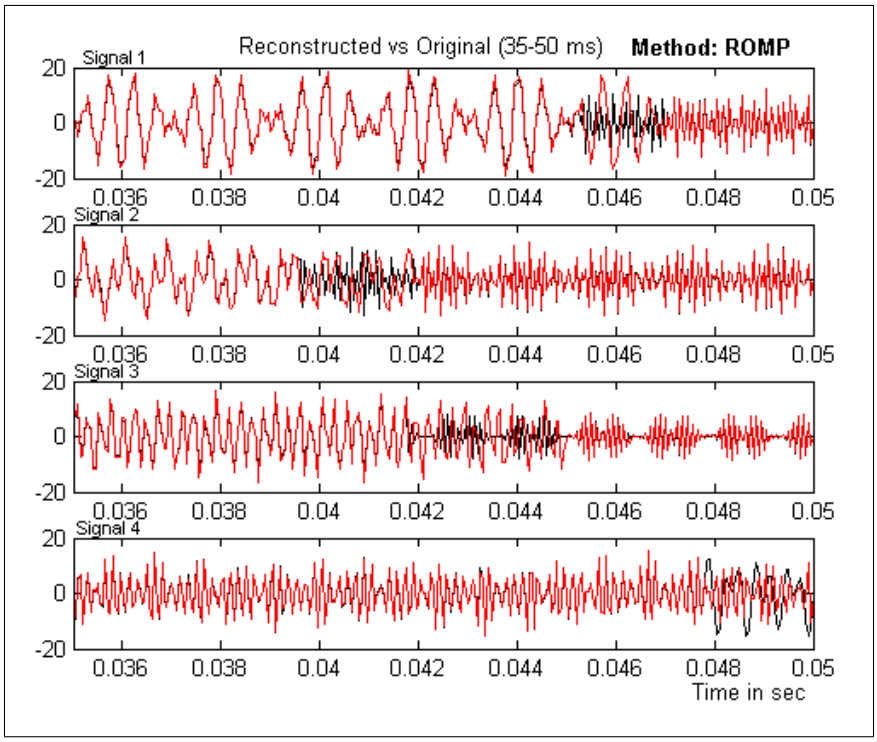


Figure 3.5: Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **ROMP**

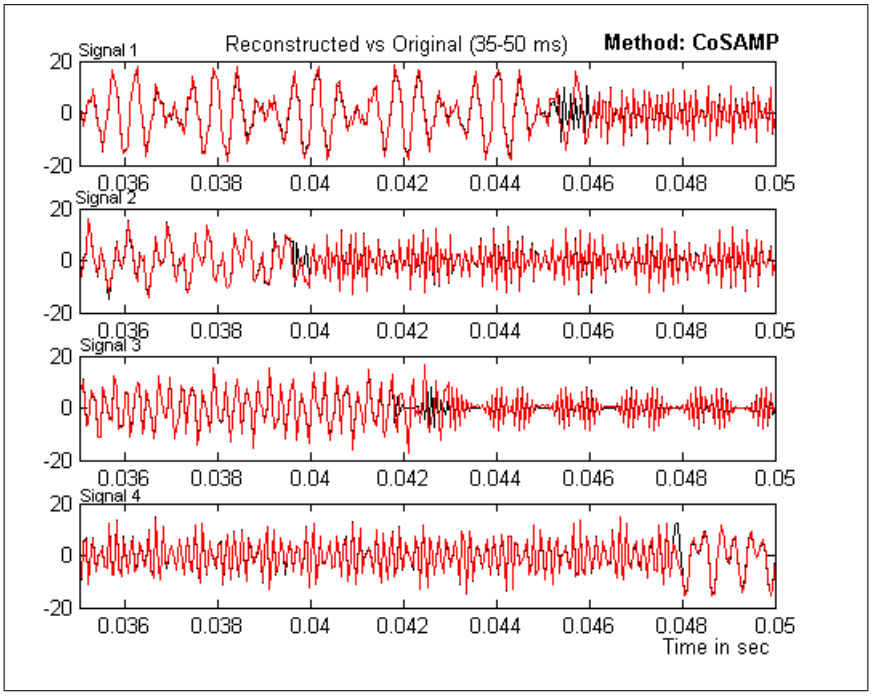


Figure 3.6: Reconstructed signal (red) over the original signal (black) for four channels with reconstruction using **CoSAMP**

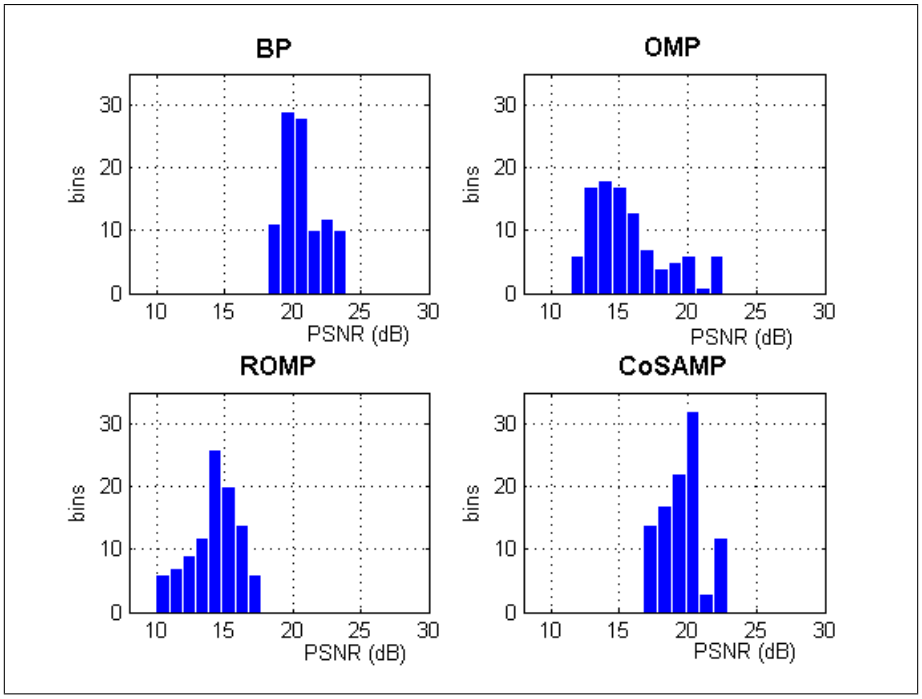


Figure 3.7: PSNR histograms for **signal 4** with four reconstruction algorithms

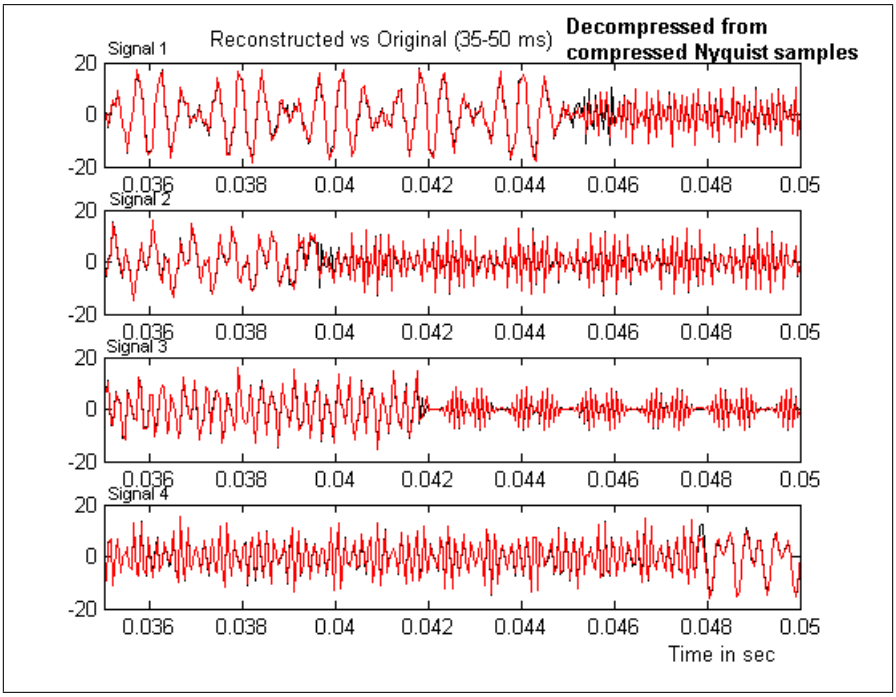


Figure 3.8: Reconstructed signal (red) over the original signal (black) for four channels of decompressed data obtained from compressed Nyquist samples

Table 3.4: COMPARISON OF MEAN AND STANDARD DEVIATION OF PSNR FROM 100 TRIALS FOR VARIOUS RECONSTRUCTION ALGORITHMS

Signal	BP	OMP	ROMP	CoSAMP
Signal 1	19.96, 1.203	15.76, 2.627	14.79, 1.841	18.73, 1.292
Signal 2	19.93, 1.134	15.01, 3.152	14.88, 1.629	18.50, 1.310
Signal 3	20.44, 0.971	15.17, 2.204	14.36, 1.852	18.95, 1.089
Signal 4	20.73, 1.426	15.68, 2.767	14.28, 1.803	19.59, 1.587
Mean exec time	251.34s	1.18s	1.40s	1.48s

3.5.5.3 Test Set 2

The second test case consists of a set of three monotone signals, in which the single constituent frequency is confined to a band away from zero frequency. Such signals are typically found in Doppler radar echoes which are mixed down to the Intermediate Frequency (IF) range (10-100 MHz). Table 3.5 gives the frequency characteristics of signals in the test set. The signals are band pass in nature and the exact location of the tones in different intervals is not known. Therefore, each signal is considered to have a bandwidth of 5 MHz, for example, the band in the case of signal 1 is 30-35 Hz. In a classical

Table 3.5: FREQUENCY CHARACTERISTICS OF SIGNALS IN TEST SET II

SIGNAL 1		SIGNAL 2		SIGNAL 3	
Time(μs)	Freq.(MHz)	Time(μs)	Freq.(MHz)	Time(μs)	Freq.(MHz)
0–60	30.5	0-70	21.5	0–190	41
60–110	31.4	70-124	23.6	190-230	44.1
110–160	32.3	124–201	21.2	230–280	43.5
160-230	33.4	201-270	24.1	280-320	42.2
$\geq$230	31.2	$\geq$270	22.8	$\geq$320	40.5

Nyquist setup one would need to use three ADCs each of sampling rate 10 MHz. Accordingly,
$F_{NYQ} = 30\,\text{MHz}$ $F_{MOS} = 10\,\text{MHz}$ $F_{MOS}^{(eff)} = 30\,\text{MHz}$ and UF[1] $= 100\%$

The reconstructed signals obtained as MOSAICS outputs are passed through a low-pass (0-5 MHz) 20-tap filter and the Power Spectral Density (PSD) of the filtered output is found. A snapshot of the PSD, during the interval 0-60 μs, is shown in figure 3.9. The peaks are indicated by red arrows. The peak frequency values for each signal during different intervals are also listed in Table 3.6. It is clear that during each SS segment the corresponding peak frequency is accurately detected. Due to the low pass filtering, the low pass frequency is detected, for example, 30.5 MHz in the original signal is manifested as 0.5 MHz in the filtered output. This result suggests that MOSAICS is well suited in applications where exact recovery of the signal after sampling is not as important as the mere detection of the prominent frequency component.

3.5.6 Hardware Architecture for Realization of MOSAICS

It is proposed that the hardware design, shown in figure 3.10 consisting of a single channel ADC and an analog multiplexer can be followed for realization of MOSAICS. The digital demultiplexer, the convex optimiza-

[1]One has to accept the fact that the utilization factor does not reflect the effectiveness of the MOSAICS setup completely, since its computation depends upon the sampling rates of the ADCs available for use in the classical Nyquist setup or in the MOSAICS setup. Since ADCs with fractional sampling rates obtained through computations are not available, the value of UF will vary. In general, as mentioned before, if the highest frequency components in the different signals are close by, the UF will be high.

Table 3.6: MOSAICS OUTPUT (FILTERED) FOR TEST SET II- HIGHEST FREQUENCY COMPONENT

SIGNAL 1		SIGNAL 2		SIGNAL 3	
Time(μs)	Freq.(MHz)	Time(μs)	Freq.(MHz)	Time(μs)	Freq.(MHz)
0–60	0.5	0-70	1.5	0–190	1.0
60–110	1.4	70-124	3.6	190-230	4.1
110–160	2.3	124–201	1.2	230–280	3.5
160-230	3.4	201-270	4.1	280-320	2.2
$\geq$230	1.2	$\geq$270	2.8	$\geq$320	0.5

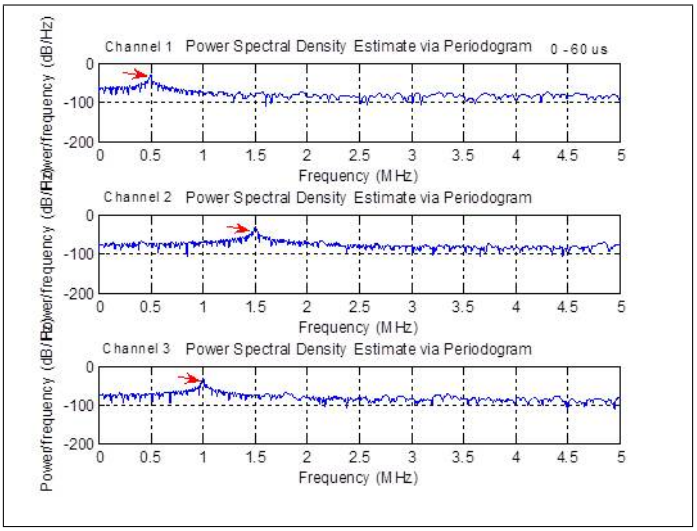


Figure 3.9: PSD of Filtered MOSAICS output during 0–60 μs

tion blocks, the sampling pattern generator and the boundary detection logic can be implemented in software running on the processor. Based on the generated sampling pattern, the channel selection input is given to the analog multiplexer through the address bus of the host processor. Many commercially available ADCs have a built in multiplexer. The buffers 1 and 2 exist in the Random Access Memory (RAM). The Direct Memory Access (DMA) is initialized by the processor to perform a direct memory transfer from the ADC to the RAM. If multiple MOSAICS blocks exist in a system, then they can share a common processor, RAM and DMA.

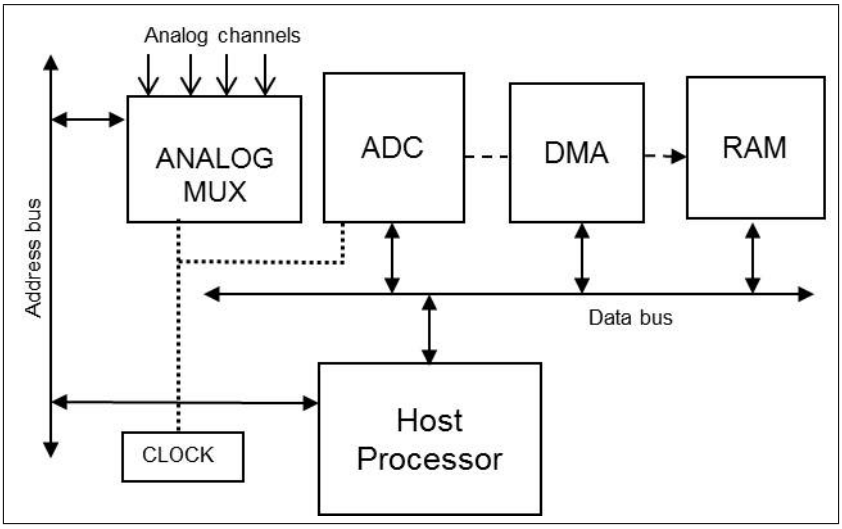


Figure 3.10: Proposed hardware architecture for realization of MOSAICS

3.5.7 Concluding Remarks on MOSAICS

In this work, MOSAICS, a setup for acquisition and reconstruction of multiple sparse signals in a multiplexed fashion under a compressed sensing framework has been proposed. The MOSAICS scheme has the following features:

i) Single analogue-to-digital converter - A single ADC is used to acquire more than one signal in a multiplexed fashion. This leads to a reduction in the number of components used - ADCs and the associated passive components, in the design of the embedded hardware. This in turn contributes to saving in cost, reduction in power consumption and dissipation and compactness of design.

ii) Reconstruction algorithm independent - Any reconstruction algorithm - Basis pursuit, Matching pursuit or their variants, can be used to reconstruct the individual signals

iii) Design based on available and realizable components - MOSAICS can be designed using commercially-off-the-shelf (COTS) components. There are no components involved which have to be redesigned from scratch. Any conventional ADC can be used to build the system.

Table 3.7: FREQUENCY CHARACTERISTICS OF SAMPLE SIGNALS WITH NON-INTEGRAL NUMBER OF CYCLES IN THE RECONSTRUCTION WINDOW

Signal	Frequencies (KHz)	η
1	3.454, 4.746	34.5
2	4.628, 6.783	46.2
3	8.5, 9.1	85
4	3.4, 5.8	64

iv) Limited computational requirement - Since the acquisition and reconstruction operate in parallel, the system can be built even with limited computational capability in the host processor.

v) No explicit measurement matrix - Unlike traditional compressed sensing hardware, the measurement involves plain random undersampling, and there is no requirement of complex arrangement for realizing the measurement matrix.

vi) Reconstruction of noisy signals - Even in the presence of measurement noise, the reconstruction is fairly good.

3.5.8 Limitation of MOSAICS

On the downside, it is worth noticing a limitation of MOSAICS - the inability to deal with those signals that do not have an integral number of cycles of one or more than one frequency components, within the reconstruction window. Consider one of the signals, $\mathbf{x}^{(i)}$ that is given as input to MOSAICS. Let the computed reconstruction segment size of $\mathbf{x}^{(i)}$ be $N^{(i)}$. Suppose $\mathbf{x}^{(i)}$ has a component with frequency f. Then the number of cycles of this component in the reconstruction segment is

$$\eta = N^{(i)} \frac{f}{F_{MOS}} \tag{3.28}$$

The DFT coefficients of the signal are no longer sparse except in the (contrived) case when the sinusoid frequencies are such that η is an inte-

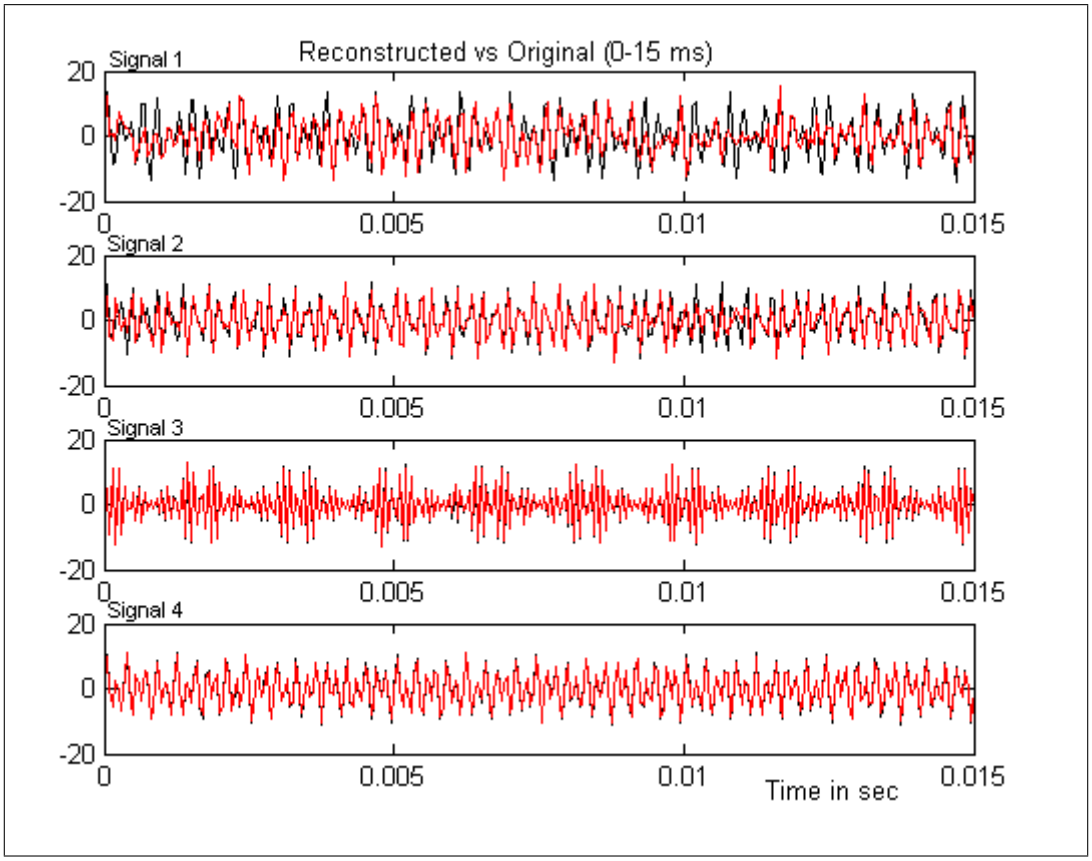


Figure 3.11: Reconstructed signal (red) over the original signal (black) for four channels. The PSNR for the reconstructed signals are: signal 1:10.0 dB, signal 2: 15.74 dB, signal 3: 28.57 dB, signal 4: 26.36 dB

ger. Otherwise, which is the usual case, there is spectral leakage introduced due to windowing (sinc convolution) and the signal representation in the frequency domain is anything but sparse. In such cases the compressed sensing reconstruction of the undersampled signal will fail. On closer inspection, it can be seen that in the case of signals whose frequency characteristics are shown in tables 3.1 and 3.5, η, as given by (3.28),is an integer for each frequency component. However, this may not be so for general frequency sparse signals.

To illustrate this point, a MOSAICS reconstruction of such signals is shown in figure 3.11. Table 3.7 gives the frequency characteristics of four signals during an arbitrary interval of time. Assuming a MOSAICS operating frequency, $F_{MOS} = 20$KHz and a uniform reconstruction window size of

200 for all the four signals, it can be seen from the table that the calculated values of the number of samples, η in a reconstruction window, is not an integer for two of the signals (1 and 2). The corresponding reconstructed signals shown in figure 3.11 are erroneous and the PSNR values for those signals are considerably lesser than for the other two. Signals, which have what shall hereafter be called as *non-integral frequency* components, will be dealt with, in the next section.

3.6 Multiplexed Signal Acquisition for General Sparse Signals

3.6.1 Introduction

Some interesting research efforts towards compressed acquisition of signals sparse in the frequency domain, without any specific constraint of integral number of cycles in the reconstruction window, have been reported in the literature. These approaches accommodate all kinds of frequency-sparse signals - even those signals with the so-called non-integral frequencies also, and make available algorithms for reconstructing the original signals from the undersampled measured vectors. In [68] this is addressed as a general compressed sensing recovery algorithm. In this paper it is shown that by simply changing the signal representation to a zero-padded DFT, the DTFT of the signal is more closely sampled. Thus the DFT basis is replaced by a redundant frame of sinusoids known as a *DFT frame*. The *DFT frame* provides a finer sampling of the DTFT coefficients for an observed signal $\mathbf{x}$. If $w \in \mathbb{N}$ denotes the frequency oversampling factor for the *DFT frame*, the frequency sampling interval is defined as:

$$\Delta := 2\pi/wN \in (0, 2\pi/N] \tag{3.29}$$

Also let

$$\mathbf{e}\left(\omega\right) := \frac{1}{\sqrt{N}}\left[1 \;\; e^{j\omega/N} \;\; e^{2j\omega/N} \;\; e^{(N-1)j\omega/N}\right]^T \tag{3.30}$$

denote a normalized vector containing regular samples of a complex sinusoid with angular frequency $\omega \in (0, 2\pi]$. The DFT frame with oversampling factor w is then defined as $\mathbf{\Psi}\left(w\right) := [\mathbf{e}(0) \; \mathbf{e}(\Delta) \; \mathbf{e}(2\Delta) \; \mathbf{e}(2\pi - \Delta)]^T$. The corresponding signal representation $\mathbf{c} = \mathbf{\Psi}(w)^T\mathbf{x}$ provides wN equispaced samples of the signal's DTFT. Note that $\mathbf{\Psi}(1) = \mathbf{F}$, the usual orthonormal DFT basis. The DFT frame $\mathbf{\Psi}(w)$ can be used to obtain sparse approximations for frequency-sparse signals with components at arbitrary frequencies; as the frequency oversampling factor w increases, the $K-$sparse approximation provided by $\mathbf{\Psi}(w)$ becomes increasingly accurate. Unfortunately, the classical algorithms, like those of compressed sensing, that aim to find the sparse approximation of the signal in the frame $\mathbf{\Psi}(w)$ do not perform well when w increases due to high coherence between the frame vectors [69; 70] , particularly for large values of w.

$$\mu(\mathbf{\Psi}(w)) = \frac{wN\sin(\pi/wN)}{\pi} \to 1 \text{ as } w \to \infty \tag{3.31}$$

Because of this the maximum frequency oversampling factor is limited by the sparsity, K. Coherence of *DFT frame* can be treated as another manifestation of the spectral leakage problem which has been classically dealt with by applying a tapered window function to the signal before applying the DFT [13; 71]. However, windowing interferes with frequency resolution, making it difficult to resolve closely spaced frequencies. A more recent alternative to classical spectrum estimation methods is a technique based on eigen analysis of the signal autocorrelation matrix which gives the line spectrum of the signal [71]. In order to find the dominant signal modes in the frequency domain, such algorithms estimate the principal components in the signal autocorrelation matrix and provide better resolution of the parameters of a frequency sparse signal. Example algorithms include

Pisarenkos method, MUltiple SIgnal Classification (MUSIC), and estimation of signal parameters via rotationally invariant techniques (ESPRIT). A line spectrum estimation algorithm $\mathbb{L}(\mathbf{x}, K)$ returns a set of dominant K frequencies for the input signal $\mathbf{x}$, with K being a controllable parameter. An algorithm based on MUSIC has been proposed in [68], for recovering a signal consisting of a sparse set of integral as well as non-integral frequencies and embedded in noise. The next sub-section gives a brief description of the algorithm.

3.6.2 Signal Reconstruction Based on MUSIC

The method proposed in [68] is a greedy algorithm in which the K prominent frequencies of a frequency sparse signal are estimated from an intermediate signal estimate, $\hat{\hat{\mathbf{x}}}$ by employing MUSIC algorithm, $\mathbb{M}(\hat{\hat{\mathbf{x}}}, K)$. Subsequently, the DTFT coefficients corresponding to the frequencies are found, from which the final estimate of the signal is obtained. K is a controllable parameter that determines the level of reconstruction accuracy. The steps of the algorithm are as follows:

3.6.2.1 Initialization

The initial estimate of the reconstruction segment, $\hat{\zeta}_0 = 0$
The initial residue, $\mathbf{r}_0 = \mathbf{f}$ [1]
Iteration count, $i = 1$

3.6.2.2 Intermediate Signal Estimate

The residue is updated.

$$\mathbf{r}_i = \mathbf{r}_{i-1} - \mathbf{\Phi}\hat{\zeta}_{i-1} \tag{3.32}$$

[1] $\mathbf{f}$ is the measured vector in (3.3)

where $\mathbf{\Phi}$ is the measurement matrix.
An intermediate estimate of the reconstruction segment is obtained.

$$\hat{\hat{\zeta}}_i = \hat{\zeta}_{i-1} + \mathbf{\Phi}^T \mathbf{r}_i \tag{3.33}$$

3.6.2.3 The *MUSIC* algorithm, $\mathbb{M}(\mathbf{x}, K)$

Given a signal vector $\mathbf{x} \in \mathbb{R}^N$, let its $P \times P$ autocorrelation matrix be denoted by $\mathbf{A}$. Suppose the eigen decomposition of $\mathbf{A}$ gives P eigen values $\lambda_1, \lambda_2, ..., \lambda_P$ and the corresponding eigen vectors, $\vartheta_1, \vartheta_2, ...\vartheta_P$. The algorithm evaluates a score function:

$$P_{MUSIC}(\omega) = \frac{1}{\sum_{p=K+1}^{P} \left| \mathbf{e}(\omega)^T \vartheta_p \right|^2} \tag{3.34}$$

and returns the locations of the K largest score function peaks as the frequencies present in the signal. A modification of *MUSIC* known as *Root MUSIC* calculates the peaks from the zeros of a polynomial that depends upon the noise subspace eigen vectors.
The intermediate estimate,$\hat{\hat{\zeta}}_i$ of the signal obtained in the previous step, is given as input to the MUSIC algorithm:

$$\{\omega_k\}_{k=1}^{K} \leftarrow \mathbb{M}(\hat{\hat{\zeta}}_i, K) \tag{3.35}$$

3.6.2.4 Estimation of DTFT Coefficients

Once the frequencies are known, the corresponding coefficients are estimated by the Periodogram method [13; 71] which gives the Maximum Likelihood Estimate (MLE) of the spectral coefficient for any known frequency, ω, as the DTFT coefficient of $\hat{\hat{\zeta}}_i$, at frequency ω:

$$\hat{\mathbf{c}}_k = \left\langle \mathbf{e}(\omega_k), \hat{\hat{\zeta}}_i \right\rangle \tag{3.36}$$

Table 3.8: FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET III

SIGNAL 1		SIGNAL 2	
Time(ms)	Freq.(KHz)	Time(ms)	Freq.(KHz)
0–23.5	3.68, 6.14, 7.23	0–28.1	1.23, 3.45, 6.10
23.5–49.3	2.45, 6.82, 7.95	28.1–38.2	2.56, 3.80, 5.72
49.3-62.8	2.96, 4.11, 8.34	38.2-56.5	2.10, 4.45, 6.53
62.8–77.4	3.64, 6.94, 8.59	56.5 74.5	5.13, 6.66, 7.60
≥77.4	1.54, 2.71, 9.2	≥74.5	3.62, 6.30, 8.34
SIGNAL 3		SIGNAL 4	
Time(ms)	Freq.(KHz)	Time(ms)	Freq.(KHz)
0–18.5	7.42, 3.75, 8.42	0–32.5	4.52, 3.34, 6.7
18.5-41.8	2.13, 5.34, 8.54	32.5-47.2	2.92, 4.16, 9.2
41.8-58.2	3.74, 6.52, 8.30	47.2-65.2	1.46, 4.2, 5.62
58.2-75.2	1.55, 5.62	65.2-79.2	3.58, 6.23
≥75.2	3.34, 7.16	≥79.2	1.89, 5.22, 6.46

3.6.2.5 Estimate the Signal

Finally the signal is estimated

$$\hat{\zeta}_i \leftarrow \sum_{k=1}^{K} \hat{\mathbf{c}}_k \mathbf{e}(\omega_k) \tag{3.37}$$

The steps are repeated in an iterative manner, until some halting criterion, like desired reconstruction accuracy, is satisfied.

3.6.3 Modification in MOSAICS

General sparse signals with the so-called non-integral frequencies are incorporated into the MOSAICS framework by substituting the convex optimization step with the greedy reconstruction algorithm based on MUSIC introduced in subsection 3.6.2. The new acquisition and reconstruction scheme shall be called **MOSAICS with MUSIC**.

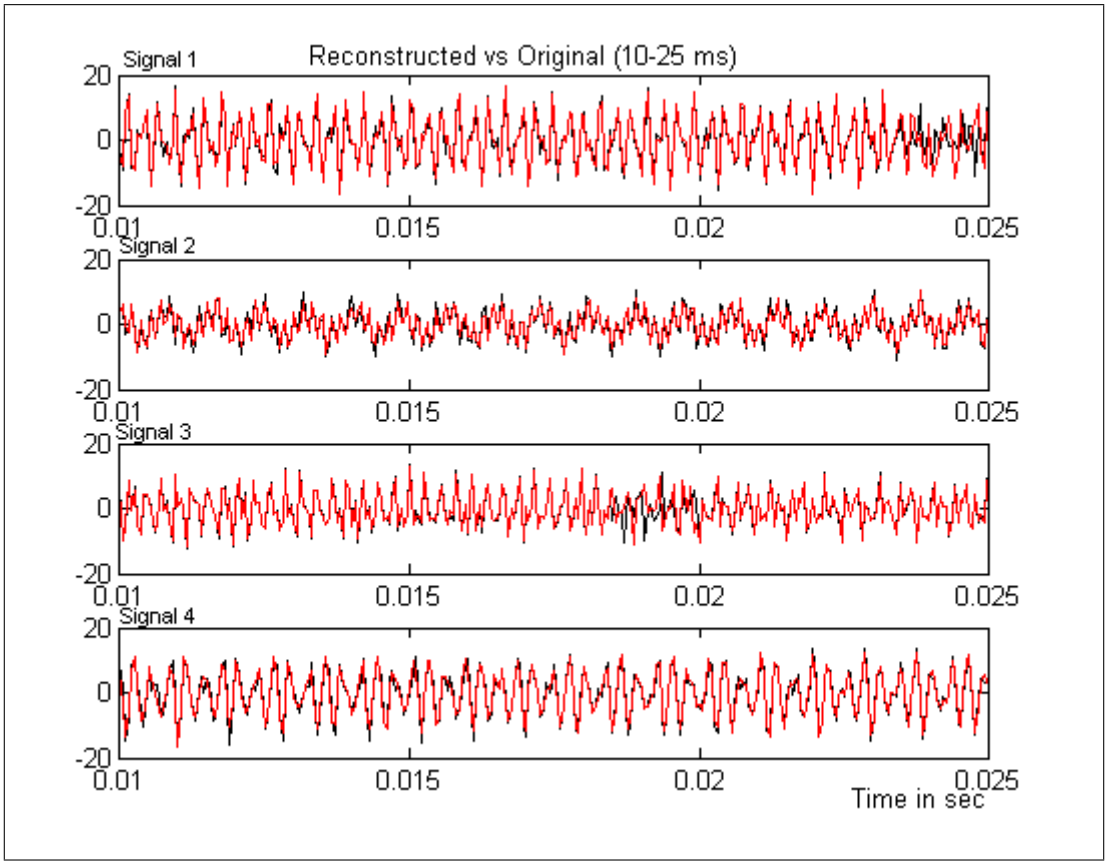


Figure 3.12: Original signals (black) of test set III and a snapshot of the reconstruction (red) using MOSAICS with MUSIC. The PSNR values for the reconstructed signals are: for signal 1:16.30 dB, for signal 2: 17.23 dB, for signal 3: 15.63 dB, for signal 4: 19.42 dB

3.6.4 Simulation and Results

Simulation is carried out for an interval of 100 ms, on a signal set similar to that in subsection 3.5.5 with the inclusion of 'non-integral' frequencies. The frequency characteristics of the signals in the set are given in Table 3.8. Since the signal set is very similar to that given in 3.5.5, except for the inclusion of the *non-integral frequencies* the input and derived parameters will be the same as in Table 3.2 and are not repeated here. White Gaussian noise at a Signal to Noise Ratio (SNR) of 20 dB is added to the signal in the buffer at the end of every acquisition cycle to simulate measurement noise. A snapshot (figure 3.12) of the reconstruction during a small interval: 10-25 ms shows how closely the reconstructed signal matches with the original signal. As before, there is a small but noticeable deviation at the

boundary between the SS segments. The PSNR values shown below the figure indicate a marked improvement in the reconstruction accuracy when compared with the reconstruction for the top two signals, in figure 3.11.

3.6.5 Concluding Remarks on MOSAICS with MUSIC

The MOSAICS framework proposed in section 3.5, offers an efficient sensing and reconstruction mechanism for sparse signals by using lesser number of components in the embedded hardware. However, its inherent limitation came to light, for the case when general sparse signals with 'non-integral' frequencies are given as input. In subsection 3.6.2, it has been shown that a greedy reconstruction algorithm based on MUSIC can alleviate the problem. The algorithm can seamlessly be incorporated into the MOSAICS architecture. The new algorithm can be used on the same hardware design that was proposed in subsection 3.5.6. Only the software that executes in the processor has to be changed. Substitution of the reconstruction algorithm by an alternative one suggests that within the MOSAICS scheme, one can have different reconstruction methods for different signals depending upon the characteristics of the signal. Thus, spectral compressive sensing methods [68] based on MUSIC algorithm have been married together with MOSAICS to deliver a better algorithm albeit more computationally intensive owing to the introduction of MUSIC based reconstruction of the signal. Implementation of the algorithm on modern fast processors running highly optimized code can address the issue of computational complexity. With the capability of handling non-integral frequency, the enhanced MOSAICS algorithm provides a general purpose acquisition system for capturing multiple signals within a compact hardware design.

3.7 The Frequency Detection Problem

3.7.1 Introduction

Many applications in science and engineering call for the detection of a single frequency tone, usually a sinusoid, buried in noise. Radar, communications, seismology and biomedical fields are a few examples. In such applications, although the unknown signal might have quite a few component frequencies, the prominent frequency is of interest, opening up the possibility of reformulating the problem as a sparse recovery problem, in particular compressed sensing. Since the emphasis is primarily on the detection of the prominent few frequencies, rather than the reconstruction of the signal itself, there are two immediate ramifications:

i) Heavy undersampling can be tolerated

ii) Detection under low SNR[1] may be possible

Thus, the MOSAICS with MUSIC scheme could be employed to detect a few frequency tones (can be non-integral also) embedded in a noisy signal. Detection of sinusoids buried in noise is an age-old problem which has been addressed in innumerable number of publications [72; 73]. The work presented in this thesis does not claim better performance than any of these approaches. The emphasis here, is mainly to treat the problem from a compressed sensing perspective in which a simultaneous undersampling is performed on multiple signals.

3.7.2 Simulation and Results

The test set taken is similar to those in the previous two sections where four signals are acquired in a multiplexed fashion. Table 3.9 gives the frequency characteristic as well as the detected frequencies for each signal.

[1] Signal-to-Noise Ratio

Table 3.9: FREQUENCY CHARACTERISTICS OF THE SIGNALS IN TEST SET IV AND DETECTED FREQUENCIES, AT DIFFERENT SNR

SIGNAL 1							
Time (s)	Freq (Hz)	Detected Frequencies at various SNR in dB					
		-5	0	5	10	15	20
0–8	4.2	4.16	4.18	4.2	4.19	4.2	4.2
8–17	8.7	8.69	8.69	8.7	8.7	8.7	8.7
17–28	6.3	10.2	6.3	6.3	6.3	6.3	6.3
28–35	16.9	16.5	16.8	16.9	16.9	16.9	16.9
≥35	13.4	13.4	13.4	14.1	13.41	13.4	13.4
SIGNAL 2							
Time (s)	Freq (Hz)	Detected Frequencies at various SNR in dB					
		-5	0	5	10	15	20
0-7	5.6	5.75	5.57	5.6	5.6	5.6	5.6
7-14.8	11.9	11.5	11.2	11.2	11.1	11.2	11.2
14.8-23	7.9	7.9	7.87	7.9	7.9	7.9	7.9
23-31	17.5	17.5	17.5	17.5	17.5	17.5	17.5
≥31	14.9	18.3	14.9	14.9	14.9	14.9	14.9
SIGNAL 3							
Time (s)	Freq (Hz)	Detected Frequencies at various SNR in dB					
		-5	0	5	10	15	20
0-11.4	10.5	11.4	10.6	10.5	10.5	10.5	10.5
11.4–19	4.4	4.3	4.45	4.4	4.39	4.4	4.4
19-25	9.1	7.05	17.9	8.4	8.1	9.0	9.0
25-32.6	15.6	15.7	15.5	15.5	12.4	15.6	15.6
SIGNAL 4							
Time (s)	Freq (Hz)	Detected Frequencies at various SNR in dB					
		-5	0	5	10	15	20
0–9.5	8.5	16.8	8.5	8.5	8.5	8.5	8.5
9.5-16.7	11.2	11.1	11.1	11.2	11.2	11.2	11.2
16.7–29	18.4	15.2	19.5	19.3	18.35	18.4	18.4
29-37.5	15.6	9.32	15.4	17.4	15.2	14.8	16

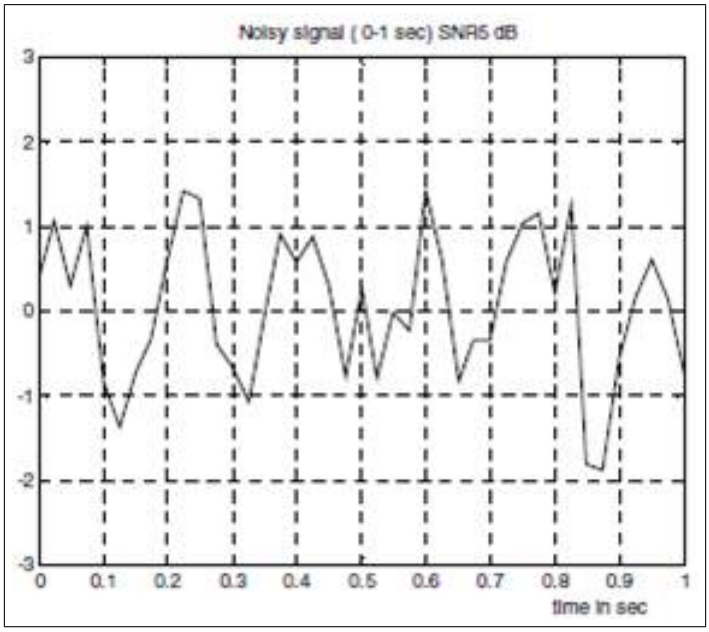


Figure 3.13: Snapshot of signal 2 of Table 3.9 at an SNR of 5 dB during 0–1 sec

The first and second columns in the table give the frequency characteristic of the signals in various time intervals. The signals are monotones, but in general MOSAICS with MUSIC will work with signals with a few more component frequencies, as long as there is sufficient sparsity. The subsequent six columns give the detected frequencies at various SNR levels in dB. Additive white Gaussian noise at various SNRs from -5 dB to 20 dB is added to signal in the buffer at the end of every acquisition cycle to simulate measurement noise. A snapshot of signal 2 at an SNR of 5 dB, is shown in figure 3.13 for the first one second of acquisition. The table clearly indicates that as the SNR increases the frequency is detected more accurately. However, even at very low SNR, the detection accuracy is reasonably high.

3.7.3 Detection of FM carrier frequency

In order to put to test the frequency detection scheme using MOSAICS to real world signals, FM modulated signals have been considered. To this end, four different audio music signals sampled at 44.1 KHz, with 16 bit

samples, have each been modulated with carrier signals centered at four different frequencies: 87.5 MHz, 92.5 MHz, 98.3 MHz and 102.7 MHz. In a simulation of 10 s, the carrier frequency in each channel is altered once in every 100 ms, to a random value within ± 100 kHz of the center frequency. Furthur, the modulated signals have been corrupted with additive noise at an SNR of 5 dB. The objective is to ascertain whether it is possible to detect the change in carrier frequency or 'tune' to it. Figure 3.14 shows the detection of carrier frequency with time, in each channel, as scatter plots (the red dots) about the changing carrier frequency (the solid blue line in the middle). Figure 3.15, which plots the first 1000 samples of the modulated signal in channel 1, has been included to give a feel of the level of noise in the signals.

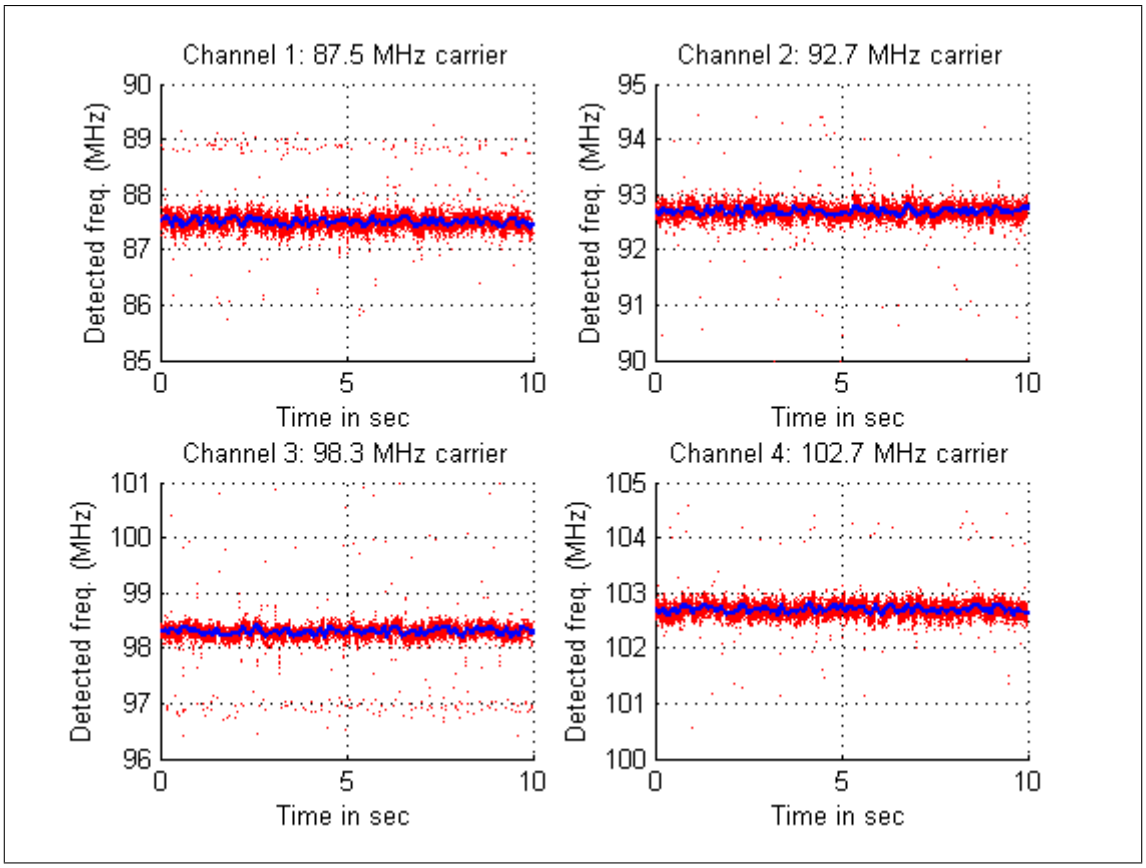


Figure 3.14: Detection of carrier frequency in FM signals with noise at 5 dB SNR. The carrier frequency changes by 100 kHz on either side of the central frequency once in every 100 ms

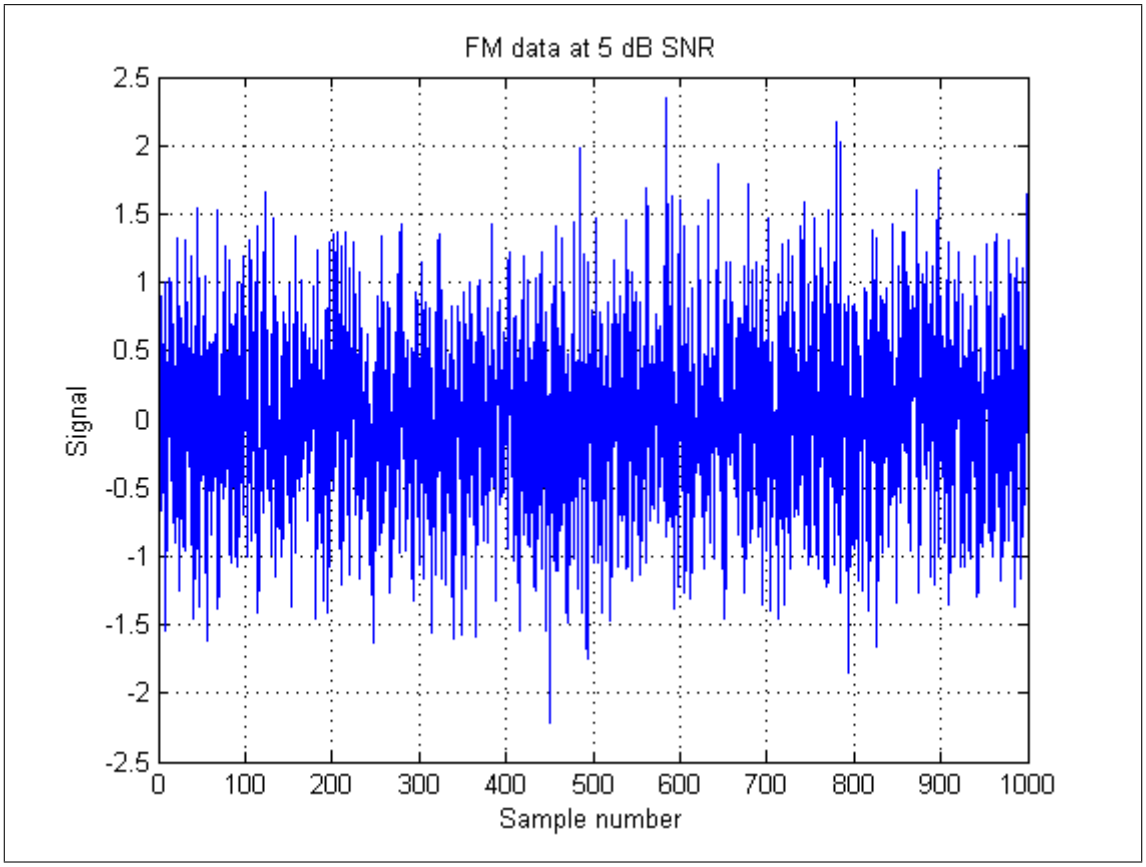


Figure 3.15: First 1000 samples of channel 1 at 5 dB SNR noise

3.8 Conclusion

Compressed sensing provides means of acquiring and reconstructing sparse signals at rates substantially lower than Nyquist rates. The entire machinery of compressed sensing works in a non-adaptive fashion. Majority of the reconstruction algorithms operate on undersampled vectors, that represent the entire signal. In this research, an attempt has been made to explore compressed acquisition and reconstruction of continuous streaming data. Full stationarity in practical signals is unrealistic and after obtaining an initial estimate of signal coefficients in the sparsity inducing basis, acquisition and reconstruction becomes superfluous, as no new knowledge is gained. Non-stationary signals cannot be captured well under a compressed sensing framework, for sparsity assumption will not hold good even for any length

of reconstruction. Assumption of piecewise stationarity is reasonable, for any real world signal can be assumed to be stationary in a finite interval of time that is small compared to the total duration of acquisition. With such an assumption, compressed acquisition of non-stationary signals can be managed, with the reconstruction error being confined to a very small duration at the point where the signal transits from one stationary segment to the other.

The issue of non-stationarity having been addressed, the next question that was put forth is whether in a classical compressed sensing apparatus, is the full sampling capacity of the analogue-to-digital converter being utilized ? Could the idle sampling instants of the ADC, resulting from undersampling, be utilized to capture more than one signal ? This gave rise to the idea of multiplexed acquisition of multiple signals, in which a single ADC samples many signals. In typical embedded data acquisition systems, this would reduce the number of components in the board design, thereby contributing to reduction in size, power requirement and cost. The MOSAICS (Multiplexed Optimal Signal Acquisition Involving Compressed Sensing) scheme that has been proposed to capture this concept, can be easily realized using COTS devices.

The MOSAICS scheme was able to handle signals with component frequencies that had integral number of cycles in a reconstruction segment. This limited the application of MOSAICS to practical real world signals. To overcome this problem, the convex optimization in the signal reconstruction was replaced by a greedy algorithm that depends on MUSIC to detect the prominent frequencies. MOSAICS with MUSIC is able to handle a wider class of signals.

Next, the improvised scheme was applied to a closely related problem of detecting sinusoids buried in heavy noise. Since the detection of sinusoid frequencies is a simpler problem as compared to signal reconstruction, it can be managed well by compressed under-sampling, even with a few mea-

surements on the signal heavily corrupted with noise. The benefits of the multiplexed acquisition schemes introduced in this chapter are summarized as follows:

1. Optimal utilization of ADCs
2. More compact embedded hardware designs with lesser number of ADCs and the associated circuitry
3. Acquisition of continuous streaming data
4. Designs based on COTS components: no special design of the ADC, conventional ADCs can be used
5. Proposed architecture is independent of the reconstruction algorithm
6. Signals with non-integral frequency components can also be handled

A pertinent question at this juncture is whether Fourier sparsity is the only type of information redundancy in a set of signals. Individually, a signal may not be sparse on any basis, yet there may exist some kind of inter-signal relationship that may have the potential to cause information redundancy. Can compressed acquisition machinery leverage upon this redundancy in information, to reconstruct the signals, after a sub-Nyquist sampling is performed on them? Exploration of such possibilities is the content of the next chapter.

Chapter 4

Compressed Acquisition of Correlated Signals

4.1 Introduction

Very often applications in science and engineering encounter signals which do not have adequate sparsity, to be candidates for compressed sensing. Nevertheless, there arise situations in which a group of signals, not necessarily sparse, but having similar characteristics, have to be sampled simultaneously. Although, individually each signal may be information-rich, forbidding the use of an under-sampling strategy, there could be considerable redundancy of information in the group as a whole. In other words, the knowledge of a few signals in the group, at any instant of time, could enable us to predict other signals at that instant. Needless to say, a set of correlated signals is the most straightforward example of such signals. Applications in which such signals are commonly found are distributed sensor networks, multi-sensor data fusion and biomedical signals to name a few. Signals from different sensors, sensing the same physical process, are interrelated by virtue of spatial or temporal correlation. In some applications sensors may be monitoring different physical parameters but may still be related due to common phenomenon driving the signals. This is true for some avionics flight data that are sent to ground telemetry stations for post flight analysis.

Distributed Compressed Sensing (DCS) is the name given to a whole range of sampling and reconstruction algorithms that deal with the compressed sensing of an ensemble of signals having intra- and inter signal correlations. The DCS theory rests on the concept of joint sparsity of a group of signals. In recent years, researchers have addressed DCS as a Multiple Measurement Vector (MMV) problem. Chen and Huo [74] have given a thorough elucidation of the MMV problem in which, given a matrix whose columns are measured vectors that are projections of the source vectors on an overcomplete dictionary, the source matrix is recovered. The authors extend the classical single measurement vector (SMV) case to MMV, with respect to recovery algorithms like convex optimization and orthogonal matching pursuit. Zhang and Rao [75] have demonstrated their approach based on a block sparse Bayesian learning framework, in cases where the source vectors are temporally correlated, with good results. However, they make a key assumption of common sparsity amongst measured vectors (MV), which is valid only if the number of MVs is small. Reconstruction of two correlated sequences, in which the second sequence is the sum of the first sequence plus an additional innovation, has been handled deftly in [76] by making use of the expectation maximization (EM) algorithm. The method depends on the a priori knowledge of the spectral characteristics of the innovation component.

In this research, the focus is on real-time acquisition of streaming data from multiple sources that possess joint sparsity. Further, the reconstruction process does not have any a priori information about the signals. The correlation structure is incrementally learnt by the system. Within a relatively short period of time, the learning process is complete and subsequently, all the N signals are acquired using only $M < N$ ADCs. Unlike the methods proposed in the cited literature, the problem is treated as one of Single Measurement Vector (SMV), in which a single MV comprises the signal samples, at any instant from multiple sources. Before a detailed

description of the proposed idea is initiated, a brief introduction to the concept of joint sparsity is given in the next section.

4.2 Joint Sparsity

The notion of signal sparsity on any basis can be generalized to the *joint sparsity* of an ensemble of signals. Joint sparsity of a set of signals, is nicely captured by means of the Joint Sparsity Models (JSMs) introduced in [77]. The authors also give examples of applications in support of each of the models. Some of the information in these papers is reproduced verbatim below, to prevent loss of clarity.

4.2.1 JSM 1: Sparse Component + Innovations

Consider N signals $\mathbf{x}^{(i)}, i = 1...N$. Let each $\mathbf{x}^{(i)} \in \mathbb{R}^D$ [1]. Let $\mathbf{\Psi}$ represent a basis for $\mathbb{R}^D$ in which $\mathbf{x}^{(i)}$ can be sparsely represented. All the N signals share a common signal component, $\mathbf{Z}$ and each individual signal has a separate innovation component, $\mathbf{I}^{(i)}$.

$$\mathbf{x}^{(i)} = \mathbf{Z} + \mathbf{I}^{(i)} \tag{4.1}$$

$\mathbf{Z}$ has a sparse representation on the basis represented by $\mathbf{\Psi}$

$$\mathbf{Z} = \mathbf{\Psi}\mathbf{c}_{\mathbf{Z}} \tag{4.2}$$

$$\|\mathbf{c}_{\mathbf{Z}}\|_0 = K_{\mathbf{Z}} \tag{4.3}$$

[1] It is to be noted that undersampling each individual signal separately is not the objective; rather the aim is to sample only a subset of the N signals in the ensemble at every sampling instant. Thus, the vector that is reconstructed through compressed sensing has N elements and not D.

$K_{\mathbf{Z}}$ is small. Similarly,

$$\mathbf{I}^{(i)} = \mathbf{\Psi}\mathbf{c}^{(i)} \tag{4.4}$$

$$\left\|\mathbf{c}^{(i)}\right\|_0 = K_i, i = 1...N \tag{4.5}$$

Each of K_i is also small. Thus while the same sparse set of coefficients, $\mathbf{c}_{\mathbf{Z}}$ represents the common component in each of the N signals, a different set of coefficients $\mathbf{c}^{(i)}$ represent the innovation component in each signal.

A practical situation well-modeled by JSM-1 is a group of sensors measuring temperatures at a number of outdoor locations throughout the day. The temperature readings $\mathbf{x}^{(i)}$ have both temporal (intra-signal) and spatial (inter-signal) correlations. Global factors, such as the sun and prevailing winds, could have an effect on $\mathbf{Z}$ that is both common to all sensors and structured enough to permit sparse representation. More local factors, such as shade, water, or animals, could contribute localized innovations $\mathbf{I}^{(i)}$ that are also structured (and hence sparse). A similar scenario could be imagined for a network of sensors recording light intensities, air pressure, or other phenomena. All of these scenarios correspond to measuring properties of physical processes that change smoothly in time and in space and thus are highly correlated.

4.2.2 JSM 2:Common Sparse Component

In JSM 2 all the N signals are constructed from the same sparse set of basis vectors but with different coefficients. That is,

$$\mathbf{x}^{(i)} = \mathbf{\Psi}\mathbf{c}^{(i)}, i = 1...N \tag{4.6}$$

Each $\mathbf{c}^{(i)}$ has non-zero elements only at the common set of indices, $\Omega \subset \{1...D\}$ with $K = |\Omega|$ being a small value. Hence, all signals have the same l_0-norm equal to K, and all are constructed from the same K basis elements but with arbitrarily different coefficients.

A practical situation well-modeled by JSM-2 is where multiple sensors acquire replicas of the same Fourier-sparse signal but with phase shifts and attenuations caused by signal propagation. Another useful application for JSM-2 is MIMO communication [78].

4.2.3 JSM 3: Non-sparse Common Component + Sparse Innovations

JSM 3 is an extension of JSM 1 in which the common component is not necessarily sparse on any basis, which is to say that $K_{\mathbf{Z}}$ may not be small but K_i for $i = 1...N$ are small. In this case individual CS reconstruction of the signals is not possible.

A practical situation well-modeled by JSM-3 is where several sources are recorded by different sensors together with a background signal that is not sparse in any basis. Consider, for example, an idealized computer vision-based verification system in a device production plant. Cameras acquire snapshots of components in the production line; a computer system then checks for failures in the devices for quality control purposes. While each image could be extremely complicated, the ensemble of images will be highly correlated, since each camera is observing the same device with minor (sparse) variations.

Many other signal models have been considered by different authors in the area of simultaneous sparse approximation [78; 79; 80]. In this setting, a collection of sparse signals share the same expansion vectors from a redundant dictionary.

Each of the JSMs proposes a basic framework for joint sparsity among an ensemble of signals and are intentionally generic. In this research a variant of JSM 3, which is named as JSM 4 is introduced.

4.2.4 JSM 4: Non-sparse Common Component + Non-Sparse Innovations

JSM 4 is an extension of JSM 3 with the additional generalization that even the innovation components need not necessarily be sparse. In other words, each of K_i need not be small. To compensate for this generalization, it is assumed that the coefficients,$\mathbf{c}^{(i)}$ of the innovation components in the basis $\mathbf{\Psi}$ are small (say 10%) compared to those of the common component.

$$\frac{\left\|\boldsymbol{c}^{(i)}\right\|_{\infty}}{\left\|\boldsymbol{c}_{\mathbf{Z}}\right\|_{\infty}} \leq 0.1, \forall i \tag{4.7}$$

where $\|\mathbf{v}\|_{\infty}$ denotes the maximum norm of a vector, $\boldsymbol{v} \in \mathbb{R}^D$ defined below:

Definition:. *The maximum norm of a vector* $\boldsymbol{v}$ *is defined as*
$\|\boldsymbol{v}\|_{\infty} := max\left(|v_1|, |v_2|, ..., |v_D|\right)$

Thus, in a set of correlated signals, the innovation components can be viewed as noise present in individual signals. Multiple correlated measurements of a physical process polluted with measurement noise at the individual sensors is a real world situation which can fit into this model.

4.3 Sparsity Inducing Basis

Inter-signal correlation having been identified as the property that could cause information redundancy, a suitable sparsifying transformation that captures this property needs to be chosen. Let $\mathbf{X} \in \mathbb{R}^{T \times N}$ be the signal matrix, whose columns are indexed by the N correlated signals and whose rows index T successive sampling time instants, $\tau_1, \tau_1 + T_S, \tau_1 + 2T_S, ..., \tau_1 + (T-1)T_S$ where $T_S = 1/F_S$. F_S is the maximum sampling rate of each of the N ADCs sampling each of the signals [1]. Since $\mathbf{X}$ is real, its covariance matrix denoted by $\Sigma_{\mathbf{X}}$ is symmetric and its eigen vectors, $\mathbf{e}_n$ are orthogonal.

[1] $F_S \geq 2F$, where F is the real interval in which the constituent frequencies are present

Consequently, one can construct an orthogonal matrix $\mathbf{\Psi} = [\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_N]$, such that

$$\Sigma_{\mathbf{X}}\mathbf{\Psi} = \mathbf{\Psi}\Lambda \tag{4.8}$$

where Λ is a diagonal matrix consisting of the eigen values, $\lambda_1, \lambda_2, ..., \lambda_N$ corresponding to the eigen vectors, $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_N$. The transpose of the matrix $\mathbf{\Psi}$ is known as the **Karhunen Loeve Transform**,

$$\kappa = \mathbf{\Psi}^T \tag{4.9}$$

that can be used to transform each vector from a set of correlated vectors onto a basis in which they are highly decorrelated. KLT [81; 82] is a well known reversible linear transformation that removes redundancy in signals by decorrelating them. It has been extensively used in image compression wherein the correlation between neighboring pixels is exploited.

Let $\mathbf{y}^{\tau_1+(p-1)T_S}$ represent the transpose of the p^{th} row of the matrix $\mathbf{X}$, corresponding to the sampling instant $\tau_1+(p-1)T_S$ comprising the samples from the signal sources. Clearly, $\mathbf{y}^{\tau_1+(p-1)T_S}$ can be expressed as

$$\mathbf{y}^{\tau_1+(p-1)T_S} = \mathbf{\Psi}\mathbf{c} \tag{4.10}$$

where $\mathbf{c}$ is the coefficient vector[1]. On account of the correlation between the signals, $\mathbf{c}$ will be sparse, or equivalently, $|\mathbf{c}|_0$ will take a small value.

Let $\mathbf{y}^{\tau_1+TT_S}$ be a column vector of samples from the N signal sources, at the sampling time instant $\tau_1 + TT_S$ which is the next sampling instant, after those that have been considered in the computation of $\Sigma_{\mathbf{X}}$. Suppose at this time instant, samples from only $M < N$ sources are taken. The measured vector $\mathbf{f}$ is given by

$$\mathbf{f} = \mathbf{\Phi}\mathbf{y}^{\tau_1+TT_S} \tag{4.11}$$

where $\mathbf{\Phi}$, as before is the $M \times N$ measurement matrix comprising rows,

[1] $\mathbf{c}$ must be written with a superscript of $\tau_1 + (p-1)T_S$, however it is omitted here for simplicity

randomly, picked up from the *identity matrix*, $\mathbf{I}_N$. Again from (4.10) and (4.11),

$$\mathbf{\Phi\Psi c} = \mathbf{f} \tag{4.12}$$

Assuming that the correlation structure of the signal ensemble does not change at the time instant $\tau_1 + TT_S$, $\mathbf{c}$ will be sparse and a compressed sensing reconstruction would be able to estimate it.

$$\hat{\mathbf{c}} = argmin\left\|\mathbf{c}^{'}\right\|_1 \text{ subject to } \mathbf{\Phi\Psi c}^{'} = \mathbf{f} \tag{4.13}$$

From $\hat{\mathbf{c}}$, $\mathbf{y}^{\tau_1+TT_S}$ can be estimated using (4.10)

$$\hat{\mathbf{y}}^{\tau_1+TT_S} = \mathbf{\Psi}\hat{\mathbf{c}} \tag{4.14}$$

4.3.1 The CS-tuple

The elements of the CS-tuple can be clearly identified as follows:

- $\mathbf{\Psi}$ is the transpose of the KLT matrix κ
- $\mathbf{\Phi}$ is the downsized identity matrix, $\mathbf{I}_N^{(M)}$
- Ξ is l_1 minimization using basis pursuit[1] (3.8)

4.4 ARCS:Acquisition and Reconstruction of Correlated Signals

In this research work, a scheme is proposed for the acquisition and reconstruction of correlated signals. Given streaming data from N signal sources as input, ARCS captures the data using only $M < N$ analogue-to-digital-converters. The acquisition starts off with only M signals and after a relatively short duration is able to acquire all the N signals as the

[1]This is not essential since the proposed method is independent of the reconstruction algorithm

correlation structure of the signals is incrementally learnt. The following subsections give the details of the algorithm execution.

4.4.1 Objective of ARCS

The objective of ARCS is to acquire and reconstruct N correlated signals using $M < N$ ADCs after an initial learning phase.

4.4.2 Input

Input to ARCS is streaming data from N correlated signal sources of the JSM-4 class. The constituent frequencies of the common component and the innovation components are real numbers in the interval $[0\ F]$. Let $F_S > 2F$ be the sampling rate of each of the M ADCs.

4.4.3 Initialization

At any sampling instant, the signals can be divided into two subsets:

- The ***familiar*** subset, $\mathscr{F}$ consists of all the signals whose correlation structure has been learnt by ARCS.
- The ***stranger*** subset, $\mathscr{S}$ consists of all the signals whose correlation structure is yet to be learnt by ARCS.

At the start of the acquisition, that is $t = 0$, $\mathscr{F} = \{\}$ and $\mathscr{S}$ consists of all signals to be acquired.

4.4.4 Signal Acquisition and Learning

i) To start with, M signals are randomly chosen from $\mathscr{S}$ to form the set $\mathscr{K}$, the set of ***entrants***. The M signals in $\mathscr{K}$ are sampled using the M ADCs at a sampling period, $T_S = 1/F_S$. After T sampling instants, a $T \times M$ signal matrix, $\mathbf{X}^{(M)}$ is formed. After several sampling cycles,

the inverse KLT (4.8, 4.9,4.10) of $\mathbf{X}^{(M)}$ converges [1]. This completes the first ***convergence cycle***. Let the inverse KLT matrix be denoted by $\mathbf{\Psi}^{(M)}$. At this point, the M entrants are moved from the stranger set to the familiar set,

$$\mathscr{S} = \mathscr{S} - \mathscr{K} \tag{4.15}$$

$$\mathscr{F} = \mathscr{F} \cup \mathscr{K} \tag{4.16}$$

and the cardinality of $|\mathscr{F}| = M$.

ii) From the immediate next sampling cycle onwards, at every sampling instant, one out of the M signals in $\mathscr{F}$ is randomly picked up to be omitted[2]. The remaining $M-1$ signals in $\mathscr{F}$ are sampled using $M-1$ ADCs. Since only $M-1$ signals are actually sampled, the full set of M signals in $\mathscr{F}$ are estimated through a compressed sensing recovery operation as given in equations 4.13 and 4.14. This is done using the inverse KLT matrix, $\mathbf{\Psi}^{(M)}$ calculated previously. The measured vector, $\mathbf{f}$ consists of the $M-1$ samples measured through ADC and the measurement matrix $\mathbf{\Phi}$ is derived from the Identity Matrix $\mathbf{I}^M$ by omitting the row corresponding to the signal that is not sampled.

iii) Since only $M-1$ ADCs are used, there is one ADC free. Let $\alpha \in \mathscr{S}$ be a new stranger randomly chosen from $\mathscr{S}$ to form the set of ***entrants***, $\mathscr{K}$, now containing only one element. The free ADC is used to sample α. Thus at each sampling instant, $M-1$ signals in $\mathscr{F}$ are obtained by direct measurement, one signal (that was omitted) in $\mathscr{F}$ is estimated through compressed sensing reconstruction and another ***stranger*** signal, α is obtained by direct measurement. In other words with only M ADCs, samples of $M+1$ signals are obtained. With suc-

[1]Inverse-KLT convergence is detected, when the RMS difference between the vectors formed out of the corresponding eigen values at successive time instants is below a small threshold.

[2]There is no hard and fast rule that one has to take only one signal out. The scheme will work even if more than one are taken out and replaced by new ones. The intention behind taking one out is to ensure that the system learns the correlation in the signals smoothly.

cessive sampling cycles, the signal matrix $\mathbf{X}^{(M+1)}$ with $M+1$ columns, and increasing number of rows, is formed.

iv) After several such sampling cycles, the inverse-KLT, $\mathbf{\Psi}^{(M+1)}$ of order $M+1$ converges and the newcomer,α is moved from the ***stranger*** set, $\mathscr{S}$ to the ***familiar*** set, $\mathscr{F}$ as in (4.15) and (4.16). At this point, $|\mathscr{F}| = M+1$. This completes another ***convergence cycle***.

v) From the next sampling cycle onwards, two out of the $M+1$ signals in $\mathscr{F}$ are omitted out of direct measurement, which means again only $M-1$ familiar signals are measured using $M-1$ ADCs. Samples in $\mathscr{F}$ are estimated through compressed sensing reconstruction. The one ADC left is used to sample another stranger. This continues until convergence of the inverse-KLT of order $M+2$ after which the stranger is again admitted into $\mathscr{F}$ as in (4.15) and (4.16).

vi) This process continues until all the signals are transferred from $\mathscr{S}$ to $\mathscr{F}$. After this, the KLT convergence step is no longer required. M out of the N signals in $\mathscr{F}$ are sampled using only M ADCs. Samples of the remaining $N-M$ signals are estimated through compressed sensing with the help of the inverse KLT matrix of order N.

A detailed listing of the algorithm is given in Appendix B.

4.5 Simulation and Results

The performance of ARCS has been evaluated through simulation. The following subsections give a description of test signals, the derived parameters required as input to ARCS and the simulation results.

4.5.1 Test Signals

The ARCS system that has been simulated has $M = 5$ ADCs. The input to the system consists of $N = 10$ signals of the JSM-4 class, each of

which have K_i frequencies in the innovation component and the common component has K_Z frequencies[1]. For the common as well as the innovation components, the frequencies are real numbers in the interval $5 - 10$ Hz. F_S is equal to 20 Hz (see footnote for section 4.3). The amplitudes of the constituent frequency components are real numbers chosen randomly such that (4.7) is satisfied.

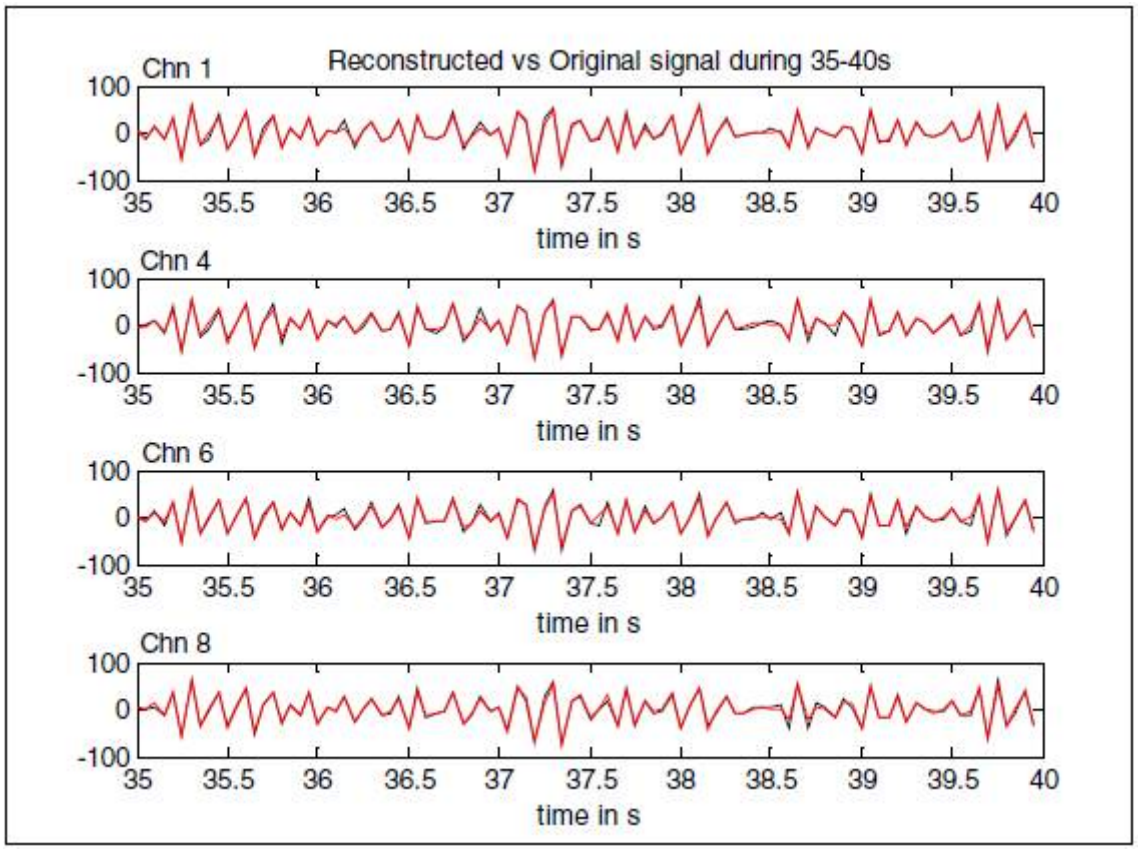


Figure 4.1: Reconstructed signal (shown in red color) is plotted against the original signal for four channels, 1, 4, 6 and 8 during the interval 35-40sec. The corresponding PSNR values in dB are: 25.6, 23.3, 22.6 and 24.0.

4.5.2 Simulation

The acquisition and reconstruction are simulated for a duration of 100 seconds. Snapshot (figure 4.1) of the reconstruction (red color) is plotted against the original (black color) for four out of the ten signals, in the

[1] K_i and K_Z are each equal to 50 in the simulations

time interval 35-40 sec. The reconstructed and the original signals show a very close match. It is asserted here, that the same degree of matching is exhibited for all the signals, throughout the simulation of 100 sec; although the plot for only four signals, in a 5 sec interval, is shown for brevity. The selection of the four signals chosen for display is also arbitrary.

In figure 4.2 is shown what can be called the 'Learning curve' of the system. This curve shows how the ten signals are incrementally introduced to ARCS and the correlation structure of the signals is learnt. Since initially, the number of signals picked up from the ***stranger*** set is equal to the number of ADCs ($M = 5$), at around 6 sec, the first inverse-KLT convergence of the first five signals takes place (each of the signals 1, 2, 3, 4, 5 are indicated by a square of different color). Subsequently, as each new signal is introduced there is an inverse-KLT convergence on a close-to linear learning curve. All ten signals are learnt by the system by 40 sec, subsequent to which all of them together are acquired with only five available ADCs. After this point, the acquisition and reconstruction continue without the inverse-KLT convergence step and there is no more learning taking place. The *cvx* toolbox has been used to carry out the l_1 minimization.

4.5.3 Robustness of ARCS

In order to probe into the robustness of ARCS, the common component in the same test data is changed twice to have a different set of constituent frequencies and amplitudes in the middle of the simulation, at sampling time instants of 60 sec and 80 sec. The innovation components remain unchanged. As a result of this non-stationarity, the Fourier sparsity is lost. However, for compressed sensing since it is the signal correlation (which is still maintained even after change in the frequency spectrum) that is made use of, the reconstruction quality does not deteriorate as indicated by the corresponding PSNR values for four signals (in the portion after introduction of the non-stationarity) in Table 4.1. Since the correlation

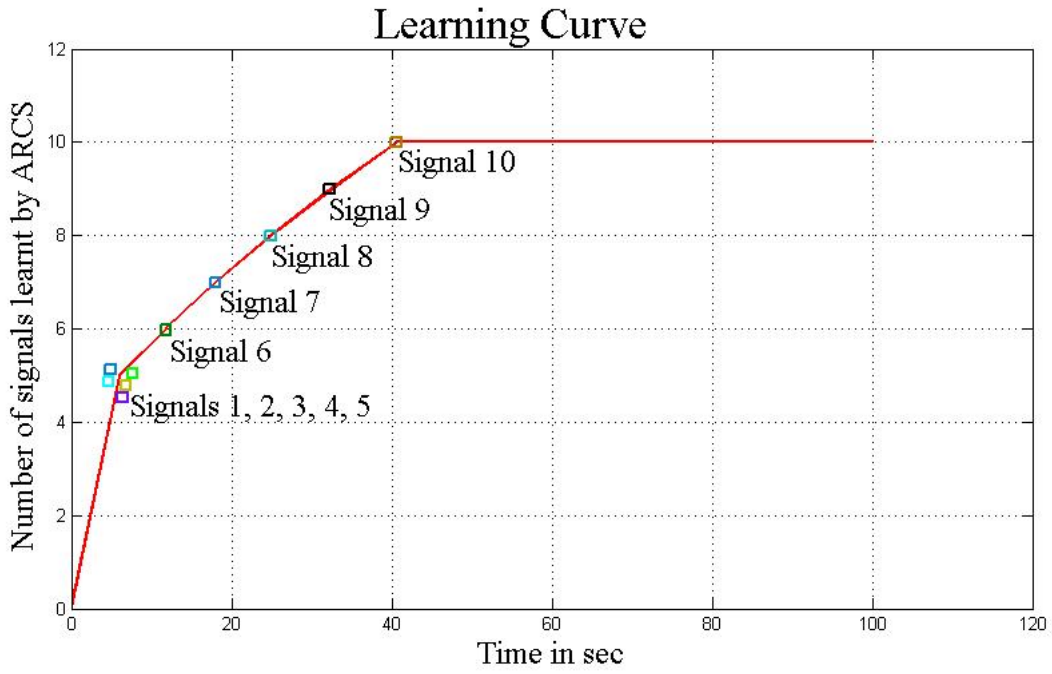


Figure 4.2: Under the ARCS scheme 10 correlated signals are acquired using 5 ADCs. The curve shows how the correlation structure is incrementally learnt by the system. The y-axis gives the total number of signals learnt at any time. The colored squares represent each of the signals. At 0 s all the signals are strangers; at around 6 s signals 1, 2, 3, 4, 5 are introduced; at around 12 s signal 6 becomes a familiar signal and the total number of familiar signals is 6. This continues until around 41 s when signal 10 becomes familiar.

structure has already been learnt by the system before the instant the non-stationarity is introduced, the reconstruction proceeds fairly well.

4.6 Conclusion

ARCS is a technique for real time capture and reconstruction of a set of streaming signals with joint correlation properties, using fewer analog to digital converters. This section is concluded with a listing of the advantages

Table 4.1: PSNR VALUES (IN dB) OF THE RECONSTRUCTED SIGNAL IF THE COMMON COMPONENT IN THE ORIGINAL SIGNAL IS REINITIALIZED

	Signal 1	**Signal 4**	**Signal 6**	**Signal 8**
After 60 sec	22.7	21.8	24.3	23.5
After 80 sec	25.8	22.7	21.1	21.5

offered by ARCS.

i) **Support for non-sparse signals** - It has been demonstrated through simulation, that a collection of signals, even though individually non-sparse, can be compressively sampled and reconstructed employing fewer ADCs. The JSM-4 model that has been introduced does not impose sparsity constraint on either the common component or the innovation components.

ii) **Robustness with respect to non-stationarity** - Once the correlation structure of the signals is learnt by ARCS, even if the frequency characteristics of the common component change, ARCS is able to reconstruct the signals.

iii) **Invariance to CS recovery algorithm** - ARCS does not depend on any particular compressed sensing recovery algorithm used. Although ARCS employs l_1-minimization with basis pursuit, CS recovery can equally well be done by other methods like matching pursuit, which could be faster but possibly less accurate. Essentially, the execution time of ARCS depends on the CS recovery method used. Thus, a faster algorithm can equally well be plugged into ARCS, to speed up its execution.

iv) **Practical measurement matrix** - As with other compressed acquisition strategies proposed in this research, the measurement matrix $\mathbf{\Phi}$ is very straightforward. At every sampling instant, it is just the downsized identity matrix, in which some of the rows are absent corresponding to the signals that are omitted from measurement. Thus the measured vector comprises samples of some of the signals obtained from direct measurement by ADCs. With this arrangement, the sampling architecture is quite simple and general, thereby having the potential to support a variety of applications.

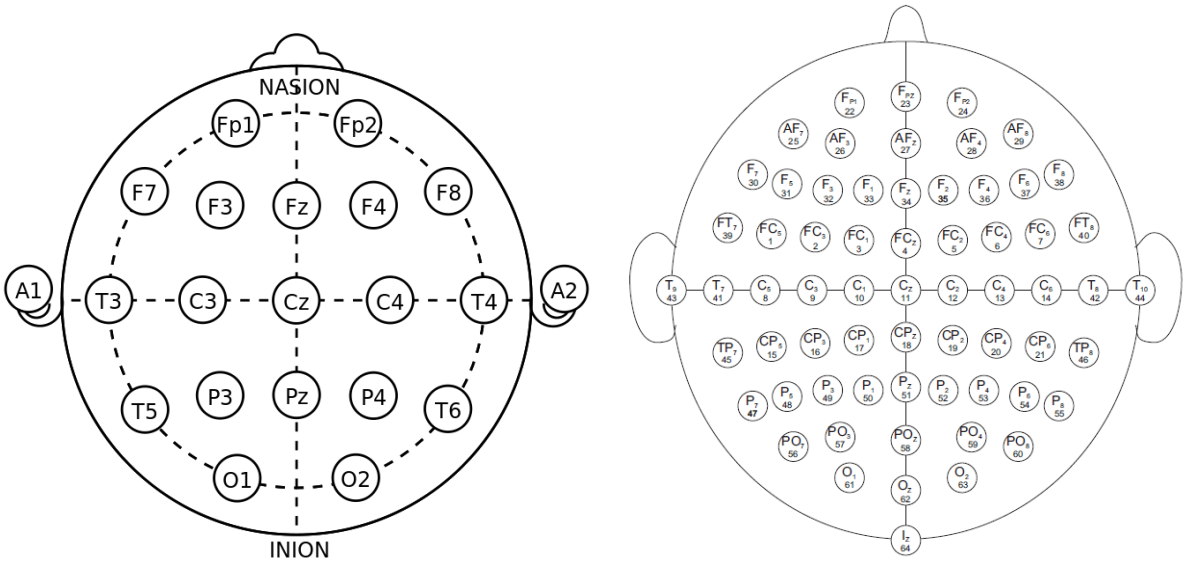
Although the performance capability of ARCS has been demonstrated with the help of simulation results, it is meaningful to explore the possibility of employing the technique to some practical application. This research is directed towards design of compact avionics data acquisition systems; however, the scope of using ARCS in an unrelated area of biomedical signal processing need not be overlooked, if there is a promise. Signals captured from the human scalp through multiple electrodes of Electro-EncephaloGraph (EEG) are known to be heavily correlated. Is it possible to perform an EEG clinical trial, using fewer EEG amplifiers than in a system that is ignorant of the inter-signal correlation ? The remaining part of this chapter is an attempt to answer this question.

4.7 Compressed EEG Acquisition using Estimated Channel Correlation

4.7.1 Electroencephalography

Electroencephalography [83; 84] is the neurophysiological measurement of the electrical activity of the brain using electrodes placed on the scalp (figure 4.3). The resulting record (signal) is known as electroencephalogram (EEG) and each EEG channel represents an electrical signal (post-synaptic potentials) from a large number of neurons. The EEG is a non-invasive procedure used for diagnostic purposes. Instead of electrical currents, the voltage differences between different parts of the scalp (brain) are observed. EEG measurements allow both time–domain and frequency analysis. The following important frequency bands have been identified in the EEG signal:

- delta: 0.1 - 3.5 Hz (deep sleep, pathological when awake)
- theta: 4 - 7.5 Hz (creativity, falling asleep)



(a) 10-20 system. Courtesy: Wikipedia **(b)** 10-10 system Courtesy: Physionet

Figure 4.3: EEG electrode placement systems

- alpha: 8 - 13 Hz (relaxation, closed eyes)
- beta: 14 - 30 Hz and more (concentration, logical and analytical thinking)
- gamma: 30 - 100 Hz (simultaneous processes)

These values have been taken from [85] and shall be used in reporting the results of simulations.

4.7.2 Standards for EEG electrode placement

The first internationally accepted standard for electrode placement is the 10/20 system (figure 4.3a) that describes head surface locations via relative distances over the scalp between cranial landmarks. The primary purpose of this standard is to provide a reproducible method for placing a relatively small number (typically 21) of EEG electrodes for various trials. With the advent of multi-channel EEG systems, the need was felt for extending the 10/20 system to higher density electrode settings for use in research and diagnosis. This led to the introduction of the 10/10 system (figure

4.3b) by Chatrian et. al. in 1985, consisting of 64 electrodes, as a logical extension of the original 10/20 system. While electrodes are placed at distances of ten and twenty percent along certain contours over the scalp in the 10/20 system, they are placed at distances of ten percent along the medial-lateral contours in the 10/10 system. Also, new contours are introduced in between the existing ones. The 10/5 system with even higher electrode density was proposed by Oostenveld and Praamstra in 2001. An elaborate description and comparison of all these systems is given in [86].

4.7.3 Background work on compressed sensing of EEG signals

Although EEG signals exhibit characteristic spectral peaks for different states of the brain, they are not in general sparse in the frequency domain. Since EEG analysis in many applications (e.g. brain electrical activity mapping) is governed by the availability of information about the spectral content in various bands, accuracy requirement of signal reconstruction per se, at such locations may be relaxed. It is important to mention here significant work, that has been reported previously, on the frequency analysis of EEG signals. Time-frequency analysis of EEG data based on adaptive periodogram technique has been proposed in [87]. Identification of the signal components through decomposition of data into time-frequency-space atoms (based on the Wigner-Ville distribution) using parallel factor analysis has been proposed in [88]. Time-frequency spectral estimation of multichannel EEG using smooth, time-frequency localized versions of the Fourier functions has been reported in [89].

In [90], the authors have exploited the joint sparsity of EEG signals on the Gabor frame and have achieved a low normalized mean square error. However in this work, the different trials have been treated as different electrodes with the assumption that in both cases the same underlying activity is measured. In [91], the authors have suggested a novel approach of structuring individual signals into groups and exploiting the group sparsity

by computing the $l_{1,2}$ norm. However their approach involves the use of the unconventional random sampling based acquisition architecture and does not exploit the joint sparsity of a group of signals. In [92], the authors have demonstrated the use of fast ICA as a preprocessing step before compressed acquisition of EEG signals to achieve a low reconstruction error. In [93], the authors have demonstrated a novel method of automated EEG analysis that segments and extracts EEG features, classifies and groups them according to various patterns, and then presents them in a compressed fashion to permit real-time viewing of several hours of EEG. In [94] the authors have exploited the sparsity of EEG on the slepian basis for low rate sampling and compressed sensing reconstruction. In [95], compressed sensing reconstruction of EEG exploiting sparsity on a Gabor frame with a Gaussian measurement matrix has been quantitatively presented. In [96], the authors propose a novel approach of application of compressed sensing in wireless telemonitoring of EEG through body are networks. In [97], the authors have presented a continuous biomedical signal acquisition system in which the sparsified measurements of signals are wirelessly transmitted to a fusion center through body sensor networks with emphasis on lower sampling rate at each sensor location. A compressed sensing framework that exploits both the temporal and spatial correlations to efficiently compress EEG signals in wireless body sensor networks is proposed in [98]. The first study [99; 100] that addressed the use of CS in EEG signal compression for telemedicine applications focused on surveying existing sparsifying dictionaries and reconstruction algorithms and testing different combinations of these elements to determine which one yielded the best results. The conclusion was that the applicability of single-channel CS for EEG signals depended on the intended application and the tolerable reconstruction error.

The objectives of the work done in these citations can be more or less classified into one of the following:

1. Compression of the EEG signal for subsequent storage or transmission efficiency
2. Reduction in sampling rate in individual channels by making use of intra-signal sparsity.

4.7.4 Inter-channel correlation in EEG

Nearby scalp channels of multi-channel EEG exhibit high correlation because EEG signals are not produced in the scalp or the neurons (brain) directly under the recording electrodes. Instead, as suggested in [101], they are a consequence of partial synchrony of local field potentials from distinct cortical domains - each domain, in the simplest case, being a patch of cortex of unknown extent. At any electrode, the EEG recording is a weighted linear mixture of underlying cortical source signals. The strong correlations observed between EEG recordings at nearby electrodes can be attributed to the spatial mixing of EEG source signals by volume conduction. Significant research effort has gone into exploring the correlation between EEG recordings at electrodes on different areas of the scalp. In [102; 103; 104], heavy correlation, sometimes as high as 0.9, has been reported between anterior-posterior EEG signals in the alpha band (8-13 Hz). Very high coherence in the delta band has been reported in [105] between posterior temporal lobe regions. Inter-hemispheric coherence in the gamma band has been studied in normal adults in [106]. Existence of very high correlation in the alpha band between temporal regions of the human brain has been reported in [107].

The primary interest in understanding inter-channel correlation in multi-channel EEG is to identify information redundancy in a measurement involving the full set of electrodes. If the channels are correlated, is it always mandatory to make recordings at all the electrodes, particularly those in close vicinity ? This conjecture is motivated by supporting literature cited above that provides evidence of correlation between nearby EEG channels.

The Karhunen-Loeve transform (KLT) has already been used in the ARCS setup (section 4.3) as a sparsifying transform for a set of correlated signals. In ARCS, the correlation structure of the signals is incrementally learnt by the system during acquisition. However, in the case of EEG signals, the KLT matrix can be computed beforehand. Consider the scenario, where the electroencephalogram of a patient undergoing treatment or a subject voluntarily involved in research, has to be frequently taken. The first few sessions can constitute the training phase in which measurements from all the defined set of electrodes are taken and used to compute the inverse KLT matrix. Once the correlation is learnt, one could do away with measurement at more than fifty percent of the electrodes and yet be able to estimate the EEG spectral signature at the locations of the missing electrodes. An immediate implication for the design of a scheme with a limited number of amplifiers is reduced cost and setup time. On the other hand, it may also be possible to reconstruct channels that are noisy or missing altogether due to electrode movement, etc.

4.7.5 Compressed Sensing of EEG Signals

The significant focus of the work done in this thesis is not to make use of the intra-signal sparsity; rather it is to exploit the inter-signal sparsity. Undersampling in this case is across the channels. In other words, at any instant, with samples from only a subset of the channels, the samples from the entire set of channels are reconstructed.

This, it is anticipated will lead to reduction in the number of electrodes used in a recording session, with the following practical implications:

1. In clinical recording of EEG, a large number of channels (21 upwards) are recorded. In recording of sleep EEG, the above recording continues for a long time also. It will be of great use, if there is a way of reconstructing one or more channels lost or not properly recorded due to electrode movement, drying up or discontinuity.

2. Ambulatory EEG [108] is carried out for an extended period (up to 72 hours) in which the patient can move about freely during the recording and data is stored in a pocket recorder. Thus, EEG recorded on a subject at rest, using a dense set of electrodes in the neurological clinic, can be used for training. Subsequently when the subject is in motion, all the channels need not be monitored. This can result in less number of electrodes and consequent reduction in weight of the ambulatory EEG recorder. Sub-section 4.7.10 demonstrates the results of simulation in which the signals at the locations of the electrodes in the 10-10 system are reconstructed using samples of the signals at the locations of 10-20 electrodes.

3. Reliability in clinical applications such as topographic mapping of brain electrical activity (BEAM) could be increased by ensuring that any missed or noisy channel can be predicted (recovered) reasonably well from the other channels, thereby creating redundancy in the number of channels recorded. Here there is no intentional sub-sampling of the channels.

4. In sleep studies, it is possible that data is missing on some channels, either due to noise or due to undesirable movement by the subject. In this case, the loss of data, treated as undersampling, can be handled by recovery through compressed sensing. Once again there is no intentional sub-sampling of the channels.

5. Deviation (due to loss of correlation) beyond a threshold between the signal values estimated through compressed sensing and the actual measurements can be used to detect the onset of seizure in epileptic subjects. In this case too, there is no intentional sub-sampling.

Before proceeding further, it is prudent to be aware of the limitations of such a proposition.

4.7.6 Limitations

One cannot ignore the fact that the inter-channel correlation may be weak and time varying. However, in applications such as EEG brain electrical activity mapping, where we are mainly concerned with the relative signal content in various frequency bands, accuracy of signal reconstruction per se, at such locations can be relaxed. The focus of this work is restricted to providing empirical evidence of obtaining approximate signal reconstruction and a reasonably good estimate of the spectral content in all the channels of EEG by recording over a limited number of channels. Some of the implications of such an effort on clinical diagnosis and research have been listed previously in subsection 4.7.5. It is definitely desirable to be able to detect small changes in the correlation pattern, which might have diagnostic significance. However, no experiments have been carried out in this work to verify that this can be accomplished by the method proposed here.

4.7.7 A note on sparse PCA

The name KLT has been synonymously used with classical principal component analysis (PCA) by the signal processing community. In the light of this, it is very pertinent to mention that the method proposed in this work is not the same as the sparse PCA approach [109; 110]. In classical PCA the matrix $\mathbf{\Psi}$ (see equation 4.8) is formed out of the eigen vectors of the autocorrelation matrix. In other words, one tries to maximize $\psi_n^T \Sigma_X \psi_n$ subject to $\|\psi_n\|_2 = 1$.

On the other hand, the sparse PCA approach seeks sparse principal components that span a low dimensional space. The matrix $\mathbf{\Psi}$ is found by solving an optimization problem with a sparsity constraint on its entries. Equivalently, one tries to maximize $\psi_n^T \Sigma_X \psi_n$ subject to both $\|\psi_n\|_2 = 1$ and $|\psi|_0 = K$ where K is the parameter that controls the sparsity. As in regular

PCA (or KLT), in this work also the sparsity constraint is not imposed on $\mathbf{\Psi}$. Instead, the sparsity constraint is applied on the vector $\mathbf{c}^{'}$ in the minimization in the equation (4.13) that also involves the matrix $\mathbf{\Psi}$ formed using the standard PCA with the help of plain matrix algebra. Thus, throughout the process of acquisition and reconstruction of the signals, the matrix $\mathbf{\Psi}$, referred to as the KLT matrix and calculated previously from the training data set, remains unaltered.

4.7.8 The Physionet database

The simulation has been carried out on signals in the publicly available *Physionet* database [111]. Detailed description of the database is given in [1]. This data set consists of over 1500 one- and two-minute EEG recordings, obtained from 109 volunteers. The subjects performed different motor/imagery tasks (see Table 4.2) while 64-channel EEG was recorded, at a sampling frequency of 160 Hz, using the BCI2000 system [112]. The placement of the electrodes is as per the international 10-10 system (excluding electrodes Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10). Each volunteer has performed, in a sequence, a set of 14 tasks: two baseline tasks followed by four different tasks repeated three times. Thus, the total number of records in the database is $109 \times 14 = 1526$.

Table 4.2: MOTOR/IMAGERY TASKS IN THE PHYSIONET DATABASE DURING WHICH THE EEG USED FOR THE STUDY HAS BEEN COLLECTED.(see [1])

TASK	DESCRIPTION OF THE TASK, DURING WHICH EEG IS RECORDED.
Baseline 1	Eyes open. **DURATION: 1 sec**
Baseline 2	Eyes closed. **DURATION: 1 sec**
Task 1	A target appears on the left or the right side of the screen. The subject opens and closes the corresponding fist until the target disappears. Then the subject relaxes. **DURATION: 2 sec**
Task 2	The stimulus is same as Task 1. However, in this case the subject imagines responding to the stimulus the same way as in Task 1 and then relaxes.**DURATION: 2 sec**
Task 3	A target appears on the top or the bottom of the screen. The subject opens and closes both fists if the target is on top and both feet if the target is on the bottom until the target disappears. Then the subject relaxes. **DURATION: 2 sec**
Task 4	The stimulus is same as Task 3. Again the subject imagines responding to the stimulus the same way as in Task 3 and then relaxes. **DURATION: 2 sec**
	After this, tasks 1 to 4 are repeated two times and stored in records 7 to 14. In other words, Records 7 and 11 correspond to task 1, records 8 and 12 correspond to task 2, records 9 and 13 correspond to task 3 and records 10 and 14 correspond to task 4.

4.7.9 The experiments

In order to compare the reconstructed signals with the original, a fractional spectral measure (FSM)[113] is used for each of the significant frequency bands - delta, theta, alpha, beta [1]. FSM is the ratio of the sum of the absolute values of the 512-point DFT coefficients within a band to the corresponding sum in the 0–30 Hz (until the end of the beta band). Thus, FSM for the i^{th} band is

$$FSM_i = \frac{\sum_{F_j \in band\, i} (abs(F_j))}{\sum_{F_t \in 0-30Hz} (abs(F_t))} \tag{4.17}$$

where i denotes one of the bands: *delta, theta, alpha* and *beta*. For example, for *theta* band, the numerator in the equation above is equal to the sum of the absolute values of DFT coefficients in the frequency range 4–7.5 Hz. A size of 512 has been chosen for the DFT so as to get a good frequency resolution of 0.3125 Hz in the DFT spectra, given that the sampling frequency of the data is 160 Hz. The FSM values for the original and the reconstructed signals are compared. The experiment comprises the following steps, that are categorized into the training and the testing phases for the sake of clarity:

4.7.9.1 The Training Phase

Step 1: At random, choose six subjects to be included in the test set – say, 1, 8, 41, 61, 77, 104

Step 2: For each subject, pick a record corresponding to one of the tasks 1, 2, 3 and 4, representing different motor/imagery tasks, to be utilized for training.

[1]Since the records in the physionet database have been sampled at 160 Hz, looking at frequencies above 80 Hz may not be of much use. Therefore, while reporting the results the gamma band has not been considered. Also the signal power in the gamma band for the simulation records was far too low to justify any consideration

Step 3: For the record that is picked, compute the inverse KLT matrix for the channels (twenty in number) - Fc5, Fc3, Fc1, Fcz, Fc2, Fc4, Fc6, C5 C3, C1, Cz, C2, C4, C6, Cp5, Cp3, Cp1, Cpz, Cp2, Cp4 for all samples in the record.

It is worth mentioning here that this way of computing the inverse KLT matrix (on all the samples in the record) led to performance improvement in the subsequent testing phase, as against the calculation of inverse KLT from the mean of inverse KLT matrices of successive (non-overlapping) windows of 1000 observations in the record. The chosen channels correspond to a set of closely spaced electrodes on the scalp (see figure 4.3b).

4.7.9.2 Testing phase

Step 4: For each of the subjects 1, 8, 41 and 61, one of the records 7 to 10, that corresponds to the same motor/imagery task as the record which was used for computing the inverse KLT matrix, is used as the test record. For example, if record 3 is used in the training phase, then record 7 (that corresponds to the same motor/imagery task, see table 4.2) is used for testing. Similarly, if the training record is 4, the testing record will be 8 and so on. In the case of subjects 77 and 104, the motor/imagery task that is chosen as the test record is different from the one used in the training phase. For example, for subject 77, records 3 and 14 are chosen for training and testing respectively. Similarly, for subject 104, records 4 and 13 are chosen for training and testing, respectively. This is done with the objective of observing the robustness of the method. At each successive time instant during 0-10 sec of the test record, a sub-sampling of channels is carried out, that is, only a subset consisting of ten channels is measured. This subset is different for each subject and the members of the subset are picked up arbitrarily. The samples from the ten channels form the measured vector, $\mathbf{f}$. The remaining ten are estimated through l_1 minimization (4.13) using the *cvx* toolbox [114]. The mean inverse KLT

matrix is used in the compressed reconstruction algorithms.
It is important to note here that the records used for training and testing are different. Thus, although all the channels of a training record are used for training, only a subset of the channels in the testing record are sampled at any instant during testing.
Step 5: Compute the DFT[1] for the original and reconstructed signals.
Step 6: Compute the fractional spectral measure for the original and the reconstructed signals and compare.

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

4.7.9.3 Results

Figure 4.4 shows the plots of the reconstructed and the original signals for nine channels (only 9 out of 10 reconstructed channels are shown due to space constraint; channel 1 is not shown) for subject 104. The original and reconstructed signals have a close match in the signal as well as the frequency domains. It is known that unlike the case of signals like ECG or nerve action potentials, the exact EEG signal amplitude at any instant is not clinically significant. Thus, the reconstruction quality obtained with the proposed method is good enough to be used for most EEG applications.

4.7.9.4 Comparison of performance using other reconstruction algorithms

In order to demonstrate that the proposed scheme is independent of the CS reconstruction algorithm employed, two other recovery algorithms: OMP (Orthogonal Matching Pursuit) and CoSAMP (Compressed Sampling Match-

[1]In order to compute the DFT, the entire signal is divided into segments of length 512, with a fifty percent overlap. The last segment is padded with sufficient number of zeros. To each segment, a 512-point Hamming window is applied. A 512-point DFT of each windowed segment is calculated. The absolute values of the DFT coefficients are averaged over all the segments.

ing Pursuit)[1] have been used in place of basis pursuit. A comparison of the FSM values of the original and reconstructed signals for record 10 of subject 61 is given in Table 4.3. The values in the table indicate that the performances of all the recovery algorithms are comparable, though the basis pursuit approach performs marginally better. The results of simulation for all the six subjects: 1, 8, 41, 61, 77 and 104 are given in graphical form in figures 4.5,4.6 and 4.7. The figures plot the FSM values, in different bands, for each of the EEG channels reconstructed through CS with a different color and line style used for each recovery method. The legend for the various plots is given at the top of each page. There is considerable overlap between the plots for the various recovery methods and they closely follow the original (black color) for a majority of the subjects.

Comparison with classical compression - These results have been compared with a classical, lossy compression scheme (see bottom right subtable in Table 4.3). For this purpose, at any instant during acquisition, the samples of the EEG channels that have been left out of measurement are not recovered with the help of compressed sensing. Instead at every instant, a vector consisting of one sample from each channel is reconstructed by applying the inverse KL transform on a vector that retains only the top half of the transform domain coefficients. It is to be noted here that the comparison is with a classical compression that is inter-channel and not intra-channel. The plot (green) of FSM values very closely follows the original in the figures 4.5,4.6 and 4.7, as expected.

Figure 4.8 gives a summary of the comparison between various recovery algorithms (the lossy compression and decompression scheme is given by the brown colored bar) with respect to the percentage error (of FSM between the original and reconstructed signals) in various EEG bands, averaged across all the subjects.

Comparison of execution time - The average execution time taken

[1]For this purpose, open source Matlab code available on the Internet has been used.

across all the subjects for a simulation of EEG acquisition of 10s is 161.7 s for basis pursuit, 0.67 s for OMP and 1.23 s for CoSAMP. The computational platform is the same as that for which the specifications have been given at the beginning of sub-section 3.5.5.

4.7.10 Reconstruction of signals at 10-10 locations from measurements at 10-20 locations

Using recordings on channels in the international 10-20 system of electrode placement, it is possible to estimate, with fair accuracy, the spectral content of channels in the denser 10-10 system. To illustrate this, the inverse KLT matrix has been calculated for subject 64 in the physionet database, using channel numbers 8 to 14 of the 10-10 system, i.e. C5, C3, C1, Cz, C2, C4 and C6 that fall in a straight line on top of the scalp. For testing, data from channel numbers 9, 11 and 13 (i.e C3, Cz and C4) at the 10-20 electrode locations only have been sampled. Even then it was possible to reconstruct the remaining four channels - C5, C1, C2 and C6 (see figures 4.3a and 4.3b) that belong to the 10-10 system. Figure 4.9 compares the FSM values of the original signal and those reconstructed using basis pursuit recovery algorithm. (The average error in the delta band is as low as 3.0% (or equivalently the spectral fidelity is about 97%). The plots of the reconstructed channels vs the original along with the corresponding DFT magnitudes are shown in figure 4.10. The results establish the effectiveness of the proposed technique to reconstruct 10-10 system signals from the recorded 10-20 system signals.

4.7.11 Conclusion

A novel approach of subsampling EEG channels, by measuring only a subset of electrodes and reconstructing the remainder through compressed sensing, has been presented. Empirically it has been demonstrated that if the correlation amongst the channels is captured with good accuracy, then

by recording at only a few locations on the scalp, it is possible to estimate the relative signal content in different frequency bands with reasonable accuracy. Thus, if it is a matter of only knowing the relative spectral content, measurement of only a few EEG channels suffices, provided the correlation is previously captured in the inverse KLT matrix using data recorded in an initial training session. However, this training requirement limits the applicability of this technique only to those cases, where repeated recordings are required from the subject for monitoring or follow up. Hence, this is not applicable for one-time testing of an occasional patient.

The method proposed is suitable for real time data capture since the computationally intensive reconstruction can be done offline. Through simulation, it has been shown that the relative frequency content can be estimated within an accuracy of 15%. It has also been shown that recordings on the 10-20 system can be used to estimate the signals on electrodes in the 10-10 system with more than 90% spectral fidelity. Further, simulation results demonstrate that the performance of the proposed method is independent of the reconstruction algorithm employed. The reconstruction accuracy with BP, OMP and CoSAMP recovery algorithms are comparable though BP is marginally superior to the rest and is considered to be more consistent. OMP and CoSAMP, on the other hand, are very fast compared to BP and are suitable for real-time acquisition.

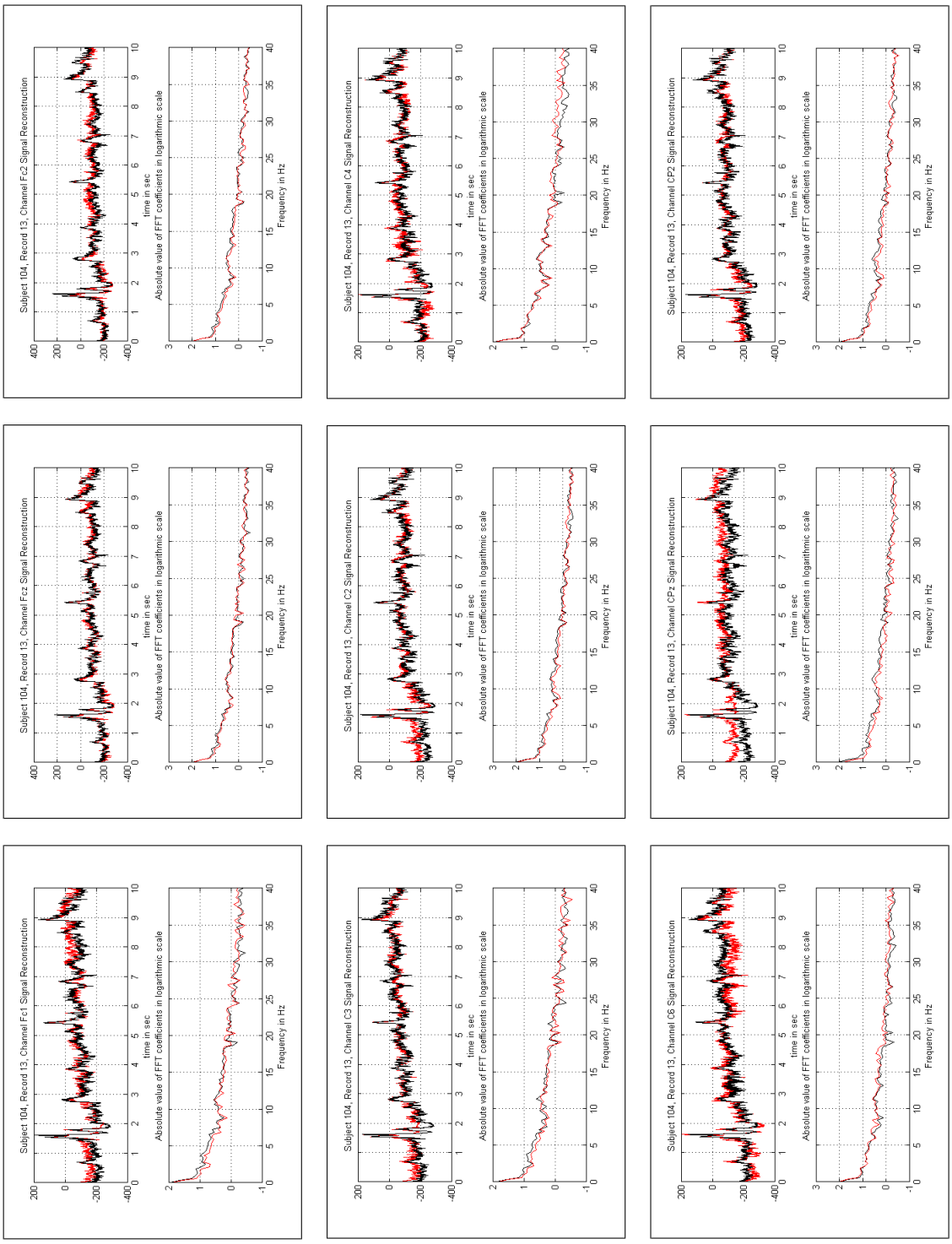


Figure 4.4: Reconstructed (red) vs original (black) signals for subject 104, record 13 using Basis Pursuit recovery algorithm

Table 4.3: COMPARISON OF FSM IN DIFFERENT BANDS OF THE ORIGINAL AND RECONSTRUCTED EEG FOR DIFFERENT CHANNELS OF SUBJECT 61, RECORD 10 USING DIFFERENT CS RECONSTRUCTION ALGORITHMS

delta band					
Chnl	Original	BP	OMP	CoSAMP	Decompr.
Fc5	0.420	0.416	0.403	0.417	0.419
Fc1	0.415	0.408	0.402	0.408	0.414
Fcz	0.407	0.403	0.400	0.403	0.406
Fc2	0.413	0.407	0.395	0.403	0.412
Fc6	0.411	0.402	0.390	0.399	0.412
Cz	0.394	0.401	0.397	0.402	0.393
C2	0.395	0.400	0.402	0.399	0.396
C4	0.401	0.389	0.407	0.377	0.402
CP3	0.386	0.388	0.395	0.385	0.384
CP4	0.391	0.395	0.388	0.390	0.390
Avg error	-	1.5%	2.6%	1.8%	0.3%

theta band					
Chnl	Original	BP	OMP	CoSAMP	Decompr.
Fc5	0.131	0.120	0.113	0.124	0.131
Fc1	0.145	0.127	0.116	0.127	0.144
Fcz	0.137	0.126	0.119	0.122	0.139
Fc2	0.137	0.118	0.115	0.121	0.137
Fc6	0.136	0.114	0.114	0.118	0.136
Cz	0.116	0.122	0.121	0.120	0.116
C2	0.118	0.119	0.121	0.119	0.119
C4	0.119	0.120	0.120	0.122	0.119
CP3	0.101	0.109	0.115	0.105	0.101
CP4	0.107	0.111	0.117	0.110	0.107
Avg error	-	7.7%	11.0%	6.6%	0.4%

alpha band					
Chnl	Original	BP	OMP	CoSAMP	Decompr.
Fc5	0.106	0.104	0.112	0.106	0.105
Fc1	0.111	0.117	0.115	0.110	0.111
Fcz	0.116	0.116	0.115	0.109	0.116
Fc2	0.107	0.116	0.110	0.112	0.107
Fc6	0.106	0.116	0.109	0.113	0.107
Cz	0.114	0.111	0.113	0.112	0.115
C2	0.116	0.112	0.109	0.109	0.115
C4	0.112	0.126	0.110	0.110	0.112
CP3	0.111	0.113	0.106	0.112	0.112
CP4	0.120	0.117	0.112	0.112	0.119
Avg error	-	4.8%	3.7%	3.6%	0.6%

beta band					
Chnl	Original	BP	OMP	CoSAMP	Decompr.
Fc5	0.198	0.197	0.212	0.199	0.199
Fc1	0.192	0.194	0.208	0.205	0.193
Fcz	0.188	0.195	0.205	0.212	0.187
Fc2	0.206	0.202	0.222	0.214	0.206
Fc6	0.206	0.207	0.227	0.218	0.205
Cz	0.185	0.187	0.192	0.186	0.189
C2	0.185	0.194	0.192	0.197	0.186
C4	0.194	0.202	0.185	0.233	0.193
CP3	0.211	0.205	0.199	0.212	0.213
CP4	0.190	0.190	0.197	0.200	0.192
Avg error	-	2.2%	6.4%	6.3%	0.6%

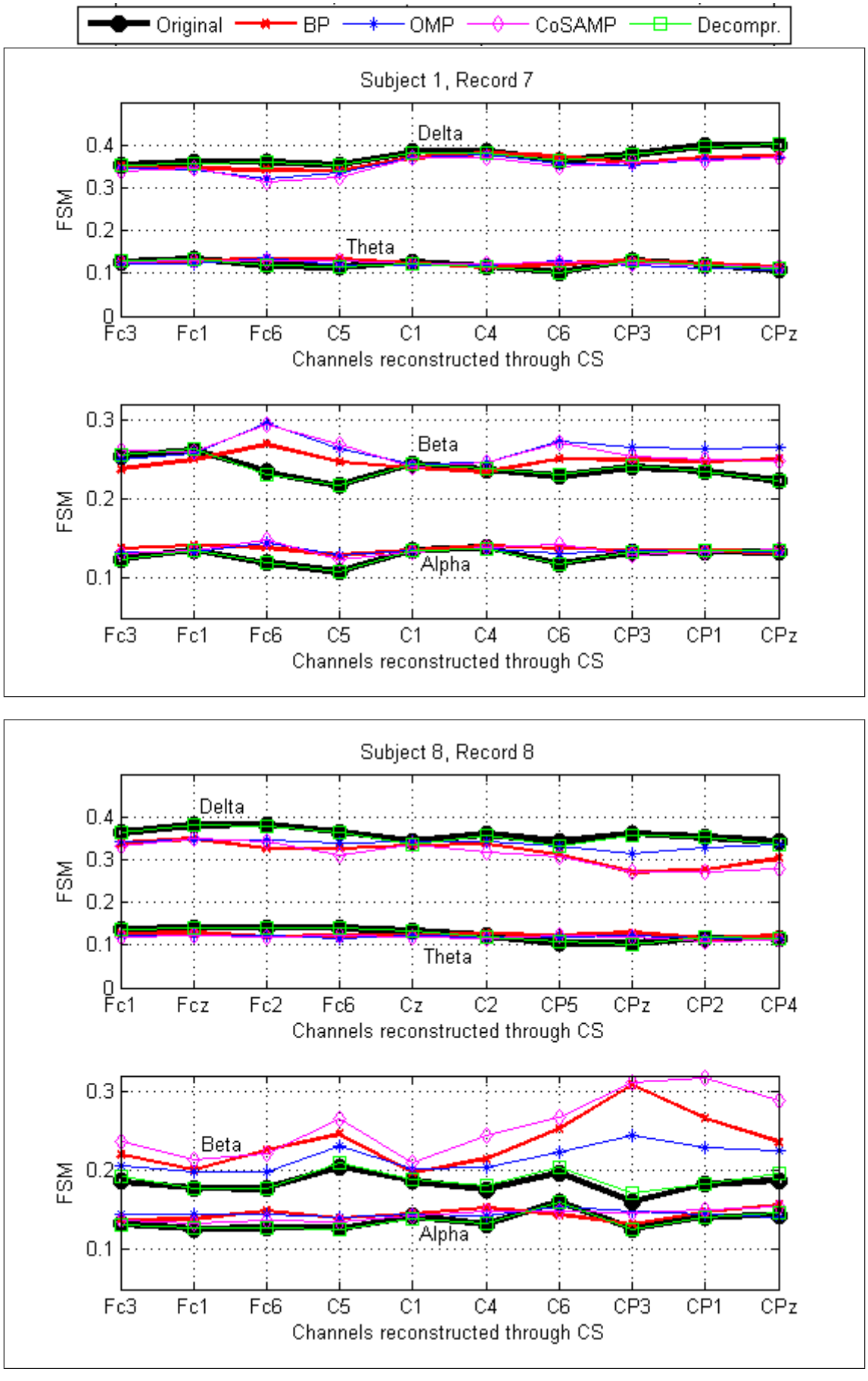


Figure 4.5: Comparison of FSM values of the 10 recovered channels for subjects 1 and 8

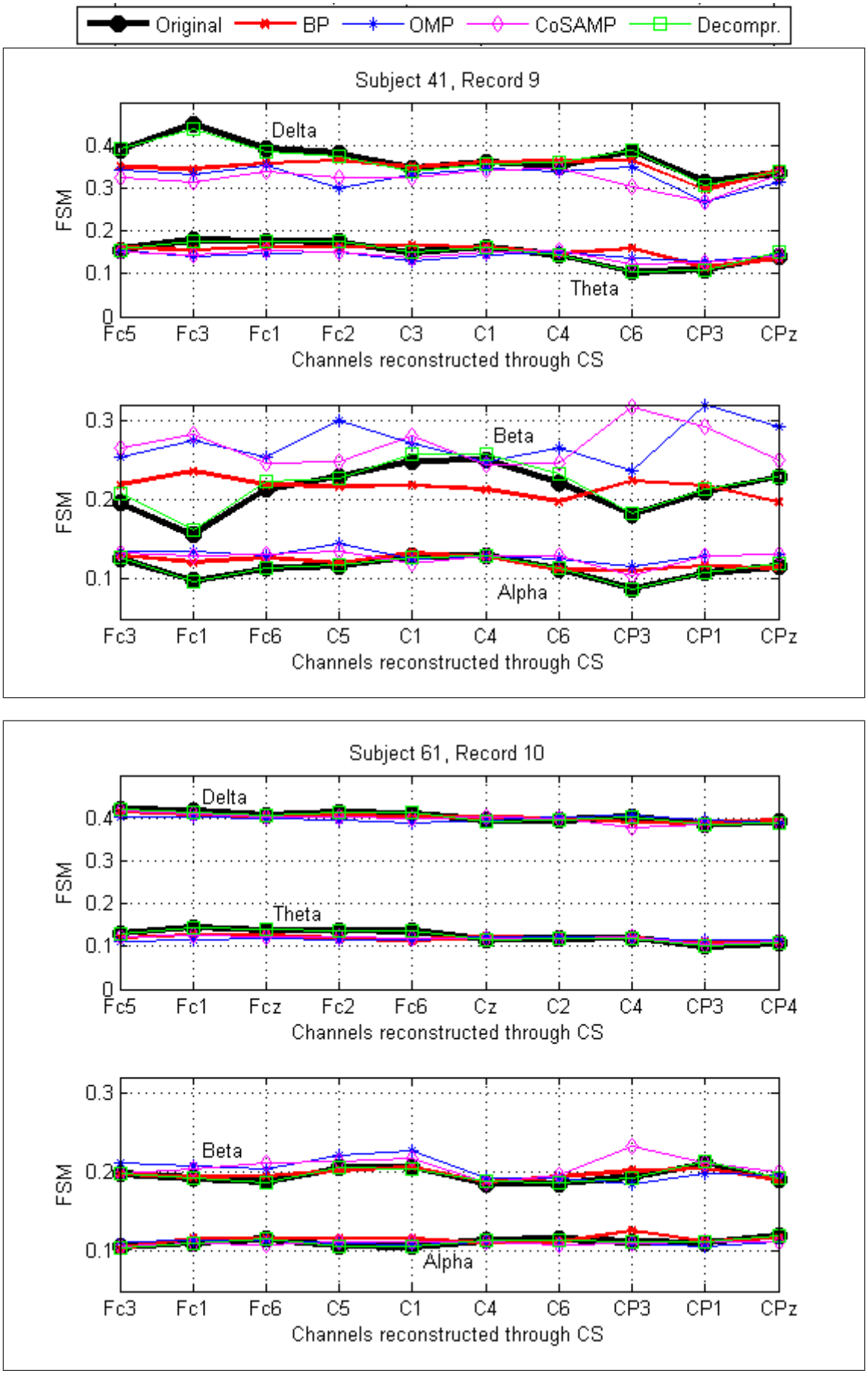


Figure 4.6: Comparison of FSM values of the 10 recovered channels for subjects 41 and 61

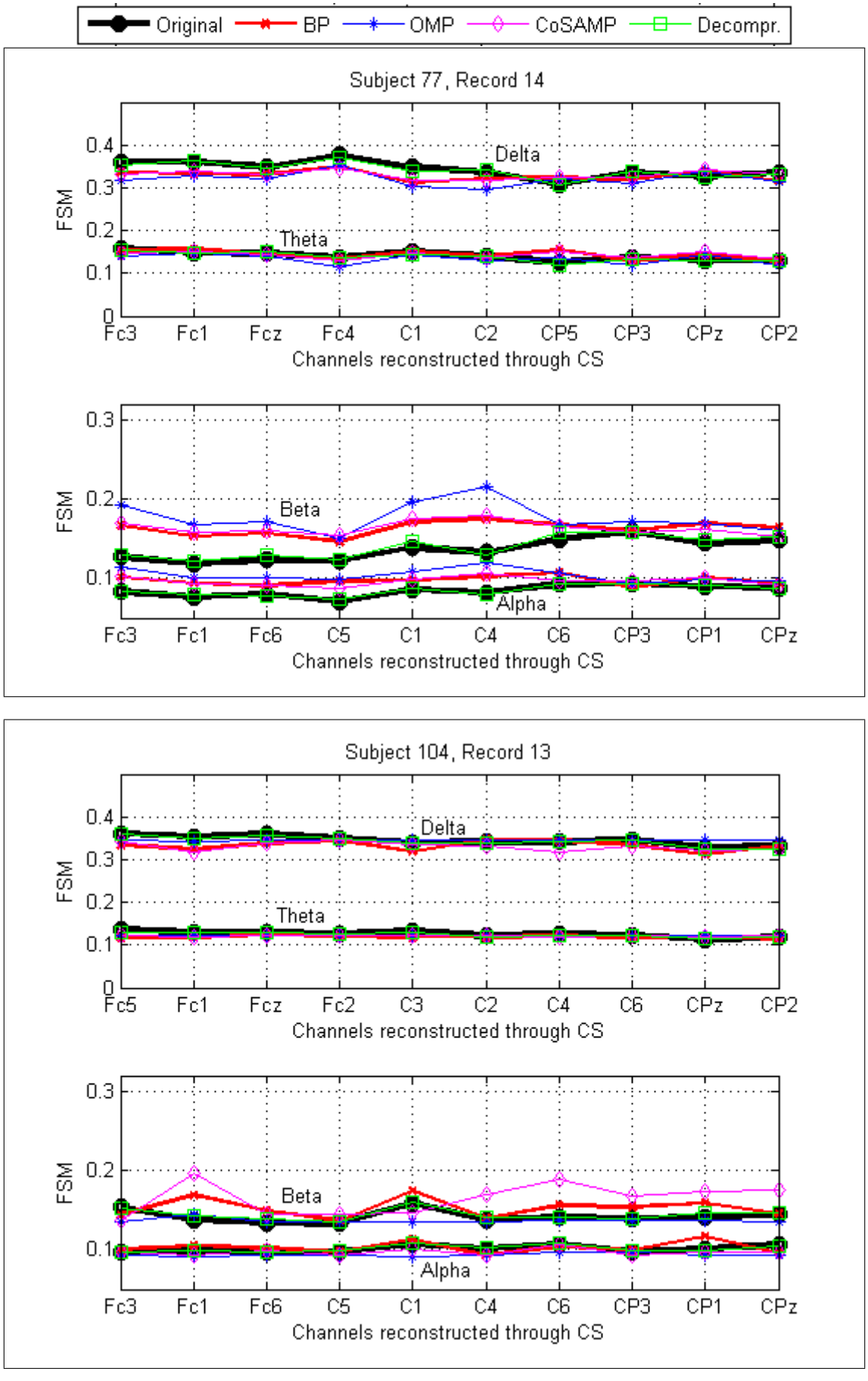


Figure 4.7: Comparison of FSM values of the 10 recovered channels for subjects 77 and 104

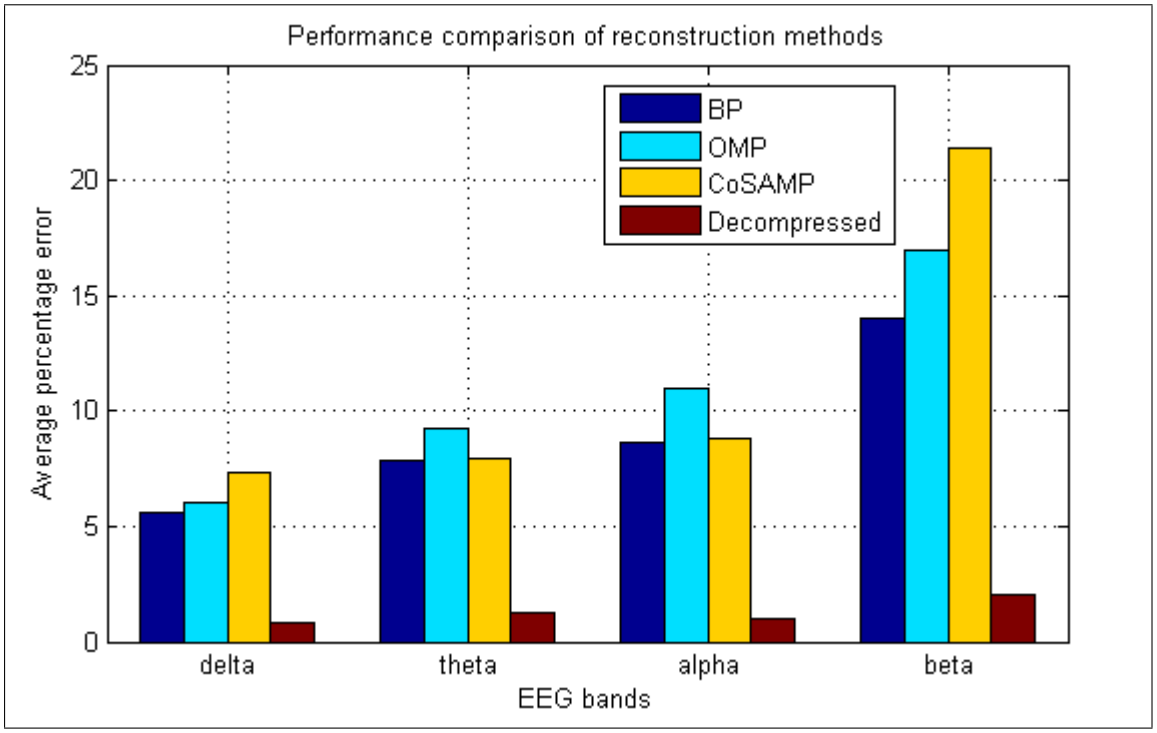


Figure 4.8: Comparison of performance with various reconstruction methods averaged over several subjects in different EEG bands

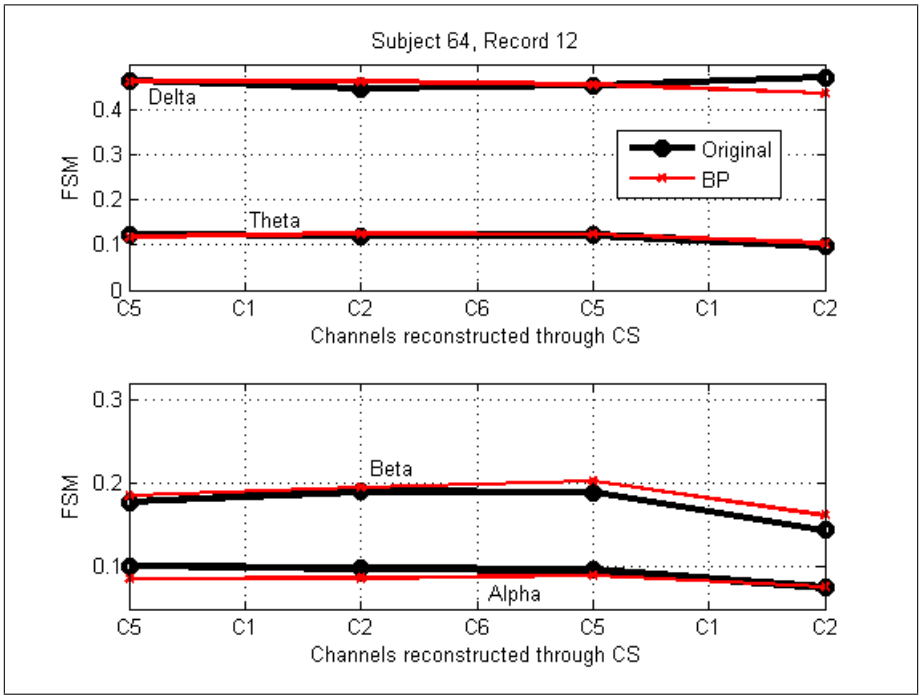


Figure 4.9: Comparison (with the original) of FSM, in different bands, of the 10-10 system EEG channels reconstructed through compressed sensing using only recordings done on the 10-20 system. Recovery algorithm used: Basis Pursuit

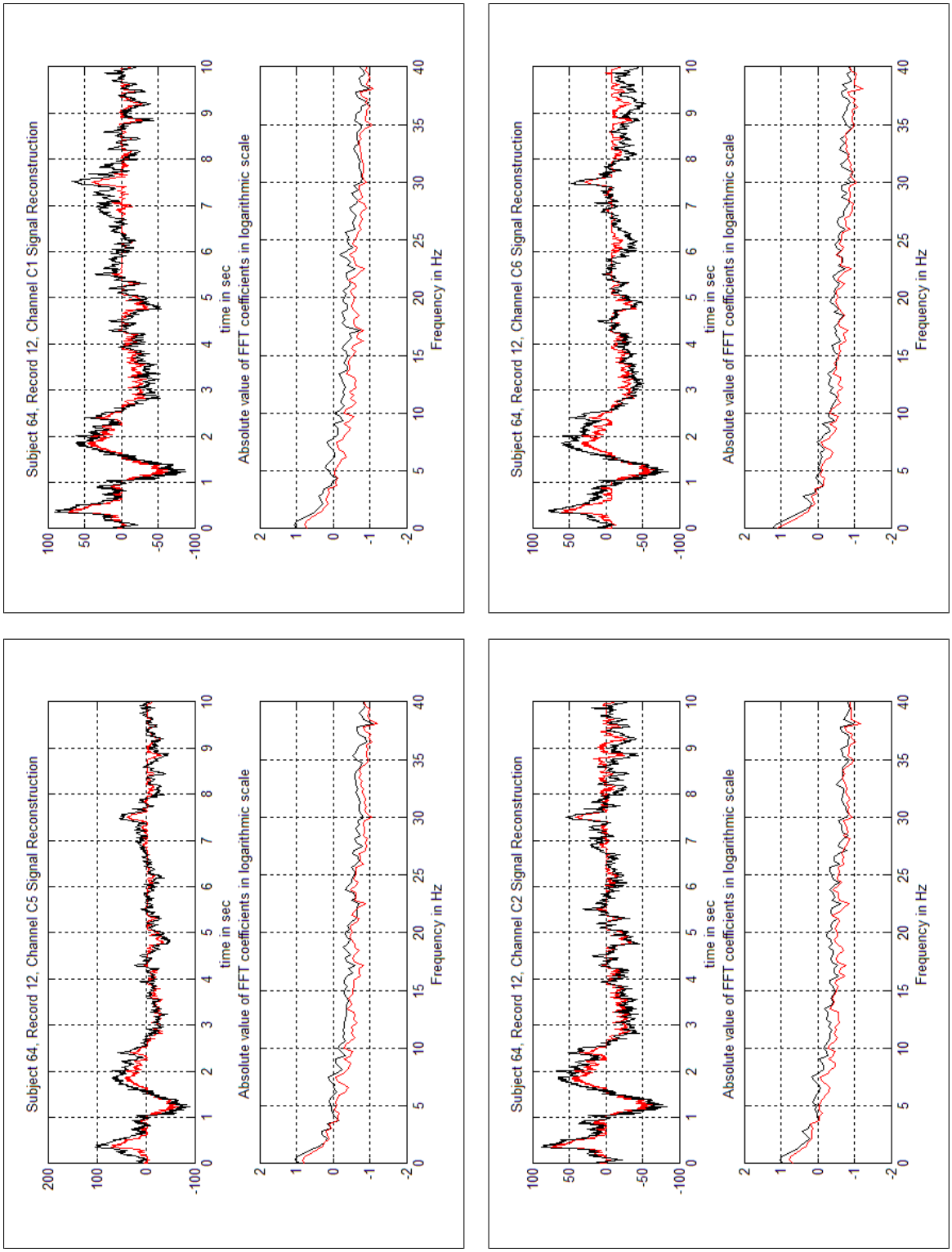


Figure 4.10: Estimation of 10-10 channels from 10-20 recordings - reconstructed (red) vs original (black) signals for subject 64, record 12 using Basis Pursuit recovery algorithm

Chapter 5

Other Compact Embedded Designs for Sparse Signals

5.1 Design with Low-Order Anti-Aliasing Filters

The ideas proposed up to this point concentrate on reduction in the number of ADCs in the design of an embedded system for the data acquisition of a collection of signals that exhibit some form of information redundancy. Decrease in the count of ADCs undoubtedly offers a multitude of benefits, viz. reduced power consumption, size and cost. Situations may arise in which an extra ADC in itself, on the embedded board, will not be so much detrimental to the benefits mentioned above as much as the associated electronics. Thus, it could be possible that even though only $M < N$ ADCs are used for the acquisition of N signals, they may require so much additional circuitry that the advantage offered by reduction in the number of ADCs appears to be marginal. A pertinent question in this regard is whether it is worthwhile to consider usage of as many ADCs as the signals, but minimize the active and passive components associated with each measurement channel ?

Traditionally, the restriction imposed by the Nyquist sampling theorem has been handled by the use of analog, low pass, anti-aliasing (AA) filters at the front-end of data acquisition to ensure that frequency components in the signals, that are above half the Nyquist rate are cut-off. These

filters, built out of passive and active analog components in most embedded designs, lead to significant utilization of space, power dissipation and increased cost. The number of components used is directly related to the filter order, which in turn depends on the sharpness of the transition needed from passband to stop band. Historically, several formulas [115; 116; 117] have been proposed and are being used to calculate the order of the filter as a function of the sampling and the cut-off frequencies. It is very clear that higher the sampling rate, the more relaxed are the restrictions on the filter.

While substantial research has already been done in designing optimal filters for signals with general frequency characteristics, what remains to be explored is if one could further optimize filter design with some additional a priori knowledge of the signal, like for example, the signal having a sparse spectral support. While employing an equal number of ADCs as the number of signals, a design which achieves a sampling rate higher than the specified sampling rate of the ADC for each signal would permit the use of a low order AA filter. In continuation to the ideas proposed in this research, two other designs that exploit signal sparsity in a different way, are proposed in this chapter.

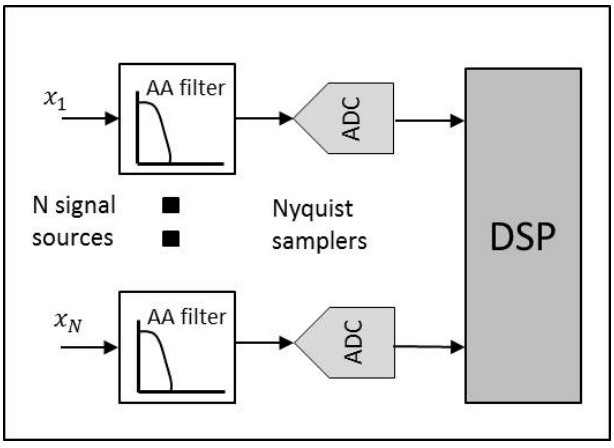


Figure 5.1: AA filters at the front-end of a data acquisition system

5.1.1 The Filtering Problem

Consider a set of signals $\mathbf{x}_i, 1 \leq i \leq S$ with constituent frequencies in the band $[0, F]$. The front-end of a Nyquist data acquisition architecture for the signal ensemble is shown in figure 5.1. The specified sampling rate, F_{ADC} of the analog to digital converters (ADC) for acquiring each signal must satisfy $F_{ADC} \geq F_{NYQ} = 2F$, F_{NYQ} denoting the Nyquist rate. If the anti-aliasing (AA) filters used have a sharp transition from passband to stop-band, then the analog signal is captured reasonably well, since no frequency higher than F will be allowed. In other words, $(F_{STOP} - F_{PASS})/F_{NYQ}$, where F_{PASS} and F_{STOP} are the pass band and stop band cut-off frequencies, must be a small positive value. However, this necessitates the use of a high order filter, a requirement which is detrimental to desirable features like compactness, minimal power consumption and lower cost, typically expected in most embedded designs. Employing the same ADCs, if it were somehow possible to sample the signals at finer sampling intervals, for example at $2F_{NYQ}$, then F_{STOP} can be greater than F, and lower order AA filters could be used. If the signals comprise only a sparse set of frequencies, it would be possible, under a CS architecture to reconstruct the signals using limited number of samples taken on a finer and uniform sampling grid.

5.1.2 Compressed Acquisition and Reconstruction

Let the N signals that are to be sampled and reconstructed be piecewise stationary and sparse of the type introduced in section 3.3. The acquisition architecture that is made use of is MOSAICS that was proposed in section 3.5 with a MUSIC based reconstruction algorithm suggested in sub-section 3.6.2. The notion of reconstruction segment (RS) introduced in section 3.4 is redefined to suit the requirement of the sampling scheme proposed in this section.

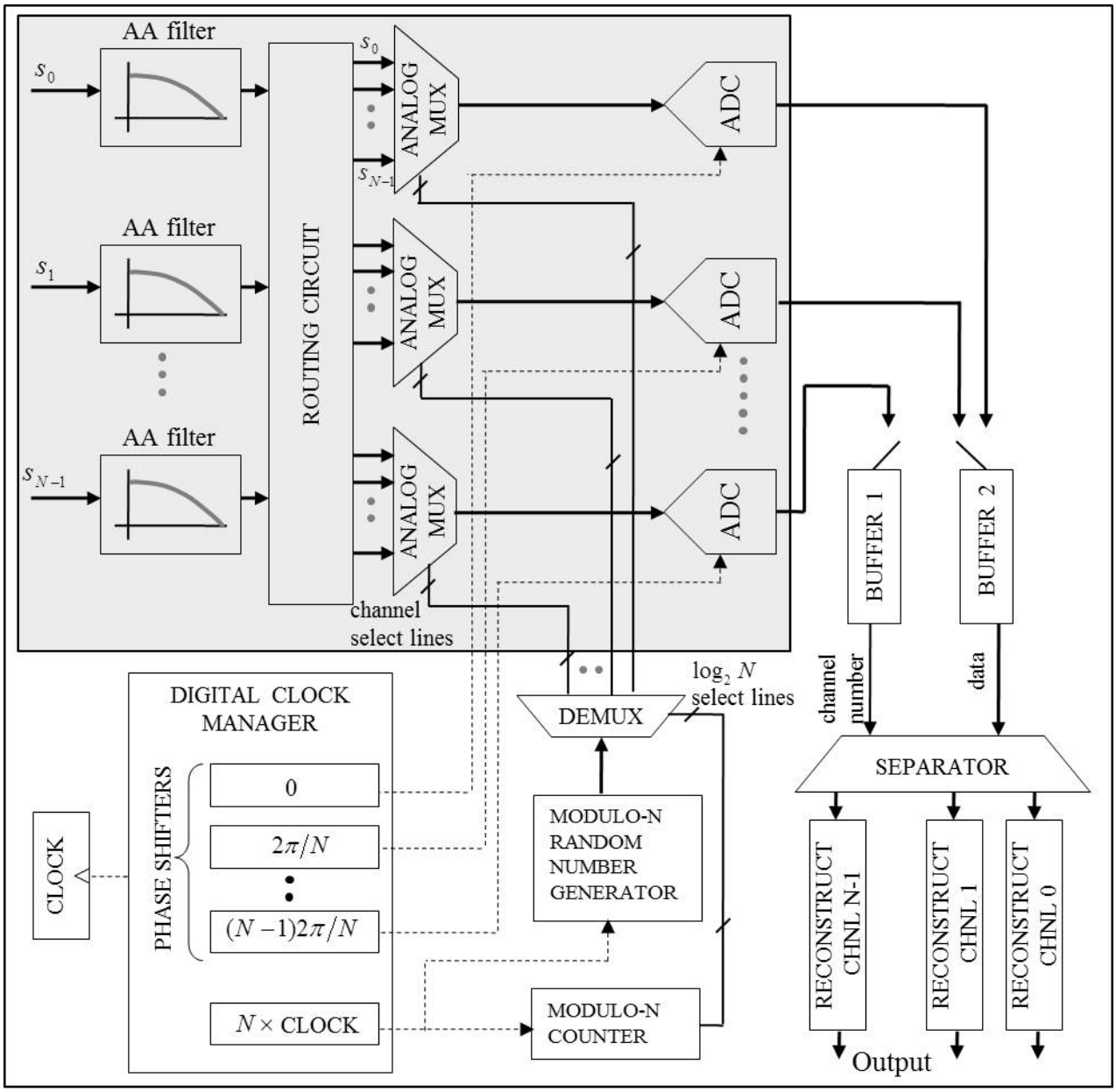


Figure 5.2: Compressed sensing architecture for acquiring signals at a higher sampling rate than the specified sampling rate of the ADC while using low order AA filter

Definition:. *A reconstruction segment (RS) of order γ with respect to the sampling of a signal is defined as the vector, $\zeta^{(\gamma)} \in \mathbb{R}^{N\times 1}$ obtained by uniformly sampling a PSS segment, at γ times the specified sampling rate, F_{ADC} of the ADC, during a finite interval $\tau_1 \leq t \leq \tau_2$ lying within the segment. Clearly, $N = \lfloor (\tau_2 - \tau_1)\, \gamma F_{ADC} \rfloor$. A PSS segment can be acquired*

and reconstructed as a series of reconstruction segments.

5.1.2.1 The CS Tuple

This is same as that given in subsection 3.4.2.

5.1.2.2 Phase Shifted Sampling

A block diagram of the proposed architecture is given in figure 5.2. Given the specified sampling rate of each of the N ADCs as F_{ADC}, let each ADC operate on clocks which have the same period $T = 1/F_{ADC}$, but are phase shifted from each other by ϑ. In other words, if the acquisition starts at time t, the i^{th} ADC, $0 \leq i \leq N-1$, operates at the time instants: $t+i\vartheta, t+i\vartheta+T, t+i\vartheta+2T...$ and so on. If we choose $\vartheta = T/N$, we have a data acquisition system, employing N ADCs, operating on a uniform sampling grid with a sampling interval of T/N or equivalently an effective sampling frequency $F_{EFF} = NF_{ADC}$. Thus all the ADCs taken together offer a finer uniform sampling grid, of order $\gamma = N$, to the N input signals which can permit samples to be taken at intervals N times smaller than that offered by the ADCs if they were not phase-shifted. During acquisition, time instants are randomly chosen from the finer grid provided by all the N ADCs together to collectively sample the N analog signals. This requires each ADC to be able to multiplex between the different analog signals in real time, which is practically realizable due to the presence of built-in multiplexers in commercially available ADCs. The shaded section in figure 5.2, which is the analog section, consists of N ADCs together with the corresponding N multiplexers. N analog signals are input to the system. Each of the signals, after passing through AA filter, is routed to every analog multiplexer.

The rest of the design (the unshaded region) is digital and can be implemented in a small size, low cost FPGA or a processor. The Digital Clock Manager (DCM) is a very standard block commonly implemented

in commercially available FPGAs. The DCM generates N phase shifted versions of the input clock of F_{ADC} Hz, in the range 0 to $(N-1)\,2\pi/N$, which are input to the ADCs. The DCM also generates another clock whose frequency is NF_{ADC}, which is input to a modulo-N counter and a modulo-N random number generator. The modulo-N random number generator outputs a random number between 0 and $N-1$ at every tick of its clock input for choosing the analog signal to be sampled. By using a proper seed, care is taken that over sufficiently long interval of acquisition, each analog signal gets an equal share of the time instants when it is sampled. The modulo-N counter releases counts from 0 to $N-1$ in succession, such that the demultiplexer routes the number of the analog signal to be sampled to the analog multiplexers of successive ADCs, in synchronization with their respective clocks. The analog multiplexer of the ADC, which gets a clock tick, routes the chosen analog signal to the ADC. Thus while each of the individual ADCs operate at their specified sampling rate of F_{ADC}, the collective acquisition takes place at NF_{ADC} Hz.

The process of acquisition and reconstruction of the signals takes place in a series of acquisition cycles. There are two digital buffers which store the samples collected by all the ADCs together. During any acquisition cycle, one of the buffers is active, into which the ADCs deposit the samples collected by them in succession. The other buffer, which contains the samples collected in the previous acquisition cycle, is read by the separator. The separator separates the samples into the individual signals with the help of the random sequence generated by the modulo-N random generator, that is fed to it at the end of an acquisition cycle. For any signal, as soon as time corresponding to an RS of order γ has elapsed, the collected samples are fed to the CS reconstruction block, the output of which is the reconstructed signal. The reconstruction algorithm is the one based on MUSIC (subsection 3.6.2) that can accommodate the so-called 'non-integral' frequencies also. The length of the acquisition cycle, which decides

Table 5.1: FREQUENCY CHARACTERISTICS OF TEST SIGNALS

SIGNAL 1		SIGNAL 2	
Time in ms	Frequencies in KHz	Time in ms	Frequencies in KHz
0–12.5	3.68, 8.14, 13.46	0-14.1	1.23, 3.4, 7.9
12.5–29.3	2.82, 3.95, 11.4	14.1-28.2	2.56, 3.8, 13.5
29.3–42.8	1.5, 4.11, 5.8	28.2-46.5	4.45, 6.53, 10.9
42.8–57.4	2.54 8.59, 14.5	46.5-59.5	3.13, 6.66, 12.54
$\geq$57.4	1.54, 2.71, 7.8	$\geq$59.5	3.62 8.34 12.81

the size of the buffers, must cater for the estimated worst case execution time of reconstruction in that cycle in which all the signals have a complete RS available for reconstruction.

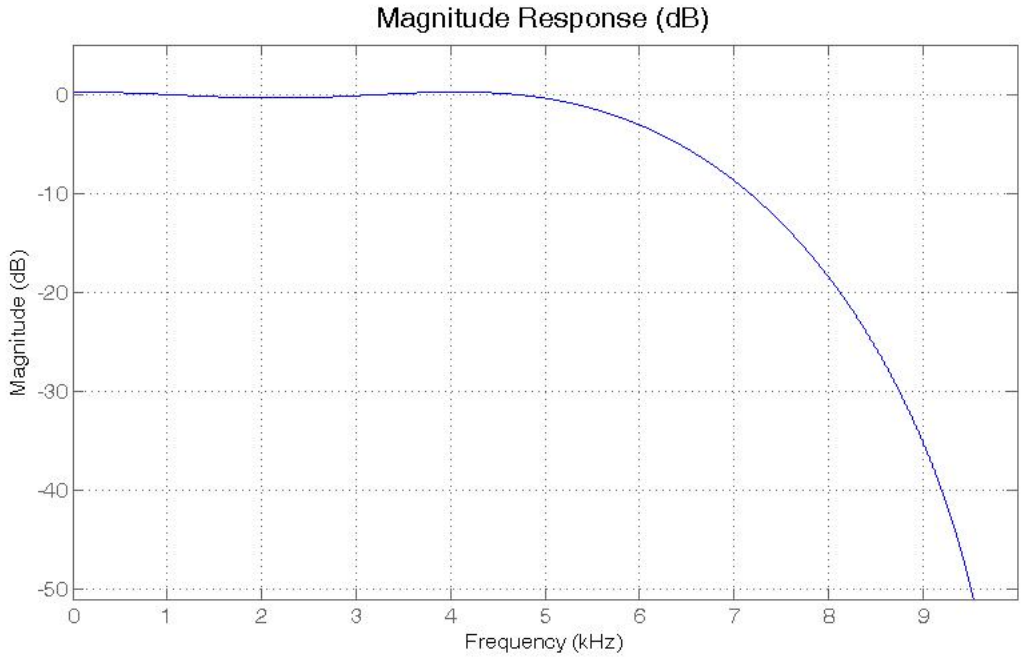


Figure 5.3: Magnitude response of FIR filter of order 8

5.1.3 Simulation and Results

For simulation, the simple test case of acquiring two signals with the frequency characteristics shown in Table 5.1 has been taken. Each signal is a concatenation of PSS segments with durations greater than 10 ms. Within each PSS segment, there are three frequency components. For both signals,

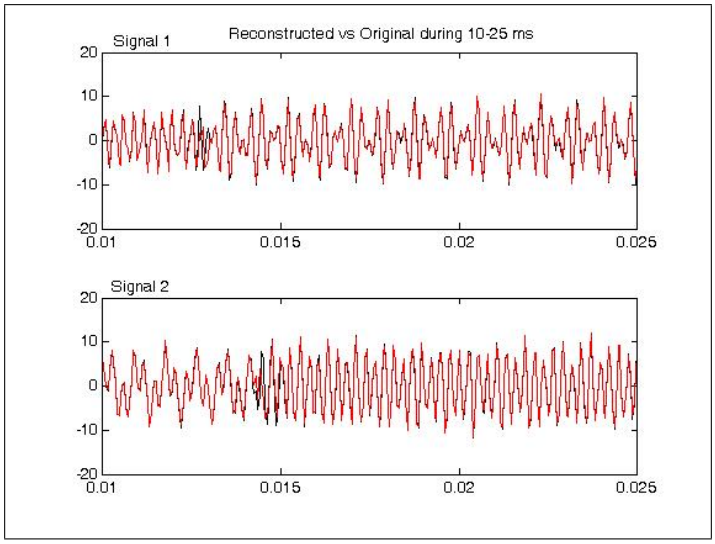


Figure 5.4: Reconstructed signal (red color) vs original signal (black color) for two signals. The deviations are 20.8 dB and 21.5 dB

the region of interest to the application is $0 - 5$ kHz, the content above which can be filtered out. In a classical data acquisition setup, we need to employ an AA filter with a cut-off at around 5 KHz and sample at a rate above the Nyquist rate of 10 KHz. For a sampling rate of 12 KHz and F_{PASS} and F_{STOP} equal to 4.5 KHz and 5.5 KHz respectively, this calls for the usage of a equiripple FIR filter of order 30. As $N = 2$, with the efficient data acquisition scheme proposed in this work, using two ADCs each with specified sampling rate of $F_{ADC} = 10$ KHz [1], we get an effective sampling rate of $F_{EFF} = 20$ KHz. This in turn implies that we can afford to choose $F_{PASS} = 4.99$ KHz and $F_{STOP} = 9.9$ KHz, while preserving the signal content below 5 KHz, without any aliasing effect. The order of the AA filter with the relaxed frequency specifications is only 8. The magnitude response of such a filter is shown in figure 5.3.

The reconstructed signal is plotted against the original signal in figure 5.4. The close match between the reconstructed and the original for both the signals is an empirical evidence of performance. The deviation in the

[1]Note that in the classical setup, the ADC sampling frequency of 12 kHz considered is slightly higher than the Nyquist frequency of 10 kHz to provide some tolerance, while in the case of compressed sensing setup this is not required as the effective sampling frequency of the two ADCs together is 20 kHz which is much higher than the Nyquist frequency

reconstruction for signal 1 at around 12.5 ms and the same for signal 2 at around 14.1 ms can be justified by the existence of PSS boundaries.

5.1.4 Performance of the proposed scheme with increase in the number of signals

The computational load increases only linearly as the number N of input signals increases, since the reconstruction of each signal is independent of the other. On the other hand, with increasing N, the effective sampling rate $F_{EFF} = NF_{ADC}$ of each signal increases due to the availability of an N times finer sampling grid. The increase in F_{EFF} far above the Nyquist rate of each signal, will only be of marginal benefit since after a point the reduction in the order of the AA filter, due to relaxed frequency specifications, will not be significant.

5.1.5 Application to Real World Signals

The above scheme has been applied to the acquisition of test data consisting of four channels of voltages proportional to the deflection of four control surfaces in the actual flight of an aerospace vehicle. The frequencies of the command signals lie in the band $[0\ 12.5\,Hz]$ demanding a minimum of 25 Hz of Nyquist rate acquisition. An anti-aliasing filter with $F_{PASS} = 10$ Hz and $F_{STOP} = 15$ Hz needs to be of order 25. Employing four ADCs each with $F_{ADC} = 25$ Hz, an effective rate of $F_{EFF} = 100$ Hz is achieved. This requires an AA filter of order 6 with $F_{PASS} = 12.5$ Hz and $F_{STOP} = 40$ Hz. Simulation has been done using the CoSAMP recovery algorithm. In general, the reconstruction accuracy achieved was better when the DCT matrix (see figure 5.5) was used as the sparsity matrix than with the DFT matrix (see figure 5.6), even though the trend reverses for the SNR of 10 dB. Table 5.2 gives the PSNR of the reconstructed signals, with DCT and DFT sparsity matrices, at various noise levels.

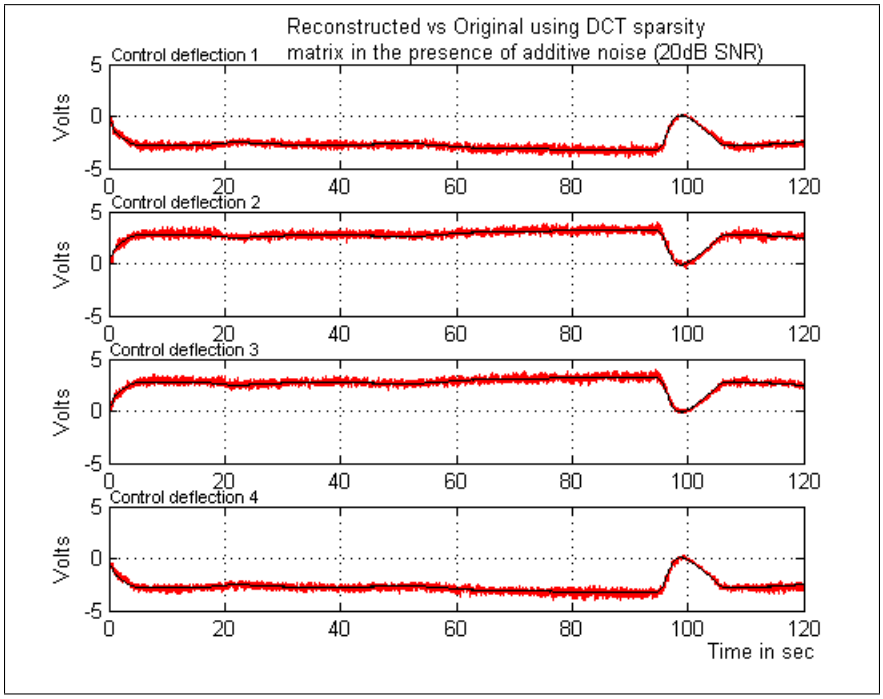


Figure 5.5: Reconstructed signal (red color) vs original signal (black color) for four channels of voltage proportional to deflection of control surfaces. **DCT was used as the sparsity matrix**. Additive noise at 20 dB SNR was added to the original signal.

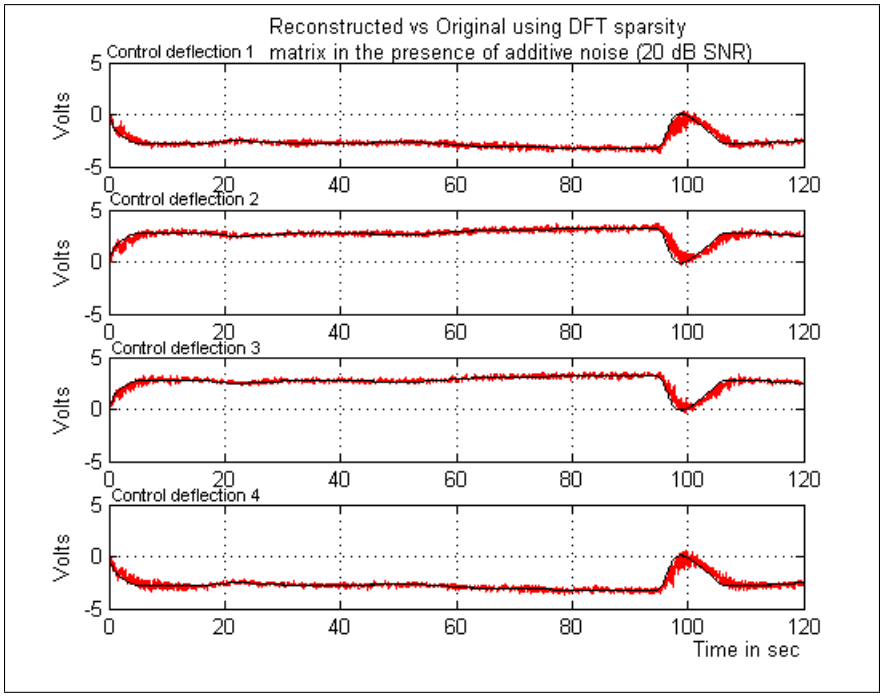


Figure 5.6: Reconstructed signal (red color) vs original signal (black color) for four channels of voltage proportional to deflection of control surfaces. **DFT was used as the sparsity matrix**. Additive noise at 20 dB SNR was added to the original signal.

Table 5.2: PSNR OF THE SIGNALS OF CONTROL SURFACE FEEDBACKS AT DIFFERENT NOISE LEVELS

	Sparsity matrix: DCT				Sparsity matrix: DFT			
SNR in dB	5	10	20	30	5	10	20	30
Deflection 1	10.17	15.46	25.21	32.81	8.38	17.85	23.67	26.34
Deflection 2	10.31	14.73	25.24	32.28	8.0	18.03	22.76	25.98
Deflection 3	10.34	15.09	24.49	32.73	8.7	18.30	23.70	26.72
Deflection 4	9.82	15.35	24.63	31.59	8.6	17.92	22.98	24.02

5.1.6 Conclusion

In this work, an architecture for capturing sparse signals has been proposed in a way that reduces the order of the AA filter at the front end. The AA filter being part of the analog circuitry, this enhancement can have a significant reduction in the number of passive components used for realizing the filter, thereby scoring on compactness, power dissipation, cost, reliability and maintainability. Although the scheme is based on the sparsity assumption, it has enormous potential to be applied in a general situation too, provided the number of frequency components of interest to the application is limited. Simulation has been carried out for a two signal ensemble and results have been reported. However, in a more general setting, one can have multiple signals, for example more than five signals, in which case, the focus is more on achieving a higher effective sampling rate rather than a reduction in the filter order. The blocks in the design have been chosen such that most of them can be realized in a low cost field programmable gate array (FPGA) that is invariably already included in most embedded designs for handling glue logic. The same is true for the multiplexers which are part of most commercially available ADCs.

5.2 Design for Better Resolution

The resolution of an ADC indicates the number of discrete values it can produce over the input range of analog values. For example, an ADC with

a resolution of 8 bits can encode an analog input to one of 256 different levels. The minimum change in voltage required to guarantee a change in the output code is called the least significant bit (LSB) voltage. Given that the maximum voltage swing in the input is V and the resolution is b bits, the LSB is given by

$$LSB = \frac{V}{2^b} \tag{5.1}$$

Thus, higher the resolution of the ADC, smaller is the LSB and better is the fidelity of the signal acquisition. An ADC is characterized by a large number of other parameters, which have considerable impact on the fidelity of the sampled signal; however the focus here is only on the resolution. The conventional approach of improving the resolution of the ADC is to employ a pipelined architecture (figure 5.7) in which there are more than one stage of conversions. Initially a coarse conversion at a lower resolution b is carried out to obtain a coarse digital estimate of the signal. The coarse estimate is converted into analog using a digital to analog converter. The difference between the input and the coarse analog estimate is amplified and the error is converted using another low resolution converter. The bits from the first and the subsequent conversions are added to get a fine estimate of the signal sample that is at a resolution as many times the resolution of a single stage, as the number of stages.

In a compressed sensing setup (figure 5.8), only sub-Nyquist number of samples of a sparse signal would suffice for capturing the signal. Here a single ADC is used in both the stages of a two stage pipeline ADC architecture. For example, if the error signal is available at the input of the single ADC, then using a built-in multiplexer (not shown in the figure), the ADC can switch over to the error input in the immediate next sampling cycle after the one in which the coarse estimate was taken. This is possible because the immediate next cycle could possibly be idle under a compressed sampling setup. Thus a high resolution sampling of the signal could be done by using a general purpose, low resolution ADC with a

multiplexer and a Digital to Analog Converter (DAC) of equal resolution.

The only aspect that needs to be taken care of is that coarse samples be not taken at successive sampling instants since the immediate cycle after a coarse sampling is done must be available for measuring the error signal. Although a true random choice of sampling instants (which is required for compressed sensing) is one that does not impose any such restriction and is ideally suited for error-free reconstruction, the restriction mentioned above requires that the next sample must be taken at any instant other than the immediate next one. This deviation from true randomness does not significantly affect the quality of reconstruction as seen through simulation.

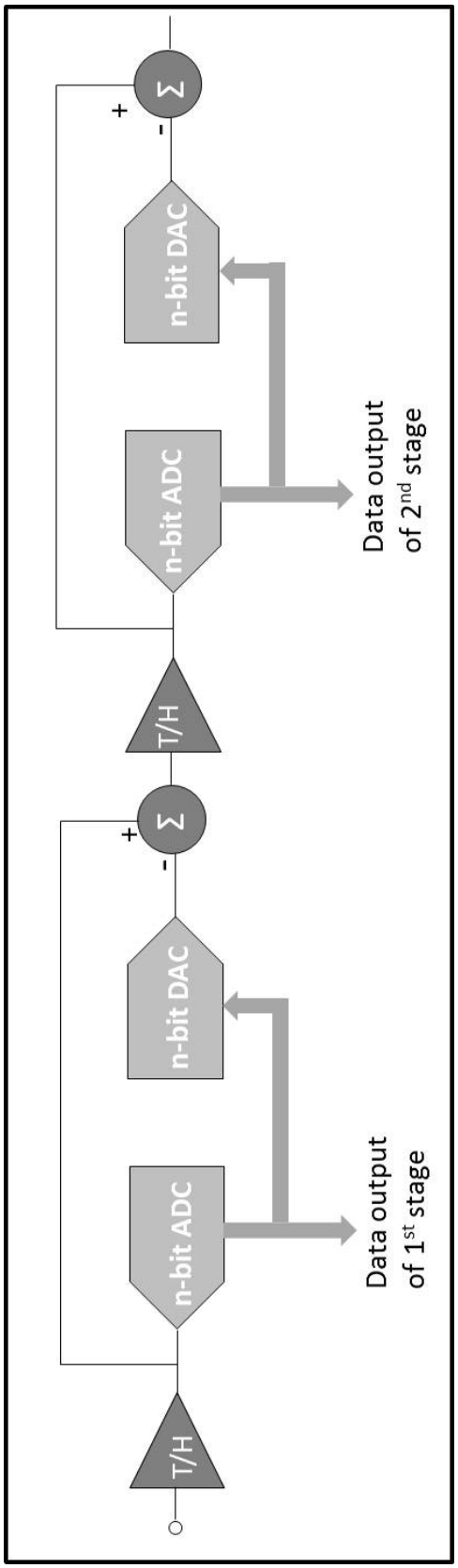


Figure 5.7: Pipeline ADC architecture

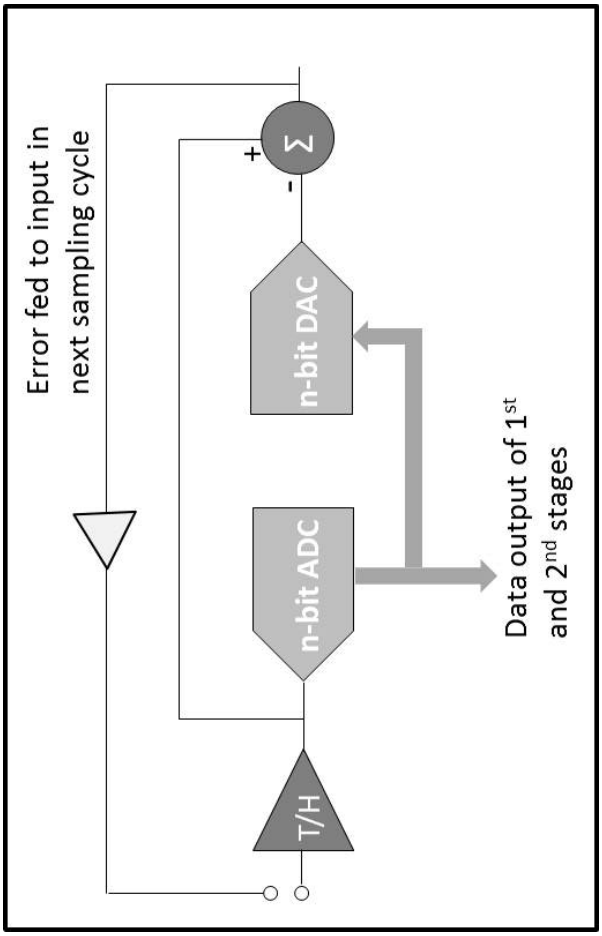


Figure 5.8: Modified pipeline ADC architecture with a single ADC and DAC

5.2.1 Simulation and Results

To simulate the idea, simplistic models of ADC and DAC have been used. The n-bit analog to digital conversion is modeled as a straightforward process of converting the input analog value (a real number between -10 and +10) to one of the 2^b discrete levels (integers). The digital to analog conversion does the reverse - the analog voltage corresponding to a discrete level is calculated.

A test signal $\mathbf{x}$ is given as input to n-bit ADC model sampling uniformly at a rate above Nyquist. The analog signal, $\mathbf{x}_{nyq}^{(n)}$ corresponding to the discrete levels at the output of the ADC are calculated and saved. The same test signal is then applied to a second model in which the following sequence of operations are carried out:

i) An n-bit ADC performs coarse compressive sampling, i.e. sampling is done in sub-Nyquist number of cycles chosen randomly without violating the condition mentioned in the previous section.

ii) The discrete output is applied to an n-bit DAC model.

iii) The analog output of the DAC is subtracted from the test input signal to get the error signal

iv) The error signal is multiplied by a gain and fed as input to the same ADC in the immediate next sampling cycle that is relieved from measuring the actual signal.

The analog samples corresponding to the discrete levels at the output of the ADC are calculated. These samples (sub-Nyquist in number) are then used to reconstruct the full signal, $\mathbf{x}_{cs}^{(2n)}$ with the help of l_1 minimization using the *cvx* toolbox. The Signal to Quantization Noise Ratio (SQNR) is

calculated as follows:

$$SQNR_{cs}^{(2n)} = 20\log_{10}\frac{rms(\mathbf{x})}{rms(\mathbf{x} - \mathbf{x}_{cs}^{(2n)})} \tag{5.2}$$

$$SQNR_{nyq}^{(n)} = 20\log_{10}\frac{rms(\mathbf{x})}{rms(\mathbf{x} - \mathbf{x}_{nyq}^{(n)})} \tag{5.3}$$

Table 5.3: SQNR VALUES (dB) AFTER n-BIT NYQUIST SAMPLING AND 2n-BIT COMPRESSIVE SAMPLING

	n-bit Nyquist	2n-bit CS
n=4	14.4	38.5
n=6	26.4	59.0
n=8	38.4	85.8
n=9	44.0	96.9
n=10	50.5	114.6

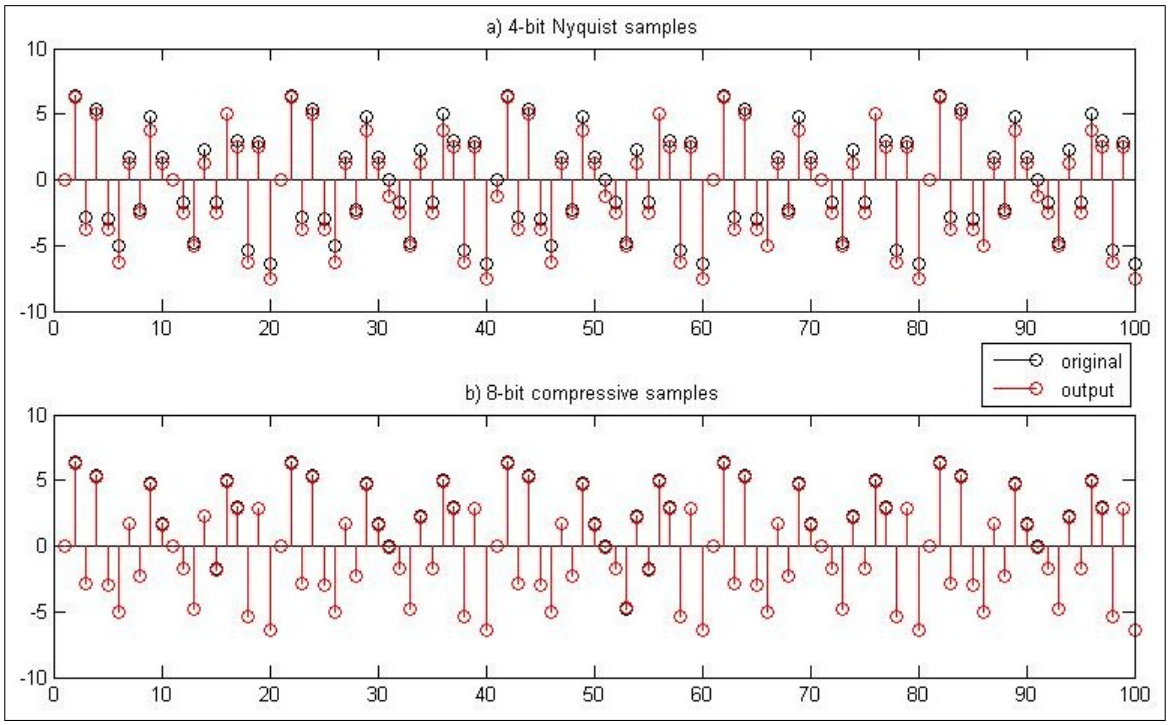


Figure 5.9: a)100 Samples from a Nyquist sampled 4-bit ADC and b) 100 Samples from a 8-bit compressively sampled ADC. In both case the reconstructed signal is shown in red and the original signal in black

A sparse signal having only three frequency tones: 3 KHz, 7 KHz and 8 KHz has been used as the test signal. The SQNR values for different values

of n for the Nyquist sampled n-bit ADC and the compressively sampled 2n-bit ADC are given in Table 5.3. In figure 5.9 are shown how the output of a 4-bit ADC and an 8-bit compressively sampled and reconstructed signal compare with the original signal. In the case of 8-bit compressively sampled ADC, there is a close match, while for the 4-bit ADC, there is a marked deviation.

5.2.2 Application to Fetal ECG Acquisition

The performance of the scheme proposed above has been additionally verified by simulation with fetal ECG captured directly from the fetal head and from the maternal abdomen taken from the Physionet database [111; 118]. The signal characteristics are given as follows:

- Bandwidth: 1 Hz - 150 Hz
- Sample resolution: 16 bit
- Sampling frequency: 1 KHz
- Voltage range: -3.2768 mV to +3.2768 mV
- Smallest voltage: 0.1 μV
- Dynamic Range: 96.33 dB

Figure 5.10 gives the plot of 1000 samples of ECG captured using an 8-bit ADC operating at Nyquist rate (top figure) and again the same 8-bit ADC operating in compressed sensing scheme at 16-bit resolution (bottom figure). Clearly, in the 8-bit acquisition, there is loss of information and the SQNR of the reconstructed signal is 7.0 dB, while in the 16-bit case it is 16.3 dB. Figure 5.11 gives the results with ECG captured on the maternal abdomen.

In this simulation experiment, an interesting observation deserves to be reported. At lower dynamic ranges (e.g. voltage range:-0.5 mV to 0.5

mV and smallest voltage: 0.1 μV), the 8-bit Nyquist acquisition performs better than 16-bit CS reconstruction because 8 bits are sufficient to fully capture the signals over the entire dynamic range, while in the case of the CS acquisition, the inherent reconstruction inaccuracy creeps in. As the dynamic range is increased up to a voltage range of -4.0 mV to +4.0 mV, there is loss of information in 8-bit Nyquist acquisition and 16-bit CS reconstruction using the same single 8-bit ADC performs better. Figure 5.12 (see footnote [1]) compares the 8-bit Nyquist acquisition with 16-bit CS acquisition at different dynamic ranges, with the voltage range varying from 1 mV to 8 mV, increased by 0.2 mV at each step.

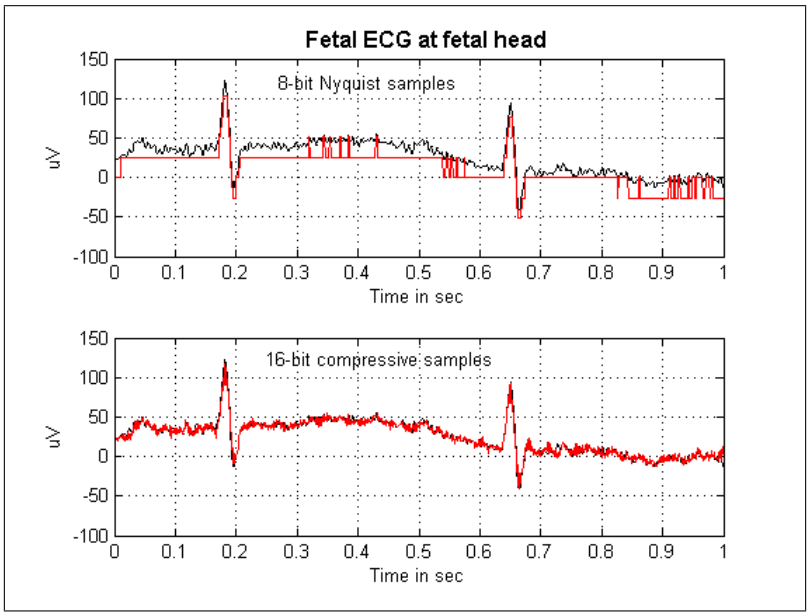


Figure 5.10: a)1000 samples of fetal ECG sensed at the fetal head using a Nyquist sampled 8-bit ADC. b) 1000 samples of fetal ECG reconstructed from a 16-bit compressively sampled ADC. In both cases the reconstructed signal is shown in red and the original signal in black. Voltage range is -3276.8 μV to +3276.8 μV.

It is pertinent to mention here, that ECG has often been regarded a signal sparse on the wavelet basis [119; 120] for compressed sensing. In the simulations, of which the results are presented here, only Fourier spar-

[1]ECG signals, in general, have a high dynamic range (voltage range:-4.0 mV to 4.0 mV and smallest voltage: 0.1 μV). The ECG records that have been used in these experiments also have a high dynamic range, as mentioned above in the signal characteristics. It is only to show the comparison of the 8-bit and the 16-bit cases that the dynamic ranges of the data in the records have been lowered.

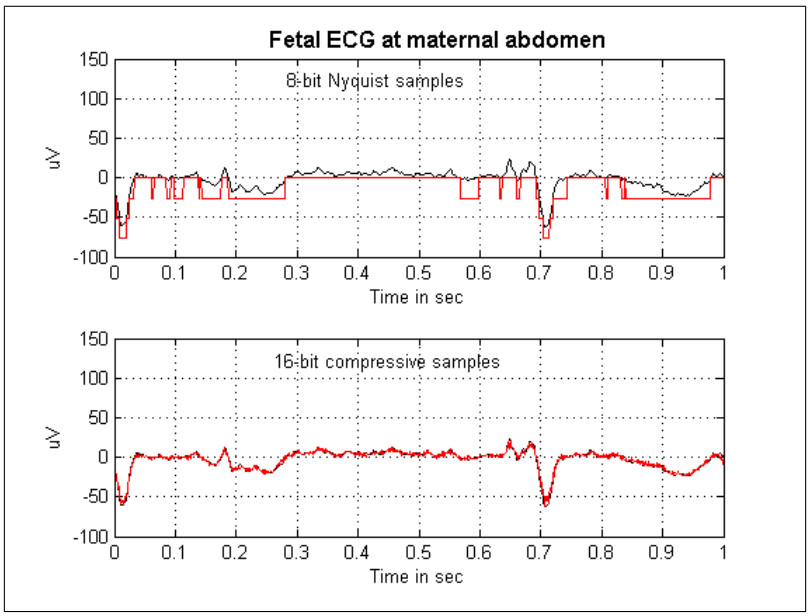


Figure 5.11: a)1000 samples of fetal ECG sensed at the maternal abdomen using a Nyquist sampled 8-bit ADC b) 1000 samples of fetal ECG reconstructed from a 16-bit compressively sampled ADC. In both cases the reconstructed signal is shown in red and the original signal in black. Voltage range is -3276.8 μV to +3276.8 μV.

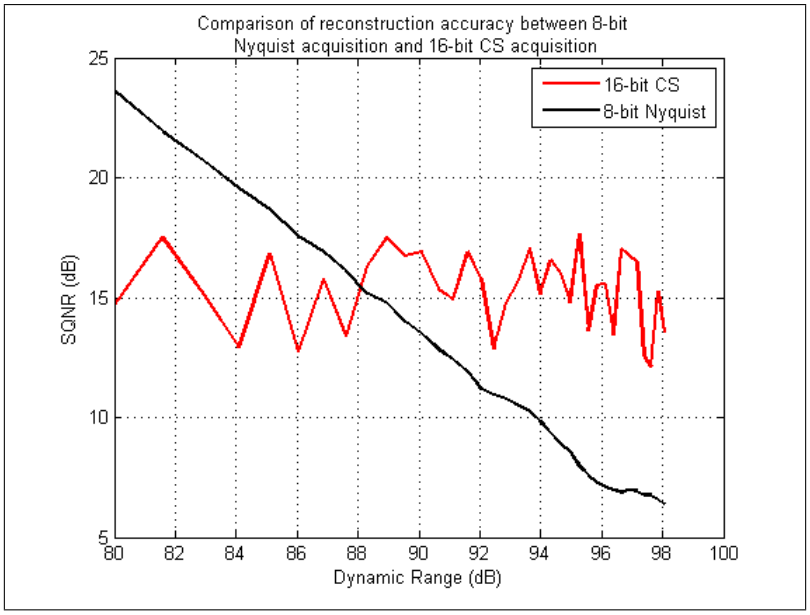


Figure 5.12: Comparison of SQNR between 8-bit Nyquist and 16-bit CS acquisition at different input dynamic ranges for fetal ECG captured directly at the fetal head

sity has been considered. The main objective here is to test the scheme that has been proposed in the previous section on a real signal rather than synthetic data. Undoubtedly, on a wavelet basis the results could be better. It is also appropriate to cite other contributions in the area of application of compressed sensing to ECG signals. Effects of quantization of the coefficients in a random measurement matrix, for compression before transmission, on the reconstructed signal in the presence of sparsity have been studied for ECG and EMG in [121]. The authors in [122] propose to use the block sparse Bayesian learning framework to compress/reconstruct nonsparse raw FECG recordings. The focus in these works is towards signal compression as against usage of a low resolution ADC that has been presented above.

5.2.3 Conclusion

To conclude, using a general purpose, low resolution ADC, a DAC of the same resolution and a summer, one can acquire a sparse signal with double the resolution of the ADC, without having to use a dedicated pipeline ADC. It is true that pipeline ADCs are currently inexpensive and can sample any general signal (sparse or non-sparse) in view of which usage of the compressed sensing based architecture may not be fully justified. However, the design proposed here could be integrated into a complete embedded system for data acquisition of many sparse signals in which one odd signal, that requires data conversion at a high resolution, could be sampled using this scheme.

Chapter 6

Conclusion of Thesis

This research has been a modest endeavor to deliver compact, power-efficient, reliable and less expensive designs for embedded data acquisition systems by leveraging upon the inherent information redundancy in the input signals. Redundancy most commonly manifests as signal sparsity on some basis. After a survey of a plethora of sparse sampling strategies, compressed sensing was identified as a promising technique suitable for being employed in embedded applications. The sensing methods and the reconstruction algorithms in compressed sensing have been studied. The l_1- minimization and the greedy orthogonal matching pursuit algorithms were found to be the ones best suited for the kind of signals encountered in embedded data acquisition. Information available in literature about issues like stability of reconstruction, choice of the appropriate measurement matrix and robust compressive sensing has been presented. The notion of CS-tuple, with which any compressed acquisition scheme can be identified has been introduced.

Next, it was suggested that if compressed sensing of a signal involves use of only sub-Nyquist number of ADC sampling cycles, why not use the idle cycles of the ADC to sample other similar signals ? This gave rise to the idea of multiplexed sensing of more than one signal. With the help of the MOSAICS architecture it was shown how this is possible. Description was given of various input and derived parameters that are involved in the

operation of MOSAICS as well as the factors affecting the performance of the system. It was shown how streaming data from pseudo stationary and sparse signal sources could be captured and reconstructed as a series of overlapping reconstruction segments. Evidence of the fact that MOSAICS is independent of the CS recovery algorithm used was provided by showing that the scheme operates equally well with different algorithms: BP, OMP, ROMP and CoSAMP. With the help of MOSAICS a comparison of the performance of these algorithms with respect to accuracy and consistency of signal recovery has been done. In addition the results have also been compared with classical compression and decompression. A scheme was proposed for implementing the MOSAICS architecture using easily available off-the-shelf computer hardware. The limitation of MOSAICS in dealing with the so-called 'non-integral' frequencies was brought out and a remedy using an algorithm based on MUSICS for signal reconstruction, in place of the erstwhile l_1- minimization, was suggested. It was then illustrated how the same recipe could be used to detect sinusoid frequencies buried in heavy noise even with high undersampling and how this could be applied to the detection of a time-varying carrier frequency in a noisy FM signal. In all the sensing mechanisms proposed, the measurement matrix used was a trivial one - samples were just randomly picked up from the source signal, something that can be practically realized without difficulty. This is unlike classical compressed sensing where the measured vector is taken as a projection of the signal vector on to a measurement matrix.

It was then pointed out that it is not necessary that sparsity be possessed by the individual signals. Inter-signal correlation between non-sparse signals also contributes to information redundancy opening up the possibility of acquiring the signal ensemble using ADCs, lesser in number than the number of signals. ARCS, a method for sensing a streaming set of signals was proposed in which there is no a priori knowledge of the inter-signal correlation and the correlation structure is incrementally learned as

stranger signals become ***familiar***, along a fast learning curve. If there is a huge database already existing of the correlated signals that have to be acquired, then the associated inverse-KLT matrix can be pre-calculated from a training set. The correlation information thus learnt can be used to sense and reconstruct a collection of similar signals using lesser number of ADCs. Making use of inter-channel correlation was shown to be effective in the case of EEG signals taken from the *Physionet* database, wherein the relative power in various spectral bands could be estimated from measurements done only on a subset of EEG channels. Here again the proof of performance of the proposed scheme was verified with three CS recovery algorithms. The reconstruction accuracies of each algorithm has been compared with the simulation results of classical lossy, compression-decompression.

Reduction in the number of ADCs is one of the ways in which an embedded design could be made compact. Associated with any ADC, there is also a lot of other electronics which occupies considerable real estate on the embedded hardware. Anti-aliasing filter is an example of such circuitry, which requires many passive or even active components for realizing the filter if it is expected to have a sharp transition from passband to stopband. It was shown that one can achieve an effective sampling rate that is equal to the individual sampling rate of each ADC multiplied by the number of ADCs, provided the ADCs sensing a set of signals are driven by phase-shifted clocks. Application of this idea to the compressed acquisition of voltages (at low to high SNRs) proportional to in-flight control surface deflections of an aerospace vehicle and subsequent recovery using DCT and DFT basis has been demonstrated. Instead of using a high-end ADC to achieve higher conversion resolution for each sample, it was demonstrated that it is possible to do the same using a general purpose, low resolution ADC, DAC and a summer, if the signal is compressively sampled. It has been shown that this design can be employed for increasing the dynamic

Table 6.1: CHARACTERISTICS OF SIGNALS TO BE ACQUIRED UNDER A COMPRESSED SENSING SETUP

Signal number	**Characteristics**	**ADC used**
1,4,5,7,11,12,21	frequency: 0–35 KHz, minimum length of PSS segment: 3 ms	ADC 1 (80 KHz) 8 channel multiplex
14,15,20	frequency: 0–35 KHz, minimum length of PSS segment: 10 ms	ADC 2* (80 KHz) 4 channel multiplex
8,9,16	frequency: 0–15 KHz, minimum length of PSS segment: 5 ms	ADC 3 (40 KHz) 4 channel multiplex
2,6,19	frequency: 0–15 KHz, minimum length of PSS segment: 1 ms	ADC 4 (40 KHz) 4 channel multiplex
3,10,13,22	frequency: 0–3 KHz, correlated signals	ADC 3 and ADC 4
17,18	frequency: 0–18 KHz, sharp cut-off from passband to stopband	ADC 1, ADC 2

*The input clock of ADC 2 is phase shifted with that of ADC 1

range in the acquisition of fetal ECG signals. Variation of the reconstruction performance as the voltage range of the ADC varies has been shown.

An informal justification of the advantages of using the downsized identity matrix as the measurement matrix in place of Gaussian or Bernoulli matrices has been given, along with citations from the literature, in Appendix C. With the help of simulations, it has been shown that the performance of the down-sized identity matrix is comparable with Gaussian or Bernoulli matrices.

6.1 Integration of proposed methods

Although the operation of the schemes suggested in this thesis have been demonstrated in separate setups for signals with specific properties, it is possible to encounter an application, which has a mix of signals of various types. To tackle such a problem, one should be able to evolve a design in which the different techniques proposed could be synergistically merged using common hardware components. An illustration of how this could be done for a fictitious application follows.

Consider a set of 22 PSS signals having different characteristics as shown in Table 6.1. Given that there exists information redundancy in this collection of signals either due to inherent sparsity or inter-signal correlation, it is desired to construct a design employing the minimum number of ADCs with the associated circuitry. Applying the CS based designs proposed in the previous chapters, it is possible to arrive at an architecture comprising just 4 ADCs. The last column of the table gives the names of the ADCs to which the signals are assigned. As suggested in chapter 3, signals with same bandwidth are placed into a single MOSAICS block. Thus there are two MOSAICS blocks one with ADC 1 and the other with ADC 2 to cater for the signals of bandwidth 0–35 KHz. Two separate ADCs are used for these signals because of the different minimum lengths of PSS segment. The same is the case for the signals: 8, 9, 16, 2, 6, 19, with bandwidth 0–15 KHz, which are divided into two separate MOSAICS blocks operating with ADC 3 and ADC 4, respectively.

There are four signals, 3, 10, 13 and 22 that are correlated. ADC 3 and ADC 4 are assigned to two out of these four signals. By measuring two out of the four signals, the other two can be found out if their correlation is known a priori or is found during real time operation as explained in Chapter 4. Although ADC 3 and ADC 4 are already part of MOSAICS blocks, one in every four sampling cycles in each of these ADCs operating at 40 KHz is reserved for two of the four correlated signals. This provides a sampling rate of 10 KHz for each of the four correlated signals sufficient to meet the requirement of their bandwidth of 0–3 KHz.

Finally there are two signals 17 and 18 of bandwidth 0–18 KHz, that need high order anti-aliasing filters due to a sharp transition between passband and stopband. One in every four sampling cycles of ADC 1 and ADC 2 is reserved for each of these signals. Thus, individually, ADC 1 and ADC 2 offer 20 KHz to these signals. However, since the clock input to ADC 2 is made to have a phase shift with respect to that of ADC 1, as explained

in chapter 5, an effective sampling rate of 40 KHz is available for each of the signals. This helps in relaxing the specifications of their respective AA filters.

Various hardware resources like computational units, memory and DMA blocks can be commonly shared by the reconstruction engines responsible for recovering the different groups of signals.

6.2 Summary of findings in the thesis

The salient ideas and findings that have emerged out of this research are listed below:

i) Signals sampled by embedded data acquisition systems have information redundancy in the form of sparsity which can be exploited to do sub-sampling.

ii) Such systems can be made more compact, less power-hungry, more reliable and cost effective in a sub-sampling scheme.

iii) ADC sampling capacity can be more efficiently utilized in a compressed acquisition by multiplexing between various signals – MOSAICS.

iv) This can be used for any general signal with arbitrary frequencies as also for detection of sinusoids buried in noise.

v) A set of correlated signals can be collectively sampled using lower number of ADCs. This idea can be used for EEG acquisition.

vi) Each of a set of sparse signals can be sampled at a higher effective sampling rate when driven by phase-shifted clocks so as to push the stop band cut-off to higher frequencies.

vii) A high resolution ADC, for sparse signals, can be designed from commonly used components.

viii) It is possible to design an acquisition system to capture an ensemble of signals that has different kinds of sparsity by combining several of the proposed compressed acquisition schemes while using shared hardware.

ix) Most of the schemes are independent of the reconstruction algorithm used. The results of simulation using different recovery algorithms are reasonably comparable.

x) The acquisition apparatus in all the cases is quite simple and straightforward and can operate on continuous streaming signals.

This thesis concludes at this point with the assertion that compressed sensing methods can be effectively used to derive maximum benefit from sub-sampling, if many signals are acquired simultaneously. Also, with mild assumptions related to stationarity, it is possible to capture in real-time, continuous and streaming signals. Compressed embedded data acquisition systems can be realized using simple, general-purpose, off-the shelf electronic components.

Appendix A

Embedded hardware design

Factor	Impact	Design techniques
Power consumption	Design of power supply, voltage regulators, dimensioning of interconnect and cooling	• Use of ASICs instead of general purpose processors • Parallelism • Dynamic Voltage Scaling • Dynamic power management • Reduction in number of components
Compactness	Reduction in weight and dimensions as part of the total system	• Miniaturization of capacitors on the PCB • Component embedded PCBs - ICs, resistors, capacitors, and inductors embedded into PCBs • Reduction in number of components
Heat dissipation and uniform temperature gradient	Cooling requirements	• Effective heat and high current routes for optimal convective heat transfer • Use of thermally conductive planes • Reduction in number of components
Reliability	Human safety and reduced MTTF (Mean Time To Fail)	• Fault-tolerant architectures • High component level reliability • Reduction in number of components leading to lesser interconnect

Table A.1: DESIGN CONSIDERATIONS FOR EFFICIENT DESIGN OF EMBEDDED SYSTEMS

Appendix B

Algorithm ARCS

Input

N	:	number of correlated signal sources
M	:	number of available ADCs
T_s	:	Sampling period

Streaming data from the N signal Sources

Output

$\mathbf{S}$	:	Signal matrix

Variables:

fam	:	set of familiar signals
str	:	set of stranger signals
entr	:	set of entrant signals
$\mathbf{f}$	:	column vector with M or $M-1$ elements
$\mathbf{\Psi}^{(n)}$	:	Inverse KLT matrix of dimension n
τ	:	threshold for KLT convergence
$\mathbf{\Phi}$	:	Measurement matrix
k	:	count of the sampling cycle during a convergence cycle. This count gets reset to 1 every time the inverse KLT converges
time	:	current acquisition time
acquisition_time	:	total time of acquisition
$\mathbf{e}$	:	vector of eigen values of the covariance matrix of $\mathbf{S}$
δ	:	l_2 norm between successive vectors of eigen values
n	:	number of signals introduced to ARCS

Method:

1. $n \leftarrow M,\ k \leftarrow 1,\ entr \leftarrow \{1, 2..M\},\ fam \leftarrow \{\}$
2. **while** $time \leq acquisition\_time$ **do**
3. **if** $n = M$ **then**
4. $\mathbf{S}(k, 1:n) \leftarrow M$ samples from each of the signals in *entr*
5. CALL INVKLTCONVERGE
6. **else**
7. **if** $n < N$ **then**
8. $\mathbf{f} \leftarrow$ samples, at the k^{th} instant, of $M-1$ randomly chosen signals from the set *fam*
9. $\mathbf{\Phi} \leftarrow$ Matrix consisting of those rows of the identity matrix of order n [1] corresponding to the chosen signal numbers
10. $\hat{\mathbf{z}} \leftarrow CONVEXOPT(\mathbf{\Phi}, \mathbf{\Psi}^{(n-1)}, \mathbf{f}, n-1)$
11. $\mathbf{S}(k, 1:n-1) \leftarrow \hat{\mathbf{z}}$
12. $\mathbf{S}(k, n) \leftarrow$ sample of the signal in *entr*
13. CALL INVKLTCONVERGE
14. **else**
15. $\mathbf{f} \leftarrow$ samples, at the k^{th} instant, of M randomly chosen signals from the set *fam*
16. $\mathbf{\Phi} \leftarrow$ Matrix consisting of those rows of the identity matrix of order n corresponding to the chosen signal numbers
17. $\hat{\mathbf{z}} \leftarrow CONVEXOPT(\mathbf{\Phi}, \mathbf{\Psi}^{(n)}, \mathbf{f}, n)$
18. $\mathbf{S}(k, 1:n) \leftarrow \hat{\mathbf{z}}$
19. **end if**
20. **end if**
21. $time \leftarrow time + T_s,\ \ k \leftarrow k + 1$
22. **end while**
23. **return**

end ARCS

[1] Note that n is incremented in INVKLTCONVERGE

Subroutine **CONVEXOPT**($\mathbf{\Phi},\ \mathbf{\Psi},\ \mathbf{f},\ n$)
Method:

1. $\hat{\mathbf{h}} \leftarrow argmin(||h||_1)$ subject to $\mathbf{\Phi\Psi}h = \mathbf{f}$
2. $\mathbf{v} \leftarrow \mathbf{\Psi}\hat{\mathbf{h}}$
3. **return** $\mathbf{v}$

end CONVEXOPT

Subroutine **INVERSEKLT**(array)
Method:

1. $\mathbf{C} \leftarrow$ Covariance Matrix of array
2. $[\mathbf{V}\ \mathbf{E}] \leftarrow$ Eigen vectors and Eigen values of C
3. **return** $(\mathbf{V},\ \mathbf{E})$

end INVERSEKLT

Procedure **INVKLTCONVERGE**
Method:

1. **if** $k > 1$ **then**
2. $\left[\mathbf{\Psi}^{(n)}, \mathbf{e}_k\right] \leftarrow INVERSEKLT(\mathbf{S})$
3. $\delta \leftarrow \left\|\mathbf{e}_k - \mathbf{e}_{k-1}\right\|_2$
4. **if** $\delta < \tau$ **then**
5. $k \leftarrow 1,\ n \leftarrow n+1,\ \mathbf{e}_k \leftarrow -\infty,\ \mathbf{S} \leftarrow NULL$
6. $fam \leftarrow fam \cup entr,\ str \leftarrow str - entr,\ entr \leftarrow \{\alpha\},\ \alpha \in str$
7. **else**
8. $\mathbf{e}_{k-1} \leftarrow \mathbf{e}_k$
9. **end if**
10. **end if**
11. **return**

end INVKLTCONVERGE

Appendix C

Justification for using the downsized identity measurement matrix

The following arguments advocate the usage of deterministic measurement matrix as against fully random matrices:

1. In [123; 124] the authors have suggested the use of a measurement matrix consisting of M rows of an N-by-N orthonormal matrix, selected uniformly at random. The main result of [123] says that given a measurement matrix $\mathbf{U}_\Omega$ formed by picking up rows, uniformly at random, from an orthogonal matrix $\mathbf{U}$, it is possible to recover a signal $\mathbf{x} \in \mathbb{R}^N$ that is sparse $|\mathbf{x}|_0 << N$ from a vector $\mathbf{f} = \mathbf{U}_\Omega \mathbf{x}$. $\mathbf{U}_\Omega$ is an $M \times N$ matrix consisting of the rows of $\mathbf{U}$ indexed by the subset $\Omega \subset \{1...N\}$ of size $|\Omega| = M$.

   Now, $\mathbf{U}_\Omega = \mathbf{I}_\Omega \mathbf{U}$, where $\mathbf{I}_\Omega$ is a matrix consisting of the rows of the identity matrix $\mathbf{I}^{(N)}$, indexed by Ω. $\mathbf{I}_\Omega$ is nothing but the down-sized identity matrix, $\mathbf{\Phi}$, that has been used all along as the measurement matrix in this research. In place of $\mathbf{U}$, the inverse of DFT, DCT or KLT matrices, identified as $\mathbf{\Psi}$ the sparsity matrix, have been used at various places in the thesis, as the need be.

   Thus, the sensing matrix $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ is a matrix consisting of rows picked up uniformly at random from an orthogonal matrix and is therefore capable of recovering a sparse vector $\mathbf{c}$ (e.g. a sparse vector

of Fourier coefficients as in Chapter 3) from limited measurements, $\mathbf{f} = \mathbf{\Phi x} = \mathbf{\Theta c}$ in the signal domain. Once $\mathbf{c}$ is recovered, $\mathbf{x}$ can be obtained from $\mathbf{x} = \mathbf{\Psi} c$.

2. Although theoretically powerful, the practical relevance of using completely random matrices like the Gaussian and Bernoulli matrices is limited since we do not have the liberty to choose the type of measurements that will be used to acquire the signals [123].

3. It is quite costly to use random matrices in practical sensing applications as they require very high computational complexity and huge memory buffering due to their completely unstructured nature [123]. The following from [125] is quoted verbatim, "For example to process a 512×512 image with 64K measurements (25% of the original sampling rate), a Bernoulli matrix requires nearly gigabytes of memory and giga-flop operations, which makes both sampling and recovery very expensive and in many cases, unrealistic". In this paper too, the authors cite the use of a sensing matrix consisting of a uniformly random subset of rows of an orthonormal matrix in which the partial Fourier matrix is a special case [126].

4. The condition specified in equation 2.25 of sub-section 2.4.2 for guarantee of obtaining the sparsest solution to equations (2.13), (2.14) and (2.15) is true from a worst-case standpoint [59]. Thus, for a given $\mu(\mathbf{\Theta})$, a small fraction of signals with K substantially above the bound specified by (2.25) can still be successfully reconstructed. Or equivalently, for a given K for all the signals under test, even if $\mu(\mathbf{\Theta})$ is somewhat higher, successful reconstruction can take place. In [59] it is suggested that considering the average performance of CS as a function of K, an average measure of coherence (possibly like the one in equation 2.24) is more likely to describe its true behavior.

   The implication of this argument is that a deterministic matrix like the

downsized identity matrix, which has been used in this research does have a fair chance of passing the test of being a good measurement matrix, in spite of the fact that there exists no proof of it satisfying the RIP or being mutually incoherent with the sparsity matrices employed.

C.1 Empirical comparison of down-sized identity matrix and gaussian matrix

This section attempts to empirically estimate the probability of the downsized identity matrix having a lower mutual coherence with the inverse KLT matrix than a random Gaussian matrix.

1. For this purpose, the inverse KLT matrix $\mathbf{\Psi}^{(8)}$ of 8 signals taken from EEG dataset used in Chapter 4 is computed.

2. Next, down-sized identity matrices of dimension 3×8 (3 rows, 8 columns) are considered. There exist 56 such matrices, $\mathbf{\Phi}_{(i)}, 1 \leq i \leq 56$. The mutual coherence $\mu_{\mathbf{\Phi}_{(i)}}$ of each matrix, $\mathbf{\Phi}_{(i)}$ with $\mathbf{\Psi}^{(8)}$ is computed using equation 2.23 in sub-section 2.4.2.

3. Each $\mu_{\mathbf{\Phi}_{(i)}}$ is compared with the mutual coherence $\mu_{\mathbf{G}_{(j)}}, j = 1...10000$ of 10000 random Gaussian matrices of dimension 3×8. A count of the number of times, $\mu_{\mathbf{\Phi}_{(i)}} < \mu_{\mathbf{G}_{(j)}}, j = 1...10000$ is found which is used to compute the probability, $p_{\mathbf{\Phi}_{(i)}}$ which is equal to this count divided by 10000.

4. $p_{\mathbf{\Phi}_{(i)}}$ is computed for each of matrices $\mathbf{\Phi}_{(i)}, 1 \leq i \leq 56$. Thus, for all the 56 matrices together, totally there are 560000 comparisons done.

Figure C.1 plots the probabilities $p_{\mathbf{\Phi}_{(i)}}, 1 \leq i \leq 56$ for each of the 56 matrices as a bar graph. It can be seen that for many of the downsized identity matrices $\mathbf{\Phi}_{(i)}$, the probability is close to 1, that is, with a great

certainty the matrix has a lower μ with the given $\mathbf{\Psi}^{(8)}$, than any Gaussian matrix. Of course, there are two matrices where the probability is less than 0.1.

The surprising observation from this experiment is that, for the inverse KLT matrix used, the number of 3×8 downsized identity matrices, that have a lower mutual coherence than their Gaussian counterparts is 38 out of 56 (more than half) from a probabilistic standpoint.

However, there is no claim made from this simple experiment that the down-sized identity matrix is a better measurement matrix than its Gaussian equivalent, which is backed by theoretical proof. The effort made here is to point out that the absence of theoretical support does not confirm the failure of a general measurement matrix in performing good CS reconstruction.

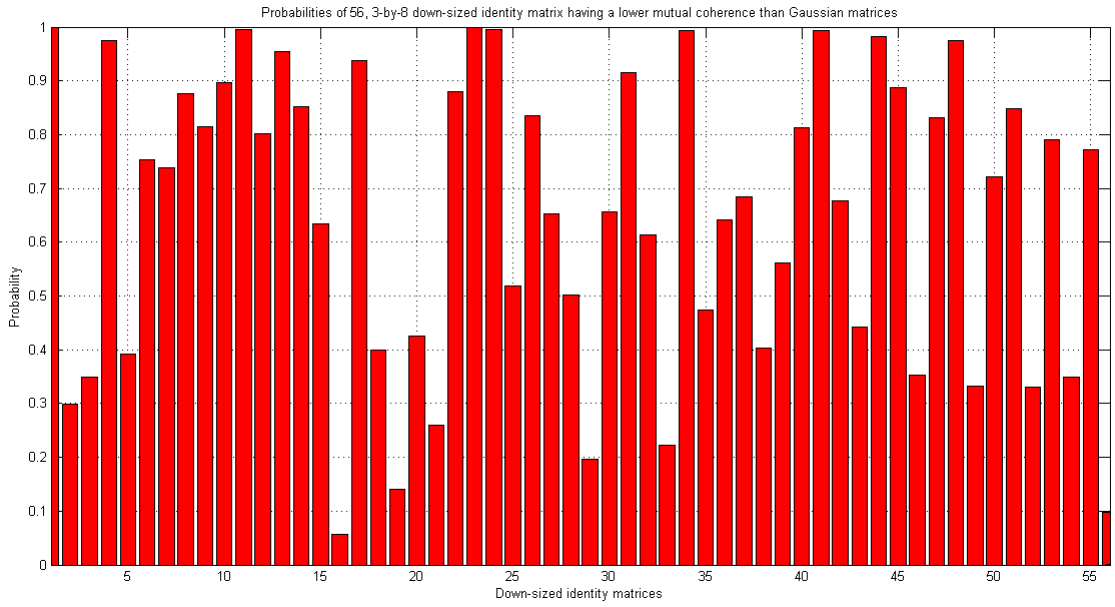


Figure C.1: Probabilities of 56 down-sized identity matrices $p_{\mathbf{\Phi}_{(i)}}$ having a lower mutual coherence with a given inverse KLT sparsity matrix, $\mathbf{\Psi}^{(8)}$ as compared to random Gaussian matrices of dimension 3×8

C.2 EEG reconstruction with Gaussian measurement matrix

For additional assurance that the performance of the down-sized identity matrix is comparable with other random matrices, the experiment on subject 1, record 7 has been repeated using the Gaussian measurement matrix. Comparison of the FSM values is shown as a plot in figure C.2. The average percentage FSM error in various EEG bands is - delta: 8.7%, theta: 6.3% alpha: 8.6% beta: 22.2%. These results show that reconstruction accuracies with the Gaussian measurement matrix and the down-sized identity matrix are comparable.

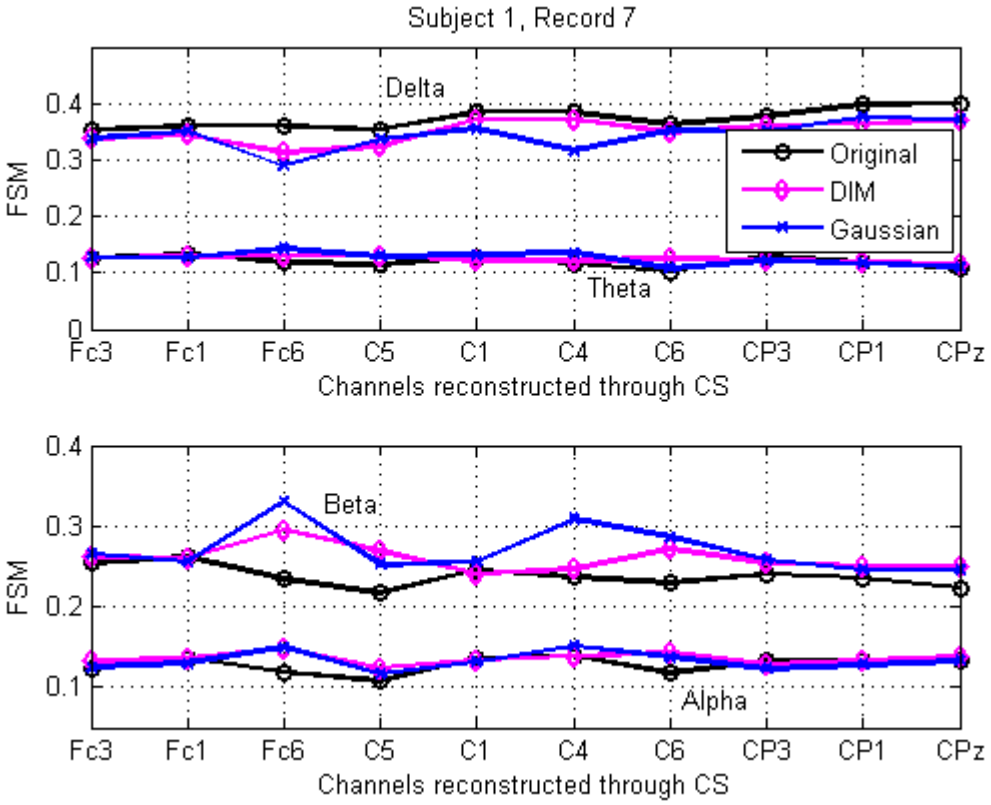


Figure C.2: Comparison of FSM values between reconstructed and original signals for Subject 1, Record 7 using down-sized identity matrices, indicated in the legend as DIM and the gaussian measurement matrix. With both the measurement matrices the CoSAMP recovery algorithm has been used.

Bibliography

[1] G.Schalk, D.J.McFarland, T.Hinterberger, N.Birbaumer, and J.R.Wolpaw, "BCI2000: A General-Purpose Brain-Computer Interface (BCI) System," *IEEE Trans. Biomed. Engg.*, vol. 51, no. 6, pp. 1034–1043, 2004. xxvi, 110, 111

[2] K. Bult, "Embedded Analog-to-Digital Converters," *ESSCIRC, 2009. Proceedings of*, pp. 52–64, Sept. 2009. 4

[3] K. Grochenig, "Acceleration of the Frame Algorithm," *IEEE Trans. on Signal Proc.*, vol. 41, no. 12, pp. 3331–3340, Dec. 1993. 9

[4] O. G. Guleryuz, "Nonlinear Approximation based Image Recovery using Adaptive Sparse Reconstructions and Iterated Denoising, parts i and ii," *IEEE Trans. on Image Proc.*, vol. 15, no. 3, pp. 539–571, March 2006. 9

[5] M. Unser, "Sampling-50 Years after Shannon," *Proc. of the IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000. 9

[6] F. Marvasti, "Random Topics in Nonuniform Sampling," in *Nonuniform Sampling: Theory and Practice.* Springer, formerly Kluwer Academic/Plenum Publishers, 2001, pp. 169–234. 9

[7] A. I. Zayed and P. L. Butzer, "Lagrange Interpolation and Sampling Theorems," in *Nonuniform Sampling: Theory and Practice.* Springer, formerly Kluwer Academic/Plenum Publishers, 2001, pp. 123–168. 9

[8] F. Marvasti, "Applications to Error Correction Codes," in *Nonuniform Sampling: Theory and Practice.* Springer, formerly Kluwer Academic/Plenum Publishers, 2001, pp. 689–738. 9

[9] J. Uriguen, Y. Eldar, P. Dragotti, and Z. Ben-Haim, "Sampling at the Rate of Innovation:Theory and Applications," in *Compressed Sensing: Theory and Applications.* Cambridge University Press, 2012, pp. 148–209. 10

[10] M. Vetterli, P. Marziliano, and T. Blu, "Sampling Signals with Finite Rate of Innovation," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002. 11

[11] I. Maravic and M. Vetterli, "Sampling and Reconstruction of Signals with Finite Rate of Innovation in the Presence of Noise," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2788–2805, August 2005. 11

[12] B. G. R. de Prony, "Essai experimental et analytique: sur les lois denla dilatabilite de fluides elastique et sur celles de la force expansive de la vapeur de lalkool, a differentes temperatures," *de lEcole Polytechnique*, vol. 1, pp. 24–76, 1795. 11

[13] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis.* Englewood Cliffs, NJ: Prentice Hall, 1997. 11, 73, 75

[14] P. Stoica and A. Nehorai, "Music, Maximum Likelihood, and Cramer-Rao Bound," *IEEE Trans. on ASSP*, vol. 37, no. 5, pp. 720–741, May 1989. 11

[15] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Trans. on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, March 1986. 11

[16] H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research: the Parametric Approach," *IEEE Signal Proc. Magazine*, vol. 13, no. 4, pp. 67–94, July 1996. 11

[17] B. D. V. Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, April 1988. 11, 12

[18] R. M. Leahy and B. D. Jeffs, "On the Design of Maximally Sparse Beamforming Arrays," *IEEE Trans. on Antennas and Propagation*, vol. 39, no. 8, pp. 1178–1187, August 1991. 12

[19] A. Kumar, P. Ishwar, and K. Ramchandran, "On Distributed Sampling of Bandlimited and Non-bandlimited Sensor Fields," *IEEE Int. Conf. on Acoustics, Speech and Signal Proc., ICASSP04, Montreal, Canada*, vol. 3, pp. 925–928, May 2004. 12

[20] W.U.Bajwa, J.Haupt, A. M. Sayeed, and R.Nowak, "Compressive Wireless Sensing," *Proc. Int. Symposium on Info. Proc. in Sensor Networks, IPSN06, Nashville, TN*, pp. 134–142, April 2006. 12

[21] S.S.Pradhan, J. Kusuma, and K.Ramchandran, "Distributed Compression in a Dense Microsensor Network," *IEEE Signal Proc. Magazine*, vol. 19, no. 2, pp. 51–60, March 2002. 12

[22] Y. Li, S. Amari, A. Chicocki, D. Ho, and S. Xie, "Underdetermined Blind Source Separation based on Sparse Representation," *IEEE Trans. on Signal Proc.*, vol. 54, no. 2, pp. 423–437, 2006. 13

[23] I. Takigawa, M. Kudo, A. Nakamura, and J. Toyama, "On the Minimum $l_1 - norm$ Signal Recovery in Underdetermined Source Separation," *Proc. of 5th Int. Conf. on Independent Component Analysis*, pp. 22–24, 2004. 13

[24] M. Aharon, M. Elad, and A. M. Bruckstein, "The k-svd: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. on Signal Proc.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006. 13

[25] F. Marvasti, A. Amini, F. Haddadi, M. Soltanolkotabi, B. H. Khalaj, A. Aldroubi, S.Holm, S. Sanei, and J. Chambers, "A Unified Approach to Sparse Signal Processing," *CoRR*, Feb. 2009. 14

[26] E. Candes, "Compressive Sampling," *Proceedings of the International Congress of Mathematicians, Madrid, Spain*, pp. 1433–1452, 2006. 19, 20, 24, 26, 28

[27] E. Candes and M. Wakin, "An Introduction to Compressive Sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, March 2008. 21, 27

[28] R. Baraniuk, "Lecture notes on Compressive Sensing," *IEEE, Signal Processing Magazine*, pp. 118–124, July 2007. 23, 24, 26, 28

[29] D.Baron, M.B.Wakin, M.F.Duarte, S.Sarvotham, and R.G.Baraniuk, "Distributed Compressed Sensing," Tech. Rep., 2005. [Online]. Available: http:www.dsp.ece.rice.edu/cs/DCS112005.pdf 23

[30] E. Candes, J. Romberg, and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006. 23, 29

[31] D. Donoho, "Compressed Sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006. 23, 29

[32] S. Chen, D. L. Donoho, and M. Saunders, "Atomic Decomposition by Basis Pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[Online]. Available: http://dx.doi.org/10.1137/S003614450037906X 24

[33] E. Cands and D. L. Donoho, "New Tight Frames of Curvelets and Optimal Representations of Objects with Piecewise C2 Singularities," *Comm. Pure Appl. Math*, vol. 57, no. 2, pp. 219–266, 2004. 24

[34] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data Compression and Harmonic Analysis," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2435–2476, 1998. 25

[35] E. Cands and T. Tao, "Decoding by Linear Programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005. 25

[36] V.Abolghasemi1, S.Ferdowsi1, B.Makkiabadi1, and S.Sanei1, "An Optimization of the Measurement Matrix for Compressive Sensing," *Proceedings of the 18th European Signal Processing Conference EUSIPCO-2010*, pp. 427–431, Aug. 2010. 27

[37] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, 2006. 29

[38] E. Candes and J. Romberg, "Practical Signal Recovery from Random Projections," *Wavelet Applications in Signal and Image Processing XI, Proc. SPIE Conf.*, vol. 5914, pp. 76–86, 2005. 29, 56

[39] K. R. Davidson and S. J. Szarek, "Local Operator Theory, Random Matrices and Banach Spaces," in *Handbook on the Geometry of Banach spaces*, W. Johnson and J. Lindenstrauss, Eds. Elsevier Science, 2003, vol. 2, p. 18191820. 29

[40] S. J. Szarek, "Condition numbers of random matrices," *J. Complexity*, vol. 7, no. 2, pp. 131–149, June 1991. 29

[41] E. J. Cands and T. Tao, "Near-optimal Signal Recovery from Random Projections and Universal Encoding Strategies," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006. 29, 30

[42] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007. 31

[43] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*, ser. Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics, 1987. [Online]. Available: http://books.google.co.in/books?id=C-MjQ98V9eoC 33

[44] D. L. Donoho, Y. Tsaig, I. Drori, and J. Starck, "Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit," *IEEE Trans. Inform. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012. 33

[45] T. Blumensath and M. E. Davies, "Gradient Pursuits," *IEEE Trans. on Signal Proc.*, vol. 56, no. 6, pp. 2370–2382, June 2008. 33

[46] D. Needell and R. Vershynin, "Signal Recovery from Incomplete and Inaccurate Measurements via Regularized Orthogonal Matching Pursuit," *IEEE journal of selected topics in signal processing*, vol. 4, no. 2, pp. 310–316, April 2010. 34

[47] D. Needell and J.A.Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis, Elsevier*, vol. 26, no. 3, pp. 301–321, May 2009. 34

[48] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Fast Reconstruction of Piecewise Smooth Signals from Random Projections," in *On-*

line Proceedings of the Workshop on Signal Processing with Adaptative Sparse Structured Representations (SPARS), Rennes, France, 2005. 34

[49] A. Gilbert, M. Strauss, J. Tropp, and R. Vershynin, "Algorithmic Linear Dimension Reduction in the l1 Norm for Sparse Vectors," in *Proc. 44th Annu. Allerton Conf. Communication, Control, Computing*, 2006. 34

[50] W. Dai, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," *IEEE Trans. Inform. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009. 35

[51] J. A. Tropp and A. C. Gilbert, "Simultaneous Sparse Approximation via Greedy Pursuit," *IEEE Trans. on Accoustics, Speech and Signal Processing*, vol. 5, pp. 721–724, 2005. 35

[52] S. Sarvotham, D. Baron, and R. G. Baraniuk, "Sudocodes: Fast Measurement and Reconstruction of Sparse Signals," in *Proc. of IEEE International Symp. on Infor. Theory*, July 2006, pp. 2804–2808. 36

[53] Compressive Sensing Resources. http://www.dsp.rice.edu/cs. 36, 40

[54] S. Ji, Y. Xue, and L. Carin, "Bayesian Compressive Sensing," *IEEE Trans. on Signal Proc.*, vol. 56, no. 6, pp. 2346–2356, June 2008. 36

[55] R. Chartrand, "Exact Reconstruction of Sparse Signals via Nonconvex Minimization," *IEEE Signal Proc. Letters*, vol. 14, no. 10, pp. 707–710, Oct. 2007. 37

[56] S. Li, F. Gao, G. Ge, and S. Zhang, "Deterministic Construction of Compressed Sensing Matrices via Algebraic Curves," *IEEE Trans. Inform. Theory*, vol. 58, no. 8, pp. 5035–5041, August 2012. 38

[57] R. A. DeVore, "Deterministic Constructions of Compressed Sensing Matrices," *J. Complexity*, vol. 23, no. 4–6, pp. 918–925, Aug. 2007. 38

[58] Bajwa, U. Waheed, J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, "Toeplitz-Structured Compressed Sensing Matrices," in *Proc. of IEEE 14th Workshop on Statistical Signal Processing*, Aug. 2007, pp. 294–298. 38

[59] M. Elad, "Optimized Projections for Compressed Sensing," *IEEE Trans. on Signal Proc.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007. 38, 158

[60] T. Petros and M. Asif, "Compressive Sensing for Streaming Signals using the Streaming Greedy Pursuit," *Proceedings of the Military Communications Conference, MILCOM-2010*, pp. 1205–1210, Oct. 2010. 39

[61] M. Asif and J. Romberg, "Streaming Measurements in Compressive Sensing:l_1 filtering," *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, pp. 1051–1058, Oct. 2008. 39

[62] J.A.Tropp, J.N.Laska, M.F.Duarte, and J.K.Romberg, "Beyond Nyquist: Efficient Sampling of Sparse Bandlimited Signals," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010. 44

[63] F.Chen, A.P.Chandrakasan, and V.M.Stojanovic, "Design and Analysis of a Hardware Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE Journal of Solid Sate Circuits*, vol. 47, no. 3, pp. 744–756, Mar. 2012. 44

[64] M.Mishali and Y.C.Eldar, "From Theory to Practice: Sub-Nyquist Sampling of Sparse Wideband Analog Signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, Apr. 2010. 44

[65] S.Kirolos, J.Laska, M.Wakin, M.Duarte, D.Baron, T.Ragheb, Y.Massoud, and R.Baraniuk, "Analog-to-information conversion via random demodulation," *IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software (DCAS)*, pp. 71–74, Oct. 2006. 44

[66] P. Boufounos and R.G.Baraniuk, "Sigma Delta Quantization for Compressive Sensing," *Proc. of Wavelets XII in SPIE International Symposium on Optical Science and Technology*, Aug. 2007. 44

[67] B.S.Atal and S. L. Hanauer, "Speech analysis and synthesis by linear prediction of the speech wave," *Journal of Acoustical Society of America*, vol. 50, no. 2, pp. 637–655, 1971. 46

[68] M. Duarte and R. G. Baraniuk, "Spectral Compressive Sensing," *Applied and Computational Harmonic Analysis*, Aug. 2012. 72, 74, 78

[69] J. A. Tropp, "Greed is good: Algorithmic Results for Sparse Approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004. 73

[70] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed Sensing and Redundant Dictionaries," *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2210–2219, May 2008. 73

[71] S. Kay, *Modern Spectral Estimation: Theory and Application.* Englewood Cliffs, NJ: Prentice Hall, 1988. 73, 75

[72] J. Gryka and I. Kale, "Detection of Multiple Sinusoids Buried in Noise via Balanced Model Truncation," *Proc. of the Conference on Instrumentation and Measurement Technology*, vol. 2, pp. 1353–1358, May 1998. 79

[73] K. H. Chon, "Accurate Identification of Periodic Oscillations Buried in White or Colored Noise using Fast Orthogonal Search," *IEEE Trans. Biomed. Engg.*, vol. 48, no. 6, pp. 622–629, June 2001. 79

[74] J. Chen and X. Huo, "Theoretical Results on Sparse Representation for Multiple Measurement Vectors (mmv) in an Over-complete Dictionary," *IEEE Trans. on Signal Proc.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006. 88

[75] Z. Zhang and B. Rao, "Sparse Signal Recovery with Temporally Correlated Source Vectors using Sparse Bayesian Learning," *IEEE Jour. of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011. 88

[76] J. Garcia-Frias and I. Esnaola, "Distributed Compression of Correlated Real Sequences using Random Projections," *Proc. of Data Compression Conference*, pp. 322–331, March 2008. 88

[77] D.Baron, M.B.Wakin, M.F.Duarte, S.Sarvotham, and R.G.Baraniuk, "Distributed Compressed Sensing of Jointly Sparse Signals," in *In Asilomar Conf. Signals, Sys., Comput*, Nov. 2005, pp. 1537–1541. 89

[78] J. Tropp, A. C. Gilbert, and M. J. Strauss, "Simulataneous Sparse Approximation via Greedy Pursuit," *IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, vol. 5, pp. 721–724, March 2005. 91

[79] V. N. Temlyakov, "A Remark on Simultaneous Sparse Approximation," *East J. Approx.*, vol. 100, pp. 17–25, March 2004. 91

[80] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse Solutions to Linear Inverse Problems with Multiple Measurement Vectors," *IEEE Trans. on Signal Proc.*, vol. 51, pp. 2477–2488, July 2005. 91

[81] H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Educ. Psychol.*, vol. 24, no. 6, pp. 417–441, Sep. 1933. 93

[82] K. Karhunen, *ber lineare methoden in der wahrscheinlichkeitsrechnung*, ser. Annales Academiae scientiarum Fennicae. Series A. 1, Mathematica-physica, 1947. [Online]. Available: http://books.google.co.in/books?id=Dn3gSAAACAAJ 93

[83] S. Sanei and J. A.Chambers, *EEG Signal Processing.* John Wiley and Sons, 2007. 102

[84] E. Niedermeyer and F. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* Lippincot Williams and Wilkins, 2004. 102

[85] E.Niedermeyer and F. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th ed. Lippincott Williams & Wilkins, 2004. 103

[86] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," *Neuroimage*, vol. 34, no. 4, pp. 1600–1611, Feb. 2007. 104

[87] Q. Zhou, "Analysis of EEG Data Using an Adaptive Periodogram Technique," *Intl. Conf. on Biomedical Engineering and Informatics*, vol. 2, pp. 353–357, May 2008. 104

[88] F. Romer, M. Haardt, D. Jannek, and P. Husar, "Multi-dimensional Space-Time-Frequency Component Analysis of Event Related EEG Data using Closed-form PARAFAC," *ICASSP*, pp. 349–352, April 2009. 104

[89] S. D. Cranstoun, H. C. Ombao, R. von Sachs, W. Guo, and B. Litt, "Time-Frequency Spectral Estimation of Multichannel EEG using

the Auto-SLEX Method," *IEEE Trans. Biomed. Engg.*, vol. 49, no. 9, pp. 988–996, Sep. 2002. 104

[90] S. Aviyente, "Compressed Sensing Framework for EEG Compression," *Proc. of IEEE 14th Workshop on Statistical Signal Processing*, pp. 181–184, Aug. 2007. 104

[91] R. D'Angelo, M. Trakimas, S. Sonkusale, and S. Aeron, "Compressed Sensing of EEG using a Random Sampling ADC in 90nm," *IEEE Int. Conf. Body Sensor Networks*, pp. 1–5, May 2013. 104

[92] S. Fauvel, A. Agarwal, and K. R. Ward, "Compressed Sensing and Energy–Aware Independent Component Analysis for Compression of EEG Signals," *IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, pp. 973–977, May 2013. 105

[93] A.K.Shah, R.Agarwal, J.R.Carhuapoma, and J.A.Loeb, "Compressed EEG Pattern Analysis for Critically Ill Neurological-Neurosurgical Patients," *Neurocritical Care*, vol. 5, no. 12, pp. 124–133, 2006. 105

[94] S.Senay, L.F.Chaparro, M.Sun, and R.J.Sclabassi, "Compressive Sensing and Random Filtering of EEG Signals using Slepian Basis," *Proc. of IEEE Transactions on Biomedical Engineering16th European Signal Processing Conference (EUSIPCO 2008)*, Aug. 2008. 105

[95] A.M.Abdulghani, A.J.Casson, and E.Rodriguez-Villegas, "Quantifying the performance of compressive sensing on scalp EEG signals," *Proc. of 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, pp. 1–5, Nov. 2010. 105

[96] 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," *IEEE Trans. Biomed. Engg.*, vol. 60, no. 1, pp. 221–224, 2013. 105

[97] S.Li, L.D.Xu, and X.Wang, "A continuous biomedical signal acquisition system based on compressed sensing in body sensor networks," *IEEE Trans. Industrial Informatics*, vol. 9, no. 3, pp. 1764–1771, Aug. 2013. 105

[98] S.Fauvel and R.K.Wardemail, "An Energy Efficient Compressed Sensing Framework for the Compression of Electroencephalogram Signals," *Sensors*, vol. 14, no. 1, pp. 1474–1496, 2014. 105

[99] A.M.Abdulghani, A.J.Casson, and E.R.Villegas, "Compressive sensing scalp EEG signals:Implementations and practical performance," *Med. Biol. Eng. Comput.*, vol. 50, pp. 1137–1145, 2012. 105

[100] A.M.Abdulghan, A.J.Casson, and E.R.Villegas, "Quantifying the performance of compressive sensing on scalp EEG signals," *Proc. of the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, pp. 1–5, Nov. 2010. 105

[101] J. Onton, M. Westerfield, J. Townsend, and S. Makeig, "Imaging Human EEG Dynamics using Independent Component Analysis," *Neuroscience and Biobehavioral Reviews*, vol. 30, pp. 808–822, 2006. 106

[102] L. Leocani, T. Locatelli, V. Martinelli, M. Rovaris, M. Falautano, M. Filippi, G. Magnani, and G. Comi, "Electroencephalographic Coherence Analysis in Multiple Sclerosis: Correlation with Clinical, Neuropsychological, and MRI Findings," *J. Neurol. Neurosurg. Psychiatry*, vol. 69, pp. 192–198, 2000. 106

[103] I. Manshanden, J. D. Munck, N. Simon, and F. da Silva, "Source Localization of MEG Sleep Spindles and the Relation to Sources of

Alpha Band Rhythms," *Clin. Neurol.*, vol. 113, pp. 1937–1947, 2002. 106

[104] P. Nunez, R. Srinivasan, A. Westdorp, R. Wijesinghe, D. Tucker, R. Silberstein, and P. Cadusch, "EEG Coherency. Statistics, Reference Electrode, Volume Conduction, Laplacians, Cortical Imaging, and Interpretation at Multiple Scales. Electroencephalography," *Clin. Neurophysiol.*, vol. 103, pp. 499–515, 1997. 106

[105] G. Winterer, M. Egan, T. Radler, T. Hyde, R. Coppola, and D. Weinberger, "An Association Between Reduced Interhemispheric EEG Coherence in the Temporal Lobe Genetic Risk for Schizophrenia," *Schizophr. Res.*, vol. 40, pp. 129–143, 2001. 106

[106] M. Knyazeva, D. Kiper, V. Vildavski, P. Despland, M. MaederIngvar, and G. Innocenti, "Visual Stimulus-Dependent Changes in Interhemispheric EEG Coherence in Humans," *J. Neurophysiol*, vol. 82, pp. 3095–3107, 1999. 106

[107] K. Cover, C. Stam, and B. van Dijk, "Detection of Very High Correlation in the Alpha Band Between Temporal Regions of the Human Brain using MEG," *Schizophr. Res.*, vol. 22, pp. 1432–1437, 2004. 106

[108] http://epilepsy.med.nyu.edu/diagnosis-treatment/eeg/ambulatory-eeg#sthash.UmDO6WON.dpbs). 108

[109] H. Zou, T. Hastie, and R. Tibshirani, "Sparse Principal Component Analysis," *Computational and Graphical Statistics*, vol. 15, no. 2, pp. 265–286, 2006. 109

[110] A. d'Aspremont, L. E. Ghaoui, M. I. Jordan, and G. R. G. Lanckriet, "A Direct Formulation for Sparse PCA using Semidefinite Programming," *SIAM Review*, vol. 49, no. 2, pp. 434–448, July 2007. 109

[111] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley. (2000, June 13) Physiobank, Physiotoolkit, and Physionet: Components of a New Research Resource for Complex Physiologic Signals. [Circulation Electronic Pages http://circ.ahajournals.org/cgi/content/full/101/23/e215]. 110, 141

[112] http://www.bci2000.org/. 110

[113] R.T.Pivik, R.J.Broughton, R.Coppola, R.J.Davidson, N.Fox, and M.R.Nuwer, "Guidelines for the recording and quantitative analysis of electroencephalographic activity in research contexts," *Psychophysiology*, vol. 30, pp. 547–558, 1993. 112

[114] (2012, December) CVX Matlab Software for Disciplined Convex Programming, Version 2.0 Build 893. http://cvxr.com/cvx/. 113

[115] J. Kaiser, "Nonrecursive Digital Filter Design using the - Sinh Window Function," *Proc. 1974 IEEE Symp. Circuits and Systems*, pp. 20–23, April 1974. 126

[116] *Selected Papers in Digital Signal Processing II.* IEEE Press, New York, 1975. 126

[117] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing.* Prentice-Hall, 1989. 126

[118] A.Matonia, J.Jezewski, T. Kupka, K.Horoba, J.Wrobel, and A.Gacek, "The Influence of Coincidence of Fetal and Maternal QRS Complexes on Fetal Heart Rate Reliability," *Medical and Biological Engg. and Computing*, vol. 44, no. 5, pp. 393–403, 2006. 141

[119] L. Polania, R.E.Carrillo, M.Blanco-Velasco, and K.E.Barner, "Compressed Sensing Based Method for ECG Compression," *Proc. of ICASSP*, pp. 761–764, 2011. 142

[120] H.Mamaghanian, N.Khaled, D.Atienza, and P.Vandergheynst, "Compressed Sensing for Real-Time Energy-Efficient ECG Compression on Wireless Body Sensor Nodes," *Proc. of IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2456–2466, Sep. 2011. 142

[121] A.Dixon, E.G.Allstot, D.Gangopadhyay, and D.J.Allstot, "Compressed Sensing System Considerations for ECG and EMG Wireless Biosensors," *Proc. of IEEE Transactions on Biomedical Circuits and Systems*, vol. 6, no. 2, pp. 156–166, Apr. 2012. 144

[122] Z.Zhang, T.P.Jung, S.Makeig, and B.D.Rao, "Compressed Sensing for Energy-Efficient Wireless Telemonitoring of Noninvasive Fetal ECG via Block Sparse Bayesian Learning," *Proc. of IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 300–309, Feb. 2013. 144

[123] E. Candes and J. Romberg, "Sparsity and Incoherence in Compressive Sampling," *Inverse problems*, vol. 23, no. 3, pp. 969–985, 2007. 157, 158

[124] L.Carin, D.Liu, and B.Guo, "Coherence, Compressive Sensing and Random Sensor Arrays," *IEEE, Antennas and Propagation Magazine*, vol. 53, no. 4, pp. 28–39, Aug. 2011. 157

[125] T.T.Do, L.Gan, N.H.Nguyen, and T.D.Tran, "Fast and Efficient Compressive Sensing using Structurally Random Matrices," *IEEE Trans. on Signal Proc.*, vol. 60, no. 1, pp. 139–154, Jan. 2012. 158

[126] S.Mendelson, A. Pajor, and N.T.Jaegermann, "Uniform Uncertainty Principle for Bernoulli and Subgaussian Ensembles," *Constructive Alg.*, vol. 28, pp. 269–283, Aug. 2008. 158