Compressive sensing reconstruction algorithms using partially correct signal information

Sekar Sathiya Narayanan

2017

Sekar Sathiya Narayanan. (2017). Compressive sensing reconstruction algorithms using partially correct signal information. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/69572>

<https://doi.org/10.32657/10356/69572>

Compressive Sensing Reconstruction Algorithms using Partially Correct Signal Information

Sekar Sathiya Narayanan

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2017

Acknowledgments

This thesis is a major milestone in my journey of research and therefore, I am feeling very happy to thank all who have supported me to reach it.

I would like to express my gratitude to my Ph.D. supervisor/advisor, Associate Professor Anamitra Makur, for his constant support and encouragement. It is his guidance and timely help that took me towards the completion of this thesis. It is indeed a great pleasure to work with a supervisor like Prof. Makur who can always accommodate the students. In many situations, I have benefited from his timely suggestions/inputs.

I am very grateful to the School of Electrical and Electronics Engineering, Nanyang Technological University (NTU) for providing me the research student scholarship as well as a conducive and comfortable environment to study. I would like to thank the Media Technology Laboratory executive, Mrs Leow How for her ever willingness to help. I end up expressing my gratitude to my friends, fellow students of my supervisor, and my loving family.

Contents

Acknowledgments	i
Contents	ii
List of Figures	v
List of Tables	x
List of Abbreviations	xii
List of General Notations	xv
Summary	xvi
1 Introduction	1
1.1 Compressive Sensing and Sparse Signal Recovery	1
1.2 Motivations	2
1.3 Contributions and Outline of the Thesis	5
2 Review of Compressive Sensing Reconstruction Algorithms	8
2.1 The Sensing Problem	8
2.2 Compressive Sensing Matrices	9
2.3 Compressive Sensing Reconstruction Algorithms	11

2.3.1	Convex Relaxation Techniques	11
2.3.2	Greedy Pursuits	12
2.3.3	Other Categories of Algorithms	14
2.4	Compressive Sensing Reconstruction Algorithms using Partially Known Support	15
2.5	Distributed Compressive Sensing	17
2.6	Fusion based Compressive Sensing Reconstruction Algorithms	19
2.7	Compressive Sensing Applications	20
2.8	Summary	21
3	Modified Adaptive Basis Pursuit	22
3.1	Shortcomings of Modified Basis Pursuit	22
3.2	Proposed Modified Adaptive Basis Pursuit	25
3.3	Comparison of MABP and Mod-BP	27
3.3.1	Comparison based on Sufficient Conditions for Exact Reconstruction	28
3.3.2	Comparison based on the High Probability Upper Bound on the RIC	32
3.3.3	Comparison based on Reconstruction Error Bound	35
3.4	Proposed Weighted Adaptive ℓ_1 -minimization	35
3.5	Reconstruction of Correlated Sparse Signals	37
3.5.1	Synthetic JSM-1 Signals	37
3.5.2	Time-sequence Reconstruction of Video Frames	43
3.6	Extension of MABP to Greedy Pursuits	44
3.7	Summary	49
4	Adaptive Backtracking Matching Pursuit	50
4.1	Drawback in PKS based Greedy Pursuits	50
4.2	Proposed Adaptive Backtracking Matching Pursuit	52

4.2.1	The AdBMP Algorithm	52
4.2.2	Complexity Analysis	54
4.2.3	Theoretical Analysis	55
4.3	Reconstruction of Synthetic JSM-1 Signals	59
4.4	Sparse Recovery of Radar Echo Signals	63
4.5	Reconstruction of Video Signals	68
4.5.1	Time-sequence Reconstruction of Video Frames	68
4.5.2	Object based Surveillance Video Coding	70
4.6	Summary	77
5	Iterative Compressive Sensing Framework for Compressive Imaging	78
5.1	Exploratory Experiment and Motivation	79
5.2	Proposed Iterative Compressive Sensing on Residual	81
5.2.1	Initial Support Set Estimation	81
5.2.2	Iterative CS on Residual	82
5.2.3	‘No-PKS’ Scenario	84
5.2.4	Complexity Analysis	84
5.3	Theoretical Analysis of ICSR	86
5.3.1	Recurrence Relation of ICSR in the Case of Noiseless Measurements	89
5.3.2	Recurrence Relation of ICSR in the Case of Noisy Measurements .	91
5.4	Application: Compressive Imaging	93
5.5	Summary	104
6	Greedy Pursuits Assisted Basis Pursuit	106
6.1	Shortcomings of Fusion-of-Algorithms	106
6.2	Proposed Greedy Pursuits Assisted Basis Pursuit	107
6.2.1	Theoretical Analysis	107

6.2.2	Complexity Analysis	109
6.3	Proposed Greedy Pursuits Assisted Weighted ℓ_1	110
6.3.1	Theoretical Analysis	111
6.3.2	Complexity Analysis	116
6.4	Reconstruction of Sparse Signals: GPABP	116
6.4.1	Synthetic Sparse Signals	116
6.4.2	Compressible ECG Signals	119
6.5	Reconstruction of Sparse Signals: GP-W ℓ_1	120
6.6	Proposed Greedy Pursuits Assisted Basis Pursuit for Multiple Measure- ment Vectors	123
6.6.1	Theoretical Analysis	126
6.7	Reconstruction of Correlated Sparse (JSM-1) Signals: GPABP-MMV . . .	129
6.7.1	Synthetic JSM-1 Signals	129
6.7.2	Time-sequence Reconstruction of Compressible ECG Signals . . .	132
6.8	Summary	133
7	Conclusions and Future Directions	135
7.1	Conclusions	135
7.2	Directions for Further Research	137
	Author's Publications	139
	References	141

List of Figures

1.1	Context diagram of the thesis contributions	5
2.1	Non-sparse signal and its sparse representation in DCT domain	9
3.1	Schematic of actual support (T) and PKS (T_k)	23
3.2	Block schematic of MABP	25
3.3	Schematic of actual support (T), PKS (T_k) and the retained support Γ_k . . .	26
3.4	High probability upper bound on the RIC versus SR for $\alpha = 0.05$	34
3.5	High probability upper bound on the RIC versus SR for $\alpha = 0.25$	34
3.6	Reconstruction probability versus MR	38
3.7	Average MSE versus MR	39
3.8	Reconstruction probability versus CSR	39
3.9	Average MSE versus CSR	40
3.10	Reconstruction probability versus MR for different μ_a : MABP	42
3.11	Reconstruction probability versus MR for different μ_a : $WA\ell_1$	42
3.12	Reconstruction probability versus MR for different ω : $WA\ell_1$ and $W\ell_1$. . .	43
3.13	Average PSNR (in dB) versus MR: YUV sequences	44
3.14	Average FSR versus MR	47
3.15	Average FSR versus CSR	48
3.16	Average FSR versus MR for different μ_a : MAGP-CoSaMP	48
3.17	Average FSR versus MR for different μ_a : MAGP-IHT	49

4.1	Schematic of actual support (T), PKS (Ω) and initial support set (T^0)	53
4.2	Average FSR versus MR	60
4.3	Average MSE versus MR	60
4.4	Average FSR versus CSR	61
4.5	Average MSE versus CSR	61
4.6	Average reconstruction time (in sec) versus CSR	62
4.7	Average MSE versus τ_0 for different τ	62
4.8	Average SRER (in dB) versus MR	65
4.9	Average reconstruction time (in sec) versus MR	66
4.10	Average SRER (in dB) versus p	66
4.11	Average PSNR (in dB) versus MR: YUV sequences	68
4.12	Average reconstruction time (in sec) versus MR: YUV sequences	69
4.13	DCVS based video object coding - encoder	70
4.14	DCVS based video object coding - decoder	72
4.15	Speed and accuracy comparison: AdBMP	77
5.1	Two major steps in ICSR framework	81
5.2	Schematic of actual support set (T), set of high probability locations (T_{hp}) and initial support set (T^0)	82
5.3	Iterations in ICSR framework	83
5.4	Average PSNR (in dB) versus MR: KODAK database	95
5.5	Average number of ICSR iterations (per block) versus MR: KODAK database	97
5.6	Effect of λ_0 on PSNR (in dB): KODAK database	98

5.7	ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim01</i> from KODAK database), reconstructed images for Mod-BP (24.33 dB), ICSR-OMP (24.05 dB), ICSR-CoSaMP (24.20 dB), ICSR-IHT (24.14 dB), ICSR-BP (24.13 dB) and ICSR-FOCUSS (24.14 dB).	98
5.8	ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim06</i> from KODAK database), reconstructed images for Mod-BP (25.50 dB), ICSR-OMP (25.14 dB), ICSR-CoSaMP (25.15 dB), ICSR-IHT (25.29 dB), ICSR-BP (25.11 dB) and ICSR-FOCUSS (25.28 dB).	99
5.9	ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim10</i> from KODAK database), reconstructed images for Mod-BP (27.19 dB), ICSR-OMP (27.08 dB), ICSR-CoSaMP (27.12 dB), ICSR-IHT (26.93 dB), ICSR-BP (27.07 dB) and ICSR-FOCUSS (27.19 dB).	99
5.10	ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim21</i> from KODAK database), reconstructed images for Mod-BP (25.60 dB), ICSR-OMP (24.95 dB), ICSR-CoSaMP (25.08 dB), ICSR-IHT (24.98 dB), ICSR-BP (25.14 dB) and ICSR-FOCUSS (25.11 dB).	100
5.11	Average PSNR (in dB) versus MR - No PKS: KODAK database	100
5.12	Average reconstruction Time (in sec) versus MR - No PKS: KODAK database	101

5.13	ICSR without PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim01</i> from KODAK database), reconstructed images for CoSaMP (16.73 dB), IFSRA-CoSaMP (17.43 dB), ICSR-CoSaMP (21.98 dB), BP (23.90 dB), IFSRA-BP (24.17 dB) and ICSR-BP (24.07 dB).	101
5.14	ICSR without PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim06</i> from KODAK database), reconstructed images for CoSaMP (17.93 dB), IFSRA-CoSaMP (18.95 dB), ICSR-CoSaMP (23.54 dB), BP (25.11 dB), IFSRA-BP (25.10) and ICSR-BP (25.21 dB).	102
5.15	ICSR without PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim10</i> from KODAK database), reconstructed images for CoSaMP (18.21 dB), IFSRA-CoSaMP (19.02 dB), ICSR-CoSaMP (24.54 dB), BP (26.03 dB), IFSRA-BP (26.24 dB) and ICSR-BP (26.27 dB).	102
5.16	ICSR without PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (<i>kodim21</i> from KODAK database), reconstructed images for CoSaMP (17.82 dB), IFSRA-CoSaMP (18.35 dB), ICSR-CoSaMP (22.15 dB), BP (23.90 dB), IFSRA-BP (24.24 dB) and ICSR-BP (24.27 dB).	103
5.17	Speed and accuracy comparison: ICSR	105
6.1	Gradual Weighting Strategy of GP- $W\ell_1$	110
6.2	Reconstruction probability versus SR	117
6.3	Average MSE versus SR	117
6.4	Reconstruction probability versus MR	118
6.5	Average MSE versus MR	118

6.6	Average MSE versus MR: Compressible ECG signals (SMNR = 20 dB) . .	120
6.7	Reconstruction probability versus MR	121
6.8	Average reconstruction time (in sec) versus MR	122
6.9	Effect of L on reconstruction probability	122
6.10	Effect of L on average MSE	123
6.11	Average MSE versus CSR	130
6.12	Average reconstruction time (in sec) versus CSR	131
6.13	Average MSE versus J	131
6.14	Average MSE versus MR	132
6.15	Average MSE versus MR: Compressible ECG signals (SMNR = 20 dB) . .	133
6.16	Speed and accuracy comparison: GPABP-MMV	134

List of Tables

4.1	Comparison of constants involved in the recurrence relation of various GPs	58
4.2	Encoding and decoding performances of DCVS-VOC	76
5.1	Fractions of correct atoms and signal energy in T_{hp}	80
5.2	Average PSNR (in dB) versus MR: KODAK database	95
5.3	Average reconstruction time (computation time in MATLAB 7.12.0 running on a 64-bit Intel(R) Core TM i5-2400 processor with 8 GB RAM) in seconds - KODAK database	96
7.1	Comparison of Proposed Algorithms	138

List of Abbreviations

AdBMP	Adaptive Backtracking Matching Pursuit
BAOMP	Backtracking based Adaptive Orthogonal Matching Pursuit
BMP	Backtracking Matching Pursuit
BP	Basis Pursuit
BPDN	Basis Pursuit Denoising
CI	Compressive Imaging
CoMACS	Committee Machine Approach for Compressive Sensing
CoSaMP	Compressive Sampling Matching Pursuit
CR	Convex Relaxation
CS	Compressive Sensing
CSR	Common Support Ratio
DCS	Distributed Compressive Sensing
DCT	Discrete Cosine Transform
DCVS-VOC	Distributed Compressive Video Sensing based Video Object Coding
DFT	Discrete Fourier Transform
FACS	Fusion of Algorithms for Compressive Sensing
FOCUSS	FOCal Underdetermined System Solver

FSR	Fraction of Support Recovered
GOP	Group-Of-Pictures
GP	Greedy Pursuit
GPABP	Greedy Pursuits Assisted Basis Pursuit
GP-Wℓ_1	Greedy Pursuits assisted Weighted ℓ_1 -minimization
ICSR	Iterative Compressive Sensing Recovery
IFSRA	Iterative Framework for Sparse Reconstruction Algorithm
IHT	Iterative Hard Thresholding
JSM	Joint-Sparsity Model
LS-CS	Least Squares Compressive Sensing
LTI	Linear Time-Invariant
MABP	Modified Adaptive Basis Pursuit
MAGP	Modified Adaptive Greedy Pursuit
MMV	Multiple Measurement Vectors
Mod-BP	Modified Basis Pursuit
Mod-GP	Modified Greedy Pursuit
MR	Measurement Ratio
MSE	Mean Square Error
M-FACS	Modified Fusion of Algorithms for Compressive Sensing
NCR	Non-Convex Relaxation
OMP	Orthogonal Matching Pursuit
PKS	Partially Known Support
PSNR	Peak Signal to Noise Ratio

P-OMP	Probability weighted Orthogonal Matching Pursuit
ReWℓ_1	Re-Weighted ℓ_1 -minimization
RIP	Restricted Isometry Property
ROMP	Regularized Orthogonal Matching Pursuit
R-MSBL	Robust Multiple Sparse Bayesian Learning
SBL	Sparse Bayesian Learning
SMNR	Signal to Measurement Noise Ratio
SMV	Single Measurement Vector
SP	Subspace Pursuit
SR	Sparsity Ratio
SRER	Signal to Reconstruction Error Ratio
StOMP	Stagewise Orthogonal Matching Pursuit
WAℓ_1	Weighted Adaptive ℓ_1 -minimization
Wℓ_1	Weighted ℓ_1 -minimization

List of General Notations

$ T $	Cardinality of set T or number of elements in vector T
T^c	Complement of set T
$\ x\ _q$	ℓ_q norm of vector x
$\ \mathbf{X}\ _F$	Frobenius norm of matrix $\mathbf{X}$
Φ_I	Sub-matrix formed by extracting the columns of matrix Φ corresponding to the indices in vector I .
$\Phi^\dagger$	Pseudo-inverse of matrix Φ
Φ'	Transpose of matrix Φ

Summary

Compressive Sensing (CS) ensures the reconstruction of a sparse signal from a set of linear measurements that are fewer compared to the signal length. The sparse signal can be reconstructed using a convex relaxation algorithm such as Basis Pursuit (BP) or a Greedy Pursuit (GP) such as Backtracking Matching Pursuit (BMP). If some information regarding the signal support (non-zero locations) is available in the form of Partially Known Support (PKS), the same sparse signal can be recovered with higher accuracy. However, the size and accuracy of the PKS varies depending upon the signal model and characteristics. A generic PKS based reconstruction algorithm might work well in a particular scenario but fail in another. This thesis focuses on developing PKS based reconstruction algorithms for different scenarios wherein they make effective use of the available PKS.

In Distributed Compressive Sensing (DCS) reconstruction of correlated sparse signals, if one signal is known a priori, remaining signals can be reconstructed using PKS based convex relaxation algorithms such as Modified Basis Pursuit (Mod-BP) or Weighted ℓ_1 -minimization ($W\ell_1$) with far fewer measurements compared to separate CS reconstruction. However, these methods are greatly influenced by the support set of the known signal that includes one or more locations that are irrelevant to the target signal. In this thesis, Modified Adaptive Basis Pursuit (MABP) and Weighted Adaptive ℓ_1 -minimization ($WA\ell_1$) algorithms are proposed wherein an adaptation process is carried out ahead of applying Mod-BP or $W\ell_1$. The purpose of this adaptation is to drop as many irrelevant locations (wrong atoms) as possible and reduce their impact on signal reconstruction. It will be theoretically

shown that if in the adaptation step there are more correct deletions than wrong deletions, then MABP/WA ℓ_1 has a more favorable requirement compared to that of the Mod-BP/W ℓ_1 . Experimental results show that the proposed MABP/WA ℓ_1 algorithms provide better signal reconstruction compared to that of the Mod-BP/W ℓ_1 algorithms.

Existing PKS based GPs have a drawback: they are totally blind to the wrong atoms present in the PKS. In this thesis, with a motive of making proper/effective use of the PKS in BMP, a PKS based adaptation mechanism is proposed which, when replaces the traditional initialization of BMP, results in a reconstruction accuracy higher than that of the existing PKS based GPs. The resulting algorithm is termed as Adaptive Backtracking Matching Pursuit (AdBMP). Experimental results show that, in recovering correlated sparse signals, the proposed AdBMP gives a better reconstruction performance compared to that of the existing PKS based greedy algorithms and the BMP itself. AdBMP is shown to be suitable for a variety of CS applications including sparse recovery of radar echo signals and time sequence reconstruction of video signals.

Many existing CS reconstruction algorithms were modified so as to make use of this PKS to improve their performance, and are termed as PKS based CS reconstruction algorithms. Most PKS based CS reconstruction algorithms are totally blind to the wrong atoms present in the PKS which degrades the reconstruction performance. Moreover, each PKS based algorithm adapts its own modification strategy. In this thesis, a generalized, algorithm-blind reconstruction framework called Iterative Compressive Sensing on Residual (ICSR) is proposed, which can accommodate any class of CS reconstruction algorithm, and is capable of improving the signal reconstruction utilizing the available PKS. Experimental results show that, in compressive imaging application, the ICSR framework improves the reconstruction accuracy of any ingredient CS reconstruction algorithm, and performs better than many of the existing PKS based algorithms. The ICSR framework can also be applied to reconstruction of sparse signals for which no PKS is available.

Fusion based CS reconstruction algorithms combine multiple CS reconstruction algorithms, which work with different principles, to obtain a better signal estimate. Examples of such algorithms include Fusion of Algorithms for Compressed Sensing (FACS) and Committee Machine Approach for Compressed Sensing (CoMACS). However, these algorithms involve solving a least squares problem which might be ill-conditioned. Since Mod-BP makes use of available signal knowledge to improve upon BP, a novel fusion technique called Greedy Pursuits Assisted Basis Pursuit (GPABP) is proposed which employs multiple GPs to derive a partial support (i.e. PKS) for Mod-BP. Experimental results show that the proposed GPABP performs better than FACS and CoMACS. Note that the proposed GPABP does not require a PKS and it is suitable for applications where no PKS is available. DCS is an extension of CS from single measurement vector problem to Multiple Measurement Vectors (MMV) problem. In DCS, several reconstruction algorithms have been proposed to reconstruct the joint-sparse signal ensemble from MMV. However, most of them are designed for signal ensembles sharing common support. Since the assumption of common sparsity pattern for 100% of the data is very restrictive, this thesis focuses on the signal ensemble containing both common and innovation components. With a goal of proposing an MMV-type algorithm that is robust to outliers (innovation components), an MMV version of GPABP is proposed, termed as GPABP-MMV. It involves MMV versions of multiple GPs and Mod-BP.

This thesis proposes several PKS based reconstruction algorithms (MABP, AdBMP and ICSR) that handle the available PKS in different ways based on the signal model. Further, the reconstruction of sparse signals with no PKS is addressed by some proposed algorithms (GPABP and ICSR).

Chapter 1

Introduction

Compressive Sensing (CS) or Compressed Sensing provides an efficient and accurate signal acquisition by combining low-rate sampling with computational power [1]. The fact that a sparse (or compressible) signal can be captured efficiently using a small number of linear incoherent measurements that are proportional to its information level has made CS suitable for a number of possible applications. This thesis focuses on developing new CS reconstruction algorithms. This chapter tells the background and the motivation behind the contributions of this thesis.

1.1 Compressive Sensing and Sparse Signal Recovery

CS ensures the reconstruction of an n -length sparse signal $x \in \mathbb{R}^n$ from its measurement vector $y \in \mathbb{R}^m$ of the form

$$y = \Phi x \in \mathbb{R}^m \tag{1.1}$$

where $m \ll n$ and $\Phi \in \mathbb{R}^{m \times n}$ is a known CS matrix [1, 2]. A signal x is said to be K -sparse if there are only $K \ll n$ significant entries in it. Typically, the measurement vector y is of length $m=O(K \log(n))$ [3, 4]. The sparse signal can be reconstructed from y using CS re-

construction algorithms. CS reconstruction algorithms can be broadly classified as Convex Relaxation (CR) methods (such as Basis Pursuit (BP) [5]) and Greedy Pursuits (GPs) [6]-[16]. Among these two classes, CR methods are known for their reconstruction accuracy and GPs are known for their faster convergence. A detailed review of CS reconstruction algorithms is given in chapter 2. Robustness of CS can be studied using the Restricted Isometry Property (RIP) of the sensing matrix. For sampling K -sparse signals, the sensing matrix Φ should obey RIP which ensures that all subsets of K columns of Φ are nearly orthogonal [17]. Note that CS requires y to be dense.

CS exploits the fact that many natural signals are sparse or compressible in the sense that they have sparse representations when expressed in an appropriate basis Ψ [4]. A signal x is said to be compressible (i.e. K -sparse in the basis Ψ) if there exists a vector $\theta \in \mathbb{R}^n$ with only $K \ll n$ significant entries such that $x = \Psi\theta$. Since CS requires y to be dense, in the case of compressible signals, the coherence between the CS matrix (also known as sensing matrix) Φ and the sparsifying basis Ψ should be low. For example, random sensing matrices with independent and identically distributed (i.i.d.) Gaussian entries will exhibit a very low coherence with any sparsifying basis [4]. For sampling compressible signals, the matrix $\Phi\Psi$ should obey RIP. If Φ is an i.i.d. Gaussian random matrix (i.e. random matrix with i.i.d. Gaussian entries), $\Phi\Psi$ satisfies RIP with overwhelming probability provided $m = O(K \log(\frac{n}{K}))$ [2]. The sensing scheme involving a Gaussian random matrix is universal in the sense that the above result holds for all sparsifying bases (including the unitary basis). In this thesis, unless specified, signal x is a sparse signal with only K non-zero entries.

1.2 Motivations

If some information regarding the signal support (non-zero locations) is available in the form of Partially Known Support (PKS), the same sparse signal can be reconstructed with higher accuracy using PKS based CS reconstruction algorithms [18]-[27]. However, the

size and accuracy of the PKS varies depending upon the signal model and characteristics. A detailed review of PKS based CS reconstruction algorithms is given in chapter 2.

In Distributed Compressive Sensing (DCS), correlated sparse signals stand for an ensemble of signals characterized by presenting a sparse correlation [28]. Joint-Sparsity Models (JSMs) generalize the notion of a signal being sparse in some orthobasis to that of an ensemble of signals being jointly sparse. The DCS theory rests on the joint sparsity of a signal ensemble where there are dramatic savings relative to the number measurements required for separate CS reconstruction. In DCS reconstruction (time sequence reconstruction) of correlated sparse signals, if one signal is known a priori, remaining signals can be reconstructed using PKS based CR algorithms such as Modified Basis Pursuit (Mod-BP) or Weighted ℓ_1 -minimization ($W\ell_1$) with far fewer measurements compared to separate CS reconstruction [18], [21]. However, these methods are greatly influenced by the PKS that includes one or more atoms/locations that are irrelevant to the target signal. The key to the success of Mod-BP/ $W\ell_1$ is the accuracy of the PKS. Since a reliable PKS could improve the reconstruction performance further, there should be some provisioning to drop the irrelevant atoms from the PKS.

DCS is an extension of CS from single measurement vector problem to Multiple Measurement Vectors (MMV) problem. In DCS, several reconstruction algorithms have been proposed to reconstruct the joint-sparse signal ensemble from MMV [29]-[39]. However, most of them are designed for signal ensembles sharing common support [29]-[34]. Since the assumption of common sparsity pattern is very restrictive, focus should be on the signal ensemble containing both common and innovation components.

Many existing GPs (such as Orthogonal Matching Pursuit (OMP), Compressive Sampling Matching Pursuit (CoSaMP) and Iterative Hard Thresholding (IHT)) were modified so as to make use of the PKS to improve their performance, and are termed as PKS based GPs. The major drawback of these algorithms is their assumption that the PKS lies in-

side the actual support, which is very restrictive. All PKS based GPs showed improvement over their regular versions provided this assumption holds. Most PKS based GPs are totally blind to the wrong atoms present in the PKS which degrades their reconstruction performances. For example, algorithms such as OMP-PKS (i.e. OMP using PKS) or CoSaMP-PKS (i.e. CoSaMP using PKS) could not drop the wrong atoms (if any) present in the PKS and their final support-set estimate will contain those wrong atoms [25]. As the backtracking based GPs such as CoSaMP or Backtracking Matching Pursuit (BMP) have the provisioning of dropping atoms, they can be modified so as to utilize the PKS in an effective manner.

Although all PKS based CS reconstruction algorithms were proposed with a common motive of improving their regular versions, each of them adapts its own modification strategy. Therefore, if a new CS reconstruction algorithm evolves, a new modification procedure needs to be devised. This is the reason why many existing CS reconstruction algorithms do not have a PKS based modified version. Therefore, there is a need for a generalized, algorithm-blind framework capable of utilizing PKS and improving any CS reconstruction algorithm plugged into it.

Fusion based CS reconstruction algorithms combine multiple CS reconstruction algorithms, which work with different principles, to obtain a better signal estimate. Examples of such algorithms include Fusion of Algorithms for Compressed Sensing (FACS) [40] and Committee Machine Approach for Compressed Sensing (CoMACS) [41]. However, these algorithms involve solving a least squares problem which might be ill-conditioned. Motivated by the effectiveness of PKS based algorithms in CS reconstruction of sparse signals, we aim at proposing a fusion framework involving some GPs and a PKS based CS reconstruction algorithm.

Based on the discussions in this section, a generic PKS based CS reconstruction algorithm might work well in a particular scenario but fail in another. This thesis focuses on

developing PKS based reconstruction algorithms for different scenarios wherein they make effective use of the available PKS.

Figure 1.1: Context diagram of the thesis contributions

1.3 Contributions and Outline of the Thesis

This thesis proposes several PKS based reconstruction algorithms that handle the available PKS in different ways based on the signal model. Further, the reconstruction of sparse signals with no PKS is addressed by some proposed algorithms. A context diagram of the thesis contributions is given as figure 1.1. Chapter 2 of this thesis reviews the sensing

mechanism, RIP of sensing matrices, CS reconstruction algorithms, PKS based CS reconstruction algorithms and some applications of CS. Following are the contributions of this thesis.

In chapter 3, Modified Adaptive Basis Pursuit (MABP) and Weighted Adaptive ℓ_1 -minimization ($WA\ell_1$) algorithms are proposed wherein an adaptation process is carried out ahead of applying Mod-BP or $W\ell_1$. The purpose of this adaptation is to drop as many irrelevant locations (wrong atoms) as possible and reduce their impact on signal reconstruction. It will be theoretically shown that if in the adaptation step there are more correct deletions than wrong deletions, then MABP/ $WA\ell_1$ has a more favorable requirement compared to that of the Mod-BP/ $W\ell_1$. Experimental results show that the proposed MABP/ $WA\ell_1$ algorithms provide better signal reconstruction compared to that of the Mod-BP/ $W\ell_1$ algorithms.

In chapter 4, with a motive of making proper/effective use of the PKS in BMP, a PKS based adaptation mechanism is proposed which when replaces the traditional initialization of BMP results in a reconstruction accuracy higher than that of the existing PKS based greedy algorithms. The resulting algorithm is termed as Adaptive Backtracking Matching Pursuit (AdBMP). Experimental results show that, in recovering correlated sparse signals, the proposed AdBMP gives a better reconstruction performance compared to that of the existing PKS based greedy algorithms and the BMP itself. AdBMP is shown to be suitable for a variety of CS applications including sparse recovery of radar echo signals and time sequence reconstruction of video signals.

In chapter 5, a generalized, algorithm-blind reconstruction framework called Iterative Compressive Sensing on Residual (ICSR) is proposed, which can accommodate any class of CS reconstruction algorithm, and is capable of improving the signal reconstruction utilizing the available PKS. Experimental results show that, in Compressive Imaging (CI) application, the ICSR framework improves the reconstruction accuracy of any ingredient

CS reconstruction algorithm, and performs better than many of the existing PKS based algorithms. The ICSR framework can also be applied to reconstruction of sparse signals for which no PKS is available.

In chapter 6, a novel fusion technique called Greedy Pursuits Assisted Basis Pursuit (GPABP) is proposed which employs multiple GPs to derive a partial support (i.e. PKS) for Mod-BP. Experimental results show that the proposed GPABP performs better than FACS and CoMACS. Note that the proposed GPABP does not require a PKS and it is suitable for applications where no PKS is available. Next, a variant of GPABP, Greedy Pursuits assisted Weighted ℓ_1 -minimization (GP-W ℓ_1), is proposed. GP-W ℓ_1 adopts a novel weighting strategy. With a goal of proposing an MMV-type algorithm that is robust to outliers (innovation components), an MMV version of GPABP, termed as GPABP-MMV, is also proposed in chapter 6. GPABP-MMV outperforms several state-of-the-art DCS reconstruction algorithms.

Finally, chapter 7 concludes the thesis and discusses some promising directions for further research.

Chapter 2

Review of Compressive Sensing

Reconstruction Algorithms

The mathematical theory underlying CS is drawn from diverse fields including convex optimization, random matrix theory, and theoretical computer science [4]. This chapter reviews the sensing mechanism, RIP of sensing matrices, CS reconstruction algorithms and PKS based CS reconstruction algorithms. Some applications of CS are also discussed at the end of this chapter.

2.1 The Sensing Problem

CS ensures the reconstruction of a K -sparse signal x from its measurement vector y (of the form given in (1.1)). CS is an efficient data acquisition protocol for low-rate signal independent sampling. CS relies on two fundamental principles: sparsity and incoherence [3]. Sparsity pertains to the signals of interest. As described in section 1.1, many natural signals are sparse or compressible. Consider a compressible signal, for example, the one-dimensional signal in figure 2.1(a) and its one-dimensional Discrete Cosine Transform (DCT) in figure 2.1(b). Although most of the signal entries are non-zero, most of its DCT

coefficients are small. Since relatively few (large) coefficients contain most of the information, remaining (small) coefficients can be discarded without much perceptual loss.

Incoherence pertains to the sensing modality. As mentioned in section 1.1, in the case of signals having sparse representation in Ψ , the coherence between Φ and Ψ should be low. In general, random matrices (e.g. Gaussian random matrix) exhibit a very low coherence with any sparsifying basis [4]. It is well known that the coherence is directly proportional to the largest correlation between any two columns of Φ and Ψ . Therefore, for any fixed Ψ , it is possible to design an efficient Φ such that the correlation between any two columns of Φ and Ψ is low [46]. The fact that CS requires the signal of interest to be sparse and its measurement vector to be dense is analogous to the duality between time and frequency (e.g. a spike in the time domain is spread out in the frequency domain).

Figure 2.1: Non-sparse signal and its sparse representation in DCT domain

2.2 Compressive Sensing Matrices

Robustness of CS can be studied using the RIP of the sensing matrix. A sensing matrix Φ is said to follow RIP if there exists a Restricted Isometry Constant (RIC) $\delta_K \in [0, 1]$

satisfying

$$(1 - \delta_K)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_K)\|x\|_2^2 \quad (2.1)$$

for all K -sparse vectors $x \in \mathbb{R}^n$. RIP implies that all subsets of K columns of Φ will be nearly orthogonal to each other. The sensing matrix has more columns than rows and therefore, their columns cannot be exactly orthogonal. When RIP holds, the Euclidean length of the K -sparse signal is nearly preserved in Φ , which in turn implies that the signal cannot be in the null space of the sensing matrix. The sensing matrix Φ should be chosen such that arbitrary subsets of column vectors are nearly orthogonal. The randomness of Φ plays a key role in satisfying RIP given by (2.1).

A sensing matrix can be formed by sampling n column vectors uniformly at random on the unit sphere. Another way of forming sensing matrix is by sampling i.i.d. entries from the Gaussian distribution with zero mean and variance $\frac{1}{m}$, or by sampling i.i.d. entries from a symmetric Bernoulli distribution, or other sub-Gaussian distributions. All these random matrices obey the RIP with overwhelming probability provided the number of measurements

$$m \geq c_0 K \log\left(\frac{n}{K}\right) \quad (2.2)$$

where c_0 is some constant depending on each instance [2, 47]. The probability of these sampling matrices not obeying the RIP when (2.2) holds is exponentially small in m .

To realize the connection between the RIP and CS, consider acquiring K -sparse signals using Φ . Suppose that δ_{2K} is sufficiently less than one. This implies that Φ obeys RIP of order $2K$ and all pairwise distances between K -sparse signals must be well preserved in the measurement space. Therefore,

$$(1 - \delta_{2K})\|x_1 - x_2\|_{\ell_2}^2 \leq \|\Phi x_1 - \Phi x_2\|_{\ell_2}^2 \leq (1 + \delta_{2K})\|x_1 - x_2\|_{\ell_2}^2 \quad (2.3)$$

holds for all K -sparse vectors x_1 and x_2 . This guarantees the existence of efficient and robust

algorithms for discriminating K -sparse signals based on their compressive measurements.

2.3 Compressive Sensing Reconstruction Algorithms

For CS reconstruction of sparse signals, the optimal strategy is solving an ℓ_0 -minimization problem:

$$\hat{x} = \arg \min_{\tilde{x}} \|\tilde{x}\|_0 \text{ s.t. } \Phi \tilde{x} = y \quad (2.4)$$

Since the above problem is Non-deterministic Polynomial-time hard (NP-hard), for practical implementations, several sub-optimal algorithms are introduced. This section discusses few of those reconstruction algorithms that are relevant to this thesis.

2.3.1 Convex Relaxation Techniques

For a K -sparse signal x , exact reconstruction using a CR method such as BP is a polynomial time problem provided there are $m = O(K \log(n/K))$ measurements [2]. The problem is formulated as

$$\hat{x}_{bp} = \arg \min_{\tilde{x}} \|\tilde{x}\|_1 \text{ s.t. } \Phi \tilde{x} = y \quad (2.5)$$

where y is obtained as given in (1.1). Assume $\delta_{2K} < \sqrt{2} - 1$. Then the solution to (2.5) obeys

$$\|\hat{x}_{bp} - x\|_2 \leq c_1 \frac{\|\hat{x}_{bp} - x_K\|_1}{\sqrt{K}} \text{ and } \|\hat{x}_{bp} - x\|_1 \leq c_1 \|\hat{x}_{bp} - x_K\|_1 \quad (2.6)$$

for some constant c_1 , where x_K is the best K -term approximation of x [46]. If x is exactly sparse, then $x = x_K$ and thus, the reconstruction is exact. If x is not exactly K -sparse, the quality of reconstruction is nearly the same as that provided by an oracle estimator [4]. In most CS applications, y is corrupted by a noise such that

$$y = \Phi x + v \in \mathbb{R}^m \quad (2.7)$$

where v is the additive noise of variance σ^2 . For those applications, the BP problem is reformulated as

$$\hat{x}_{bp} = \arg \min_{\tilde{x}} \|\tilde{x}\|_1 \text{ s.t. } \|\Phi \tilde{x} - y\|_2 \leq \epsilon \quad (2.8)$$

where $\epsilon = \sigma$ (if the noise variance is known) or $\epsilon = \zeta$ (ζ is the error tolerance). Certain articles in literature mention the problem in (2.8) as Basis Pursuit DeNoising (BPDN) [48]. However, most articles state a different problem as BPDN [5, 19]. In this thesis, (2.8) is dubbed as a BP problem for the noisy measurements case whereas (2.5) is a BP problem for the clean measurements case. If $\delta_{2K} < \sqrt{2} - 1$, then the solution to (2.8) obeys

$$\|\hat{x}_{bp} - x\|_2 \leq c_1 \frac{\|\hat{x}_{bp} - x_K\|_1}{\sqrt{K}} + c_2 \epsilon \quad (2.9)$$

for some constants c_1 and c_2 [46, 49].

2.3.2 Greedy Pursuits

Although the CR techniques could solve the CS reconstruction problem, their computational cost is high enough to make them impractical for some CS applications. Hence GPs received significant attention because of their low computational cost and simple geometric interpretation [6]-[16]. A brief overview of some existing GPs is presented in this subsection. The measurement matrix, Φ , is considered as a dictionary and each of its column vectors is considered as an atom. GPs are iterative algorithms wherein each iteration involves two major steps: 1) *Atom selection step* in which one or more atoms having the highest correlations with the measurement residual is added to the estimated support set; 2) *Updation step* in which the approximate coefficients corresponding to the selected set is updated to minimize the residual error, and the updated residual is used for the next iteration.

The basic GP, the Matching Pursuit (MP) algorithm [6], provides a practical solution by

iteratively decomposing a measurement vector into a linear combination of atoms that are selected from a redundant dictionary. MP first finds the atom that has the highest correlation with the measurement vector. The correlated part is then subtracted from the measurement signal and then MP iterates this procedure on the resulting residual signal. Though MP guarantees asymptotic convergence, the resulting approximate coefficients after any finite number of iterations will be sub-optimal. The Orthogonal Matching Pursuit (OMP) algorithm [7], a variation of MP method, adds a least-squares minimization step to each MP iteration to obtain the best approximation over the atoms in the estimated support set. This revision significantly improves the behaviour of the MP algorithm resulting in minimum error with the set of all selected atoms. The main benefit of OMP over MP is the fact that it is guaranteed to converge in K iterations. All additional computations needed for the orthogonalization step of OMP may be performed recursively [50]. Several works analyzing the performance of OMP have been reported in literature [51]-[56].

The MP and OMP methods choose just one atom in each iteration, but some greedy algorithms like Stagewise Orthogonal Matching Pursuit (StOMP), Regularized Orthogonal Matching Pursuit (ROMP), Subspace Pursuit (SP) and CoSaMP choose several atoms at a time. In every iteration, several atoms of the dictionary (Φ) are selected using the correlation between the columns of Φ and the residual vector from the previous iteration. The selection step of StOMP forms a matched filter and identifies atoms with magnitudes exceeding a chosen threshold (hard thresholding) [8]. ROMP method applies a preliminary selection followed by a final selection to select the atoms, and adds them to the estimated support set [9]. The preliminary selection step chooses a set of atoms as in the case of StOMP method. The final selection step computes the energy for all possible subsets of the chosen set having comparable amplitudes and chooses the subset with the maximal energy. The selected subset is added to the estimated support set, and the residual is updated. ROMP is the first GP to provide strong guarantees similar to that of the CR techniques.

Backtracking based GPs such as the SP and CoSaMP algorithms select the atoms (to be added at each iteration) using a two-step selection technique: preliminary selection and final selection. In SP algorithm, the preliminary selection step of SP chooses a K -atom supplement set (atoms having the K largest correlations between the columns of Φ and the measurement residual), and a candidate set is created by the union of the chosen supplement set and the K -atom set obtained at the end of previous iteration [10]. The final selection step solves a least squares problem, chooses a subset of K atoms corresponding to K largest values in magnitude of the least square solution, and then updates the estimated support set. In CoSaMP [11], unlike SP, the preliminary selection step chooses a $2K$ -atom supplement set to create the candidate set, and the final selection step is similar to that of SP. RIP-based performance guarantees for SP and CoSaMP are reported in [57]

Based on SP and CoSaMP, several algorithms have been proposed to further improve the reconstruction accuracy. Sparsity Adaptive Matching Pursuit [13, 14], Backtracking based Adaptive Orthogonal Matching Pursuit (BAOMP) [15] and BMP [16] are notable among them. In BMP, unlike SP or CoSaMP, the preliminary selection step of BMP chooses an arbitrary-sized ($\lfloor \tau K \rfloor$) atom set to form the candidate set where τ is the selection parameter. The final selection is same as that of SP or CoSaMP. Hence, BMP can be considered as a generalization with SP and CoSaMP being special cases. It was shown that BMP with $\tau = 0.25$ results in a faster convergence [16]. Algorithm 1 shows the step-by-step procedure involved in the BMP algorithm.

2.3.3 Other Categories of Algorithms

In addition to GPs and CR methods, there are other categories of algorithms such as combinatorial methods and Non-Convex Relaxation (NCR) methods. Combinatorial methods reconstruct sparse signals following the principles of group testing [58]. They are fast and efficient than both CR methods and GPs. Popular examples include chaining pursuits

Algorithm 1 Backtracking Matching Pursuit

InputMeasurement matrix: Φ Measurement vector: y Sparsity level: K Preliminary selection parameter: τ **Initialization**

- 1) Initial support: $T^0 = \{\text{indices corresponding to the } K \text{ largest entries in } \Phi'y\}$.
- 2) Updated initial measurement residual: $y_r^0 = y - \Phi_{T^0}\Phi_{T^0}^\dagger y$.

BMP iterationsAt the i^{th} iteration do the following

- 3) $\tilde{T}^i = T^{i-1} \cup \{\text{indices corresponding to the } \lfloor \tau K \rfloor \text{ largest entries in } \Phi'y_r^{i-1}\}$.
- 4) $T^i = \{\text{indices corresponding to the } K \text{ largest entries in } \Phi_{\tilde{T}^i}^\dagger y\}$.
- 5) Updated measurement residual: $y_r^i = y - \Phi_{T^i}\Phi_{T^i}^\dagger y$.
- 6) If $\|y_r^i\|_2 > \|y_r^{i-1}\|_2$, assign $T^i = T^{i-1}$ and quit the iteration.

OutputSignal estimate: $\hat{x}^{(t)}$, such that $\hat{x}_{T^i}^{(t)} = \Phi_{T^i}^\dagger y$ and $\hat{x}_{T^{ic}}^{(t)} = 0$.

[59], Fourier sampling algorithm [60], and sudocodes [61]. Instead of relaxing ℓ_0 minimization problem to an ℓ_1 minimization problem, NCR methods solve ℓ_q -minimization problem where $0 < q < 1$. Another NCR strategy is to use a Bayesian framework. Popular examples of NCR methods include Re-Weighted ℓ_1 -minimization (ReW ℓ_1) [62], iterative support detection [63], and Sparse Bayesian Learning (SBL) [64]-[66].

2.4 Compressive Sensing Reconstruction Algorithms using Partially Known Support

Vaswani et al modified BP for problems where signal support is partially known in the form of PKS, say T_k . Note that $T_k \subset \{1, 2, \dots, n\}$ and it is the set of values of i where $x(i)$ is

expected to be large and hence expected to be a non-zero entry of x . A CS reconstruction procedure called Mod-BP was proposed in [18]. Mod-BP aimed at estimating the sparsest possible signal estimate outside T_k . The known part is either available from prior knowledge (as in static problems) or it can be estimated from the known signal in the ensemble (as in time sequence problems). [18] applied Mod-BP for time sequence reconstruction problem. The RIP requirements of Mod-BP was shown to be much weaker compared to that of the BP. The key to this algorithm is the accuracy of T_k .

Another BP based modified CS algorithm, Least Squares Compressive Sensing (LS-CS), used signal knowledge in a different manner compared to that of Mod-BP [20]. LS-CS replaced CS on the measurement vector by CS on the least squares measurement residual vector computed using T_k . Reason why LS-CS significantly outperformed regular CS was the signal minus its LS estimate contains much fewer non-zero elements than the signal itself. Though LS-CS is faster compared to Mod-BP, reconstruction accuracy of latter is better than that of the former.

In [21], $W\ell_1$ was proposed for reconstructing signals whose prior information is available in the form of T_k . The weighted ℓ_1 problem is formulated as

$$\hat{x} = \arg \min_{\tilde{x}} \|\tilde{x}\|_{1,w} \text{ s.t. } \Phi \tilde{x} = y. \quad (2.10)$$

where $w \in [0, 1]^n$ and $\|\tilde{x}\|_{1,w} := \sum_i w(i)|\tilde{x}(i)|$ is the weighted ℓ_1 norm with $w(i) = \omega \in [0, 1)$ whenever $i \in T_k$, and $w(i) = 1$ otherwise. The idea behind such an assignment is to choose w such that the entries of x that are expected to be large are penalized less in this weighted objective function. It was shown that, if at least 50% of the PKS is accurate, then $W\ell_1$ is stable and robust under weaker sufficient conditions than the analogous conditions for standard ℓ_1 -minimization.

Carrillo et al extended the modification idea to GPs resulting in Modified GPs (Mod-GPs) such as OMP-PKS [25], CoSaMP-PKS [25] and IHT-PKS [26]. Let us define an

algorithmic function

$$\hat{x} = \text{Mod-GP}(\Phi, y, \text{PKS}, K)$$

which denotes the general form of any Mod-GP algorithm. Mod-GPs made use of T_k to modify the initialization step and/or iterations of the GPs. Consider CoSaMP-PKS [25] which is basically a CoSaMP that exploits PKS (say T_k). During the preliminary selection step (at each iteration), a candidate set is created by the union of Ω and the $2(K - |T_k|)$ -sized supplement set. Final selection step chooses $K - |T_k|$ atoms that are not in T_k and preserves all components in T_k . Carrillo, then proposed IHT-PKS [26], an extension of IHT to include PKS. The major drawback of these algorithms is their assumption: $T_k \subset T$, which is very restrictive. All Mod-GPs showed improvement over their regular versions provided the PKS lies inside the actual support T . On the other hand, Mod-GPs are totally blind to the wrong atoms present in the PKS. Algorithms such as CoSaMP-PKS and IHT-PKS could not drop the wrong atoms (if any) in the PKS and their final support-set estimate will contain those wrong atoms. Since Mod-GPs made use of PKS to show improvement over their regular versions, they are also termed as PKS based GPs.

2.5 Distributed Compressive Sensing

DCS exploits both intra-signal and inter-signal correlation structures [28]. For a typical DCS scenario, imagine a microphone array recording a sound field at several points in space. The time-series measurements acquired at a given sensor have considerable temporal (intra-signal) correlation and have sparse representation in a local Fourier basis. Since all microphones listen to the same sources, the ensemble of measurements acquired at all sensors might have considerable spatial (inter-signal) correlation. Each sensor independently encodes its signal by projecting it onto another, incoherent, random basis and then transmits just a few of the resulting coefficients to the decoder. The decoder jointly recon-

structs all the signals precisely.

In [67], two JSMs have been introduced and specially tailored reconstruction algorithms were proposed. Consider a network of sensors measuring temperatures in a forest throughout the day. Factors like the sun and prevailing winds will have a global effect that is common to all sensors and has sparse representation in some basis. Factors such as shade and water will contribute localized innovations that are also structured (and hence sparse). This scenario can be modelled by JSM-1. For this signal model, there exists a measurement rate region analogous to the Slepian-Wolf rate region [68, 69]. Due to the joint sparsity of a signal ensemble, the savings in the required number of measurements can be substantial over separate CS reconstruction. If the common sparse component dominates, DCS witnesses dramatic savings in the number of measurements compared to that of separate CS encoding/decoding. In JSM-2, all K -sparse signals are constructed with arbitrarily different coefficients, but the support is the same. In sensor networks, multiple sensors acquire the same signal but with phase shifts and attenuations caused by signal propagation. In such cases, it becomes critical to recover each of the signals and this scenario can be modelled by JSM-2. This joint sparsity model is applicable to many real world scenarios including acoustic localization, array processing and multiple-input and multiple-output communication. Tight and achievable measurement rate bounds for JSM-2 were given in [70]. The number of required measurements (per sensor) for exact reconstruction reduces with the increase in the number of sensors.

The data matrix for an ensemble of J signals is denoted as follows,

$$\mathbf{Y} = \Phi \mathbf{X} + \mathbf{V} \in \mathbb{R}^{J \times m} \quad (2.11)$$

where $\mathbf{X} = [x_1, x_2, \dots, x_J] \in \mathbb{R}^{J \times n}$, $\mathbf{V} = [v_1, v_2, \dots, v_J] \in \mathbb{R}^{J \times m}$ and $\mathbf{Y} = [y_1, y_2, \dots, y_J]$.

2.6 Fusion based Compressive Sensing Reconstruction Algorithms

In this section, FACS and its variants are discussed. FACS employed multiple CS reconstruction algorithms and fused their resultant signal estimates to obtain a better signal estimate [40, 41, 71, 72, 73]. As T denotes the actual support, let $\hat{T}_i$ be the support estimated by i^{th} algorithm. Notice that $|\hat{T}_i| = K$. Let the union of the estimated support sets be denoted as $\tilde{T}_J$ (*joint support*) and the intersection of the estimated support sets be denoted as $\tilde{T}_C$ (*common support*). FACS showed that the probability of estimating more correct atoms from the union set is higher than that individually estimated by ingredient algorithms [40]. Assume BP, OMP and SP are involved in the fusion framework, and BP is the best performing algorithm. Then, the performance of FACS will be better than that of the BP. Algorithm 2 shows the step-by-step procedure in FACS involving three CS algorithms (BP and two GPs).

A modification to FACS, M-FACS proposed in [73], included a step prior to least squares to ensure the least squares problem is well-conditioned. It kept removing the index of the smallest coefficient of $\Phi_\Gamma^\dagger y$ from the set Γ until the 2-norm condition number of Φ_Γ falls below a predefined condition number threshold. M-FACS gave better reconstruction performance than FACS in the low dimension measurement regime. Another variant of FACS, the CoMACS algorithm [41], used both $\tilde{T}_J$ and $\tilde{T}_C$ to obtain $\hat{T}$. It had $\tilde{T}_C \subset \hat{T}$. Then it estimated the remaining $K - |\tilde{T}_C|$ indices in the same manner as in steps 5 and 6 of algorithm 2. The main assumption for all these fusion based CS algorithms is that $|\tilde{T}_J| \leq m$. However, these algorithms involve solving a least squares problem which may be ill-conditioned.

Algorithm 2 Fusion of Algorithms for Compressive Sensing

InputMeasurement matrix: Φ Measurement vector: y Sparsity level: K **Procedure**Step 1: $\hat{T}_1 = \text{GP}_1(\Phi, y, K)$ Step 2: $\hat{T}_2 = \text{GP}_2(\Phi, y, K)$ Step 3: $\hat{T}_3 = \text{BP}(\Phi, y)$ Step 4: *Joint support* $\tilde{T}_J = \hat{T}_1 \cup \hat{T}_2 \cup \hat{T}_3$ Step 5: $\check{x}$, such that $\check{x}_\Gamma = \Phi_{\tilde{T}_J}^\dagger y$ and $\check{x}_{\tilde{T}_J^c} = 0$.Step 6: $\hat{T} =$ indices corresponding to the K largest magnitude entries in $\check{x}$.Step 7: $\hat{x}$, such that $\hat{x}_{\hat{T}} = \check{x}_{\hat{T}}$ and $\hat{x}_{\hat{T}^c} = 0$.**Output** $\hat{x}$: Signal estimate

2.7 Compressive Sensing Applications

In data compression, the sparse basis may be unknown at the encoder or impractical to implement. However, a randomly designed sensing matrix can be considered for a universal encoding strategy, as it is highly incoherent with any sparsifying basis. The knowledge of the sparse basis is needed only for the decoding/recovery of the signal. This property of CS makes it particularly helpful for applications such as CI [74]. CI involves transforming the image into an appropriate basis and then coding only the important expansion coefficients. In general, the interest lies in sensing/reconstructing a signal that is sparse in some known transform Ψ (say, DCT). At the encoder, the sensing mechanism ((i.e., CS acquisition as in (1.1)) can be applied directly on the compressible signal and the resulting CS measurement vector is transmitted to the decoder. At the decoder, the sparse signal can be reconstructed from the CS measurement vector and Φ using a CS reconstruction algorithm provided $\Phi\Psi$

satisfies RIP with a high probability [74]. Actual non-sparse signal can be obtained from the inverse transform of the reconstructed sparse signal.

Fast error correction codes can be designed by applying CS principles like sparsity and randomness [47]. A single pixel camera designed based on a digital micro-mirror device randomly modulates the light from the scene to compute random CS measurements [75]. In Magnetic Resonance (MR) setups such as MR angiography, CS reduces the scan time and provides sharper images of living tissues [76]. In data acquisition, CS principles could be applied to design a sampling device that can record discrete, incoherent, low-rate measurements of the incident analog signal. Structured CS matrices can be employed in certain application areas such as identification of a Linear Time-Invariant (LTI) system [77, 78]. Haupt et al described a channel estimation problem where Toeplitz or circulant CS matrices are used to identify the impulse response of a LTI system [77]. Circulant CS matrices can be employed in video processing applications such as camera motion estimation [79] and compressive coded video compression [80].

Though the algorithms proposed in this thesis are suitable for many practical scenarios, focuss will be on these four applications: video reconstruction (frame/object based), image reconstruction, radar echo signal recovery and ECG signal reconstruction.

2.8 Summary

Several CS reconstruction algorithms are briefly described in this chapter. PKS based CS reconstruction algorithms such as Mod-BP, CoSaMP-PKS and IHT-PKS are discussed in detail. Fusion based CS reconstruction algorithms such as FACS and CoMACS are also described. In the following chapters, the algorithms discussed in this chapter will be extensively used.

Chapter 3

Modified Adaptive Basis Pursuit

In DCS reconstruction of correlated sparse signals, if one signal is known apriori, remaining signals can be reconstructed using PKS based convex relaxation algorithms (such as Mod-BP or $W\ell_1$) with far fewer measurements compared to separate CS reconstruction. However, these methods are greatly influenced by the support set of the known signal that includes one or more locations that are irrelevant to the target signal. In this chapter, MABP and $WA\ell_1$ algorithms are proposed wherein an adaptation process is carried out ahead of applying Mod-BP or $W\ell_1$. The purpose of adaptation is to drop as many irrelevant locations (wrong atoms) as possible and reduce their impact on signal reconstruction. It will be theoretically shown that if in the adaptation step there are more correct deletions than wrong deletions, then MABP/ $WA\ell_1$ has a weaker RIP requirement compared to that of the Mod-BP/ $W\ell_1$. Experimental results show that the proposed MABP/ $WA\ell_1$ algorithms provide better signal recovery compared to that of the Mod-BP/ $W\ell_1$ algorithms.

3.1 Shortcomings of Modified Basis Pursuit

For problems with signal support partially known in the form of T_k (i.e. reconstruction of x from y using T_k), Vaswani et al proposed Mod-BP [18]. The Mod-BP problem (ℓ_1 version)

is formulated as

$$\hat{x}_{modbp} = \arg \min_{\tilde{x}} \|\tilde{x}_{T_k^c}\|_1 \text{ s.t. } \Phi \tilde{x} = y \quad (3.1)$$

where $T_k^c := \{1, 2, \dots, n\} \setminus T_k$ (i.e., T_k^c is the set complement of T_k). An ℓ_0 version of the above problem is also available. The support of x is denoted as T and it can be split as $T := T_k \cup T_u \setminus T_e$ where $T_u := T \setminus T_k$ is the unknown part to be estimated and $T_e := T_k \setminus T$ is the error in T_k . A schematic describing T , T_k , T_u , and T_e is given in figure 3.1. For problems where y is corrupted by v such that $y = \Phi x + v$ and $\|v\|_2 \leq \sigma$, the Mod-BP problem is restated as follows:

$$\hat{x}_{modbp} = \arg \min_{\tilde{x}} \|\tilde{x}_{T_k^c}\|_1 \text{ s.t. } \|\Phi \tilde{x} - y\|_2 \leq \epsilon \quad (3.2)$$

where $\epsilon = \sigma$ (if σ is known) or $\epsilon = \zeta$ (ζ is the error tolerance).

Figure 3.1: Schematic of actual support (T) and PKS (T_k)

In DCS applications, each signal in the ensemble is independently sensed using Φ , and then the resulting measurement vectors are transmitted to the decoder. As the signals are sparse, one could reconstruct each of them separately using a CS reconstruction algorithm. If one signal is known apriori, motivated by the idea of using prior information, the remaining signals in the ensemble can be reconstructed using modified CS algorithms with far fewer measurements compared to separate CS reconstruction. In the correlated sparse signal ensemble (JSM-1 analyzed in [67]), each signal consists of two components: a com-

mon sparse component that is present in all the signals and a sparse innovation component that is unique to each signal. Let us denote the J signals in the ensemble by $x^{(t)} \in \mathbb{R}^n$, $t \in 1, 2, \dots, J$. Let $\Phi \in \mathbb{R}^{m \times n}$ be the i.i.d. Gaussian measurement matrix for signal $x^{(t)}$ such that

$$y^{(t)} = \Phi x^{(t)} \quad (3.3)$$

gives $m \ll n$ incoherent measurements of $x^{(t)}$. At the decoder, every n -length sparse signal $x^{(t)}$ is reconstructed from its m -length measurement vector. Vaswani (in [18]) applied Mod-BP for time sequence reconstruction problem where $y \equiv y^{(t)}$ and $x \equiv x^{(t)}$ with support $T \equiv T^{(t)}$. As mentioned earlier, the support of $x^{(t)}$ can be expressed as $T = T_k \cup T_u \setminus T_e$. The error in the known part, T_e , comes from those non-zero locations in $x^{(t-1)}$ that becomes zero in $x^{(t)}$. The sparsity patterns change slowly and therefore, $|T_u|$ and $|T_e|$ are assumed to be much smaller than $|T_k|$. In Mod-BP based time sequence reconstruction of correlated sparse signals, at time $t = 1$, $\hat{x}_{modbp}^{(1)}$ is obtained using (3.1) or (3.2) with $T_k = \emptyset$. For the subsequent signals, the PKS for obtaining $\hat{x}_{modbp}^{(t)}$ via (3.1) or (3.2) is formed as follows,

$$T_k = \{i = [1 : n] : |\hat{x}_{modbp}^{(t-1)}(i)| > \mu_0\} \quad (3.4)$$

where μ_0 is the initial selection parameter.

Reconstruction of correlated sparse signals via Mod-BP using T_k has a drawback: The PKS, T_k , not only contains a part of correct support but also wrong locations. Solving (3.1) or (3.2) might lead to complication and inaccuracy. To overcome this drawback, for each signal $x^{(t)}$, prior to Mod-BP, an adaptation step is included which tries to retain only those locations in $T_k := \hat{T}^{(t-1)}$ that are relevant to the target signal $x^{(t)}$. In other words, the adaptation step drops/prunes those atoms that are irrelevant to $x^{(t)}$ in order to improve the accuracy of the signal approximation (via Mod-BP). The resulting algorithm is termed as Modified Adaptive Basis Pursuit (MABP). *The main contribution of this chapter is to show*

that if in the adaptation step there are more correct deletions than wrong deletions, then MABP has a weaker (more favorable) RIP requirement compared to that of the Mod-BP. In other words, if at least half of the deleted atoms are wrong atoms then the MABP gives better reconstruction compared to that of the Mod-BP.

3.2 Proposed Modified Adaptive Basis Pursuit

Figure 3.2: Block schematic of MABP

Figure 3.2 shows the schematic of MABP. After estimating T_k , an adaptation procedure follows: First, the approximate coefficients ($\tilde{x}_{T_k} = \Phi_{T_k}^\dagger y$) are computed and the atoms with approximate coefficients smaller than $\mu_a \|\tilde{x}_{T_k}\|_\infty$ are deleted in order to initialize the support set more effectively. Note that μ_a is the adaptation parameter controlling the number of atoms to be deleted. Since T_k obeys $T := T_k \cup T_u \setminus T_e$, after adaptation, the retained support set $\Gamma_k (\subset T_k)$ will be such that $T := \Gamma_k \cup T'_u \setminus T'_e$ where $T'_e := \Gamma_k \setminus T$ is the error in Γ_k and $T'_u := T \setminus \Gamma_k$ is the unknown part to be estimated. Note that $|T| = |T_k| + |T_u| - |T_e| = |\Gamma_k| + |T'_u| - |T'_e|$. A schematic describing T , T_k , Γ_k , T'_u , and T'_e is given in figure 3.3.

The MABP problem (ℓ_0 version) is formulated as,

$$\hat{x}_{mabp} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Gamma_k^c}\|_0 \text{ s.t. } \Phi \tilde{x} = y \quad (3.5)$$

where $\Gamma_k^c := [1 : n] \setminus \Gamma_k$ is the complement of Γ_k . As the minimization (3.5) is NP-hard, the

Figure 3.3: Schematic of actual support (T), PKS (T_k) and the retained support Γ_k

ℓ_1 version of the problem is stated as

$$\hat{x}_{mabp} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Gamma_k^c}\|_1 \text{ s.t. } \Phi \tilde{x} = y. \quad (3.6)$$

For problems where y is corrupted by v and $\|v\|_2 \leq \sigma$, the MABP problem (in (3.6)) is restated as follows:

$$\hat{x}_{mabp} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Gamma_k^c}\|_1 \text{ s.t. } \|\Phi \tilde{x} - y\|_2 \leq \epsilon \quad (3.7)$$

where $\epsilon = \sigma$ or ζ .

Algorithm 3 shows the step-by-step procedure in MABP based reconstruction of correlated sparse signals. There are some intuitions for proposing this kind of modifications to the Mod-BP. In the known support estimation step, when μ_0 is small, it chooses more atoms. Since a refining step follows, a coarse selection of known support will not affect the reconstruction performance. Care should be taken while deciding the adaptation parameter. If μ_a is too close to zero, then MABP behaves like Mod-BP. On the other hand, if μ_a is too close to one, it behaves like traditional BP. The adaptation procedure has a signal approximation step requiring $O(|T_k|^2 m)$ computations which is insignificant compared to the complexity of BP [81, 82]. Therefore, inclusion of adaptation step will not affect the overall complexity of MABP. Note that the algorithm 3 does not need to know the sparsity level K .

Algorithm 3 Modified Adaptive Basis Pursuit

InputMeasurement matrix: Φ Measurement vector: $y^{(t)}$ Previous reconstructed signal: $\hat{x}^{(t-1)}$ Initial selection parameter: μ_0 Adaptation parameter: μ_a **Initialization**

Estimate the known support as

$$T_k = \{i = [1 : n] : |\hat{x}^{(t-1)}(i)|^2 > \mu_0\}.$$

AdaptationStep 1: The approximate coefficients ($\tilde{x}_{T_k} = \Phi_{T_k}^\dagger y^{(t)}$) are computed.Step 2: $\Gamma_k = T_k - \{\text{indices corresponding to coefficients with magnitude smaller than } \mu_a \|\tilde{x}_{T_k}\|_\infty\}$.**Modified Basis Pursuit**Solve the Mod-BP problem to obtain $\hat{x}^{(t)}$:

$$\hat{x}^{(t)} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Gamma_k^c}\|_1 \text{ s.t. } \Phi \tilde{x} = y^{(t)}$$

Output $\hat{x}^{(t)}$: Signal estimate

3.3 Comparison of MABP and Mod-BP

In this section, Mod-BP and MABP are compared based on three criteria: sufficient conditions for exact reconstruction, high probability upper bounds on RIP constant and reconstruction error bound. Let $K = |T|$ denotes the size of the actual support (also known as signal sparsity level), $k = |T_k|$ denotes the size of the PKS, $e = |T_e|$ denotes the size of the error in the PKS and $u = |T_u|$ denotes the size of the unknown part of the support. Similarly, $k' = |\Gamma_k|$ denotes the size of the retained support, $e' = |T_e'|$ denotes the size of the error in the retained support and $u' = |T_u'|$ denotes the size of the unknown part of the support (after pruning/adaptation).

3.3.1 Comparison based on Sufficient Conditions for Exact Reconstruction

In this subsection, first, the exact reconstruction conditions are given for the ℓ_0 and ℓ_1 versions of MABP. Then, the conditions for MABP to outperform Mod-BP are derived.

Theorem 3.1: Let x be a sparse vector with support, $T = \Gamma_k \cup T_{u'} \setminus T_{e'}$, where Γ_k and $T_{u'}$ are disjoint and $T_{e'} \subset \Gamma_k$. Consider reconstructing it from $y = \Phi x$ by solving (3.5). $\hat{x}_{mabp}$ is the unique minimizer of (3.5) if

$$\delta_{k'+2u'} < 1. \quad (3.8)$$

Proof of this theorem can be drawn in similar lines as that of the proof of proposition 1 in [18].

Theorem 3.2: Let x be a sparse vector with support, $T = \Gamma_k \cup T_{u'} \setminus T_{e'}$, where Γ_k and $T_{u'}$ are disjoint and $T_{e'} \subset \Gamma_k$. Consider reconstructing it from $y = \Phi x$ by solving (3.6). $\hat{x}_{mabp}$ is the unique minimizer of (3.6) if

$$\delta_{\max\{k'+2u', 3u'\}} < \frac{1}{5}. \quad (3.9)$$

Proof: From [18], $\hat{x}_{modbp}$ is the unique minimizer of (3.1) if

$$2\delta_{2u} + \delta_{3u} + \delta_k + \delta_{k+u}^2 + 2\delta_{k+2u}^2 < 1. \quad (3.10)$$

This, in turn, holds if

$$u \leq k \text{ and } \delta_{k+2u} < \frac{1}{5}.$$

These conditions can be rewritten for MABP by replacing u and k with u' and k' respec-

tively. Therefore, $\hat{x}_{mabp}$ is the unique minimizer of (3.6) if

$$u' \leq k' \text{ and } \delta_{k'+2u'} < \frac{1}{5}.$$

These two conditions can be combined as one condition: $\delta_{\max\{k'+2u', 3u'\}} < \frac{1}{5}$. $\square$

Two parameters, γ and ρ , are defined as follows,

$$\gamma = \frac{k}{K} \text{ and } \rho = \frac{|T \cap T_k|}{K}. \quad (3.11)$$

Using the definitions in (3.11), k , u , and e are expressed as $k = \gamma K$, $u = (1 - \rho)K$ and $e = (\gamma - \rho)K$ respectively. Note that ρ denotes the fraction of the actual support present in T_k . Two more parameters, γ' and α are defined as follows,

$$\gamma' = \frac{k'}{K} \text{ and } \alpha = \frac{|T \cap T_k| - |T \cap \Gamma_k|}{k - k'}. \quad (3.12)$$

Using the definitions in (3.12), k' , u' , and e' are expressed as $k' = \gamma' K$, $u' = (1 - \rho)K + \alpha(\gamma - \gamma')K$ and $e' = (\gamma - \rho)K - (1 - \alpha)(\gamma - \gamma')K$ respectively. Note that α denotes the fraction of the falsely deleted atoms in $T_k \setminus \Gamma_k$. The condition on α for which the ℓ_0 version of MABP has a weaker RIP requirement compared to that of the ℓ_0 version of Mod-BP is derived. For MABP, the sufficient condition is given in (3.8). Using $k' = K + e' - u'$, the condition becomes

$$\delta_{K+e'+u'} < 1. \quad (3.13)$$

Using the definitions of e' and u' , and the fact that $K = \frac{u}{1-\rho}$, the above condition becomes

$$\delta \left[\frac{2(1-\rho)+2\alpha(\gamma-\gamma')+\gamma'}{1-\rho} \right] u < 1. \quad (3.14)$$

Using the relation

$$\delta_{cr} \leq c\delta_{2r}, \quad (3.15)$$

the condition in (3.14) can be written as

$$\frac{2(1-\rho) + 2\alpha(\gamma - \gamma') + \gamma'}{1-\rho} \delta_{2u} < 1 \quad (3.16)$$

or

$$\delta_{2u} < \frac{1-\rho}{2(1-\rho) + 2\alpha(\gamma - \gamma') + \gamma'}. \quad (3.17)$$

For Mod-BP, the sufficient condition is $\delta_{K+e+u} < 1$. Using the definitions of e and u , this condition becomes

$$\delta \left[\frac{2(1-\rho)+\gamma}{1-\rho} \right]_u < 1 \quad (3.18)$$

which can also be written as

$$\delta_{2u} < \frac{1-\rho}{2(1-\rho) + \gamma}. \quad (3.19)$$

Comparing (3.17) and (3.19), the ℓ_0 version of MABP will have a weaker RIP requirement than that of its Mod-BP counterpart if

$$\gamma' + 2\alpha(\gamma - \gamma') < \gamma \quad (3.20)$$

which holds if

$$\alpha < \frac{1}{2}. \quad (3.21)$$

Consider a typical time sequence reconstruction with $\gamma = 1$, $\gamma' = 0.5$ and $\rho = 0.75$. The sufficient condition for Mod-BP, (3.19), becomes $\delta_{2u} < \frac{1}{6}$. On the other hand, the condition for MABP, (3.17), becomes $\delta_{2u} < \frac{0.25}{1+\alpha}$ which is clearly weaker (i.e. more favorable) when $\alpha < 0.5$. Fixing $\alpha = 0.25$, the condition for MABP becomes $\delta_{2u} < \frac{1}{5}$. Increasing α from 0.25 to 0.5 will tighten the condition from $\delta_{2u} < \frac{1}{5}$ to $\delta_{2u} < \frac{1}{6}$.

Next, the conditions on ρ and α for which the ℓ_1 version of MABP has a weaker RIP requirement compared to that of the ℓ_1 version of Mod-BP is derived. For MABP, the sufficient condition is $\delta_{\max\{k'+2u',3u'\}} < \frac{1}{5}$. Using (3.15) and the definitions of k' and u' , this condition becomes

$$\delta_{2u} < \frac{1-\rho}{5\{f(\rho, \gamma, \gamma', \alpha)\}} \quad \text{where} \quad (3.22)$$

$$f(\rho, \gamma, \gamma', \alpha) = \max \left\{ 2(1-\rho) + 2\alpha(\gamma - \gamma') + \gamma', 3(1-\rho) + 3\alpha(\gamma - \gamma') \right\}.$$

If $\rho > (1 - \gamma') + \alpha(\gamma - \gamma')$, (3.22) can be expressed as

$$\delta_{2u} < \frac{1-\rho}{5\{2(1-\rho) + 2\alpha(\gamma - \gamma') + \gamma'\}}. \quad (3.23)$$

For Mod-BP, the sufficient condition is $\delta_{\max\{k+2u,3u\}} < \frac{1}{5}$. Using (3.15) and the definitions of k and u , this condition becomes

$$\delta_{2u} < \frac{1-\rho}{5\{\max\{2(1-\rho) + \gamma, 3(1-\rho)\}\}}. \quad (3.24)$$

If $\rho > (1 - \gamma') + \alpha(\gamma - \gamma')$, due to the fact that $\rho > (1 - \gamma') + \alpha(\gamma - \gamma')$ is inclusive of $\rho > 1 - \gamma$, the sufficient condition will be

$$\delta_{2u} < \frac{1-\rho}{5\{2(1-\rho) + \gamma\}}. \quad (3.25)$$

Therefore, provided ρ satisfies $\rho > (1-\gamma')+\alpha(\gamma-\gamma')$, comparing (3.23) and (3.25), the ℓ_1 version of MABP will have a weaker RIP requirement than that of its Mod-BP counterpart if

$$\alpha < \frac{1}{2}. \quad (3.26)$$

In other words, if in the adaptation step there are more correct deletions than wrong deletions, then MABP has a weaker (more favorable) RIP requirement compared to that of the

Mod-BP. Continuing with the same time sequence reconstruction example for comparing (3.23) and (3.25), condition (3.25) becomes $\delta_{2u} < \frac{1}{30}$. On the other hand, condition (3.23) becomes $\delta_{2u} < \frac{1}{25}$ and $\delta_{2u} < \frac{1}{30}$ for $\alpha = 0.25$ and $\alpha = 0.5$ respectively.

For the limiting case of γ where $0 < \gamma \leq 1$, the condition on ρ becomes

$$\rho > (1 - \gamma')(1 + \alpha). \quad (3.27)$$

Invoking $\alpha < \frac{1}{2}$, the condition on ρ becomes

$$\rho > \frac{3}{2}(1 - \gamma'), \quad (3.28)$$

which holds if

$$(1 - \gamma') < \frac{2}{3}\rho. \quad (3.29)$$

Therefore,

$$(\gamma - \gamma') \leq (1 - \gamma') < \frac{2}{3}\rho \quad (\text{or}) \quad (3.30)$$

$$|T_k \setminus \Gamma_k| = (\gamma - \gamma')K < \frac{2}{3}\rho K = \frac{2}{3}|T \cap T_k|. \quad (3.31)$$

The above expression relates the number of pruned atoms to the accuracy of the PKS. For example, if $|T \cap T_k| = \frac{3}{4}K$, the number of pruned atoms should be less than $\frac{K}{2}$.

3.3.2 Comparison based on the High Probability Upper Bound on the RIC

Next, MABP and Mod-BP are compared based on the high probability upper bounds on the RIC. As given in [47], for an $m \times n$ Gaussian random measurement matrix, with high probability (w.h.p.),

$$\delta_S < g_{n/m}\left(\frac{S}{n}\right),$$

where

$$g_{n/m}\left(\frac{S}{n}\right) := -1 + \left[1 + f'\left(\frac{S}{n}, \frac{n}{m}\right)\right]^2,$$

$$f'\left(\frac{S}{n}, \frac{n}{m}\right) := \sqrt{\frac{n}{m}}\left(\sqrt{\frac{S}{n}} + \sqrt{2H\left(\frac{S}{n}\right)}\right)$$

and binary entropy $H(r) := -r \log r - (1-r) \log (1-r)$ for $0 \leq r \leq 1$. Thus, w.h.p., MABP achieves exact reconstruction from Gaussian random measurements if the high probability upper bound on the RIC (denoted as η_{mabp}) obeys

$$\begin{aligned} \eta_{mabp} = & 2g_{n/m}\left(\frac{2u'}{n}\right) + g_{n/m}\left(\frac{3u'}{n}\right) + g_{n/m}\left(\frac{K+e'-u'}{n}\right) \\ & + \left[g_{n/m}\left(\frac{K+e'}{n}\right)\right]^2 + 2\left[g_{n/m}\left(\frac{K+e'+u'}{n}\right)\right]^2 < 1. \end{aligned} \quad (3.32)$$

Similarly, for Mod-BP, exact reconstruction is possible from Gaussian random measurements if the high probability upper bound on the RIC (denoted as η_{modbp}) obeys

$$\begin{aligned} \eta_{modbp} = & 2g_{n/m}\left(\frac{2u}{n}\right) + g_{n/m}\left(\frac{3u}{n}\right) + g_{n/m}\left(\frac{K+e-u}{n}\right) \\ & + \left[g_{n/m}\left(\frac{K+e-u}{n}\right)\right]^2 + 2\left[g_{n/m}\left(\frac{K+e+u}{n}\right)\right]^2 < 1. \end{aligned} \quad (3.33)$$

Figures 3.4 and 3.5 show the variation of the high probability upper bound on the RIC (η_{modbp} and η_{mabp}) against the Sparsity Ratio ($SR = \frac{K}{n}$) for different choices of Measurement Ratio ($MR = \frac{m}{n}$). The same γ , γ' and ρ that are used in the previous subsection (time sequence reconstruction example) are used here. Figure 3.4 and figure 3.5 correspond to $\alpha = 0.05$ and $\alpha = 0.25$ respectively. As can be seen in figure 3.4, for a small α , for any given MR, w.h.p., MABP gives a better bound on RIP constant compared to that of Mod-BP. For a higher α (as shown in figure 3.5), there is no significant difference between η_{mabp} and η_{modbp} . It is worth noting that, due to the looseness of the bounds, the maximum allowed SR is much smaller than what we observed in simulations. For the same reason, the difference between η_{mabp} and η_{modbp} is not much significant.

Figure 3.4: High probability upper bound on the RIC versus SR for $\alpha = 0.05$

Figure 3.5: High probability upper bound on the RIC versus SR for $\alpha = 0.25$

3.3.3 Comparison based on Reconstruction Error Bound

Finally, MABP and Mod-BP are compared based on their reconstruction error bounds, in the case of noisy measurements. Assume that the measurement vector y is corrupted by a noise v such that $y = \Phi x + v$. Consider reconstructing x from y by solving (3.2). As reported in [83], if $\|v\|_2 \leq \epsilon$ and $\delta_{K+u+e} < \sqrt{2} - 1$, then

$$\|x - \hat{x}_{modbp}\|_2 \leq A(K + u + e)\epsilon, \quad \text{where} \quad (3.34)$$

$$A(S) \triangleq \frac{4\sqrt{1+\delta_S}}{1 - (\sqrt{2} + 1)\delta_S}.$$

Similarly, in reconstruction of x from y by solving (3.7), the error bound for MABP can be expressed as

$$\|x - \hat{x}_{mabp}\|_2 \leq A(K + u' + e')\epsilon. \quad (3.35)$$

Comparing (3.34) and (3.35), owing to the fact that $A(S)$ is an increasing function in S , MABP's error bound will be smaller than that of the Mod-BP if $K + u' + e' < K + u + e$ which was shown to hold if $\alpha < \frac{1}{2}$. In other words, MABP will certainly have a smaller reconstruction error bound compared to that of the Mod-BP if $\alpha < \frac{1}{2}$. For the time sequence reconstruction example discussed in previous subsections, the error bound for Mod-BP is $\|x - \hat{x}_{modbp}\|_2 \leq A(6u)\epsilon$. For the same, MABP's error bounds are $\|x - \hat{x}_{mabp}\|_2 \leq A(4.2u)\epsilon$ and $\|x - \hat{x}_{mabp}\|_2 \leq A(5u)\epsilon$ for $\alpha = 0.05$ and $\alpha = 0.25$ respectively.

3.4 Proposed Weighted Adaptive ℓ_1 -minimization

As in the case of Mod-BP, $W\ell_1$ is also influenced by the PKS. Motivated by the effectiveness of adaptation/prunning step in MABP, the idea is applied to $W\ell_1$, resulting in Weighted

Adaptive ℓ_1 (WAL_1). The WAL_1 problem is formulated as

$$\hat{x}_{\text{wal1}} = \arg \min_{\tilde{x}} \|\tilde{x}\|_{1,w} \text{ s.t. } \Phi \tilde{x} = y. \quad (3.36)$$

where $w \in [0, 1]^n$ and $\|\tilde{x}\|_{1,w} := \sum_i w(i)|\tilde{x}(i)|$ is the weighted ℓ_1 norm with $w(i) = \omega \in [0, 1]$ whenever $i \in \Gamma_k$, and $w(i) = 1$ otherwise.

Next, the condition for which the WAL_1 has a weaker RIP requirement compared to that of the $\text{W}\ell_1$ is derived. For $\text{W}\ell_1$ [21], the sufficient condition that guarantees stable and robust recovery is

$$\delta_{2K} < \frac{1}{\sqrt{2}(\omega + (1 - \omega)\sqrt{1 + \gamma - 2\rho}) + 1}. \quad (3.37)$$

It is known that $\rho = \frac{|T \cap T_k|}{K}$. Similarly, for WAL_1 ,

$$\delta_{2K} < \frac{1}{\sqrt{2}(\omega + (1 - \omega)\sqrt{1 + \gamma' - 2\rho'}) + 1} \quad (3.38)$$

where $\rho' = \frac{|T \cap \Gamma_k|}{K}$, guarantees stable and robust recovery. On comparing the above two conditions, WAL_1 will have a weaker RIP requirement than that of $\text{W}\ell_1$ if

$$\gamma' - 2\rho' < \gamma - 2\rho.$$

As per the definitions of γ , γ' , ρ , and ρ' , the above condition can be written as

$$\begin{aligned} \frac{k'}{K} - 2\frac{|T \cap \Gamma_k|}{K} &< \frac{k}{K} - 2\frac{|T \cap T_k|}{K} \\ 2(|T \cap T_k| - |T \cap \Gamma_k|) &< k - k' \\ \frac{|T \cap T_k| - |T \cap \Gamma_k|}{k - k'} &< \frac{1}{2} \end{aligned}$$

or

$$\alpha < \frac{1}{2}. \quad (3.39)$$

The above condition for $W\ell_1$ (to perform better than $W\ell_1$) is the same as that for the MABP (to perform better than Mod-BP). Therefore, the Mod-BP problem in algorithm 3 could well be replaced with the $W\ell_1$ problem.

3.5 Reconstruction of Correlated Sparse Signals

In this section, the reconstruction performances of the proposed techniques (MABP and $W\ell_1$) are presented with appropriate comparisons. Signal reconstruction is performed using Mod-BP [18], $W\ell_1$ [21], MABP and $W\ell_1$. The *cvx* solver [84] is used for the implementation of all four techniques.

3.5.1 Synthetic JSM-1 Signals

Four correlated sparse signals (Gaussian random signals as in JSM-1) each of length $n = 250$ are generated. The total sparsity of each signal in the ensemble is fixed as $K = 20$ (i.e., the common sparsity (K_c) and the innovation sparsity ($K - K_c$) adding up to 20). In each of the following experiments (presented in this subsection), 250 independent trials are conducted to obtain the average results. In each trial, an $m \times n$ Gaussian random matrix (Φ) is generated for sensing. y is noiseless and therefore, the minimizer in each of the four algorithms uses $\Phi\tilde{x} = y$ as its constraint. The initial selection parameter, μ_0 , is fixed as 0.02.

Reconstruction accuracy versus Measurement Ratio

In the first experiment, the effect of MR on reconstruction accuracy is presented. The number of measurements, m , is varied from 50 to 90 in steps of 5 and hence the MR varies from 0.2 to 0.36. The common sparsity, K_c , is fixed as 15. For MABP and $W\ell_1$, the adaptation parameter, μ_a , is set to be 0.1. For $W\ell_1$ and $W\ell_1$, ω is fixed as 0.05. If the maximum magnitude difference between the original signal and the reconstructed signal is

smaller than 0.01, the reconstruction is considered to be exact. This is due to the fact that the generated sparse signals $x^{(t)}$ obey $\frac{\|x^{(t)}\|_2}{n} > 0.01$ with overwhelming probability.

Figure 3.6: Reconstruction probability versus MR

Figure 3.6 shows the probability of exact reconstruction versus MR. It can be seen that the performances of Mod-BP and $W\ell_1$ are similar. Also, the performances of MABP and $WA\ell_1$ are much closer. It is evident that the MABP/ $WA\ell_1$ has a better probability of exact reconstruction compared to that of Mod-BP/ $W\ell_1$. For example, at MR = 0.28, MABP/ $WA\ell_1$ gives 95% probability of exact reconstruction whereas Mod-BP/ $W\ell_1$ recovers with 72% probability. Figure 3.7 shows the average Mean Square Error (MSE) versus MR. The MSE is computed as follows,

$$\text{MSE} = \frac{1}{nJ} \sum_{i=1}^J \|x_i - \hat{x}_i\|_2^2. \quad (3.40)$$

For any MR, MABP has a lesser reconstruction error compared to that of the Mod-BP. Also, $WA\ell_1$ has a lesser error compared to that of the $W\ell_1$. In the region where the reconstruction

Figure 3.7: Average MSE versus MR

Figure 3.8: Reconstruction probability versus CSR

Figure 3.9: Average MSE versus CSR

occurs with higher probability, say $MR \geq 0.3$, the average MSE is almost the same for both MABP and $WA\ell_1$. However, in the region where the reconstruction probabilities are not so high, say $MR \leq 0.24$, $WA\ell_1$ gives significantly lesser MSE compared to that of the MABP. Notice that the $WA\ell_1$ gives the least MSE among all the four techniques.

Reconstruction accuracy versus Common Support Ratio

In the next experiment, the effect of Common Support Ratio ($CSR = \frac{K_c}{K}$) on reconstruction accuracy is presented. The common support size, K_c , is varied from 5 to 20 in steps of 3 and hence the CSR varies from 0.25 to 1. Values of MR , μ_a and ω are fixed as 0.3, 0.1 and 0.05 respectively. Plot showing the exact reconstruction probability versus CSR is given in figure 3.8. For $CSR \geq 0.7$, MABP's performance is similar to that of the $WA\ell_1$ and the Mod-BP's performance is similar to that of the $W\ell_1$. For $CSR \geq 0.8$, MABP/ $WA\ell_1$ gives 100% exact reconstruction whereas Mod-BP/ $W\ell_1$ gives only 91%. On the other hand, for $CSR < 0.35$, MABP and $WA\ell_1$ give better reconstruction probability compared to that of

the Mod-BP and $W\ell_1$ respectively. Figure 3.9 shows the effect of CSR on the average MSE. For any CSR value, MABP and $WA\ell_1$ give lesser MSE compared to that of the Mod-BP and $W\ell_1$ respectively. As in the previous experiment, $WA\ell_1$ gives the least MSE among all the four techniques.

Choice of μ_a

Next, MABP's probability of exact reconstruction (versus MR) is presented for different μ_a values. K_c is fixed as 15 and the MR is varied from 0.2 to 0.36. ω is fixed as 0.05. Figure 3.10 shows the reconstruction probability curves of MABP for four μ_a values (0.05, 0.1, 0.2 and 0.4). As can be seen in the figure, MABP gives its best performance for $\mu_a = 0.1$. When μ_a is increased from 0.05 to 0.1, MABP's performance improves. However, when μ_a is increased further, the performance starts getting worse. The same experiment is repeated for $WA\ell_1$ and figure 3.11 shows the reconstruction probability curves of $WA\ell_1$ for different μ_a values. As in the case of MABP, $WA\ell_1$ gives its best reconstruction performance for $\mu_a = 0.1$.

Choice of ω

Finally, $WA\ell_1$'s probability of exact reconstruction (versus MR) is presented for different ω values. K_c is fixed as 15 and the MR is chosen from 0.2 to 0.36. μ_a is fixed as 0.1. Since $W\ell_1$ also uses ω , comparison with $W\ell_1$ is included. Figure 3.12 shows the reconstruction probability curves of $WA\ell_1$ and $W\ell_1$, for four ω values (0.05, 0.1, 0.2 and 0.4). As can be seen in the figure, $WA\ell_1$ gives its best performance when $\omega \leq 0.2$. When ω is increased from 0.2 to 0.4, $WA\ell_1$'s performance gets worse. On the other hand, the reconstruction performance of $W\ell_1$ is almost the same for all four ω values. It can be seen that, for any ω value, $WA\ell_1$'s performance is better than that of the $W\ell_1$.

Figure 3.10: Reconstruction probability versus MR for different μ_a : MABP

Figure 3.11: Reconstruction probability versus MR for different μ_a : WAl_1

Figure 3.12: Reconstruction probability versus MR for different ω : $WA\ell_1$ and $W\ell_1$

3.5.2 Time-sequence Reconstruction of Video Frames

In order to validate the performance of MABP in a practical scenario, time-sequence reconstruction of video frames is considered. It is well known that, each video frame (or each block of a frame) is temporally correlated to its previous frame/block, and these are sparse under 2-D DCT. Therefore, a signal ensemble formed using the DCT coefficients of a sequence of video frames resembles JSM-1. At the decoder, the first frame is reconstructed using a regular CS reconstruction algorithm (e.g. BP), and the remaining frames are reconstructed using its PKS version (e.g. Mod-BP). Sensing and reconstruction procedures are carried out block wise (block size of 32×32 pixels) on first 100 frames from each of the 21 YUV video sequences available at <http://videocoders.com/yuv.html>. An i.i.d. Gaussian random matrix is used as the sensing matrix. y is corrupted by a noise such that its SMNR is 20 dB. Since y is noisy, the minimizer in each of the four algorithms uses $\|\Phi\tilde{x} - y\|_2 \leq 0.001$ as its constraint. Notice that $n = 1024$. Values of μ_0 and μ_a are fixed

as 0.05 and 0.15 respectively. For $W\ell_1$ and $WA\ell_1$, ω is fixed as 0.05. The reconstruction quality is measured in terms of PSNR which is defined as

$$\text{PSNR} = 10\log_{10}\left(\frac{255^2}{\text{MSE}}\right) \quad (3.41)$$

where MSE is the mean square error computed using the original and reconstructed video frames. Figure 3.13 shows the PSNR/frame (average over all video sequences) as a function of MR. It can be seen that the performances of MABP and $WA\ell_1$ are almost similar. For any MR, the reconstruction accuracy of MABP/ $WA\ell_1$ is better than that of the Mod-BP/ $W\ell_1$.

Figure 3.13: Average PSNR (in dB) versus MR: YUV sequences

3.6 Extension of MABP to Greedy Pursuits

As mentioned in earlier chapters, Carrillo et al extended the modification idea to GPs, resulting in Mod-GP algorithms such as CoSaMP-PKS [25] and IHT-PKS [26]. The Mod-

Algorithm 4 Modified Adaptive Greedy Pursuit

InputMeasurement matrix: Φ Measurement vector: $y^{(t)}$ Previous reconstructed signal: $\hat{x}^{(t-1)}$ Sparsity level: K Initial selection parameter: μ_0 Adaptation parameter: μ_a **Initialization**

Estimate the known support as

 $T_k = \{\text{indices corresponding to } \lfloor \mu_0 K \rfloor \text{ largest entries in } \hat{x}^{(t-1)}\}.$ **Adaptation:**Step 1: The approximate coefficients ($\tilde{x}_{T_k} = \Phi_{T_k}^\dagger y^{(t)}$) are computed.Step 2: $\Gamma_k = T_k - \{\text{indices corresponding to coefficients with magnitude smaller than } \mu_a \|\tilde{x}_{T_k}\|_\infty\}.$ **Modified Greedy Pursuit**Solve the Mod-GP problem to obtain $\hat{x}^{(t)}$:

$$\hat{x}^{(t)} = \text{Mod-GP}(\Phi, y^{(t)}, \Gamma_k, K)$$

Output $\hat{x}^{(t)}$: Signal estimate.

GP algorithms showed improvement over their regular versions provided $T_k \subset T$. These algorithms can perform much better if the PKS is pruned in the same manner as in MABP. Therefore, with a motivation of improving Mod-GPs, the pruning idea of MABP (i.e., carving out a better PKS out of T_k) is applied in Mod-GPs. The resulting algorithms are termed as Modified Adaptive Greedy Pursuits (MAGPs). Algorithm 4 shows the step-by-step procedure involved in MAGP based recovery of correlated sparse signals.

Given the previous reconstructed signal $\hat{x}^{(t-1)}$, the PKS, T_k , is formed using the $\lfloor \mu_0 K \rfloor$ indices corresponding to the largest entries in $\hat{x}^{(t-1)}$ where $\mu_0 < 1$ is the initial selection parameter. If the signals are highly correlated μ_0 can be chosen close to 1. Since most GPs

require K to be known, the initial support estimation step can utilize K . The adaptation step is similar to that in MABP. Upon obtaining Γ_k via adaptation, the signal is estimated using a Mod-GP. As mentioned earlier, the adaptation step requires $O(K^2m)$ computations. Mod-GPs have the complexity in the same order as that of their GP versions. Consider CoSaMP for example, its complexity is $O(mn \log K)$ and therefore, the computational complexity of MAGP-CoSaMP will be $O(mn \log K + K^2m)$. If $K \ll n$, the complexity becomes $O(mnK)$. In general, GPs are faster than CR techniques and hence, MAGPs are generally faster than MABP or $W\mathcal{A}\ell_1$.

Next, the experimental results of MAGPs (MAGP-CoSaMP and MAGP-IHT) are presented, comparing them with that of the Mod-GPs (CoSaMP-PKS and IHT-PKS). As given in subsection 3.5.1, four correlated sparse signals are generated. 250 independent trials are conducted to obtain the average results. The initial selection parameter, μ_0 , is fixed as 0.75.

Reconstruction accuracy versus MR: In the first experiment, the effect of MR on reconstruction accuracy is presented. MR is varied from 0.2 to 0.48. K_c is fixed as 15. For MAGP-CoSaMP, μ_a is set to be 0.5. For MAGP-IHT, μ_a is fixed as 0.3. These choices of μ_a will be justified in the later experiments. Figure 3.14 shows the average Fraction of Support Recovered (FSR) as a function of MR. The FSR is defined as

$$\text{FSR} = \frac{|T \cap \hat{T}|}{K} \quad (3.42)$$

where $\hat{T}$ is the estimated support set. For any MR greater than 0.3, MAGPs result in a significantly larger FSR compared to that of their Mod-GP versions. For $\text{MR} \geq 0.4$, MAGP-CoSaMP and MAGP-IHT have FSRs higher than 0.95 whereas CoSaMP-PKS and IHT-PKS have FSRs lesser than 0.85.

Reconstruction accuracy versus CSR: Next, the effect of CSR on reconstruction accuracy is presented. CSR is varied from 0.25 to 1. MR is fixed as 0.4. Values of μ_a for MAGP-CoSaMP and MAGP-IHT are fixed as 0.5 and 0.3 respectively. Figure 3.15 shows

the average FSR as a function of CSR. For any CSR in the chosen range, MAGPs give a better reconstruction compared to that of Mod-GPs. This indicates the robustness of MAGPs with respect to innovation components. Among MAGPs, MAGP-CoSaMP results in a larger FSR compared to that of the MAGP-IHT.

Figure 3.14: Average FSR versus MR

Choice of μ_a for MAGPs: In the last experiment, the effect of μ_a on the reconstruction performances of MAGPs is presented. The MR is varied from 0.2 to 0.48. K_c is fixed as 15. Figure 3.16 shows the FSR curves of MAGP-CoSaMP for four different μ_a values (0.1, 0.3, 0.5 and 0.7). As can be seen in the figure, for $MR > 0.3$, MAGP-CoSaMP gives its best performance when $\mu_a \geq 0.5$. The same experiment is repeated for MAGP-IHT and figure 3.17 shows the FSR curves of MAGP-IHT for different μ_a values. In the low MR region, say $MR < 0.3$, FSR increases with a decrease in μ_a . However, for MRs greater than 0.35, MAGP-IHT's performance for $\mu_a = 0.1$ is much worse than that for $\mu_a \geq 0.3$.

Figure 3.15: Average FSR versus CSR

Figure 3.16: Average FSR versus MR for different μ_a : MAGP-CoSaMP

Figure 3.17: Average FSR versus MR for different μ_a : MAGP-IHT

3.7 Summary

In this chapter, MABP and $WA\ell_1$ algorithms were proposed for reconstructing correlated sparse signals. In MABP/ $WA\ell_1$, an adaptation process was carried out ahead of applying Mod-BP/ $W\ell_1$. It was theoretically shown that if in the adaptation step there are more correct deletions than wrong deletions, then MABP/ $WA\ell_1$ has a weaker (more favorable) RIP requirement compared to that of the Mod-BP/ $W\ell_1$. As the adaptation step requires only $O(|T_k|^2 m)$ computations, the computational complexity of MABP/ $WA\ell_1$ is of the same order as that of the Mod-BP/ $W\ell_1$. Experimental results showed that the proposed MABP/ $WA\ell_1$ algorithms provide better signal reconstruction compared to that of the Mod-BP/ $W\ell_1$ algorithms. Motivated by the effectiveness of adaptation step in MABP/ $WA\ell_1$, the adaptation idea was extended to Mod-GPs, resulting in MAGPs. Experimental results showed that the proposed MAGPs perform better than the Mod-GPs.

Chapter 4

Adaptive Backtracking Matching Pursuit

Existing PKS based GPs have a drawback: they are totally blind to the wrong atoms present in the PKS. In this chapter, with a motive of making proper/effective use of the PKS in BMP, a PKS based adaptation mechanism is proposed which, when replaces the traditional initialization of BMP, results in a reconstruction accuracy higher than that of the existing PKS based GPs. The resulting algorithm is termed as Adaptive Backtracking Matching Pursuit (AdBMP). Experimental results show that, in recovering correlated sparse signals, the proposed AdBMP gives a better reconstruction performance compared to that of the existing PKS based GPs and the BMP itself.

4.1 Drawback in PKS based Greedy Pursuits

As mentioned in chapter 3, for multiple signals such as correlated sparse signals following JSM-1 model, CS can be extended to DCS. A signal ensemble similar to the one discussed in chapter 3 is considered. The J signals in the ensemble are denoted as $x^{(t)} \in \mathbb{R}^n$, $t \in [1, J]$. Let $\Phi \in \mathbb{R}^{m \times n}$ be the i.i.d. Gaussian measurement matrix such that $y^{(t)} = \Phi x^{(t)} \in \mathbb{R}^m$

gives m incoherent measurements of $x^{(t)}$. The CS problem wishes to recover $x^{(t)}$ from $y^{(t)}$, sequentially. In DCS, each signal in the ensemble is independently sensed (using Φ) and transmitted to the decoder. While reconstructing the correlated sparse signals at the decoder, the support of each signal in the joint-sparse ensemble is partially known (in the form of PKS). Vaswani et al proposed Modified Basis Pursuit (Mod-BP) which made use of the PKS to recover the sparse signal [18]. Several variations of Mod-BP (including MABP) have been proposed. Though these algorithms recover the JSM-1 signals with high accuracy, they take a considerable amount of time to converge. Since GPs converge faster than the CR methods, Carrillo et al extended the modification idea to GPs to obtain OMP-PKS, CoSaMP-PKS and IHT-PKS [25, 26]. These modified GPs are collectively dubbed as PKS based GPs.

Consider CoSaMP-PKS [25] which is basically a CoSaMP that exploits PKS (say Ω). During the preliminary selection step (at each iteration), a candidate set is created by the union of Ω and the $2(K - |\Omega|)$ -sized supplement set. Final selection step chooses $K - |\Omega|$ atoms that are not in Ω and preserves all components in Ω . Therefore, the final support estimate of CoSaMP-PKS contains Ω . Carrillo then proposed IHT-PKS [26], an extension of IHT to include PKS. As in the case of CoSaMP-PKS, the final estimate of IHT-PKS contains Ω . These PKS based GPs showed improvement over their regular versions provided the PKS lies inside the actual support T . The major drawback of these algorithms is their assumption: $\Omega \subset T$, which is very restrictive. In general, the PKS might contain some wrong atoms as well, which is likely for JSM-1 signals such as video frames. However, most PKS based GPs are totally blind to the wrong atoms present in the PKS. Their final support-set estimate contains the entire PKS and therefore it also contains those wrong atoms. This might affect the reconstruction accuracy.

In this chapter, with a motive of making proper/effective use of the PKS in BMP, a PKS based adaptation mechanism is devised which, when replaces the traditional initialization

of BMP, reconstructs the sparse signal with an accuracy higher than that of the existing PKS based GPs. The resulting algorithm is termed as Adaptive Backtracking Matching Pursuit (AdBMP). Moreover, the theoretical upper bound on the unrecovered signal energy at the end of the adaptation mechanism, and hence the upper bound on the unrecovered signal energy after every iteration of AdBMP, is formulated. Experimental results show that, in the case of synthetic JSM-1 signals, AdBMP gives a better reconstruction accuracy compared to that of the state-of-the-art PKS based GPs (CoSaMP-PKS and IHT-PKS). In order to validate the performance of AdBMP in a practical scenario, three applications are discussed: (i) recovery of radar echo signals, (ii) time-sequence reconstruction of video frames, and (iii) object based surveillance video coding.

4.2 Proposed Adaptive Backtracking Matching Pursuit

In practice, PKS may not satisfy being a subset of the actual support set T . Therefore, backtracking based greedy algorithms can perform much better if the PKS is utilized in a more effective manner compared to that in CoSaMP-PKS. Since BMP is considered as a generalization with SP ($\tau=1$) and CoSaMP ($\tau=2$) being special cases, and it was shown that BMP with $\tau = 0.25$ results in a faster convergence compared to SP/CoSaMP, BMP is modified. In the initialization step of BMP, the initial support set is the set of indices corresponding to K largest entries in $\Phi^T y$ ($\prime$ represents matrix transposition) and the initial measurement residue y_r^0 is equal to y . This initialization step of BMP is modified to obtain AdBMP.

4.2.1 The AdBMP Algorithm

Algorithm 5 shows the step-by-step procedure involved in AdBMP recovery of correlated sparse signals. In the adaptation process, given the previous reconstructed signal ($\hat{x}^{(t-1)}$)

Algorithm 5 Adaptive Backtracking Matching Pursuit

Input

 Measurement matrix: Φ

 Measurement vector: $y = y^{(t)}$

 Previous reconstructed signal: $\hat{x}^{(t-1)}$

 Sparsity level: K

 Preliminary selection parameter: τ

 Initial adaptation parameter: $\tau_0 (< 1)$
Adaptation (modified initialization)

- 1) PKS: $\Omega =$ indices corresponding to $\lfloor \tau_0 K \rfloor$ largest entries in $\hat{x}^{(t-1)}$.
- 2) Initial measurement residual: $y_\Omega = y - \Phi_\Omega \Phi_\Omega^\dagger y$.
- 3) Initial support: $T^0 = \Omega \cup \{\text{indices corresponding to the } K - |\Omega| \text{ largest entries in } \Phi' y_\Omega\}$.
- 4) Updated initial measurement residual: $y_r^0 = y - \Phi_{T^0} \Phi_{T^0}^\dagger y$.

BMP iterations

 At the i^{th} iteration do the following

- 5) $\tilde{T}^i = T^{i-1} \cup \{\text{indices corresponding to the } \lfloor \tau K \rfloor \text{ largest entries in } \Phi' y_r^{i-1}\}$.
- 6) $T^i = \{\text{indices corresponding to the } K \text{ largest entries in } \Phi_{\tilde{T}^i}^\dagger y\}$.
- 7) Updated measurement residual: $y_r^i = y - \Phi_{T^i} \Phi_{T^i}^\dagger y$.
- 8) If $\|y_r^i\|_2 > \|y_r^{i-1}\|_2$, assign $T^i = T^{i-1}$ and quit the iteration.

Output

 Signal estimate: $\hat{x}^{(t)}$, such that $\hat{x}_{T^i}^{(t)} = \Phi_{T^i}^\dagger y$ and $\hat{x}_{T^i^c}^{(t)} = 0$.

 Figure 4.1: Schematic of actual support (T), PKS (Ω) and initial support set (T^0)

and the initial adaptation parameter ($\tau_0 < 1$), the PKS Ω is formed using the indices corresponding to the $\lfloor \tau_0 K \rfloor$ largest entries in $\hat{x}^{(t-1)}$. Using Ω , initial measurement residual y_Ω is

computed as follows,

$$y_\Omega = y - \Phi_\Omega \Phi_\Omega^\dagger y. \quad (4.1)$$

Note that $\dagger$ represents matrix pseudo inversion. As in step 3 of algorithm 5, the initial support set T^0 is formed using Ω and the indices corresponding to the $K - |\Omega|$ largest entries in $\Phi' y_\Omega$. A schematic describing the actual support set T , Ω and T^0 is shown in figure 4.1. The initial measurement residual is updated as follows,

$$y_r^0 = y - \Phi_{T^0} \Phi_{T^0}^\dagger y. \quad (4.2)$$

The adaptation process is followed up with BMP iterations. In BMP, unlike SP or CoSaMP, the preliminary selection step chooses an arbitrary-sized ($\lfloor \tau K \rfloor$) atom set to form the candidate set where α is the selection parameter. A larger value of τ results in a larger size of the supplement set formed at the preliminary selection step. BMP and the proposed AdBMP use this parameter. This parameter is significant in the sense that it directly contributes to the RIP requirement of the CS algorithm. The final selection is same as that of the SP or CoSaMP. Note that in AdBMP, Ω need not be a subset of the final support estimate. This allows the algorithm to drop wrong atoms (if any) present in Ω .

No PKS Scenario: AdBMP algorithm is capable of reconstructing signals with no prior information. In such cases, $\Omega = \emptyset$ or equivalently $\tau_0 = 0$. Therefore, the adaptation process is similar to that of the BMP initialization, and hence the AdBMP tends to regular BMP.

4.2.2 Complexity Analysis

GPs require lesser computations compared to linear programming based techniques. For example, SP algorithm requires $O(mn \log K)$ as compared to BP's requirement of $O(m^2 n^{\frac{3}{2}})$. BMP will have the same complexity as that of the SP. Signal approximations in the adaptation process of AdBMP require $O(K^2 m)$ computations [81, 82]. Therefore, the compu-

tational complexity of the AdBMP will be $O(mn \log K + K^2 m)$. If $K \ll n$, the complexity becomes $O(mnK)$.

4.2.3 Theoretical Analysis

For analyzing the effect of adaptation process, let us define $\rho = \frac{|T \cap T^0|}{K}$ such that $0 \leq \rho \leq 1$. Next, T^0 is partitioned into two subsets: the set $T^0 \cap T$, containing ρK indices in the actual support set, and $T^0 - T$, the set of $(1 - \rho)K$ incorrectly selected indices. Then

$$\begin{aligned} \|\Phi'_{T^0} y\|_2 &\leq \|\Phi'_{T^0 \cap T} y\|_2 + \|\Phi'_{T^0 - T} y\|_2 \\ &\leq \|\Phi'_{T^0 \cap T} y\|_2 + \|\Phi'_{T^0 - T} \Phi_T x_T\|_2 \\ &\leq \|\Phi'_{T^0 \cap T} y\|_2 + \delta' \|x\|_2 \end{aligned} \quad (4.3)$$

where the last inequality follows from the near-orthogonality of columns (lemma 1 in [10]) and $\delta' = \delta_{|T^0 - T| + |T|} = \delta_{(2 - \rho)K}$ is the RIC (since $|T^0 - T| = (1 - \rho)K$ and $|T| = K$). Furthermore,

$$\|\Phi'_{T^0 \cap T} y\|_2 \leq \|\Phi'_{T^0 \cap T} \Phi_{T^0 \cap T} x_{T^0 \cap T}\|_2 + \|\Phi'_{T^0 \cap T} \Phi_{T - T^0} x_{T - T^0}\|_2.$$

Using the restricted isometry property on $\|\Phi'_{T^0 \cap T} \Phi_{T^0 \cap T} x_{T^0 \cap T}\|_2$ and since $|T^0 \cap T| = \rho K$, it is bounded by $(1 + \delta_{\rho K}) \|x_{T^0 \cap T}\|_2$. Also, using the near-orthogonality property, the second term can be upper bounded by $\delta_K \|x_{T - T^0}\|_2$, which can further be relaxed to $\delta_K \|x\|_2$. Therefore

$$\|\Phi'_{T^0 \cap T} y\|_2 \leq (1 + \delta_{\rho K}) \|x_{T^0 \cap T}\|_2 + \delta_K \|x\|_2. \quad (4.4)$$

Combining (4.3) and (4.4),

$$\|\Phi'_{T^0} y\|_2 \leq (1 + \delta_{\rho K}) \|x_{T^0 \cap T}\|_2 + (\delta_K + \delta') \|x\|_2. \quad (4.5)$$

Let the initial support set T^0 obey

$$\|\Phi'_{T^0} y\|_2 \geq b\|x\|_2 \quad (4.6)$$

where b is a non-negative constant. Combining (4.5) and (4.6) gives

$$b\|x\|_2 \leq (1 + \delta_{\rho K})\|x_{T^0 \cap T}\|_2 + (\delta_K + \delta')\|x\|_2. \quad (4.7)$$

Hence,

$$\|x_{T^0 \cap T}\|_2 \geq \frac{b - (\delta_K + \delta')}{(1 + \delta_{\rho K})} \|x\|_2. \quad (4.8)$$

The unrecovered signal energy after adaptation can be expressed as

$$\|x_{T-T^0}\|_2^2 = \|x\|_2^2 - \|x_{T^0 \cap T}\|_2^2 \leq B_0 \|x\|_2^2 \quad (4.9)$$

where $B_0 = 1 - \left[\frac{b - (\delta_K + \delta')}{(1 + \delta_{\rho K})} \right]^2$. Note that $B_0 \leq 1$. As given in [16], the unrecovered signal energy after i iterations of BMP can be expressed as,

$$\|x_{T-T^i}\|_2^2 \leq B_1 \|x_{T-T^{i-1}}\|_2^2 \leq (B_1)^i \|x_{T-T^0}\|_2^2 \quad (4.10)$$

where

$$B_1 = \begin{cases} \left(\frac{1+\delta''}{1-\delta''} \right)^2 \left(1 - \tau + \frac{8\delta''}{(1-\delta'')^4} \right), & \text{if } 0 < \tau < 1 \\ \left[\frac{(1 + \sqrt{\frac{1}{\tau}})\delta''(1+\delta'')}{(1-\delta'')^3} \right]^2, & \text{if } \tau \geq 1 \end{cases} \quad (4.11)$$

and $\delta'' = \delta_{(2+\tau)K}$. Furthermore, B_1 is less than 1 if δ'' satisfies

$$0 < \delta'' \leq \begin{cases} \frac{\tau}{12+4\tau}, & \text{if } 0 < \tau < 1 \\ \frac{\tau}{1+4\tau}, & \text{if } \tau \geq 1 \end{cases}. \quad (4.12)$$

Invoking (4.9), the last inequality in (4.10) becomes

$$\|x_{T-T^i}\|_2^2 \leq B_0(B_1)^i \|x\|_2^2 \quad (4.13)$$

which is the recurrence relation of AdBMP. Therefore, at start (after adaptation), the unrecovered signal energy is bounded by $B_0\|x\|_2^2$, and usually decreases thereafter depending on B_1 . The recurrence relation in (4.13) is common to all GP algorithms in the sense that only the constants B_0 and B_1 are different for different algorithms. Table 4.1 gives the constants B_0 and B_1 for SP, CoSaMP, IHT, BMP and AdBMP. AdBMP's advantage over BMP is obvious; for example, for $\tau = 0.25$, $\rho = 0.5$, $b = 0.98$ and $\delta_{2K} = 0.015$ ($\delta_{\rho K} = 0.004$, $\delta_K = 0.008$, $\delta' = 0.011$ and $\delta'' = 0.017$), the bound $B_0\|x\|_2^2$ (at start of AdBMP) is comparable to the bound $B_0(B_1)^7\|x\|_2^2$ (after 7 iterations of regular BMP).

Table 4.1: Comparison of constants involved in the recurrence relation of various GPs

Algorithm	B_0	B_1	RIP condition
SP [10]	$\frac{8\delta_{2K}(1-\delta_{2K})}{(1+\delta_{2K})^2}$	$\left[\frac{2\delta_{3K}(1+\delta_{3K})}{(1-\delta_{3K})^3} \right]^2$	$\delta_{3K} \leq 0.139$
CoSaMP [11]	$\frac{1-\delta_{2K}}{1-2\delta_{2K}}$	0.25	$\delta_{4K} \leq 0.1$
IHT [12]	$\frac{1-\delta_{2K}}{1-2\delta_{2K}}$	0.25	$\delta_{3K} \leq \frac{1}{\sqrt{32}}$
BMP [16]	$\frac{8\delta_{2K}(1-\delta_{2K})}{(1+\delta_{2K})^2}$	$\begin{cases} \left(\frac{1+\delta''}{1-\delta''} \right)^2 \left(1 - \tau + \frac{8\delta''}{(1-\delta'')^4} \right), & \text{if } 0 < \tau < 1 \\ \left[\frac{(1+\sqrt{\frac{1}{\tau}})\delta''(1+\delta'')}{(1-\delta'')^3} \right]^2, & \text{if } \tau \geq 1 \end{cases}$	$\delta'' \leq \begin{cases} \frac{\tau}{12+4\tau}, & \text{if } 0 < \tau < 1 \\ \frac{\tau}{1+4\tau}, & \text{if } \tau \geq 1 \end{cases}$
Proposed AdBMP	$1 - \left[\frac{b-(\delta_K+\delta')}{(1+\delta_{\rho K})} \right]^2$	$\begin{cases} \left(\frac{1+\delta''}{1-\delta''} \right)^2 \left(1 - \tau + \frac{8\delta''}{(1-\delta'')^4} \right), & \text{if } 0 < \tau < 1 \\ \left[\frac{(1+\sqrt{\frac{1}{\tau}})\delta''(1+\delta'')}{(1-\delta'')^3} \right]^2, & \text{if } \tau \geq 1 \end{cases}$	$\delta'' \leq \begin{cases} \frac{\tau}{12+4\tau}, & \text{if } 0 < \tau < 1 \\ \frac{\tau}{1+4\tau}, & \text{if } \tau \geq 1 \end{cases}$

4.3 Reconstruction of Synthetic JSM-1 Signals

Four correlated sparse signals (Gaussian random signals as in JSM-1) each of length $n = 250$ and $K = 20$ are generated. For the first experiment, the sparsities of common and innovation components are fixed as 15 and 5 respectively. For different MR values, signal recovery is performed using BMP [16], CoSaMP-PKS [25], IHT-PKS [26] and AdBMP. For each MR, 500 independent trials are performed to obtain the average FSR and the average MSE. In each trial, an $m \times n$ Gaussian random measurement matrix is generated. The PKS (for CoSaMP-PKS, IHT-PKS and AdBMP) is formed as shown in step 1 of algorithm 5, fixing $\tau_0 = 0.6$. For BMP and AdBMP, performances are evaluated for three τ values: 0.25, 1 and 1.5. Figure 4.2 shows the average FSR for different MR values and figure 4.3 shows the corresponding average MSE. It can be seen that the proposed AdBMP (with $\tau = 0.25$) gives the best reconstruction accuracy among all the methods. For MRs less than 0.25, AdBMP with $\tau = 0.25$ or 1.0 gives a better performance compared to that of AdBMP with $\tau = 1.5$. For any MR, AdBMP with $\tau = 0.25$ or 1.0 gives much better performance compared to that of the CoSaMP-PKS and the IHT-PKS.

For the next experiment, the number of measurements is fixed as 75 and the sparsity of common component is varied from 5 to 20 in steps of 3. For each CSR, 500 independent trials are performed to obtain the average FSR, the average MSE and the average reconstruction time. Figure 4.4 shows the average FSR for different CSR values and figure 4.5 shows the corresponding average MSE. Again, it can be seen that the proposed AdBMP gives the best reconstruction accuracy among all the methods. Figure 4.6 shows the average reconstruction time (computation time in MATLAB 7.12.0 running on a 64-bit Intel(R) CoreTM i5-2400 processor with 8 GB RAM) for different CSR values. AdBMP clearly outperforms CoSaMP-PKS and IHT-PKS in terms of reconstruction speed.

Figure 4.7 shows the effect of τ_0 on AdBMP's reconstruction accuracy. For this experi-

Figure 4.2: Average FSR versus MR

Figure 4.3: Average MSE versus MR

Figure 4.4: Average FSR versus CSR

Figure 4.5: Average MSE versus CSR

Figure 4.6: Average reconstruction time (in sec) versus CSR

Figure 4.7: Average MSE versus τ_0 for different τ

ment, the number of measurements is fixed as 75 and the sparsity of common component is fixed as 15. Regardless of τ value chosen, the reconstruction accuracy of AdBMP increases with an increase in τ_0 and it saturates beyond $\tau_0 = 0.8$.

4.4 Sparse Recovery of Radar Echo Signals

In radar applications, it is not always possible to sample the radar signal ideally [85]-[89]. Further, consecutive radar echo signals show some correlation which may be exploited [90]. In this section, the radar echo signal is modelled and a sensing mechanism is adopted to acquire it. For CS reconstruction, AdBMP (which makes use of the PKS to reconstruct the sparse version of radar echo signal) is used. AdBMP's reconstruction performance is compared to that of the BMP, CoSaMP-PKS and IHT-PKS.

Let the radar pulse consist of a chirp of width W_{ch} at time t_0 over the center frequency of f_c ,

$$s(t) = e^{j2\pi f_c t} \text{rect}\left(\frac{t - t_0}{W_{ch}}\right) e^{j2\pi \left[\frac{1}{2}\gamma_{ch}(t-t_0)^2\right]} \quad (4.14)$$

where γ_{ch} is the chirp rate and $\text{rect}(\cdot)$ stands for the rectangular function. Let the target distance be $R(t)$ giving rise to an echo at time $t_0 + (2R(t)/c)$ where c is the wave propagation speed. Then the echo is,

$$r_1(t) = B' \text{rect}\left(\frac{t - t_0 - 2\frac{R(t)}{c}}{W_{ch}}\right) e^{j\pi\gamma_{ch}\left[t - t_0 - 2\frac{R(t)}{c}\right]^2} \quad (4.15)$$

where $B' = B e^{-j4\pi f_c \frac{R(t)}{c}}$ and B denotes the scattering coefficient. Under the assumption of constant velocity of the target, $R(t)$ may be approximated as

$$R(t) = R(t_0) + \nu t \quad (4.16)$$

where ν is the radial velocity. Matched filtering on $r_1(t)$ and subsequent transformation to

range-frequency domain gets rid of the third term on the RHS of (4.15) and results in the echo located around frequency f [85]. Since W_{ch} is very small compared to t , the slowly varying echo signal may be written as

$$\begin{aligned} r_2(f, t) &= B'' e^{-j4\pi(f+f_c)\frac{R(t)}{c}} \\ &= B'' e^{-j2\pi(f+f_c)2\frac{v}{c}t + j\theta} \\ &= B'' e^{-j2\pi f_d t + j\theta} \end{aligned}$$

where Doppler frequency $f_d = (f + f_c)(2\frac{v}{c})$ and $\theta = -4\pi(f + f_c)\frac{R(t_0)}{c} = -2\pi f_d \frac{R(t_0)}{v}$. It is obvious that the slowly varying echo signal is sparse under Discrete Fourier Transform (DFT) basis. Extending to multiple targets, the slowly varying echo is expressed as

$$r(f, t) = \sum_{i=1}^{J_{Tar}} B_i'' e^{-j2\pi f_{di}t + j\theta_i} \quad (4.17)$$

where J_{Tar} targets may have different real coefficients B_i'' , Doppler frequencies f_{di} , and phases θ_i (assumed between 0 and 2π). The location of each target's echo in the DFT basis is decided only by f_{di} . Since PKS supplies only the locations of the non-zero components but not the non-zero values, among the three parameters (B_i'' , f_{di} and θ_i) in $r(f, t)$, the PKS is related only to f_{di} , and not to B_i'' and θ_i . In the case of targets at constant or slowly varying velocity, f_{di} is constant or changes slowly over time. Therefore it is possible to predict f_{di} from the previous pulse. It is worth noting that this prediction is independent from one target to another.

Reconstruction Performance

Next, AdBMP's reconstruction performance in recovery of sparse radar echo signals is evaluated. The precision of pulse transmitting time is fixed as 0.001 and random signals r (as in (4.17)) of length $n = 1000(=\frac{1}{0.001})$ are generated. As these signals are complex, they

Figure 4.8: Average SRER (in dB) versus MR

are represented as $2n$ -length real signals. Resulting sparse signals are represented as,

$$x = \begin{bmatrix} Re(\Psi) & -Im(\Psi) \\ Im(\Psi) & Re(\Psi) \end{bmatrix} \star \begin{bmatrix} Re(r) \\ Im(r) \end{bmatrix}$$

where Ψ is the $n \times n$ DFT matrix, $Re(\cdot)$ and $Im(\cdot)$ represents the real and imaginary parts of the complex entries, and $\star$ represents matrix multiplication. Signals may also be handled directly in the complex domain with proper modifications to the BMP algorithm. However, to keep all the processing of radar echo signal in ‘real’ domain, they are split into real and imaginary parts.

In each experiment, 2000 trials are conducted to obtain the average reconstruction performance of all the techniques. The number of targets $J_{Tar} = 3$, where the targets are set as frequencies chosen between 200 and 400 Hertz. In trial 1, the targets (f_{di}) are fixed as 225 Hz, 275 Hz and 325 Hz. In the subsequent trials, each target is same as that of the

Figure 4.9: Average reconstruction time (in sec) versus MR

Figure 4.10: Average SRER (in dB) versus p

previous trial with probability p . This probability is termed as frequency repetition probability. The PKS (for CoSaMP-PKS, IHT-PKS and AdBMP) is formed as shown in step 1 of algorithm 5, fixing $\tau_0 = 0.6$. In each trial, an $m \times 2n$ Gaussian random measurement matrix is generated. The CS measurement vector y is corrupted by noise such that its SMNR is 20 dB. The sparsity is fixed as $K = 2J_{Tar}$. The reconstruction accuracy is measured in terms of Signal-to-Reconstruction-Error-Ratio (SRER) defined as follows,

$$SRER = 10 \log_{10} \frac{\|x\|_2^2}{\|x - \hat{x}\|_2^2} \quad (4.18)$$

where $\hat{x}$ represents the reconstructed sparse signals. Targets are assumed to be only on discrete grids. Even though off-grid target is not considered, due to the fact that this section considers the CS problem with observation noise, and that spectral leakage due to off-grid target is similar to observation noise, off-grid performance is not expected to be very different.

As first experiment, the reconstruction performance of AdBMP is compared with that of CoSaMP-PKS, IHT-PKS and BMP. Figure 4.8 shows the average SRER as a function of MR. The frequency repetition probability p is fixed as 0.6. Reconstruction accuracy of AdBMP is given for three different values of τ (0.25, 1.0 and 1.5). It can be inferred from figure 4.8 that the AdBMP with $\tau = 0.25$ gives the best performance among all the techniques. Figure 4.9 shows the reconstruction time (computation time in MATLAB 7.12.0 running on a 64-bit Intel(R) CoreTM i5-2400 processor with 8 GB RAM) as a function of MR. Proposed AdBMP algorithm clearly outperforms CoSaMP-PKS and IHT-PKS in terms of time taken for reconstruction. Also, AdBMP's computation time is comparable to that of the regular BMP. In the next experiment, the effect of p is presented. Figure 4.10 shows the variation of SRER as a function of p . The number of measurements m is fixed as 32. Again, the AdBMP algorithm clearly outperforms CoSaMP-PKS and IHT-PKS. It is evident from figure 4.10 that, for any PKS based GP, there will be an improvement in

performance with an increase in p .

Figure 4.11: Average PSNR (in dB) versus MR: YUV sequences

4.5 Reconstruction of Video Signals

In order to validate the performance of AdBMP in a practical scenario, two video reconstruction frameworks are considered: (i) block-based reconstruction of video frames and (ii) object-based reconstruction of surveillance video. The former is a general video reconstruction framework suitable for any video, whereas the latter is specifically designed for video frames with small moving objects and constant background.

4.5.1 Time-sequence Reconstruction of Video Frames

As mentioned in subsection 3.5.2, a signal ensemble formed using the DCT coefficients of a sequence of video frames resembles JSM-1. Sensing and reconstruction procedures are carried out block wise (block size of 32×32 pixels) on first 100 frames from each of the

Figure 4.12: Average reconstruction time (in sec) versus MR: YUV sequences

21 YUV video sequences available at <http://videocoders.com/yuv.html>. At the decoder, the first frame is reconstructed using a regular CS algorithm (e.g. BMP), and the remaining frames are reconstructed using its PKS version (e.g. AdBMP given as algorithm 5). Gaussian random matrix is used as the sensing matrix and y is corrupted by a noise such that its SMNR is 20 dB. Since the sparsity level K is not available, it is fixed as $\hat{K} = \frac{m}{\log(n)}$ (for all the signals in the ensemble). Notice that $n = 1024$. τ_0 is fixed as 0.6. AdBMP's performance is reported for three τ values (0.25, 1.0 and 1.5). Figure 4.11 shows the PSNR/frame (average over all video sequences) as a function of MR for four methods (including BMP). Figure 4.12 shows their corresponding decoding times. For each MR, the reconstruction performance (both accuracy and speed) of AdBMP with $\tau = 0.25$ is the best among all the four methods.

4.5.2 Object based Surveillance Video Coding

CS based video object coding schemes have a simpler encoder compared to that of the non-CS techniques [91]-[92]. In this subsection, a compression scheme called Distributed Compressive Video Sensing based Video Object Coding (DCVS-VOC) is proposed. Upon describing the encoder and decoder of DCVS-VOC, a PKS formation method is proposed in order to apply AdBMP (and other PKS based GPs) at the decoder. Then, the reconstruction performance of AdBMP is compared with that of the other algorithms.

Encoder

Figure 4.13 depicts the encoder of DCVS-VOC. Motivated by the frame classification idea in Distributed Compressive Video Sensing (DCVS) [93], in the encoder of DCVS-VOC, object segmentation is carried out only for the key frames. As in DCVS, video frames are classified into key frames and CS frames. For key frames, object segmentation is performed. The foreground objects are segmented from the background using an edge-based object segmentation (as in [94]). These frames are then coded using conventional video compression such as MPEG/H.264 or using CS with more number of measurements.

Figure 4.13: DCVS based video object coding - encoder

For CS frames following the key frame, for i^{th} object, a tentative object mask M_i is

obtained by dilating the corresponding object mask in the preceeding key frame, $M_i^{(key)}$. The tentative object mask can be expressed as

$$M_i = M_i^{(key)} \oplus Z \quad (4.19)$$

where Z is a disk-shaped structuring element of radius $2G$, and G is the Group-Of-Pictures (GOP) length/size. In scenarios where the object segmentation is not accurate, the object mask might not include some of the object pixels, but the dilated mask includes those missing pixels. Masked object block $f_{M_i}^{(t)}(x_i)$ is obtained for each segmented object by multiplying the object block $f^{(t)}(x_i)$ with the tentative object mask as follows,

$$f_{M_i}^{(t)}(x_i) = M_i \star f^{(t)}(x_i). \quad (4.20)$$

The vectorized (masked) object block $F_{M_i}^{(t)}(x_i)$ is sensed using a random CS matrix Φ_i as follows,

$$y_i = \Phi_i F_{M_i}^{(t)}(x_i) \in R^{m_i} \quad (4.21)$$

where $\Phi_i \in R^{m_i \times n_i}$ is a CS sensing matrix for object i and y_i is its measurement vector. This sensing procedure is carried out for all the objects. Note that f (lower case) denotes the 2-D frame/object whereas F (upper case) denotes the vectorized f . The parameters to be transmitted to the decoder are: location and shape of the objects, object masks, dimensions of the measurement matrices, and measurement vectors.

Decoder

At the decoder, key frames are first reconstructed using an appropriate decoding technique. Given the measurement vector y_i and key frames (succeeding and preceeding), the problem of reconstructing the i^{th} object block is considered. The measurement matrix Φ_i and the object mask M_i are also available at the decoder. As shown in figure 4.14, DCVS-VOC de-

coder has two parts: motion estimation and compensation and CS reconstruction of object error.

Figure 4.14: DCVS based video object coding - decoder

Motion Estimation and Compensation: First step is a novel object motion estimation (using CS measurements). Motion estimation is performed with respect to both preceding key frame $f^{(key,p)}$ and succeeding key frame $f^{(key,s)}$ separately. Two motion vectors are obtained as follows,

$$MV^{(key,p)} = \arg \min_{u \in W} \|\Phi_i(\hat{F}_{M_i}^{(key,p)}(u + x_i)) - y_i\|_2^2 \quad (4.22)$$

$$MV^{(key,s)} = \arg \min_{u \in W} \|\Phi_i(\hat{F}_{M_i}^{(key,s)}(u + x_i)) - y_i\|_2^2. \quad (4.23)$$

The motion vector indicating the smaller motion

$$MV_i = \arg \min_{j \in p,s} \|MV^{(key,j)}\|_2 \quad (4.24)$$

and the corresponding key frame $\hat{f}^{(key)}$ are used for motion compensation. Upon obtaining

MV_i , to predict the object block $f^{(t)}(x_i)$, motion compensation is performed as follows,

$$\tilde{f}^{(t)}(x_i) = M_i^c \star Bg(x_i) + M_i \star (\hat{f}^{(key)}(x_i + MV_i)) \quad (4.25)$$

where M_i^c denotes the binary complement of the mask of object i , and $Bg(x_i)$ denotes the background pixels in the object block x_i .

CS Reconstruction of Object Error: The measurement vector of the motion compensated object block is obtained as follows,

$$\tilde{y}_i = \Phi_i \tilde{F}_{M_i}^{(t)}(x_i). \quad (4.26)$$

The motion compensated object error is sparse, and therefore by expressing the difference between the measurement vectors y_i and $\tilde{y}_i$ as

$$y_i - \tilde{y}_i = \Phi_i [F_{M_i}^{(t)}(x_i) - \tilde{F}_{M_i}^{(t)}(x_i)] = \Phi_i [\Delta \hat{F}_{M_i}^{(t)}(x_i)], \quad (4.27)$$

the vectorized object error $\Delta \hat{F}_{M_i}^{(t)}(x_i)$ is efficiently reconstructed using a CS reconstruction algorithm. The signal to be reconstructed is a sparse signal with sparsity $\|\Delta \hat{F}_{M_i}^{(t)}(x_i)\|_0 \leq \sum_{pixels} M_i$. The vectored version of the object block $\hat{f}^{(t)}(x_i)$ is then obtained using

$$\hat{F}_{M_i}^{(t)}(x_i) = \Delta \hat{F}_{M_i}^{(t)}(x_i) + \tilde{F}_{M_i}^{(t)}(x_i). \quad (4.28)$$

The video object block is recovered by adding the obtained object error to the motion compensated object block. Finally, the video frame can be recovered using the reconstructed object blocks, object mask and the available background.

PKS Formation

In order to replace the CS algorithm with a PKS based algorithm, PKS formation is required. The key frame that was used for motion compensation, $\hat{f}^{(key)}$, is also used for PKS formation. The PKS, Ω , is the set of indices corresponding to $\lfloor \tau_0 K_i \rfloor$ largest entries in $\hat{F}_{M_i}^{(key)}(x_i)$ where K_i is the sparsity of the target signal.

Reconstruction Performance

The encoding performance of DCVS-VOC is evaluated in terms of average encoding time (T_{en}) and bits per pixel required (BPP). The decoding performance is evaluated in terms of average object-PSNR ($PSNR_{Ob}$ - PSNR considering object pixels only) and average decoding time (T_{de}). Two surveillance video sequences are considered for evaluation: Walk sequence (300 frames - 288×384 pixels) and Hall monitor sequence (300 frames - 240×352 pixels). Among these two sequences, Walk sequence has relatively small moving objects compared to that of the Hall monitor sequence. T_{en} is obtained by dividing the total time taken for encoding the sequence by the number of frames in the sequence. T_{de} is also computed in a similar manner as that of T_{en} . $PSNR_{Ob}$ indicates the object-PSNR averaged over the number of frames, respectively. Original video frames along with the object masks (obtained from object segmentation) are used as ground truth for $PSNR_{Ob}$ computation.

In DCVS-VOC, BPP can be varied by varying the MR. Here, $MR = \frac{m_i}{n_i}$. As the sparsity information is not available, the target signal's sparsity is fixed as $K_i = \frac{m_i}{4}$. The search window for motion estimation is fixed as ± 7 pixels. Key frames are encoded and decoded as in [92]. The GOP size is fixed as 4. Four CS reconstruction algorithms are applied. Table 4.2 shows the encoding and decoding performances of DCVS-VOC, for all four algorithms, for Walk and Hall monitor sequences. Values of τ_0 and τ are fixed as 0.6 and 0.25 respectively. For any MR, AdBMP gives the best reconstruction accuracy among the four algorithms. At the same time, BMP and AdBMP algorithms result in a faster

reconstruction compared to that of the CoSaMP-PKS and the IHT-PKS. In other words, AdBMP gives the fastest reconstruction among the three PKS based algorithms. Therefore, at the DCVS-VOC decoder, AdBMP is an appropriate choice for a fast CS reconstruction algorithm.

Table 4.2: Encoding and decoding performances of DCVS-VOC

Sequence				BMP		CoSaMP-PKS		IHT-PKS		AdBMP	
	MR	BPP	T_{en} (sec)	$PSNR_{Ob}$ (dB)	T_{de} (sec)	$PSNR_{Ob}$ (dB)	T_{de} (sec)	$PSNR_{Ob}$ (dB)	T_{de} (sec)	$PSNR_{Ob}$ (dB)	T_{de} (sec)
Walk	0.12	0.0386	0.61	17.41	0.76	20.92	4.92	18.58	1.77	21.21	0.75
	0.15	0.0415	0.65	19.12	0.93	21.04	6.02	19.93	1.97	21.39	0.89
	0.18	0.0442	0.67	19.63	1.12	21.23	7.72	20.29	2.26	21.98	1.06
	0.21	0.0467	0.68	20.41	1.33	21.33	8.11	20.81	3.61	22.43	1.29
	0.24	0.0493	0.67	20.92	1.58	21.56	10.08	21.31	4.88	22.66	1.59
Hall Monitor	0.06	0.0624	0.70	17.98	1.03	22.11	6.81	18.94	2.99	22.36	0.98
	0.12	0.0751	0.77	18.72	1.71	22.23	10.66	19.83	5.12	22.87	1.70
	0.18	0.0865	0.82	19.81	2.97	22.51	19.44	21.68	7.34	23.16	2.98
	0.24	0.0971	0.88	21.21	3.88	22.82	23.93	22.29	11.09	23.71	3.81
	0.30	0.1071	0.90	21.94	5.03	23.21	26.78	22.85	16.06	24.19	4.99

4.6 Summary

In this chapter, a PKS based backtracking matching pursuit algorithm, AdBMP, is proposed for recovery of correlated sparse signals, and the theoretical upper bound on the unrecovered signal energy after each of its iteration is formulated. Proposed AdBMP could well be seen as a PKS version of BMP. Experimental results show that the proposed AdBMP makes use of PKS in a more efficient manner compared to that of existing PKS based GPs. Then, a radar echo signal is modelled and a sensing mechanism is adopted to acquire it. AdBMP is used for reconstructing the sparse version of the radar echo signal. This will serve as a solution for current challenges in acquisition of radar echo signals. As in the case of synthetic JSM-1 signals and synthetically generated radar echo signals, AdBMP also exhibits superior performance compared to that of the CoSaMP-PKS, IHT-PKS and the regular BMP in the case of video reconstruction. As summarized in figure 4.15, AdBMP (with $\tau \leq 1$) reconstructs the correlated sparse signals with an accuracy higher than that of the CoSaMP-PKS and IHT-PKS. Also, AdBMP converges much faster than CoSaMP-PKS.

Figure 4.15: Speed and accuracy comparison: AdBMP

Chapter 5

Iterative Compressive Sensing

Framework for Compressive Imaging

In many CS applications, partial information about the non-zero locations of the sparse signal of interest may be available in the form of PKS. Many existing CS reconstruction algorithms were modified so as to make use of this PKS to improve their performance, and are termed as PKS based algorithms. As mentioned in earlier chapters, most PKS based algorithms are totally blind to the wrong atoms present in the PKS which degrades the reconstruction performance. Moreover, each PKS based algorithm adapts its own modification strategy. In this chapter, a generalized, algorithm-blind reconstruction framework called Iterative Compressive Sensing on Residual (ICSR) is proposed, which can accommodate any class of CS reconstruction algorithm, and is capable of improving the signal recovery utilizing the available PKS. Experimental results show that, in compressive imaging application, the ICSR framework improves the reconstruction accuracy of any ingredient CS reconstruction algorithm, and performs better than many of the existing PKS based algorithms. The ICSR framework can also be applied for reconstruction of sparse signals whose PKS is not available.

5.1 Exploratory Experiment and Motivation

In many applications, partial information about the non-zero locations of the sparse signal of interest may be available *a priori*, in the form of PKS. Vaswani et al proposed to modify BP so as to use PKS resulting in Mod-BP, LS-CS, etc. Carrillo et al extended the modification idea to GP algorithms resulting in PKS based GPs. Aforementioned modified CS algorithms are collectively dubbed as PKS based CS reconstruction algorithms. All PKS based CS reconstruction algorithms showed improvement over their regular versions and the Mod-BP is the best performing algorithm among them in terms of reconstruction accuracy.

The application of interest, CI, sketches an image by measuring the K largest transform coefficients and then switches to i.i.d. Gaussian sensing matrix to acquire the details [74]. For CI reconstruction using a PKS based algorithm, the PKS can be formed based on the sparsifying transformation. If 2-D DCT is used as a sparsifying transform, it is worth assuming that most of the significant transform coefficients fall within the first few zig-zag locations (say $K = \lfloor \frac{m}{\log n} \rfloor$ such coefficients are present in the first $2K$ zig-zag locations). This set of high probability locations is denoted as T_{hp} . To validate this assumption, an experiment is conducted on natural images (24 images from KODAK database). For each image, block-wise (32×32) DCT transform coefficients are obtained and the fraction of the indices corresponding to K largest magnitudes present in the first $2K$ zig-zag locations is computed. This fraction is denoted as λ_{hp} . In addition, the fraction of signal energy (E_{hp}) contained in those correct locations is computed. Table 5.1 shows the averaged λ_{hp} and E_{hp} values for different MR values. It can be seen that around 70 – 75% of the correct atoms (containing around 99.5% of the signal energy) fall within the assumed high probable locations T_{hp} . The statistics did not vary much with the change in block size. This knowledge of the signal’s support can be efficiently used either to bring down the number of required

measurements or to improve the reconstruction accuracy. Any sparsifying transform can be applied but T_{hp} needs to be found accordingly. For example, if DWT is used as the sparsifying transform, its lowest subband can be treated as T_{hp} [18].

Recovery conditions and error bounds of PKS based algorithms revealed the fact that a good PKS will give a better recovery [18, 19, 20, 22]. However, reconstruction via these algorithms has two drawbacks: First, most PKS based algorithms are totally blind to the wrong atoms present in the PKS and their final support-set estimate might contain those wrong atoms. Second, although all PKS based algorithms were proposed with a common motive of improving their regular versions, each of them adapts its own modification strategy. Therefore, if a new CS algorithm evolves, a new modification procedure needs to be devised. For some existing algorithms (such as FOCal Underdetermined System Solver (FOCUSS) [95]), there is no PKS based modified version. Therefore, we are more interested in proposing a generalized algorithm-blind framework capable of utilizing PKS and improving any CS algorithm plugged into it.

In this chapter, an algorithm-blind CS reconstruction framework (depicted in figure 5.1) called ICSR is proposed, which can accommodate any class of CS reconstruction algorithm and improve its signal recovery utilizing the available PKS. ICSR framework first estimates the initial support set using the available PKS, and then iteratively performs CS on the measurement residual to update the support set estimate. ICSR provides enough provisioning to drop wrong atoms present in the PKS. It is expected that two or more algorithms with differing reconstruction performances result in much closer performances after being plugged

Table 5.1: Fractions of correct atoms and signal energy in T_{hp}

MR	0.05	0.1	0.15	0.2	0.25
λ_{hp}	0.7102	0.7211	0.7297	0.7387	0.7446
E_{hp}	0.9932	0.9941	0.9955	0.9961	0.9967

into ICSR. Another feature of ICSR is, if PKS is not available, the initial support set can be formed in an alternate way, thereby increasing the flexibility of the ICSR framework with respect to the availability of PKS.

Figure 5.1: Two major steps in ICSR framework

5.2 Proposed Iterative Compressive Sensing on Residual

In this section, ICSR framework is proposed. It has two major steps: 1. Initial support set estimation and 2. Iterative CS on residual. The main steps of the ICSR framework are summarized in algorithm 6.

5.2.1 Initial Support Set Estimation

First, the initial support set T^0 is formed using the PKS. The set T^0 is chosen to be the set of indices corresponding to the $\lfloor \lambda_0 K \rfloor$ largest locations in $\Phi_{T_{hp}}^\dagger y$ where $\lambda_0 < \min(1, \frac{|T_{hp}|}{K})$ is the initial selection parameter and T_{hp} is the set of high probability locations. In other words, considering only the high probability locations, an LS estimate of the signal is obtained and the locations corresponding to the $\lfloor \lambda_0 K \rfloor$ largest entries are retained. This ensures that the assumed PKS is appropriately pruned. λ_0 can be chosen based on the knowledge of the support. In CI, it can be chosen based on the sparsifying transform. A schematic describing T_{hp} , T^0 , and the actual support set T is shown in figure 5.2.

Figure 5.2: Schematic of actual support set (T), set of high probability locations (T_{hp}) and initial support set (T^0)

5.2.2 Iterative CS on Residual

During each iteration, a schematic of which is depicted in figure 5.3, the signal approximation x_a^{i-1} and the measurement residual y_r^i are first computed using y and T^{i-1} as follows,

$$x_a^{i-1} = \Phi_{T^{i-1}}^\dagger y \quad \text{and} \quad y_r^{i-1} = y - \Phi_{T^{i-1}} x_a^{i-1}. \quad (5.1)$$

Then the residual signal x_r^i is reconstructed by applying a CS reconstruction algorithm on the computed measurement residual y_r^{i-1} . Let us define an algorithmic function

$$[\hat{x}, \hat{T}] = CSA(\Phi, y, K) \quad (5.2)$$

which denotes the general form of any CS reconstruction algorithm. Here, $\hat{T}$ is the estimated support and $\hat{x}$ is the estimated signal. GP algorithms like OMP and SP make use of the sparsity level K , whereas CR techniques need not require it.

In the initial iteration, T_r^i comes from CSA, and T^{i-1} comes from the signal knowledge. ICSR iterations progress by forming a reliable T^i each time. Upon obtaining x_r^i using CSA, refinement of the support estimated at the end of previous iteration T^{i-1} is done. A pruned support set, T_p^{i-1} , is formed out of T^{i-1} by retaining only those indices in x_a^{i-1} having magnitude larger than $\|x_r^i\|_\infty$. In any iteration, T_r^i is trusted and included in T^i . In

Figure 5.3: Iterations in ICSR framework

Algorithm 6 Iterative Compressive Sensing on Residual (With PKS)

Input

Measurement matrix: Φ

Measurement vector: y

High probability locations: T_{hp}

Sparsity level: K

Initial selection parameter: $\lambda_0 (< \min(1, \frac{|T_{hp}|}{K}))$

Initial support set formation:

$i=0$ and $T^i = \{\text{indices corresponding to the } \lfloor \lambda_0 K \rfloor \text{ largest entries in } \Phi_{T_{hp}}^\dagger y\}$

CS on Residual

Repeat

Step 1: Residual formation: $i=i+1$; $x_a^{i-1} = \Phi_{T^{i-1}}^\dagger y$ and $y_r^{i-1} = y - \Phi_{T^{i-1}} x_a^{i-1}$

Step 2: CS on residual: $[x_r^i, T_r^i] = \text{CSA}(\Phi, y_r^{i-1}, K - |T^{i-1}|)$

Step 3: Pruning: $T_p^{i-1} = T^{i-1} - \{l \in T^{i-1} : |x_a^{i-1}(l)| \leq \|x_r^i\|_\infty\}$

Step 4: Support merger: $T^i = T_p^{i-1} \cup T_r^i$

Until stopping criterion

Output

Signal estimate: $\hat{x}$, such that $\hat{x}_{T^i} = \Phi_{T^i}^\dagger y$ and $\hat{x}_{T^c} = 0$.

the succeeding iteration, indices in T_r^{i-1} , which is a now a part of T^{i-1} , will be dropped (if required). Then the support set T^i is formed¹ by merging the pruned support set T_p^{i-1} with

¹Note that T_r^i is the set of indices corresponding to $K - |T^{i-1}|$ largest magnitude entries in x_r^i , resulting from CSA. In the case of CR techniques such as BP, x_r^i contains more than $K - |T^{i-1}|$ non-zero entries, and therefore, T^i can include indices corresponding to $K - |T^{i-1}|$ largest magnitudes in x_r^i . However, to maintain

T_r^i . Unlike existing backtracking based algorithms, ICSR do not form a preliminary support set of size more than K , which will lead to convergence in lesser number of iterations. In ICSR, an estimate of size $|T^i| < K$ is maintained throughout the iterations until the stopping criterion, $|T^i| = K$, is met. In other words, ICSR stops when all $|T^{i-1}|$ non-zero entries of x_a^{i-1} have magnitude larger than $\|x_r^i\|_\infty$.

The ICSR framework is designed in such a way that it can accommodate any class of CSA: CR, NCR or GP. If the CS algorithm is BP, then the framework is called ICSR-BP. Similarly, ICSR-OMP, ICSR-CoSaMP, etc., are obtained.

5.2.3 ‘No-PKS’ Scenario

If PKS (i.e. T_{hp}) is not available, a regular CS reconstruction algorithm or a generalized framework such as Iterative Framework for Sparse Reconstruction Algorithm (IFSRA) [96] can be applied. For implementing ICSR in such scenarios (as summarized in algorithm 7), due to the absence of T_{hp} , T^0 can be obtained via the algorithmic function,

$$[x^0, T^0] = \text{CSA}(\Phi, y, \lfloor \lambda_0 K \rfloor). \quad (5.3)$$

Therefore, ICSR can reconstruct any sparse signal regardless of the availability of the PKS.

5.2.4 Complexity Analysis

If the participating CS reconstruction algorithm has a complexity, say $C(m, n, K)$, then the ICSR framework has a reconstruction complexity upper bounded by $d \times C(m, n, (1 - \lambda_0)K)$ where d is the number of ICSR iterations. This follows the fact that each iteration of ICSR framework has complexity of $C(m, n, K - |T^{i-1}|)$ which is upper bounded by $C(m, n, (1 - \lambda_0)K)$. For example, typical GP algorithms like OMP and SP have a complexity

our framework algorithm-blind, T^i includes only those indices corresponding to $K - |T^{i-1}|$ largest entries.

Algorithm 7 Iterative Compressive Sensing on Residual (Without PKS)

Input

Measurement matrix: Φ

Measurement vector: y

Sparsity level: K

Initial selection parameter: $\lambda_0 (< 1)$
Initial support set formation
 $i=0$ and $[x^i, T^i] = \text{CSA}(\Phi, y, \lfloor \lambda_0 K \rfloor)$
CS on Residual
Repeat

Step 1: Residual formation: $i=i+1$; $x_a^{i-1} = \Phi_{T^{i-1}}^\dagger y$ and $y_r^{i-1} = y - \Phi_{T^{i-1}} x_a^{i-1}$

Step 2: CS on residual: $[x_r^i, T_r^i] = \text{CSA}(\Phi, y_r^{i-1}, K - |T^{i-1}|)$

Step 3: Pruning: $T_p^{i-1} = T^{i-1} - \{l \in T^{i-1} : |x_a^{i-1}(l)| \leq \|x_r^i\|_\infty\}$

Step 4: Support merger: $T^i = T_p^{i-1} \cup T_r^i$
Until stopping criterion
Output

Signal estimate: $\hat{x}$, such that $\hat{x}_{T^i} = \Phi_{T^i}^\dagger y$ and $\hat{x}_{T^{ic}} = 0$.

$C(m, n, K) = O(f(m, n).K)$. Each iteration of ICSR requires atmost $O(f(m, n).(1 - \lambda_0)K)$ computations. Therefore, the complexity of ICSR is $O(d.f(m, n).(1 - \lambda_0)K)$. For the proposed framework to be computationally efficient, $d < \frac{1}{1-\lambda_0}$, which ensures its complexity is in the same order as that of the GP. Algorithms where $C(m, n, K)$ is not a function of K do not provide any computational gain due to the $1 - \lambda_0$ factor.

Choice of λ_0 : The initial selection parameter λ_0 could well be chosen based on the statistics of the signal support. As can be seen from Table I, around 70-75% of the K locations fall within the first $2K$ zig-zag locations. Hence, the typical range for λ_0 (in CI application) will be 0.7-0.9. However, there can be an extreme: $\lambda_0 = 0$. λ_0 is set to zero if no prior knowledge of high probability locations is available. In this case, $d = 1$ and the ICSR framework tends to behave like a regular CS reconstruction algorithm, so they both have equal complexity.

5.3 Theoretical Analysis of ICSR

The following result will be used few times. If $I, J \subset \{1, 2, \dots, n\}$ are two disjoint sets and $\delta_{|I|+|J|} < 1$, then for any arbitrary vector g ,

$$\|\Phi_I^\dagger \Phi_J g\|_2 \leq \frac{\delta_{|I|+|J|}}{1 - \delta_{|I|+|J|}} \|g\|_2. \quad (5.4)$$

It may be shown as follows:

$$\begin{aligned} \|\Phi_I^\dagger \Phi_J g\|_2 &= \|(\Phi_I^\star \Phi_I)^{-1} \Phi_I^\star \Phi_J g\|_2 \stackrel{(i)}{\leq} \frac{1}{1 - \delta_{|I|}} \|\Phi_I^\star \Phi_J g\|_2 \\ &\stackrel{(ii)}{\leq} \frac{\delta_{|I|+|J|}}{1 - \delta_{|I|}} \|g\|_2 \\ &\stackrel{(iii)}{\leq} \frac{\delta_{|I|+|J|}}{1 - \delta_{|I|+|J|}} \|g\|_2 \end{aligned}$$

where (i) follows RIP definition, (ii) follows the near orthogonality of columns (lemma 1 in [10]), and (iii) follows the monotonicity of RIP.

Lemma 5.1: As given in step 1 of algorithm 6, let $x_a^{i-1} = \Phi_{T^{i-1}}^\dagger y$ be the projection coefficient vector obtained using T^{i-1} . It can be shown that

$$\|x_a^{i-1} - x_{T^{i-1}}\|_2 \leq \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \|x_{T-T^{i-1}}\|_2 \quad (5.5)$$

where $\rho^{i-1} = \frac{|T \cap T^{i-1}|}{|T|}$.

Proof: Note that,

$$\begin{aligned} x_a^{i-1} &= \Phi_{T^{i-1}}^\dagger y = \Phi_{T^{i-1}}^\dagger \Phi_T x_T \\ &= \Phi_{T^{i-1}}^\dagger \Phi_T \cap T^{i-1} x_T \cap T^{i-1} + \Phi_{T^{i-1}}^\dagger \Phi_{T-T^{i-1}} x_{T-T^{i-1}} \\ &= \Phi_{T^{i-1}}^\dagger \Phi_{T^{i-1}} x_{T^{i-1}} + \Phi_{T^{i-1}}^\dagger \Phi_{T-T^{i-1}} x_{T-T^{i-1}} \\ &= x_{T^{i-1}} + \Phi_{T^{i-1}}^\dagger \Phi_{T-T^{i-1}} x_{T-T^{i-1}} \end{aligned}$$

where the last equality follows the definition of $\Phi_{T^{i-1}}^\dagger$. Therefore,

$$\begin{aligned}\|x_a^{i-1} - x_{T^{i-1}}\|_2 &= \|\Phi_{T^{i-1}}^\dagger \Phi_{T-T^{i-1}} x_{T-T^{i-1}}\|_2 \\ &\leq \frac{\delta_{|T^{i-1}|+|T-T^{i-1}|}}{1 - \delta_{|T^{i-1}|+|T-T^{i-1}|}} \|x_{T-T^{i-1}}\|_2\end{aligned}$$

where the above inequality follows (5.4). Using the definition of ρ^{i-1} and the fact that $|T^{i-1}| < K$, the bound above can be expressed as

$$\|x_a^{i-1} - x_{T^{i-1}}\|_2 \leq \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \|x_{T-T^{i-1}}\|_2. \quad \square$$

Lemma 5.2: Let x_r^i be the residual signal obtained using a CS reconstruction algorithm (as in step 2 of algorithm 6) and T_r^i be the estimated support. It can be shown that

$$\|x_r^i - x_{T_r^i}\|_2 \leq \frac{2\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} \|x_{T-T^i}\|_2 + \left[\frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right]^2 \|x_{T-T^{i-1}}\|_2 \quad (5.6)$$

where $\rho^i = \frac{|T \cap T^i|}{|T|}$.

Proof: As given in eq (24) of [10], the measurement residual can be expressed as

$$y_r^{i-1} = \Phi_{T-T^{i-1}} x_{T-T^{i-1}} + \Phi_{T^{i-1}} x_{p,T^{i-1}} \quad (5.7)$$

where $x_{p,T^{i-1}} = -\Phi_{T^{i-1}}^\dagger (\Phi_{T-T^{i-1}} x_{T-T^{i-1}})$. Since

$$\Phi_{T-T^{i-1}} x_{T-T^{i-1}} = \Phi_{T \cap T_r^i} x_{T \cap T_r^i} + \Phi_{T-T^i} x_{T-T^i} - \Phi_{T \cap \Delta T^i} x_{T \cap \Delta T^i}$$

where $\Delta T^i = T^{i-1} - T_p^{i-1}$, (5.7) becomes

$$y_r^{i-1} = \Phi_{T \cap T_r^i} x_{T \cap T_r^i} + \Phi_{T-T^i} x_{T-T^i} + \Phi_{T^{i-1}} x_{p,T^{i-1}} - \Phi_{T \cap \Delta T^i} x_{T \cap \Delta T^i}.$$

Now, let us write

$$\begin{aligned}
x_r^i &= \Phi_{T_r^i}^\dagger y_r^{i-1} \\
&= \Phi_{T_r^i}^\dagger \Phi_{T \cap T_r^i} x_T \cap T_r^i + \Phi_{T_r^i}^\dagger \Phi_{T-T^i} x_{T-T^i} + \Phi_{T_r^i}^\dagger \Phi_{T^{i-1}} x_{p, T^{i-1}} - \Phi_{T_r^i}^\dagger \Phi_{T \cap \Delta T^i} x_T \cap \Delta T^i \\
&= \Phi_{T_r^i}^\dagger \Phi_{T_r^i} x_{T_r^i} + \Phi_{T_r^i}^\dagger \Phi_{T-T^i} x_{T-T^i} + \Phi_{T_r^i}^\dagger \Phi_{T^{i-1}} x_{p, T^{i-1}} - \Phi_{T_r^i}^\dagger \Phi_{T \cap \Delta T^i} x_T \cap \Delta T^i \\
&= x_{T_r^i}^i + \Phi_{T_r^i}^\dagger \Phi_{T-T^i} x_{T-T^i} + \Phi_{T_r^i}^\dagger \Phi_{T^{i-1}} x_{p, T^{i-1}} - \Phi_{T_r^i}^\dagger \Phi_{T \cap \Delta T^i} x_T \cap \Delta T^i
\end{aligned}$$

where the last equality follows the definition of $\Phi_{T_r^i}^\dagger$. Therefore,

$$\begin{aligned}
\|x_r^i - x_{T_r^i}^i\|_2 &\leq \|\Phi_{T_r^i}^\dagger \Phi_{T-T^i} x_{T-T^i}\|_2 + \|\Phi_{T_r^i}^\dagger \Phi_{T^{i-1}} x_{p, T^{i-1}}\|_2 + \|\Phi_{T_r^i}^\dagger \Phi_{T \cap \Delta T^i} x_T \cap \Delta T^i\|_2 \\
&\stackrel{(i)}{\leq} \frac{\delta_{|T_r^i|+|T-T^i|}}{1 - \delta_{|T_r^i|+|T-T^i|}} \|x_{T-T^i}\|_2 + \frac{\delta_{|T_r^i|+|T^{i-1}|}}{1 - \delta_{|T_r^i|+|T^{i-1}|}} \|x_{p, T^{i-1}}\|_2 \\
&\quad + \frac{\delta_{|T_r^i|+|T \cap \Delta T^i|}}{1 - \delta_{|T_r^i|+|T \cap \Delta T^i|}} \|x_T \cap \Delta T^i\|_2 \\
&\stackrel{(ii)}{\leq} \frac{\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} \|x_{T-T^i}\|_2 + \frac{\delta_K}{1 - \delta_K} \|x_{p, T^{i-1}}\|_2 + \frac{\delta_K}{1 - \delta_K} \|x_T \cap \Delta T^i\|_2
\end{aligned}$$

where (i) follows from (5.4) and (ii) follows the facts $|T_r^i| + |T - T^i| = K + (1 - \rho^i)K$, $|T_r^i| + |T \cap \Delta T^i| \leq K$, and $|T_r^i| + |T^{i-1}| = K$. Since $T \cap \Delta T^i$ is a subset of $T - T^i$ and $\|x_{p, T^{i-1}}\|_2$ can be bounded using (5.4),

$$\begin{aligned}
\|x_r^i - x_{T_r^i}^i\|_2 &\leq \left(\frac{\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} + \frac{\delta_K}{1 - \delta_K} \right) \|x_{T-T^i}\|_2 + \frac{\delta_K}{1 - \delta_K} \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \|x_{T-T^{i-1}}\|_2 \\
&\leq \frac{2\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} \|x_{T-T^i}\|_2 + \left[\frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right]^2 \|x_{T-T^{i-1}}\|_2
\end{aligned} \tag{5.8}$$

where the last inequality follows the monotonicity of RIP. $\square$

5.3.1 Recurrence Relation of ICSR in the Case of Noiseless Measurements

Lemma 5.3 (BP on residual [20] - noiseless measurements): Let $y = \Phi x$. Assume that $\Phi \in \mathbb{R}^{m \times n}$ obeys RIP of order $2K$. Now, consider ICSR (algorithm 6) in which BP is used (in the place of CSA) to reconstruct $x - x_a^{i-1}$ from y_r^{i-1} . If in the i^{th} iteration, $\delta_{(2-\rho^{i-1})K} < \sqrt{2}-1$, then it can be shown that

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2$$

for some C_1^i given in the proof.

Proof: For BP on residual [20], if $\delta_{2|T-T^{i-1}|} < \sqrt{2} - 1$ and $\delta_{|T^{i-1}|} < \frac{1}{2}$, the reconstruction error can be bounded as

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_2^i \|x_{T-T^{i-1}}\|_2 \quad (5.9)$$

where the constant

$$C_2^i = 2\Theta_{|T^{i-1}|, K-|T \cap T^{i-1}|} \sqrt{\frac{|T^{i-1}|}{K - |T \cap T^{i-1}|} - 1} \frac{1 + 0.414\delta_{2(K-|T \cap T^{i-1}|)}}{1 - 2.414\delta_{2(K-|T \cap T^{i-1}|)}}$$

and $\Theta_{|T^{i-1}|, K-|T \cap T^{i-1}|}$ is the restricted orthogonality constant (defined in [47]). Using the definition of ρ^{i-1} , and using the facts $\Theta_{S, S'} \leq \delta_{S+S'}$, $2(K - |T \cap T^{i-1}|) = 2(1 - \rho^{i-1})K \leq (2 - \rho^{i-1})K$, and $|T^{i-1}| < K$, the bound in (5.9) can be written as,

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2 \quad (5.10)$$

where

$$C_1^i = \sqrt{\frac{4\rho^{i-1}}{1 - \rho^{i-1}}} \left(\frac{\delta_{(2-\rho^{i-1})K}(1 + 0.414\delta_{(2-\rho^{i-1})K})}{1 - 2.414\delta_{(2-\rho^{i-1})K}} \right). \quad \square$$

Theorem 5.1 (ICSR's recurrence relation - noiseless measurements): Let $y = \Phi x$. As-

sume that $\Phi \in \mathbb{R}^{m \times n}$ obeys RIP of order $2K$. Now, consider ICSR (algorithm 6) which uses BP (in the place of CSA). In the i^{th} iteration of ICSR, if $\delta_{(2-\min(\rho^{i-1}, \rho^i))K} < \frac{1}{3}$ and

$$\|x_{T^{i-1}-T_p^{i-1}}\|_2 \leq \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \|x_{T-T^{i-1}}\|_2, \quad (5.11)$$

then, it can be shown that

$$\|x_{T-T^i}\|_2 \leq C_3^i \|x_{T-T^{i-1}}\|_2$$

for some constant C_3^i given in the proof.

Proof: The reconstruction error resulting from CS on residual can be expressed as,

$$\begin{aligned} \|x - (x_a^{i-1} + x_r^i)\|_2 &= \|x - x_a^{i-1} - x_r^i\|_2 \\ &= \|x - (x_{T^{i-1}} + x_a^{i-1} - x_{T^{i-1}}) - (x_{T_r^i} + x_r^i - x_{T_r^i})\|_2 \\ &= \|x - x_{T^{i-1}} - x_{T_r^i} - (x_a^{i-1} - x_{T^{i-1}}) - (x_r^i - x_{T_r^i})\|_2 \\ &= \|x - (x_{T_p^{i-1}} + x_{T^{i-1}} - x_{T_p^{i-1}}) - x_{T_r^i} - (x_a^{i-1} - x_{T^{i-1}}) - (x_r^i - x_{T_r^i})\|_2 \\ &= \|x - x_{T_p^{i-1}} - x_{T_r^i} - (x_a^{i-1} - x_{T^{i-1}}) - (x_r^i - x_{T_r^i}) - (x_{T^{i-1}-T_p^{i-1}})\|_2 \\ &= \|x_{T-T^i} - (x_a^{i-1} - x_{T^{i-1}}) - (x_r^i - x_{T_r^i}) - (x_{T^{i-1}-T_p^{i-1}})\|_2 \end{aligned}$$

where the last equality follows from the definition of T^i . Using reverse triangle inequality,

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \geq \|x_{T-T^i}\|_2 - \|x_a^{i-1} - x_{T^{i-1}}\|_2 - \|x_r^i - x_{T_r^i}\|_2 - \|x_{T^{i-1}-T_p^{i-1}}\|_2. \quad (5.12)$$

Combining (5.12) with the result of lemma 5.3,

$$\|x_{T-T^i}\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2 + \|x_a^{i-1} - x_{T^{i-1}}\|_2 + \|x_r^i - x_{T_r^i}\|_2 + \|x_{T^{i-1}-T_p^{i-1}}\|_2. \quad (5.13)$$

Comparing (5.13) with the results of lemmas 5.1 and 5.2, and using the assumption (5.11),

$$\begin{aligned}
\|x_{T-T^i}\|_2 &\leq \left[C_1^i + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \left(2 + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right) \right] \|x_{T-T^{i-1}}\|_2 + \frac{2\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} \|x_{T-T^i}\|_2 \\
&\leq \frac{1 - \delta_{(2-\rho^i)K}}{1 - 3\delta_{(2-\rho^i)K}} \left[C_1^i + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \left(2 + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right) \right] \|x_{T-T^{i-1}}\|_2 \\
&\leq C_3^i \|x_{T-T^{i-1}}\|_2
\end{aligned}$$

where

$$C_3^i = \frac{1 - \delta^*}{1 - 3\delta^*} \left[C_1^i + \frac{\delta^*}{1 - \delta^*} \left(2 + \frac{\delta^*}{1 - \delta^*} \right) \right]$$

and $\delta^* = \delta_{(2-\min(\rho^{i-1}, \rho^i))K}$. □

The above result shows that, in each iteration of ICSR, the recurrence relation depends on δ^* which in turn depends on ρ^{i-1} and ρ^i . As ICSR progresses, with high probability, $\min(\rho^{i-1}, \rho^i)$ increases, which in turn reduces the order of the RIP requirement. If $C_3^i < 1$, then the ICSR satisfies the recurrence inequality that converges over iterations. For example, for $\delta^* = 0.1$ and $\rho^{i-1} = 0.5$, the value of C_3^i is 0.6547 which is less than 1.

5.3.2 Recurrence Relation of ICSR in the Case of Noisy Measurements

Lemma 5.4 (BP on residual [20] - noisy measurements): Let $y = \Phi x + v$ with $\|v\|_2 \leq \epsilon$. Assume that $\Phi \in \mathbb{R}^{m \times n}$ obeys RIP of order $2K$. Now, consider ICSR (algorithm 6) in which BP is used (in the place of CSA) to reconstruct $x - x_a^{i-1}$ from y_r^{i-1} . If in the i^{th} iteration, $\delta_{(2-\rho^{i-1})K} < \sqrt{2} - 1$, then it can be shown that

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2 + C_4^i \epsilon$$

for C_1^i given in lemma 5.3 and for some C_4^i given in the proof.

Proof: For BP on residual [20], if $\|v\|_2 \leq \epsilon$, $\delta_{2|T-T^{i-1}|} < \sqrt{2} - 1$, and $\delta_{|T^{i-1}|} < \frac{1}{2}$, the

reconstruction error can be bounded as

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_2^i \|x_{T-T^{i-1}}\|_2 + C_5^i \epsilon \quad (5.14)$$

where C_2^i is defined in lemma 5.3 and

$$C_5^i = \sqrt{\frac{2|T^{i-1}|}{K - |T \cap T^{i-1}|} - 1} \frac{1 + 0.414\delta_{2(K-|T \cap T^{i-1}|)}}{1 - 2.414\delta_{2(K-|T \cap T^{i-1}|)}} + \frac{4\sqrt{1 + \delta_{2(K-|T \cap T^{i-1}|)}}}{1 - 2.414\delta_{2(K-|T \cap T^{i-1}|)}}.$$

Using the definition of ρ^{i-1} , and using the facts $\Theta_{S,S'} \leq \delta_{S+S'}$, $2(K - |T \cap T^{i-1}|) = 2(1 - \rho^{i-1})K \leq (2 - \rho^{i-1})K$, and $|T^{i-1}| < K$, the bound in (5.14) can be written as,

$$\|x - (x_a^{i-1} + x_r^i)\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2 + C_4^i \epsilon \quad (5.15)$$

where C_1^i is defined in lemma 5.3 and

$$C_4^i = \sqrt{\frac{2\rho^{i-1}}{1 - \rho^{i-1}}} \left(\frac{1 + 0.414\delta_{(2-\rho^{i-1})K}}{1 - 2.414\delta_{(2-\rho^{i-1})K}} \right) + \frac{4\sqrt{1 + \delta_{(2-\rho^{i-1})K}}}{1 - 2.414\delta_{(2-\rho^{i-1})K}}. \quad \square$$

Theorem 5.2 (ICSR's recurrence relation - noisy measurements): Let $y = \Phi x + v$ with $\|v\|_2 \leq \epsilon$. Assume that $\Phi \in \mathbb{R}^{m \times n}$ obeys RIP of order $2K$. Now, consider ICSR (algorithm 6) which uses BP (in the place of CSA). Then, in the i^{th} iteration of ICSR, if $\delta_{(2-\min(\rho^{i-1}, \rho^i))K} < \frac{1}{3}$, it can be shown that

$$\|x_{T-T^i}\|_2 \leq C_3^i \|x_{T-T^{i-1}}\|_2 + C_6^i \epsilon$$

for C_3^i given in theorem 5.1 and for some C_6^i given in the proof.

Proof: Comparing (5.12) with the result of lemma 5.4 gives,

$$\|x_{T-T^i}\|_2 \leq C_1^i \|x_{T-T^{i-1}}\|_2 + \|x_{T^{i-1}-T_p^{i-1}}\|_2 + \|x_a^{i-1} - x_{T^{i-1}}\|_2 + \|x_r^i - x_{T_i^i}\|_2 + C_4^i \epsilon. \quad (5.16)$$

Using the results of lemmas 5.1 and 5.2, and the assumption (5.11),

$$\begin{aligned}
\|x_{T-T^i}\|_2 &\leq \left[C_1^i + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \left(2 + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right) \right] \|x_{T-T^{i-1}}\|_2 + \frac{2\delta_{(2-\rho^i)K}}{1 - \delta_{(2-\rho^i)K}} \|x_{T-T^i}\|_2 + C_4^i \epsilon \\
&\leq \frac{1 - \delta_{(2-\rho^i)K}}{1 - 3\delta_{(2-\rho^i)K}} \left(\left[C_1^i + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \left(2 + \frac{\delta_{(2-\rho^{i-1})K}}{1 - \delta_{(2-\rho^{i-1})K}} \right) \right] \|x_{T-T^{i-1}}\|_2 + C_4^i \epsilon \right) \\
&\leq C_3^i \|x_{T-T^{i-1}}\|_2 + C_6^i \epsilon
\end{aligned}$$

where C_3^i is defined in theorem 5.1 and

$$C_6^i = \frac{1 - \delta^*}{1 - 3\delta^*} C_4^i. \quad \square$$

If $C_3^i < 1$ and C_6^i are known constants, then the ICSR satisfies the recurrence inequality that converges over iterations. Continuing the same example, the value of C_3^i is 0.6547 which is less than 1, and the value of C_6^i is 9.6262.

5.4 Application: Compressive Imaging

In this section, image reconstruction performance of ICSR is evaluated. CS acquisition and reconstruction procedures are carried out block-wise (block size of 32×32) on each of the 24 images taken from KODAK database. An i.i.d. Gaussian random matrix is used for acquiring the vectored image block of length $n = 1024$. The 2-D DCT transformation matrix is used as the sparsifying dictionary. The sensing mechanism captures m observations of the vectorized image block (containing $K = \lfloor \frac{m}{\log n} \rfloor$ largest DCT coefficients). The measurement vector, y , is corrupted by a noise such that its SMNR is 20 dB. Since y is noisy, CR based methods (BP, Mod-BP, etc) have $\|\Phi\tilde{x} - y\|_2 \leq 0.001$ as their constraint. Since ICSR's image reconstruction performance is evaluated, for comparison purpose, Sparsity Averaged Reweighted Analysis (SARA) [97], a recently proposed algorithm for CI is also implemented. SARA, based on analysis reweighted l_1 -formulation, made use of the obser-

vation that natural images exhibit strong average sparsity over multiple frames.

Reconstruction Accuracy

First, the reconstruction quality of ICSR is presented as a function of MR. Table 5.2 shows the PSNR for different CS algorithms, their ICSR versions and for various PKS based algorithms. Since PKS based algorithms do not have any specific guideline to form a PKS, for CI application, we used T_0 (obtained in the initial support set estimation step) as PKS, to ensure an equal platform for all the algorithms. For ICSR, λ_0 is fixed as 0.8. The PSNR difference between a regular CS algorithm and its ICSR version is dramatic. In the case of BP, the proposed ICSR achieves a PSNR gain of at least 0.98 dB with the largest gain observed for MR=0.05. In the case of CoSaMP, ICSR-CoSaMP achieves a PSNR gain of at least 6 dB with the largest gain observed for MR=0.25. Notably, GPs (OMP, CoSaMP and IHT) achieve a huge PSNR gain. ICSR gives improvement in the case of FOCUSS too. As expected, all five CS algorithms plugged into ICSR framework give almost similar performance (as can be seen in figure 5.4) in terms of reconstruction accuracy. All five ICSR techniques give a better reconstruction accuracy compared to that of the OMP-PKS, CoSaMP-PKS, IHT-PKS, LS-CS and SARA. Among all the reported techniques, Mod-BP gives the best reconstruction accuracy. Its reconstruction accuracy is slightly better compared to that of the ICSR techniques.

Reconstruction Speed

The reconstruction/decoding time for all the above reported techniques is given in table 5.3. Since for GP, complexity decreases with decrease in sparsity level, one expects ICSR on low sparsity residual should take less time compared to regular versions. This is verified in table 5.3. Note that PKS versions of GP sometimes increase the complexity. However, ICSR with a GP is faster than the PKS version of the same GP. For example, ICSR-CoSaMP

Table 5.2: Average PSNR (in dB) versus MR: KODAK database

Method \ MR	0.05	0.1	0.15	0.2	0.25
OMP	19.23	20.85	21.79	22.89	23.71
CoSaMP	15.20	15.96	16.36	16.93	17.43
IHT	15.30	16.18	17.43	18.54	19.98
BP	19.96	21.94	23.06	24.12	24.78
FOCUSS	17.58	18.95	20.56	21.68	22.54
OMP-PKS	21.91	23.20	24.30	25.06	25.63
CoSaMP-PKS	22.06	23.10	23.85	24.47	24.92
IHT-PKS	21.78	23.18	24.18	24.82	25.23
Mod-BP	22.44	23.80	24.90	25.64	26.02
LS-CS	22.14	23.34	24.17	24.72	25.04
SARA	20.73	22.45	23.79	24.73	25.45
ICSR-OMP	22.12	23.46	24.36	25.23	25.70
ICSR-CoSaMP	22.19	23.57	24.59	25.28	25.79
ICSR-IHT	22.16	23.57	24.57	25.31	25.73
ICSR-BP	22.15	23.48	24.51	25.24	25.76
ICSR-FOCUSS	22.10	23.52	24.57	25.34	25.84

Figure 5.4: Average PSNR (in dB) versus MR: KODAK database

is around 30-40 times faster than the CoSaMP-PKS. The fact that the optimization based algorithms (Mod-BP, ICSR-FOCUSS and ICSR-BP) take longer time to converge is verified in table 5.3. *Sparselab* solver [98] is used for the implementation of BP in ICSR-BP. *cvx* solver [84] is used for the implementation of Mod-BP. Though Mod-BP gives a slightly better PSNR than the best performing ICSR technique, its decoding time is longer compared to that of the ICSR techniques. Among ICSR techniques, those involving a GP is much faster than those involving optimization based techniques. Average number of ICSR iterations (per block) is shown in figure 5.5.

Table 5.3: Average reconstruction time (computation time in MATLAB 7.12.0 running on a 64-bit Intel(R) CoreTM i5-2400 processor with 8 GB RAM) in seconds - KODAK database

Method \ MR	0.05	0.1	0.15	0.2	0.25
OMP	0.93	2.81	7.64	32.01	44.47
CoSaMP	0.25	0.49	0.84	0.74	3.19
IHT	1.77	2.16	3.06	3.77	3.11
BP	8.08	10.97	17.37	23.07	29.05
FOCUSS	34.23	51.68	72.75	100.94	126.98
OMP-PKS	1.2	3.39	8.92	18.20	32.48
CoSaMP-PKS	30.55	45.26	75.81	120.34	171.86
IHT-PKS	1.57	1.97	2.34	2.38	3.09
Mod-BP	60.32	328.41	573.80	805.87	908.91
LS-CS	6.29	8.10	11.87	15.68	19.56
SARA	4247.4	4307.9	4330.6	4588.1	4641.6
ICSR-OMP	0.71	1.31	2.21	12.25	21.25
ICSR-CoSaMP	0.88	1.06	1.70	2.47	4.04
ICSR-IHT	0.65	0.96	1.10	1.67	2.35
ICSR-BP	8.30	16.34	31.69	50.55	76.46
ICSR-FOCUSS	77.10	179.62	320.00	489.86	674.77

Effect of λ_0

Next, the effect of λ_0 on reconstruction accuracy is studied. MR is fixed as 0.25. $\lambda_0 = 0$ corresponds to regular CS (i.e., PKS is not utilized). An increase in λ_0 indicates more

Figure 5.5: Average number of ICSR iterations (per block) versus MR: KODAK database

correct locations are utilized. As can be seen in figure 5.6, ICSR framework gives better reconstruction when λ_0 is chosen between 0.8 and 0.9. This range of λ_0 seems to be effective for most natural images.

Visual Assessment

Next, visual assessment of the reconstruction quality is done, fixing MR=0.25. As can be seen in figures 5.7 till 5.10, all ICSR techniques and Mod-BP exhibit qualitatively closer performance. Though their reconstruction accuracies (in terms of PSNR) are not the same, the reconstructed images look similar.

‘No-PKS’ Scenario

For last experiment, ‘no-PKS’ scenario is considered. Reconstruction performance of ICSR is compared to that of the IFSRA, fixing $\lambda_0 = 0.4$. An implementation of IFSRA (for CoSaMP and BP) is available at <http://www.ece.iisc.ernet.in/~ssplab/Public/IFSRA.tar.gz>.

Figure 5.6: Effect of λ_0 on PSNR (in dB): KODAK database

Figure 5.7: ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (*kodim01* from KODAK database), reconstructed images for Mod-BP (24.33 dB), ICSR-OMP (24.05 dB), ICSR-CoSaMP (24.20 dB), ICSR-IHT (24.14 dB), ICSR-BP (24.13 dB) and ICSR-FOCUSS (24.14 dB).

Figure 5.8: ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (*kodim06* from KODAK database), reconstructed images for Mod-BP (25.50 dB), ICSR-OMP (25.14 dB), ICSR-CoSaMP (25.15 dB), ICSR-IHT (25.29 dB), ICSR-BP (25.11 dB) and ICSR-FOCUSS (25.28 dB).

Figure 5.9: ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (*kodim10* from KODAK database), reconstructed images for Mod-BP (27.19 dB), ICSR-OMP (27.08 dB), ICSR-CoSaMP (27.12 dB), ICSR-IHT (26.93 dB), ICSR-BP (27.07 dB) and ICSR-FOCUSS (27.19 dB).

Figure 5.10: ICSR with PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (*kodim21* from KODAK database), reconstructed images for Mod-BP (25.60 dB), ICSR-OMP (24.95 dB), ICSR-CoSaMP (25.08 dB), ICSR-IHT (24.98 dB), ICSR-BP (25.14 dB) and ICSR-FOCUSS (25.11 dB).

Figure 5.11: Average PSNR (in dB) versus MR - No PKS: KODAK database

Figure 5.12: Average reconstruction Time (in sec) versus MR - No PKS: KODAK database

Figure 5.13: ICSR without PKS (MR=0.25 and SMNR=20 dB): From left to right and top to bottom: original image (*kodim01* from KODAK database), reconstructed images for CoSaMP (16.73 dB), IFSRA-CoSaMP (17.43 dB), ICSR-CoSaMP (21.98 dB), BP (23.90 dB), IFSRA-BP (24.17 dB) and ICSR-BP (24.07 dB).

Figure 5.14: ICSR without PKS ($MR=0.25$ and $SMNR=20$ dB): From left to right and top to bottom: original image (*kodim06* from KODAK database), reconstructed images for CoSaMP (17.93 dB), IFSRA-CoSaMP (18.95 dB), ICSR-CoSaMP (23.54 dB), BP (25.11 dB), IFSRA-BP (25.10) and ICSR-BP (25.21 dB).

Figure 5.15: ICSR without PKS ($MR=0.25$ and $SMNR=20$ dB): From left to right and top to bottom: original image (*kodim10* from KODAK database), reconstructed images for CoSaMP (18.21 dB), IFSRA-CoSaMP (19.02 dB), ICSR-CoSaMP (24.54 dB), BP (26.03 dB), IFSRA-BP (26.24 dB) and ICSR-BP (26.27 dB).

Figure 5.16: ICSR without PKS ($MR=0.25$ and $SMNR=20$ dB): From left to right and top to bottom: original image (*kodim21* from KODAK database), reconstructed images for CoSaMP (17.82 dB), IFSRA-CoSaMP (18.35 dB), ICSR-CoSaMP (22.15 dB), BP (23.90 dB), IFSRA-BP (24.24 dB) and ICSR-BP (24.27 dB).

Figure 5.11 shows the reconstruction accuracy (in terms of average PSNR) and figure 5.12 shows the average reconstruction time. The reconstruction accuracy of ICSR-BP is slightly better than that of the IFSRA-BP. ICSR-BP is much faster compared to IFSRA-BP. On the other hand, the reconstruction accuracy of ICSR-CoSaMP is much better compared to that of the IFSRA-CoSaMP. There is a PSNR difference of at least 2 dB. ICSR-CoSaMP is slightly faster than IFSRA-CoSaMP. Visual assesment of the reconstructed images is done using figures 5.13 till 5.16. ICSR-CoSaMP exhibits a qualitatively much better performance compared to that of the IFSRA-CoSaMP and the regular CoSaMP. However, the optimization based algorithms (BP, ICSR-BP and IFSRA-BP) give a qualitatively similar performance.

5.5 Summary

In this chapter, an algorithm-blind CS reconstruction framework called ICSR is proposed. It first estimates the initial support set using the available PKS, and then iteratively performs CS on the measurement residual to update the support set estimate. The initial support set captures most of the signal energy. It contains the locations corresponding to larger transform coefficients and therefore, ICSR recovers the remaining locations accurately. Any CS reconstruction algorithm (CR, NCR or GP) can be plugged into ICSR framework. Therefore, ICSR is universal. Performance-wise, ICSR improves every CS reconstruction algorithm plugged into it. As expected, CS reconstruction algorithms with varying performance (in terms of reconstruction accuracy), when plugged into ICSR, gives a much similar performance. As summarized in figure 5.17, the key feature of the ICSR framework is the PSNR gain-execution speed combination of greedy algorithms (ICSR-CoSaMP, ICSR-OMP and ICSR-IHT), which makes them applicable in scenarios where fast recovery is required. In the case of signals whose PKS is not available, the initial support set can be formed as in (5.3), thereby increasing ICSR's flexibility with respect to the availability of PKS.

Figure 5.17: Speed and accuracy comparison: ICSR

Chapter 6

Greedy Pursuits Assisted Basis Pursuit

Fusion based CS reconstruction algorithms (such as FACS and CoMACS) combine multiple CS reconstruction algorithms, which work with different principles, to obtain a better signal estimate. However, these algorithms involve solving a least squares problem which might be ill-conditioned. In this chapter, a novel fusion technique called Greedy Pursuits Assisted Basis Pursuit (GPABP) is proposed which employs multiple GPs to derive a PKS for Mod-BP. Experimental results show that the proposed GPABP performs better than FACS and CoMACS. The proposed GPABP does not require a PKS and it is suitable for applications where no PKS is available. This chapter also focuses on reconstructing the JSM-1 signal ensemble (containing both common and innovation components) from MMV. With a goal of proposing an MMV-type algorithm that is robust to outliers (innovation components), an MMV version of GPABP is proposed, and termed as GPABP-MMV.

6.1 Shortcomings of Fusion-of-Algorithms

In CS, different reconstruction algorithms can be combined to improve sparse recovery. A fusion framework, FACS proposed in [40], employs two or more CS reconstruction algorithms and fuses the individual estimates to obtain a better sparse signal estimate.

Though FACS performs better than the best participating algorithm, it involves solving a least squares problem (step 5 in algorithm 2) which may be ill-conditioned. Two variants of FACS were proposed namely M-FACS [73] and CoMACS [41]. Though M-FACS attempts to ensure the least squares problem is well-conditioned, M-FACS could outperform FACS in the low measurement regime only. In CoMACS, since $\tilde{T}_C \subset \hat{T}$, any wrongly chosen atom in $\tilde{T}_C$ will be retained in $\hat{T}$. For every wrong index in $\tilde{T}_C$, there will be a correct index undetected. Therefore, this chapter focuses on proposing a fusion framework that could perform better than the FACS (and its variants).

6.2 Proposed Greedy Pursuits Assisted Basis Pursuit

PKS based algorithms (such as Mod-BP) ensure efficient reconstruction of a sparse signal provided some knowledge about the signal is available in the form of PKS. Recovery conditions and error bounds of Mod-BP revealed the fact that a good PKS will give a much better recovery. Since Mod-BP makes use of available signal knowledge to improve upon BP, a novel fusion technique called Greedy Pursuits Assisted Basis Pursuit (GPABP) is proposed which employs multiple GPs to derive a PKS, Λ , for Mod-BP. As Mod-BP makes use of PKS derived using GPs, the proposed algorithm is termed as GPABP. The RIP requirement for GPABP is expected to be favourable compared to that of FACS and CoMACS. Algorithm 8 shows the step-by-step procedure in GPABP.

6.2.1 Theoretical Analysis

Lemmas 6.1 and 6.2 give exact reconstruction condition and reconstruction error bound for GPABP, respectively. The RIC, δ_K , for an $n \times m$ matrix, Φ , is as defined in (2.1). The support of x , T , can be split as $T := \Lambda \cup \Delta_u \setminus \Delta_e$ where $\Delta_u := T \setminus \Lambda$ is the unknown part to be estimated and $\Delta_e := \Lambda \setminus T$ is the error in Λ .

Algorithm 8 Greedy Pursuits Assisted Basis Pursuit

Input

 Measurement matrix: Φ

 Measurement vector: y

 Sparsity level: K

 Number of GPs involved: L
Procedure

 Step 1: *for* $i=1:L$; $\hat{T}_i = \text{GP}_i(\Phi, y, K)$; *end*

 Step 2: *Common support*: $\Lambda = \bigcap_{i=1:L} \hat{T}_i$

 Step 3: Obtain $\hat{x}$ using Mod-BP:

$$\hat{x} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Lambda^c}\|_1 \text{ s.t. } \Phi \tilde{x} = y$$

Output
 $\hat{x}$: Signal estimate

Lemma 6.1 (GPABP Exact Reconstruction): Let $y = \Phi x$. If Λ is the derived *common support* such that $T = \Lambda \cup \Delta_u \setminus \Delta_e$, $\hat{x}$ is the unique minimizer of Mod-BP (in algorithm 8) if

$$|\Delta_u| \leq |\Lambda| \text{ and } \delta_{K+|\Delta_u|+|\Delta_e|} < \frac{1}{5}. \quad (6.1)$$

Above lemma is a modified version of the theorem 1 in [18]. Assuming $|\Delta_u| \leq |\Lambda|$, the worst case RIP requirement of GPABP may be obtained as follows. The condition in (6.1) is satisfied when $\delta_{|\Lambda|+2|\Delta_u|} < \frac{1}{5}$, which simplifies to $\delta_{3|\Lambda|} < \frac{1}{5}$. Since $|\Lambda| \leq K$, the RIP requirement becomes $\delta_{3K} < \frac{1}{5}$.

Lemma 6.2 (GPABP Error Bound): Let $y = \Phi x + v$. If $\|v\|_2 \leq \epsilon$ and $\delta_{K+|\Delta_u|+|\Delta_e|} < \sqrt{2} - 1$, then

$$\|x - \hat{x}\|_2 \leq D(K + |\Delta_u| + |\Delta_e|)\epsilon, \text{ where}$$

$$D(S) \triangleq \frac{4\sqrt{1+\delta_S}}{1-(\sqrt{2}+1)\delta_S}. \quad (6.2)$$

Above lemma is a modified version of the theorem 1 in [83]. Substituting $\delta_{K+|\Delta_u|+|\Delta_e|} < \frac{1}{5}$ in (6.2) will give, $\|x - \hat{x}\|_2 \leq 8.472\epsilon$.

Prior knowledge of the signal support may not be available in many scenarios, and hence Mod-BP cannot be applied directly. Also the general CS theory do not assume any such prior knowledge. For reconstructing such signals using Mod-BP, one might use $\Phi^\dagger y$ to obtain a PKS. But the reliability of such estimate is crucial to satisfy the conditions in lemma 6.1. The correct locations in the estimated support will improve recovery performance and at the same time, the wrong ones will degrade it as well. Eventually, GPABP shows how to use a PKS based reconstruction algorithm in scenarios where prior knowledge is not available.

GPABP can be applied not only for the sparse signals whose prior knowledge is unavailable but also for those signals whose partially known support is unreliable. Consider, for example, the time sequence reconstruction of JSM-1 signals having a common support size much lesser than K . In such cases, fusion step of GPABP is expected to give a better PKS compared to that obtained using the previous reconstructed signal. The *common support*, Λ , of GPABP is highly reliable as each of its entry is chosen by every ingredient GP.

6.2.2 Complexity Analysis

If each ingredient GP algorithm has a complexity, say $C(m, n, K)$, then the fusion framework has a complexity $\approx L \times C(m, n, K)$. It is well known that the computational complexity of GP will be much lower compared to that of a CR algorithm (BP, Mod-BP or LS-CS) [99]. For example, OMP and SP algorithms require $O(mnK)$ and $O(mn \log K)$ computations respectively, and on the other hand, convex relation algorithm requires $O(m^2 n^{\frac{3}{2}})$ computa-

Figure 6.1: Gradual Weighting Strategy of GP-W ℓ_1

tions. Therefore, the computational complexity of GPABP will be in the same order as that of the chosen modified BP algorithm. Low computational complexity of GP allows us to include more number of GPs to derive Λ . However, the size of Λ is non-increasing in L .

6.3 Proposed Greedy Pursuits Assisted Weighted ℓ_1

Mod-BP is chosen for the PKS based algorithm mainly due to its reconstruction accuracy. Mod-BP utilizes the PKS but does not retain it as a part of the support estimate (unlike PKS based greedy algorithms in [25]). Algorithms such as MABP, W ℓ_1 , and WA ℓ_1 can be used in the place of Mod-BP. However, GPABP has a limitation: the weighting is only binary (0 or 1 in the case of Mod-BP and ω or 1 in the case of W ℓ_1). In order to have more number of weights, a novel weighting strategy is proposed which assigns the weight for each location based on the number of GPs picking that location. This gradual weighting strategy is expected to give more precise weights compared to that of the binary weighting strategy. The proposed weighting strategy will be effective for large L (i.e. for more number of GPs). The resulting algorithm is termed as GP-W ℓ_1 . Algorithm 9 shows the step-by-step procedure in GP-W ℓ_1 . The weights are computed as given in step 2 of algorithm 9. Figure 6.1 shows the schematic of the gradual weighting strategy of GP-W ℓ_1 . The set, $[1, 2, \dots, n]$, is partitioned into $L + 1$ disjoint subsets N_l , $l \in 0, 1, \dots, L$ where N_l denotes the set of locations picked by exactly l (out of L) GPs.

Algorithm 9 Greedy Pursuits Assisted Weighted ℓ_1 -minimization

Input

Measurement matrix: Φ
 Measurement vector: y
 Sparsity level: K
 Number of GPs involved: L
 Error tolerance: ϵ

Procedure

Step 1: *for* $i=1:L$; $\hat{T}_i = \text{GP}_i(\Phi, y, K)$; *end*
 Step 2: *Weight vector*: $w' = [w'(1) \ w'(2) \ \dots \ w'(n)]$ where

$$w'(i) = 1 - \frac{G_i}{L}, \quad i \in 1, 2, \dots, n \quad (6.3)$$

and G_i is the number of GPs picking the location i

Step 3: Obtain $\hat{x}$ using Weighted- ℓ_1 :

$$\hat{x} = \arg \min_{\tilde{x}} \|\tilde{x}\|_{1, w'} \text{ s.t. } \|\Phi \tilde{x} - y\|_2 \leq \epsilon \quad (6.4)$$

Output

$\hat{x}$: Signal estimate

6.3.1 Theoretical Analysis

Theorem 6.1: Let $x \in \mathbb{R}^n$ and x_K be its best K -term approximation, supported on T . Assume L GPs are used for weights estimation and let $[1, 2, \dots, n]$ be partitioned into $L + 1$ disjoint subsets N_l , $l \in 0, 1, \dots, L$ where N_l mentioned as before denotes the set of locations picked by exactly l (out of L) algorithms. Suppose that there exists an $a \in \frac{1}{K}\mathbb{Z}$ that obeys $|N_1 \cup N_2 \cup \dots \cup N_L| = aK$, and the matrix Φ obeys RIP with

$$\begin{aligned} \delta_{aK} + \frac{a}{\sqrt{1 - \beta' + a(1 - \alpha')}} \delta_{(a+1)K} \\ \leq \frac{a}{\sqrt{1 - \beta' + a(1 - \alpha')}} - 1, \end{aligned}$$

for $\alpha' = \frac{|(N_1 \cup N_2 \cup \dots \cup N_L) \cap T|}{aK}$ and $\beta' = \frac{|N_L \cap T|}{K}$. Then the solution $\hat{x}$ to (6.4) obeys

$$\|\hat{x} - x\|_2 \leq D_1 \epsilon + D_2 \sum_{l=0}^{L-1} w_l \|x_{N_l \cap T^c}\|_1,$$

where D_1 and D_2 are constants (given in the proof below) that depend on Φ , a , α' and β' .

Proof: Let $\hat{x} = x + h$ be the minimizer of (6.4). Then

$$\|x + h\|_{1, w'} \leq \|x\|_{1, w'}.$$

Moreover, due to the choice of weights (given by (6.3)),

$$\begin{aligned} \sum_{l=0}^L w(l) \|x_{N_l} + h_{N_l}\|_1 &\leq \sum_{l=0}^L w(l) \|x_{N_l}\|_1. \\ \sum_{l=0}^L w(l) \|x_{N_l \cap T} + h_{N_l \cap T}\|_1 + \sum_{l=0}^L w(l) \|x_{N_l \cap T^c} + h_{N_l \cap T^c}\|_1 \\ &\leq \sum_{l=0}^L w(l) \|x_{N_l \cap T}\|_1 + \sum_{l=0}^L w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned}$$

Using forward and reverse triangle inequalities, the above expression becomes,

$$\sum_{l=0}^L w(l) \|h_{N_l \cap T^c}\|_1 \leq \sum_{l=0}^L w(l) \|h_{N_l \cap T}\|_1 + 2 \sum_{l=0}^L w(l) \|x_{N_l \cap T^c}\|_1.$$

By adding $\sum_{l=0}^L (1 - w(l)) \|h_{N_l \cap T^c}\|_1$ on both sides,

$$\begin{aligned} \|h_{T^c}\|_1 &\leq \sum_{l=0}^L w(l) \|h_{N_l \cap T}\|_1 + \sum_{l=0}^L (1 - w(l)) \|h_{N_l \cap T^c}\|_1 \\ &\quad + 2 \sum_{l=0}^L w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned}$$

Since $w(L) = 1 - w(0) = 0$

$$\begin{aligned} \|h_{T^c}\|_1 &\leq \sum_{l=0}^{L-1} w(l) \|h_{N_l \cap T}\|_1 + \sum_{l=1}^L (1 - w(l)) \|h_{N_l \cap T^c}\|_1 \\ &\quad + 2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned}$$

Noting that the weights range from 0 to 1,

$$\begin{aligned} \|h_{T^c}\|_1 &\leq \|h_{(N_0 \cup N_1 \cup \dots \cup N_{L-1}) \cap T}\|_1 \\ &\quad + \|h_{(N_1 \cup N_2 \cup \dots \cup N_L) \cap T^c}\|_1 + 2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned}$$

Let $\check{T} = (T - N_L) \cup (T^c - N_0)$. Then,

$$\|h_{T^c}\|_1 \leq \|h_{\check{T}}\|_1 + 2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1. \quad (6.5)$$

Let us define two parameters $\alpha' = \frac{|(N_1 \cup N_2 \cup \dots \cup N_L) \cap T|}{aK}$ and $\beta' = \frac{|N_L \cap T|}{K}$ such that $|\check{T}| = (1 - \beta' + a(1 - \alpha'))K$. Now, using the relation between the ℓ_1 -norm and ℓ_2 -norm,

$$\begin{aligned} \|h_{T^c}\|_1 &\leq \sqrt{(1 - \beta' + a(1 - \alpha'))K} \|h_{\check{T}}\|_2 \\ &\quad + 2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned} \quad (6.6)$$

As in [21], the coefficients of h_{T^c} are sorted in the order of decreasing magnitude partitioning T^c into disjoint sets T_j , $j \in \{1, 2, \dots\}$ each of size aK , where $a > 1$. That is, T_1 indexes the aK largest magnitude coefficients of h_{T^c} , T_2 indexes the second aK largest magnitude coefficients of h_{T^c} , and so on. This gives

$$\|h_{T_j}\|_2 \leq (aK)^{-1/2} \|h_{T_{j-1}}\|_1. \quad (6.7)$$

Let $T_{01} = T \cup T_1$, then using the triangle inequality in the above gives

$$\|h_{T_{01}^c}\|_2 \leq (aK)^{-1/2} \|h_{T^c}\|_1. \quad (6.8)$$

Since x and $\hat{x}$ are feasible, $\|\Phi h\|_2 \leq 2\epsilon$ and

$$\|\Phi h_{T_{01}}\|_2 \leq 2\epsilon + \sum_{j>1} \|\Phi h_{T_j}\|_2 \leq 2\epsilon + \sqrt{1 + \delta_{aK}} \sum_{j>1} \|h_{T_j}\|_2$$

where the last inequality follows RIP. Since $\sum_{j>1} \|h_{T_j}\|_2 = \|h_{T_{01}^c}\|_2$, using (6.8) in the above equation gives

$$\|\Phi h_{T_{01}}\|_2 \leq 2\epsilon + \frac{\sqrt{1 + \delta_{aK}}}{\sqrt{aK}} \|h_{T^c}\|_1. \quad (6.9)$$

Using RIP on the R.H.S. of the above we get

$$\sqrt{1 - \delta_{(a+1)K}} \|h_{T_{01}}\|_2 \leq 2\epsilon + \frac{\sqrt{1 + \delta_{aK}}}{\sqrt{aK}} \|h_{T^c}\|_1. \quad (6.10)$$

Since T_1 contains the largest aK coefficients of h_{T^c} , $\|h_{\bar{T}}\|_2 \leq \|h_{T_{01}}\|_2$, and therefore (6.6) becomes

$$\begin{aligned} \|h_{T^c}\|_1 &\leq \sqrt{(1 - \beta' + a(1 - \alpha'))K} \|h_{T_{01}}\|_2 \\ &\quad + 2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1. \end{aligned} \quad (6.11)$$

Combining (6.10) and (6.11) gives

$$\|h_{T_{01}}\|_2 \leq \frac{2\epsilon + 2 \frac{\sqrt{1 + \delta_{aK}}}{\sqrt{aK}} \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1}{\sqrt{1 - \delta_{(a+1)K}} - \frac{\sqrt{1 - \beta' + a(1 - \alpha')}}{\sqrt{a}} \cdot \sqrt{1 + \delta_{aK}}} \quad (6.12)$$

Finally, using the relation $\|h\|_2 \leq \|h_{T_{01}}\|_2 + \|h_{T_{01}^c}\|_2$, and substituting (6.12), (6.8) and

(6.11) in it gives

$$\|h\|_2 \leq D_1 \epsilon + D_2 \sum_{l=0}^{L-1} w(l) \|x_{N_l \cap T^c}\|_1 \quad (6.13)$$

where the constants

$$D_1 = \frac{2 \left(1 + \frac{\sqrt{1-\beta'+a(1-\alpha')}}{\sqrt{a}} \right)}{\sqrt{1-\delta_{(a+1)K}} - \frac{\sqrt{1-\beta'+a(1-\alpha')}}{\sqrt{a}} \sqrt{1+\delta_{aK}}} \quad (6.14)$$

and

$$D_2 = \frac{2 \frac{\sqrt{1-\delta_{(a+1)K}} + \sqrt{1+\delta_{aK}}}{\sqrt{aK}}}{\sqrt{1-\delta_{(a+1)K}} - \frac{\sqrt{1-\beta'+a(1-\alpha')}}{\sqrt{a}} \sqrt{1+\delta_{aK}}}, \quad (6.15)$$

with the condition that the denominator is positive, equivalently

$$\begin{aligned} \delta_{aK} + \frac{a}{\sqrt{1-\beta'+a(1-\alpha')}} \delta_{(a+1)K} \\ \leq \frac{a}{\sqrt{1-\beta'+a(1-\alpha')}} - 1. \quad \square \end{aligned}$$

Remark 1: If x is exactly sparse (i.e. $\|x_{N_l \cap T^c}\|_1 = 0$) and y is noise-free (i.e. $\epsilon = 0$), then $\|h\|_2 = 0$ which implies the reconstruction is exact.

Remark 2: For theorem 6.1 to hold, it is sufficient if Φ satisfies

$$\delta_{(a+1)K} < \frac{a - (1 - \beta' + a(1 - \alpha'))}{a + (1 - \beta' + a(1 - \alpha'))}. \quad (6.16)$$

Compared to the standard ℓ_1 -minimization [2] which requires Φ to satisfy $\delta_{(a+1)K} < \frac{a-1}{a+1}$, GP-W ℓ_1 has a weaker requirement provided

$$\beta' > a(1 - \alpha'). \quad (6.17)$$

As per the definitions of a , α' , and β' , the above condition holds when the ingredient GP

algorithms in GP- $W\ell_1$ (i.e., GP_i in step 1 of algorithm 9) estimate the signal support such that $|(N_1 \cup N_2 \cup \dots \cup N_L) \cap T| + |N_L \cap T| > |N_1 \cup N_2 \cup \dots \cup N_L|$.

6.3.2 Complexity Analysis

The computational complexity of GPs will be much lower compared to that of the $W\ell_1$ [99]. Therefore, the computational complexity of GP- $W\ell_1$ will be in the same order as that of the $W\ell_1$. Low computational complexity of GPs allows us to include more GPs (large L) to derive w' . The fact that the reconstruction accuracy of GP- $W\ell_1$ increases with an increase in L can be verified from figure 6.9 and figure 6.10.

6.4 Reconstruction of Sparse Signals: GPABP

In this section, experimental results are presented for GPABP based reconstruction of synthetic sparse signals and real compressible signals such as ECG signals. Signal reconstruction is performed using four methods: FACS, M-FACS, CoMACS and proposed GPABP. Except for GPABP, in all other methods, fusion involved three algorithms, OMP, SP and BP. In the case of GPABP, fusion step involved two algorithms: OMP and SP for GP_1 and GP_2 . For M-FACS, the condition number threshold is fixed as 10 (as in [73]). The *Sparselab* solver [98] is used for the implementation of BP in FACS, M-FACS and CoMACS. The *cvx* solver [84] is used for the implementation of Mod-BP in GPABP.

6.4.1 Synthetic Sparse Signals

For our experiments, Gaussian sparse signals of length $n = 250$ are generated. Note that y is noiseless and therefore, BP (in FACS, CoMACS and M-FACS) and Mod-BP (in GPABP) have $\Phi\tilde{x} = y$ as their constraint. First, the reconstruction performance is presented as a function of SR. For each value of K , 250 independent trials are performed to obtain the

Figure 6.2: Reconstruction probability versus SR

Figure 6.3: Average MSE versus SR

Figure 6.4: Reconstruction probability versus MR

Figure 6.5: Average MSE versus MR

average results. In each trial, an $m \times n$ Gaussian random measurement matrix is generated. Number of measurements m is fixed to be 125 and the sparsity levels are chosen from $K=35$ to $K=75$ in steps of 5. If the maximum magnitude difference between the original signal and the reconstructed signal is smaller than 0.01, the reconstruction is considered to be perfect. Figure 6.2 shows that our proposed GPABP has the best probability of exact reconstruction among all four methods. Surprisingly, the performances of FACS and CoMACS are indistinguishable (as can be seen in figure 6.2). Also, the performance of M-FACS is worse compared to that of FACS. For the same parameters, average MSE plot is shown in figure 6.3. It can be noticed that GPABP has the least MSE among all the methods and is stable even for high SR values. For SRs greater than 0.25, CoMACS and M-FACS perform better than FACS.

Next, the reconstruction performance is presented as a function of MR. K is fixed as 30 and the MR is varied between 0.2 and 0.48. Figure 6.4 shows the probability of exact reconstruction as a function of MR. For $MR > 0.3$, GPABP gives the best reconstruction among all four methods. Figure 6.5 shows the average MSE as a function of MR. For any MR, GPABP's reconstruction error is the least among all four methods. Though all four methods used the same number (and type) of ingredient algorithms, proposed GPABP gives the best performance.

6.4.2 Compressible ECG Signals

Next experiment illustrates the performance for real world compressible signals. The leads (ECG signals) are extracted from records 100, 101, 102 and 103 from the MIT-BIH Arrhythmia database [100]. They are processed in chunks of 256 samples (with amplitudes ranging from 0 to 255). These signals will serve as the ground truth for MSE computation. DCT is used to obtain sparse representations of the signals. The sparsity level K is fixed as $\lfloor \frac{m}{\log n} \rfloor$. For each chunk, the measurement vector y is corrupted by a noise such that its

Figure 6.6: Average MSE versus MR: Compressible ECG signals (SMNR = 20 dB)

SMNR is 20 dB. Since y is noisy, BP (in FACS, CoMACS and M-FACS) and Mod-BP (in GPABP) have $\|\Phi\tilde{x} - y\|_2 \leq 0.001$ as their constraint. Figure 6.6 shows the average MSE (MSE per chunk) as a function of MR (varied between 0.2 and 0.4). It is evident that the GPABP gives better recovery performance (compared to FACS and its variants) in the case of real world compressible signals too. Again, the performances of FACS and CoMACS are indistinguishable.

6.5 Reconstruction of Sparse Signals: $\text{GP-W}\ell_1$

In this section, experimental results revealing the significance of $\text{GP-W}\ell_1$ are presented. Experiment set-up is the same as that used for generating figure 6.4 and figure 6.5. Signal reconstruction is performed using three methods: $\text{ReW}\ell_1$ [62] (for various number of iterations), GPABP [45] and the proposed $\text{GP-W}\ell_1$. Both GPABP and $\text{GP-W}\ell_1$ involve four GPs (OMP, SP, IHT and BAOMP). The *cvx* solver [84] is used for the implementation of

$W\ell_1$ in all three methods. y is noiseless and therefore, $W\ell_1$ has $\Phi\tilde{x} = y$ as its constraint. Figure 6.7 shows that the GP- $W\ell_1$ has the best probability of exact reconstruction among all three methods provided $MR > 0.3$. Time taken by each method for reconstruction is shown in figure 6.8. As expected, GP- $W\ell_1$ converges much faster compared to Re $W\ell_1$.

Figure 6.7: Reconstruction probability versus MR

Next, the effect of L on the reconstruction performance of GP- $W\ell_1$ is presented. Experiment set-up is the same as that of the previous experiment and the GP- $W\ell_1$'s performance is recorded for $L = 2, 3$ and 4. In each trial, support set estimates are obtained for all four GPs mentioned above. In the case of $L < 4$, only L randomly chosen support sets are used for weights estimation. It can be seen from figure 6.9 and figure 6.10 that the reconstruction accuracy of GP- $W\ell_1$ increases with an increase in L .

Figure 6.8: Average reconstruction time (in sec) versus MR

Figure 6.9: Effect of L on reconstruction probability

Figure 6.10: Effect of L on average MSE

6.6 Proposed Greedy Pursuits Assisted Basis Pursuit for Multiple Measurement Vectors

DCS is an extension of CS from SMV problem to MMV problem. The data matrix for an ensemble of J signals is given in (2.11). For joint sparse signals sharing common support (i.e. JSM-2 formulation), several algorithms have been developed to reconstruct $\mathbf{X}$ from $\mathbf{Y}$, utilizing the common sparsity condition [29]-[35]. However, in most cases this knowledge does not exist and the assumption of common sparsity pattern for 100% of the data becomes far from ideal. Therefore, we are more interested in JSM-1 formulation. In [36], a SMV-type solution was proposed for reconstructing JSM-1 signals which solves a single linear program. However, due to concatenation of measurement vectors and concatenation of measurement matrices, complexity of the algorithm is high. An MMV-type recovery method for JSM-1, dubbed the Texas Hold 'Em algorithm [37], separated the recovery of

the common and innovation components into two stages. The computational complexity of the Texas Hold 'Em algorithm is linear in number of signals J . However, its reconstruction accuracy is a concern. To recover two correlated signals with partially disjoint supports, a modified Orthogonal Matching Pursuit (OMP) algorithm called Probability weighted OMP (P-OMP) was proposed [38]. However, it has a drawback: it requires prior information of the signals. In [39], a Robust Multiple Sparse Bayesian Learning (R-MSBL) algorithm was proposed which captures the support set of the majority of data vectors. Though R-MSBL, unlike P-OMP, does not require any probabilistic information about the signal, the time taken for reconstruction is a concern. To fix these concerns, an MMV algorithm that is robust to outliers (innovation components) is sought.

With a goal of proposing a robust MMV-type algorithm for reconstruction of JSM (JSM-1 in particular) signals, GPABP is extended to GPABP-MMV. As given in algorithm 10, in the first step of GPABP-MMV, multiple (say L) MMV-GPs (i.e. MMV versions of multiple greedy CS algorithms [101]) are applied, and their support estimates are obtained. As in committee machine approach [41], a common support Λ_1 is formed by taking the intersection of all these support estimates. Then, Λ_1 is utilized to compute the residual data matrix which in turn is used to compute the residual norm vector (steps 3 and 4 of algorithm 10). Note that Φ_I denotes the sub-matrix obtained by extracting the columns of Φ corresponding to the indices in I . $\Phi_I^\dagger$ denotes the pseudo-inverse of Φ_I .

The residual norm vector is sorted in ascending order (as in step 5 of algorithm 10) to form the set of sorted indices, φ . In other words, φ contains the order of signals with increasing residual vector norm. Then, the JSM-1 signals are reconstructed sequentially by applying Mod-BP, in the order given by φ . While reconstructing the first signal, $x_{\varphi(1)}$, Λ_1 is used as the PKS. For reconstruction of subsequent signals, $x_{\varphi(i)}$, the PKS is formed by taking the intersection of Λ_{i-1} and the previous support estimate, $\hat{T}_{\varphi(i-1)}$. This choice of PKS, with high probability, increases the accuracy of PKS which is very crucial for Mod-

Algorithm 10 Greedy Pursuits Assisted Basis Pursuit for Multiple Measurement Vectors

Input

Measurement matrix: Φ
 Concatenation of Measurement vectors: Y
 Sparsity level: K
 Number of GPs involved: L
 Number of signals in the ensemble: J
 Error tolerance: ϵ

Procedure

Step 1: *for* $i=1:L$; $\hat{T}_i = \text{MMV-GP}_i(\Phi, Y, K)$; *end*
 Step 2: *Common support*: $\Lambda_1 = \bigcap_{i=1:L} \hat{T}_i$
 Step 3: Residual data matrix: $Y_{res} = Y - \Phi_{\Lambda_1} \Phi_{\Lambda_1}^\dagger Y$
 Step 4: Residual norm vector: $[y_{r,1}, y_{r,2}, \dots, y_{r,J}]$ where $y_{r,i}$ denotes the ℓ_2 -norm of the i^{th} column of Y_{res}
 Step 5: Sort the residual norm vector:

$$\varphi = \text{sort-ascend}([y_{r,1}, y_{r,2}, \dots, y_{r,J}])$$

and initialize $i = 1$

repeat

Step 6: Obtain $\hat{x}_{\varphi(i)}$ using Mod-BP:

$$\hat{x}_{\varphi(i)} = \arg \min_{\tilde{x}} \|\tilde{x}_{\Lambda_i^c}\|_1 \text{ s.t. } \|\Phi \tilde{x} - y_{\varphi(i)}\|_2 \leq \epsilon$$

Step 7: $\hat{T}_{\varphi(i)}$ = indices corresponding to the K largest magnitude entries in $\hat{x}_{\varphi(i)}$

Step 8: Update: $\Lambda_{i+1} = \Lambda_i \cup \hat{T}_{\varphi(i)}$ and $i = i + 1$

until $i \leq J$

Output

$\hat{X}$: Concatenation of signal estimates

BP. The purpose of sorting signals based on their residual vector norm is to reconstruct them in the order of effectiveness of the PKS. Since Δ_u and Δ_e are defined such that $|\Lambda| = K - |\Delta_u| + |\Delta_e|$, Δ_u^j and Δ_e^j are defined such that $|\Lambda_j| = K - |\Delta_u^j| + |\Delta_e^j|$. Performance of GPABP-MMV is theoretically analysed via exact reconstruction condition (theorem 6.2) and reconstruction error bound (theorem 6.3).

6.6.1 Theoretical Analysis

Lemma 6.3: If $\forall i \in [2, J]$, the PKS used for reconstructing $x_{\varphi(i)}$, Λ_i , obeys

$$\Lambda_i \subset \Lambda_{i-1} \quad \text{and} \quad |T_{\varphi(i)} \cap (\Lambda_{i-1} \setminus \Lambda_i)| < \frac{|\Lambda_{i-1} \setminus \Lambda_i|}{2} \quad (6.18)$$

where $\Lambda_i = T_{\varphi(i)} \cup \Delta_e^i \setminus \Delta_u^i$ and $\Lambda_{i-1} = T_{\varphi(i-1)} \cup \Delta_e^{i-1} \setminus \Delta_u^{i-1}$, the restricted isometry constant $\delta_{K+|\Delta_u^i|+|\Delta_e^i|}$ is a decreasing function of i i.e.,

$$\delta_{K+|\Delta_u^i|+|\Delta_e^i|} < \delta_{K+|\Delta_u^{i-1}|+|\Delta_e^{i-1}|} < \delta_{K+|\Delta_u^1|+|\Delta_e^1|}.$$

Proof: While reconstructing $x_{\varphi(i)}$, the size of the unknown support (misses) can be expressed as,

$$|\Delta_u^i| = |\Delta_u^{i-1}| + |T_{\varphi(i)} \cap (\Lambda_{i-1} \setminus \Lambda_i)|.$$

If $|T_{\varphi(i)} \cap (\Lambda_{i-1} \setminus \Lambda_i)| < \frac{|\Lambda_{i-1} \setminus \Lambda_i|}{2}$, the above expression becomes

$$|\Delta_u^i| < |\Delta_u^{i-1}| + \frac{|\Lambda_{i-1} \setminus \Lambda_i|}{2}. \quad (6.19)$$

This condition on $T_{\varphi(i)} \cap (\Lambda_{i-1} \setminus \Lambda_i)$ also ensures that,

$$|\Delta_e^i| < |\Delta_e^{i-1}| - \frac{|\Lambda_{i-1} \setminus \Lambda_i|}{2}. \quad (6.20)$$

Combining (6.19) and (6.20) gives

$$|\Delta_u^i| + |\Delta_e^i| < |\Delta_u^{i-1}| + |\Delta_e^{i-1}|. \quad (6.21)$$

Since δ_S is a monotonically increasing function of S , using (6.21), it can be said that

$$\delta_{K+|\Delta_u^i|+|\Delta_e^i|} < \delta_{K+|\Delta_u^{i-1}|+|\Delta_e^{i-1}|} < \delta_{K+|\Delta_u^1|+|\Delta_e^1|},$$

or in general, $\delta_{K+|\Delta_u^i|+|\Delta_e^i|}$ is a decreasing function of i . $\square$

Theorem 6.2 (GPABP-MMV Exact Reconstruction): Let $\mathbf{Y} = \Phi \mathbf{X}$. Assume that $\forall i \in [2, J]$, Λ_i obeys (6.18). If Λ_1 is such that $T_{\varphi(1)} = \Lambda_1 \cup \Delta_u^1 \setminus \Delta_e^1$ is the *common support* obtained via MMV-GPs, then $\forall i \in [1, J]$ $\hat{x}_{\varphi(i)}$ is the unique minimizer of Mod-BP if

$$|\Delta_u^J| \leq |\Lambda_J| \text{ and } \delta_{K+|\Delta_u^1|+|\Delta_e^1|} < \frac{1}{5}. \quad (6.22)$$

Proof: Using (6.1), conditions for exact reconstruction of $x_{\varphi(i)}$ can be written as

$$|\Delta_u^i| \leq |\Lambda_i| \text{ and } \delta_{K+|\Delta_u^i|+|\Delta_e^i|} < \frac{1}{5}.$$

It is well known that $|\Delta_u^i|$ is a non-decreasing function of i . On the other hand, if Λ_i obeys (6.18), as given by lemma 6.3, $\delta_{K+|\Delta_u^i|+|\Delta_e^i|}$ is a decreasing function of i . Therefore, exact reconstruction conditions for the entire signal ensemble can be expressed as

$$|\Delta_u^J| \leq |\Lambda_J| \text{ and } \delta_{K+|\Delta_u^1|+|\Delta_e^1|} < \frac{1}{5}. \quad \square$$

Theorem 6.3 (GPABP-MMV error bound): Let $\mathbf{Y} = \Phi \mathbf{X} + \mathbf{V}$. Assume that $\forall i \in [2, J]$, Λ_i obeys (6.18). If $\forall i \in [1, J]$ $\|v_i\|_2 \leq \epsilon$, and $\delta_{K+|\Delta_u^1|+|\Delta_e^1|} < \sqrt{2} - 1$, then

$$\|X - \hat{X}\|_F \leq \sqrt{J}D(K + |\Delta_u^1| + |\Delta_e^1|)\epsilon,$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $D(S)$ is as defined in (6.2).

Proof: Using (6.2), the reconstruction error bound for each signal, $x_{\varphi(i)}$, is given by

$$\|x_{\varphi(i)} - \hat{x}_{\varphi(i)}\|_2 \leq D(K + |\Delta_u^i| + |\Delta_e^i|)\epsilon,$$

where $D(S)$ is as defined in (6.2). The reconstruction error for the entire signal ensemble can be expressed as

$$\begin{aligned} \|X - \hat{X}\|_F &= \sqrt{\sum_{i=1}^J \|x_{\varphi(i)} - \hat{x}_{\varphi(i)}\|_2^2} \\ &\leq \sqrt{\sum_{i=1}^J [D(K + |\Delta_u^i| + |\Delta_e^i|)]^2 \epsilon^2} \end{aligned} \quad (6.23)$$

Since $\forall i \in [2, J]$, Λ_i obeys (6.18), using lemma 6.3, the error bound can be written as

$$\begin{aligned} \|X - \hat{X}\|_F &\leq \sqrt{J[D(K + |\Delta_u^1| + |\Delta_e^1|)]^2 \epsilon^2} \\ &\leq \sqrt{J} D(K + |\Delta_u^1| + |\Delta_e^1|) \epsilon. \quad \square \end{aligned}$$

As Mod-BP has the provision of picking any atom regardless of its presence in PKS, GPABP-MMV will be robust to outliers (innovation components). Moreover, being sequential in nature, GPABP-MMV reconstruction will also be stable for large J . As mentioned earlier, GPABP-MMV does not require the size of the common support, K_c , to be known. Only K is required. K , being marginal statistics, can be estimated easily. On the other hand, K_c , being joint statistics, cannot be estimated that easily. GPABP-MMV is designed for JSM-1 model and it can handle JSM-2 as well.

Computation Complexity: If the MMV-GP algorithms have a complexity, say $C(m, n, K, J)$, then the fusion framework has a complexity $L \times C(m, n, K, J)$. Therefore, if Mod-BP has a complexity, say $C'(m, n)$, the computational complexity of GPABP-MMV will be

$$C_{\text{GPABP-MMV}} = L \times C(m, n, K, J) + J \times C'(m, n).$$

Since $C(m, n, K, J) \ll C'(m, n)$, and typically $L \ll J$, GPABP-MMV's complexity becomes $\approx J \times C'(m, n)$. Low computational complexity of MMV-GPs allows us to have a large L . However, the accuracy of Λ_1 does not vary much when L is increased beyond 2 [45].

6.7 Reconstruction of Correlated Sparse (JSM-1) Signals: GPABP-MMV

For the experiments reported in this section, JSM recovery is performed using four methods: Robust-MSBL, Texas Hold 'Em algorithm, GPABP-separate and the proposed GPABP-MMV. For Texas Hold 'Em, the fraction of community measurements is fixed as 75, and BP is used for the CS reconstruction algorithm. GPABP-separate refers to the sequential but independent recovery of each signal using GPABP. For both GPABP and GPABP-MMV, L is fixed as 2. In the case of GPABP, fusion step involves OMP and Subspace Pursuit (SP) for GP_1 and GP_2 respectively. For GPABP-MMV, analogous to OMP and SP used in GPABP, MMV-OMP and MMV-SP algorithms are used in the place of MMV- GP_1 and MMV- GP_2 respectively [102]. The *Sparselab* solver [98] is used for the implementation of BP in Texas Hold 'Em. The *cvx* solver [84] is used for the implementation of Mod-BP in GPABP and GPABP-MMV.

6.7.1 Synthetic JSM-1 Signals

Correlated sparse signals (as in JSM-1) of length $n = 250$ are generated. The sparsity (common + innovation components) of each signal is fixed as 20. In each experiment, 250 independent trials are performed to obtain the average results. In each trial, an $m \times n$ Gaussian random measurement matrix is generated. y is noiseless and therefore, BP (in Texas Hold 'Em) and Mod-BP (in GPABP-MMV and GPABP-separate) have $\Phi\tilde{x} = y$ as their constraint. In the first experiment, the effect of CSR on reconstruction accuracy

(expressed in terms of MSE) is presented. m and J are fixed as 64 and 10 respectively. By choosing different K_c values, signal recovery is performed using all four techniques. Average MSE plot is shown in figure 6.11. GPABP-MMV gives the least MSE among all the methods and is also stable for low K_c values. Figure 6.12 shows the reconstruction time (computation time in MATLAB 7.12.0 running on a 64-bit Intel(R) CoreTM i5-2400 processor with 8 GB RAM) for the same experiment. For $CSR < 0.8$, GPABP-MMV clearly outperforms Robust-MSBL (GPABP-MMV's closest rival in terms of accuracy) in terms of reconstruction speed.

Figure 6.11: Average MSE versus CSR

In the next experiment, the effect of J on MSE is presented. K_c and m are fixed as 15 and 64 respectively. Average MSE versus J plot is shown in figure 6.13. It can be seen that GPABP-MMV gives the least MSE for J chosen above 5. It is interesting to note that, as J increases, R-MSBL becomes worse in terms of reconstruction accuracy. Next, the effect of MR on MSE is shown. K_c and J are fixed as 15 and 10 respectively. Signal recovery is

Figure 6.12: Average reconstruction time (in sec) versus CSR

Figure 6.13: Average MSE versus J

Figure 6.14: Average MSE versus MR

performed for different values of MR (chosen between 0.15 and 0.4). From figure 6.14, it is evident that GPABP-MMV gives the least MSE for any MR greater than 0.2.

6.7.2 Time-sequence Reconstruction of Compressible ECG Signals

The leads (ECG signals) are extracted from records 100, 101, 102 and 103 from the MIT-BIH Arrhythmia database [100]. They are divided into chunks of $n = 900$ samples (with amplitudes ranging from 0 to 255). This choice of n is based on the periodicity of the signals obtained from their annotations. These signals will serve as the ground truth for MSE computation. Discrete cosine transform is used to obtain sparse representations of the signals. Joint-sparse signal ensemble (similar to JSM-1) is formed using J chunks of these sparse signals. J and K are fixed as 50 and $\lfloor \frac{m}{\log n} \rfloor$ respectively. For each signal in the ensemble, y is obtained using a Gaussian random measurement matrix. y is then corrupted by a noise such that its SMNR is 20 dB. Since y is noisy, BP (in Texas Hold

Figure 6.15: Average MSE versus MR: Compressible ECG signals (SMNR = 20 dB)

'Em) and Mod-BP (in GPABP-MMV and GPABP-separate) have $\|\Phi\tilde{x} - y\|_2 \leq 0.001$ as their constraint. Figure 6.15 shows the variation of average MSE as a function of MR (varied between 0.2 and 0.6). At MR= 0.2, R-MSBL gives a slightly lesser MSE compared to that of the GPABP. However, for $MR \geq 0.3$, GPABP-MMV gives the least reconstruction error (among all four techniques).

6.8 Summary

In this chapter, an efficient fusion approach (involving GPs and Mod-BP) is proposed for reconstructing sparse signals. Experimental results show that the proposed GPABP gives better reconstruction accuracy compared to that of the existing fusion based algorithms (FACS and its variants). GPABP is a PKS based recovery technique for scenarios where PKS is not available. Another technique called GP- $W\ell_1$ is proposed. GP- $W\ell_1$ uses a novel weighting strategy and it is found to be effective for large L . For reconstructing a cor-

related sparse signal ensemble, GPABP-MMV (an MMV-type algorithm for JSM-1 and JSM-2) is proposed. Experimental results show that the GPABP-MMV is robust to outliers. Moreover, being sequential reconstruction, GPABP-MMV remains stable for large number of signals. As depicted in figure 6.16, GPABP-MMV clearly outperforms state-of-the-art DCS reconstruction algorithms in terms of reconstruction accuracy. Performances of GPABP and GPABP-MMV are validated via their reconstruction condition and reconstruction error bound.

Figure 6.16: Speed and accuracy comparison: GPABP-MMV

Chapter 7

Conclusions and Future Directions

7.1 Conclusions

In this thesis, several PKS based CS reconstruction algorithms have been proposed. In time sequence reconstruction of correlated sparse signals, if one signal is known a priori, remaining signals can be reconstructed using existing PKS based convex relaxation algorithms such as Mod-BP or $W\ell_1$. However, these methods are greatly influenced by the support set of the known signal that includes one or more locations that are irrelevant to the target signal. In chapter 3, MABP and $WA\ell_1$ algorithms have been proposed for reconstructing correlated sparse signals. In MABP/ $WA\ell_1$, an adaptation/prunning process is carried out ahead of applying Mod-BP/ $W\ell_1$. If in the prunning step there are more correct deletions than wrong deletions, then MABP/ $WA\ell_1$ has a more favorable RIP requirement compared to that of the Mod-BP/ $W\ell_1$. Motivated by the effectiveness of adaptation step in MABP/ $WA\ell_1$, the adaptation idea is extended to Mod-GPs, resulting in MAGPs. In chapter 4, the AdBMP algorithm, a PKS version of BMP, is proposed for recovery of correlated sparse signals, and the theoretical upper bound on the unrecovered signal energy after each of its iteration is formulated. AdBMP makes use of PKS in a more efficient manner compared to that of existing PKS based GPs such as CoSaMP-PKS and IHT-PKS. In applica-

tions such as sparse recovery of radar echo signal, frame-based video reconstruction, and object-based video reconstruction, AdBMP exhibits superior reconstruction performance compared to that of the CoSaMP-PKS, IHT-PKS and the regular BMP.

Most PKS based CS reconstruction algorithms are totally blind to the wrong atoms present in the PKS. Moreover, each PKS based algorithm adapts its own modification strategy. A generalized, algorithm-blind reconstruction framework called ICSR is proposed in chapter 5. ICSR first estimates the initial support set using the available PKS, and then iteratively performs CS on the measurement residual to update the support set estimate. In applications such as reconstruction of natural images, the initial support set captures/contains the locations corresponding to larger transform coefficients and therefore, ICSR recovers the remaining locations accurately. Any CS reconstruction algorithm can be plugged into ICSR framework. Performance-wise, ICSR improves every CS reconstruction algorithm plugged into it. CS reconstruction algorithms with varying performance (in terms of reconstruction accuracy), when plugged into ICSR, gives a much similar performance. As summarized in figure 5.17, the key feature of the ICSR framework is the PSNR gain-execution speed combination of GPs (ICSR-CoSaMP, ICSR-OMP and ICSR-IHT), which makes them applicable in scenarios where fast recovery is required. Also, ICSR is flexible with respect to the availability of PKS.

Fusion based CS reconstruction algorithms such as FACS and CoMACS combine multiple CS reconstruction algorithms, which work with different principles, to obtain a better signal estimate. However, these algorithms involve solving a least squares problem which might be ill-conditioned. In chapter 6, an efficient fusion approach called GPABP is proposed for reconstructing sparse signals. Experimental results show that the proposed GPABP gives better reconstruction accuracy compared to that of the FACS and CoMACS. GPABP does not require a PKS and therefore, it is suitable for applications where no PKS is available. Another technique called GP- $W\ell_1$ is proposed. GP- $W\ell_1$ uses a novel weight-

ing strategy and it is found to be effective for large L . For reconstructing a correlated sparse signal ensemble, GPABP-MMV (an MMV-type algorithm for JSM-1 and JSM-2) is proposed. GPABP-MMV is robust to outliers and it remains stable for large number of signals. As depicted in figure 6.16, GPABP-MMV clearly outperforms state-of-the-art DCS reconstruction algorithms in terms of reconstruction accuracy.

As summarized in table 7.1, for time sequence reconstruction of correlated sparse signals, algorithms such as MABP, $WA\ell_1$, MAGPs, AdBMP, ICSR or GPABP-MMV can be applied. For reconstructing a single sparse signal, GPABP, $GP-W\ell_1$ or ICSR can be used. Algorithms such as MABP, $WA\ell_1$, MAGPs, and AdBMP need signal information in the form of PKS whereas ICSR, GPABP, $GP-W\ell_1$ and GPABP-MMV algorithms do not require PKS. With that, we believe that algorithms proposed in this thesis add more flexibility to the CS reconstruction of sparse signals with respect to the signal model and the availability of PKS. Further, they motivate the development of several elegant PKS based CS reconstruction algorithms.

7.2 Directions for Further Research

Based on our study of CS reconstruction algorithms, some possible directions for further research are briefly described in this subsection.

Most CS reconstruction algorithms are capable of handling real signals only. Only few algorithms (e.g. Complex Approximate Message Passing) are designed to handle both real and complex signals. Therefore, acquisition/reconstruction of complex signals is a challenging task. If CS reconstruction algorithms such as SP, CoSaMP, and BMP are modified such that they can handle complex signals, then they can be easily extended to AdBMP, which will be a PKS based CS reconstruction algorithm for complex signals.

In GPABP and GPABP-MMV, Mod-BP could be replaced with any PKS based CS reconstruction algorithm. For applications where reconstruction speed becomes a concern,

Table 7.1: Comparison of Proposed Algorithms

Proposed Algorithm	Class/Type	Signal Model	Is PKS Mandatory?
MABP $W\ell_1$ MAGPs AdBMP	CR CR GP GP	Correlated sparse signals (Time-sequence reconstruction)	YES
ICSR	General	(i) Single sparse signal (ii) Correlated sparse signals (Time-sequence reconstruction)	NO
GPABP GP- $W\ell_1$	Fusion Fusion	Single sparse signal	NO
GPABP-MMV	Fusion	Correlated sparse signals (Joint-then-sequential reconstruction)	NO

PKS based GP such as IHT-PKS or AdBMP can be used in the place of Mod-BP. Further research on this choice of algorithm would give more insights on the applicability of GPABP and GPABP-MMV.

Author's Publications

1. **S. Narayanan**, S. K. Sahoo and A. Makur, "Modified Adaptive Basis Pursuits for Recovery of Correlated Sparse Signals," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 4136-4140, Italy, May 2014.
2. **S. Narayanan** and A. Makur, "Compressive Sensing based Video Object Compression Schemes for Surveillance Systems," *SPIE Electronic Imaging: Video Surveillance and Transportation Imaging Applications*, San Francisco, Mar. 2015.
3. **S. Narayanan**, S. K. Sahoo, and A. Makur, "Greedy Pursuits Assisted Basis Pursuit for Compressive Sensing," *EUropean Signal Processing Conference (EUSIPCO)*, pp. 694-698, Nice, Aug. 2015.
4. **S. Narayanan**, S. K. Sahoo, and A. Makur, "Sparse Recovery of Radar Echo Signals using Adaptive Backtracking Matching Pursuit," *IEEE Radar Conference*, pp. 339-343, Johannesburg, Oct. 2015.
5. **S. Narayanan**, S. K. Sahoo, and A. Makur, "Recovery of Correlated Sparse Signals using Adaptive Backtracking Matching Pursuit," *IEEE Visual Communication and Image Processing (VCIP)*, pp. 1-4, Singapore, Dec. 2015.
6. **S. Narayanan** and A. Makur, "Low Complexity Surveillance Video Coding based on Distributed Compressive Video Sensing," *IEEE Region 10 Conference (TENCON)*, Singapore, Dec. 2016.

7. **S. Narayanan**, S. K. Sahoo and A. Makur, “Greedy Pursuits Assisted Basis Pursuit for Reconstruction of Joint-Sparse Signals,” submitted to *Signal Processing* (second round of review).
8. **S. Narayanan**, S. K. Sahoo, and A. Makur, “An Iterative Reconstruction Framework for Compressive Sensing with Partially Known Support,” submitted to *IEEE Transactions on Signal Processing* (first round of review).
9. **S. Narayanan**, S. K. Sahoo and A. Makur, “Greedy Pursuits based Gradual Weighting Strategy for Weighted ℓ_1 -Minimization,” submitted to *IEEE Signal Processing Letters*.

References

- [1] D. Donoho, “Compressed Sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289-1306, Apr. 2006.
- [2] E. Candes, J. Romberg, and T. Tao, “Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information,” *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.
- [3] E. Candes and J. Romberg, “Sparsity and Incoherence in Compressive Sampling,” *Inverse Problems*, vol. 23, no. 3, pp. 969-985, 2007.
- [4] E. Candes and M. Wakin, “An Introduction To Compressive Sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21-30, Mar. 2008.
- [5] S. Chen, D. Donoho, and M. Saunders, “Atomic Decomposition by Basis Pursuit,” *SIAM Journal on Scientific Computing*, vol. 20, pp. 3361, 1998.
- [6] S. Mallat and Z. Zhang, “Matching Pursuit with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, no.12, pp. 3397-3415, Dec. 1993.
- [7] J. A. Tropp, “Greed is Good: Algorithmic Results for Sparse Approximation,” *IEEE Transactions on Information Theory*, vol. 50, pp. 2231-2242, Oct. 2004.
- [8] D. L. Donoho, Y. Tsaig, I. Drori, and J. C. Starck, “Sparse Solution of Underdetermined Linear Equations by Stagewise Orthogonal Matching Pursuit,” Technical Report, Mar. 2006. Available Online: <http://www.dsp.ece.rice.edu/cs/>.
- [9] D. Needell and R. Vershynin, “Uniform Uncertainty Principle and Signal Recovery via Regularized Orthogonal Matching Pursuit,” *Found. Comput. Math.*, vol. 9, pp. 317-334, 2009.

- [10] W. Dai and O. Milenkovic, "Subspace Pursuit for Compressive Sensing Signal Reconstruction," *IEEE Transactions on Information Theory*, vol. 55, pp. 2230-2249, May. 2009.
- [11] D. Needell and J. A. Tropp, "Cosamp: Iterative Signal Recovery from Incomplete and Inaccurate Samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301-321, May. 2009.
- [12] T. Blumensath and M. E. Davies, "Iterative Hard Thresholding for Compressed Sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265-274, Nov. 2009.
- [13] T. Do, L. Gan, N. Nguyen, and T. Tran, "Sparsity Adaptive Matching Pursuit Algorithm for Practical Compressed Sensing, *Asilomar Conference on Signals, Systems and Computers*, pp. 581-587, Oct. 2008.
- [14] C. Ma, C. Xu, L. Shen, and S. Zheng, "A Fast Sparsity Adaptive Matching Pursuit Algorithm for Compressed Sensing, *Advances in Automation and Robotics*, vol. 1, pp. 363-368, 2012.
- [15] H. Huang and A. Makur, "Backtracking-Based Matching Pursuit Method for Sparse Signal Reconstruction," *IEEE Signal Processing Letters*, vol. 18, no. 7, pp. 391-394, Jul. 2011.
- [16] L. Chen and Y. Gu, "Backtracking Matching Pursuit with Supplement Set of Arbitrary Size," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6541-6545, May 2013.
- [17] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A Simple Proof of the Restricted Isometry Property for Random Matrices," *Constructive Approximation*, vol. 28, 2008.
- [18] N. Vaswani and W. Lu, "Modified-cs: Modifying Compressive Sensing for Problems with Partially Known Support," *IEEE Transactions on Signal Processing*, vol. 58, no. 9, pp. 4595-4607, Sep. 2010.
- [19] W. Lu and N. Vaswani, "Modified BPDN for noisy compressive sensing with partially known support, *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 3926-3929, 2010.

- [20] N. Vaswani, "LS-CS: Compressive Sensing on Least Squares Residual," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4108-4120, Aug. 2010.
- [21] M. P. Friedlander, H. Mansour, R. Saab and O. Yilmaz, "Recovering Compressively Sampled Signals Using Partial Support Information," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp.1122-1134, Feb. 2012.
- [22] A. Khajehnejad, W. Xu, A. Avestimehr, and B. Hassibi, "Weighted L1 Minimization for Sparse Recovery with Prior Information, *IEEE International Symposium on Information Theory (ISIT)*, pp. 483487, 2009.
- [23] Wei Lu; Vaswani, N., "Exact Reconstruction Conditions for Regularized Modified Basis Pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2634-2640, May 2012.
- [24] W. Lu and N. Vaswani, "Regularized Modified BPDN for Noisy Sparse Reconstruction with Partially Erroneous Support and Signal Value Knowledge," *IEEE Transactions on Signal Processing*, vol. 60, no. 1, pp. 182-196, 2012.
- [25] R. E. Carrillo, L. F. Polania, and K. E. Barner, "Iterative Algorithms for Compressed Sensing with Partially Known Support," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 3654-3657, Mar. 2010.
- [26] R. E. Carrillo, L. F. Polania, and K. E. Barner, "Iterative Hard Thresholding for Compressed Sensing with Partially Known Support," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 4028-4031, May 2011.
- [27] N. Vaswani, "Stability (over Time) of Modified-CS for Recursive Causal Sparse Reconstruction," *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp.1722-1729, Sep. 2010.
- [28] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "Distributed Compressed Sensing of Jointly Sparse Signals," *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2005.
- [29] M. Mishali and Y. C. Eldar, "The Continuous Joint Sparsity Prior for Sparse Representations: Theory and Applications," *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMPASAP)*, pp. 125-128, Dec. 2007.

- [30] E. van den Berg and M. P. Friedlander, "Theoretical and Empirical Results for Recovery From Multiple Measurements," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2516-2527, May 2010.
- [31] V. H. Tang, A. Bouzerdoun, S. L. Phung, and F. H. C. Tivive, "Enhanced Wall Clutter Mitigation for Compressed Through-the-Wall Radar Imaging using Joint Bayesian Sparse Signal Recovery," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7804-7808, May 2014.
- [32] A. Gogna, A. Shukla, H. K. Agarwal, and A. Majumdar, "Split Bregman Algorithms for Sparse/Joint-Sparse and Low-Rank Signal Recovery: Application in Compressive Hyperspectral Imaging," *IEEE International Conference on Image Processing (ICIP)*, pp. 1302-1306, Oct. 2014.
- [33] L. Kiryung Lee, Y. Bresler, and M. Junge, "Subspace Methods for Joint Sparse Recovery," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3613-3641, Jun. 2012.
- [34] F. Joe-Mei and L. Chia-Han, "Generalized Subspace Pursuit for Signal Recovery from Multiple-Measurement Vectors," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2874-2878, Apr. 2013.
- [35] J. Yuzhea and B. D. Rao, "Support Recovery of Sparse Signals in the Presence of Multiple Measurement Vectors," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 3139-3157, May 2013.
- [36] P. Vinuelas-Peris and A. Artes-Rodriguez, "Sensing Matrix Optimization in Distributed Compressed Sensing," *IEEE/SP Workshop on Statistical Signal Processing*, pp. 638-641, Aug. 2009.
- [37] S. R. Schnelle, J. N. Laska, C. Hegde, M. F. Duarte, M. A. Davenport, and R. G. Baraniuk, "Texas Hold 'Em Algorithms for Distributed Compressive Sensing," *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010.
- [38] I. Esnaola, R. E. Carrillo, J. Garcia-Frias, and K. E. Barner, "Orthogonal Matching Pursuit based Recovery for Correlated Sources with Partially Disjoint Supports," *Annual Conference on Information Sciences and Systems (CISS)*, pp. 1-6, Mar. 2010.

- [39] O. Balkan, K. Kreutz-Delgado, and S. Makeig, "Robust Joint Sparse Recovery on Data with Outliers," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3821-3825, May 2013.
- [40] S. K. Ambat, S. Chatterjee, and K.V.S. Hari, "Fusion of Algorithms for Compressed Sensing," *IEEE Transactions on Signal Processing*, vol. 61, no. 14, pp. 3699-3704, Jul. 2013.
- [41] S. K. Ambat, S. Chatterjee, and K.V.S. Hari, "A Committee Machine Approach for Compressed Sensing Signal Reconstruction," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1705-1717, Apr. 2014.
- [42] **S. Narayanan**, S. K. Sahoo and A. Makur, "Modified Adaptive Basis Pursuits for Recovery of Correlated Sparse Signals," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 4136-4140, Italy, May 2014.
- [43] **S. Narayanan**, S. K. Sahoo, and A. Makur, "Sparse Recovery of Radar Echo Signals using Adaptive Backtracking Matching Pursuit," *IEEE Radar Conference*, pp. 339-343, Johannesburg, Oct. 2015.
- [44] **S. Narayanan** and A. Makur, "Compressive Sensing based Video Object Compression Schemes for Surveillance Systems," *SPIE Electronic Imaging: Video Surveillance and Transportation Imaging Applications*, San Francisco, Mar. 2015.
- [45] **S. Narayanan**, S. K. Sahoo, and A. Makur, "Greedy Pursuits Assisted Basis Pursuit for Compressive Sensing," *European Signal Processing Conference (EUSIPCO)*, pp. 694-698, Nice, Aug. 2015.
- [46] E. Candes, J. Romberg, and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp.1207-1223, Aug. 2006.
- [47] E. Candes and T. Tao, "Decoding by Linear Programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203-4215, Dec. 2005.
- [48] *Basis Pursuit Denoising*. Available Online: https://en.wikipedia.org/wiki/Basis_pursuit_denoising.
- [49] E. Candes and T. Tao, "The Dantzig Selector: Statistical Estimation When p is Much Larger Than n ," *Annals of Statistics*, vol. 35, no. 6, 2007.

- [50] Y. C. Pati, R. Rezaeiifar, P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition," *Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 40-44, Nov. 1993.
- [51] M. A. Davenport and M. B. Wakin, "Analysis of Orthogonal Matching Pursuit Using the Restricted Isometry Property," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4395-4401, Sep. 2010.
- [52] E. Livshitz, "On Efficiency of Orthogonal Matching Pursuit," 2010. Available Online: <http://arxiv.org/abs/1004.3946>.
- [53] T. Zhang, "Sparse Recovery with Orthogonal Matching Pursuit under RIP," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6215-6221, Sep. 2011.
- [54] J. Ding, L. Chen, and Y. Gu, "Perturbation Analysis of Orthogonal Matching Pursuit and Random Dictionaries," *IEEE Transactions on Signal Processing*, vol. 61, no. 2, pp. 398-410, Jan 2013.
- [55] S. Chatterjee, D. Sundman, M. Vehkaperä, and M. Skoglund, "Projection Based and Look-Ahead Strategies for Atom Selection" *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 634-647, Feb. 2012.
- [56] S. Chatterjee, D. Sundman, and M. Skoglund, "Look Ahead Orthogonal Matching Pursuit," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4024-4027, May 2011.
- [57] R. Giryes and M. Elad, "RIP-Based Near-Oracle Performance Guarantees for SP, CoSaMP, and IHT," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1465-1468, Mar. 2012.
- [58] R. Dorfman, "The Detection of Defective Members of Large Populations," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 436-440, 1943.
- [59] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "Algorithmic Linear Dimension Reduction in the ℓ_1 Norm of Sparse Vectors," *Annual Allerton Conference on Communication, Control, and Computing*, 2006.
- [60] A. C. Gilbert, S. Muthukrishnan, and M. Strauss, "Improved Time Bounds for Near-Optimal Sparse Fourier Representations," *SPIE Wavelets*, 2003.

- [61] S. Sarvotham, D. Baron, and R. G. Baraniuk, "Sudocodes - Fast Measurement and Reconstruction of Sparse Signals," *International Symposium on Information Theory (ISIT)*, pp. 2804-2808, Jul. 2006.
- [62] E. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing Sparsity by Reweighted L1 Minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5-6, pp. 877-907, 2008.
- [63] Y. Wang and W. Yin, "Sparse Signal Reconstruction via Iterative Support Detection," *SIAM Journal of Imaging Sciences*, vol. 3, no. 3, pp. 462-491, 2010.
- [64] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint Approximately Sparse Channel Estimation and Data Detection in OFDM Systems using Sparse Bayesian Learning," *IEEE Transactions on Signal Processing*, vol. 62, no. 14, pp. 3591-3603, Jul 2014.
- [65] R. Prasad and C. R. Murthy, "Cramer-Rao Type Bounds for Sparse Bayesian Learning," *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 622-632, Feb. 2013.
- [66] Z. Zhang and B. D. Rao, "Sparse Signal Recovery with Temporally Correlated Source Vectors using Sparse Bayesian Learning," *IEEE Journal on Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912-926, Sep. 2011.
- [67] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "An Information-Theoretic Approach to Distributed Compressed Sensing," *Allerton Conference on Communication, Control, and Computing*, Monticello, Sep. 2005.
- [68] T. M. Cover and J. A. Thomas, "Elements of Information Theory," Wiley, New York, 1991.
- [69] D. Slepian and J. K. Wolf, "Noiseless Coding of Correlated Information Sources," *IEEE Transactions on Information Theory*, vol. 19, pp. 471-480, Jul. 1973.
- [70] J. Tropp, A.C. Gilbert, and M.J. Strauss, "Simultaneous Sparse Approximation via Greedy Pursuit," *IEEE International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, Philadelphia, Mar. 2005.
- [71] S. K. Ambat, S. Chatterjee, and K. V. S. Hari, "Subspace Pursuit Embedded in Orthogonal Matching Pursuit," *IEEE Region 10 Conference (TENCON)*, pp. 1-5, 2012.

- [72] S. K. Ambat, S. Chatterjee, and K. V. S. Hari, "Fusion of Greedy Pursuits for Compressed Sensing Signal Reconstruction," *EUropean Signal Processing Conference (EUSIPCO)*, Romania, Aug. 2012.
- [73] K.G. Deepa, S. K. Ambat and K.V.S. Hari, "Modified Greedy Pursuits for Improving Sparse Recovery," *National Conference on Communications (NCC)*, pp.1-5, India, Mar. 2014.
- [74] J. Romberg, "Imaging via Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14-20, Mar. 2008.
- [75] D. Takhar, V. Bansal, M. Wakin, M. Duarte, D. Baron, K.F. Kelly, and R.G. Baraniuk, "A Compressed Sensing Camera: New Theory and an Implementation using Digital Micromirrors," *SPIE Electronic Imaging*, San Jose, CA, 2006.
- [76] M. Lustig, D.L. Donoho, and J.M. Pauly, "Rapid MR Imaging with Compressed Sensing and Randomly Under-Sampled 3DFT Trajectories," *ISMRM Annual Meeting*, Seattle, WA, May 2006.
- [77] W. U. Bajwa, J. Haupt, G. Raz, S. J. Wright, and R. Nowak, "Toeplitz-Structured Compressed Sensing Matrices," *IEEE/SP Workshop on Statistical Signal Processing*, pp. 294-298, Madison, WI, Aug. 2007.
- [78] W. Yin, S. P. Morgan, J. Yang, and Y. Zhang, "Practical Compressive Sensing with Toeplitz and Circulant Matrices," *Visual Communications and Image Processing (VCIP)*, 2010.
- [79] S. Narayanan and A. Makur, "Camera Motion Estimation using Circulant Compressive Sensing Matrices," *International Conference on Information, Communications and Signal Processing (ICICS)*, pp.1-5, Taiwan, Dec. 2013.
- [80] S. Narayanan and A. Makur, "Compressive Coded Video Compression using Measurement Domain Motion Estimation," *IEEE International Conference on Electronics, Computing and Communication Technologies (IEEE CONECCT)*, pp.1-6, Bangalore, Jan. 2014.
- [81] Björck Åke, "Numerical Methods for Least Squares Problems," *Society for Industrial and Applied Mathematics*, 1996.

- [82] Y. C. Eldar and G. Kutyniok, "Compressed Sensing: Theory and Applications," Cambridge University Press, 2012.
- [83] L. Jacques, "A Short Note on Compressed Sensing with Partially Known Signal Support," *Signal Processing*, vol. 90, pp. 3308-3312, Dec. 2010.
- [84] cvx Solver. Available online: <http://cvxr.com/cvx/>.
- [85] A. Ferrari, C. Berenguer, and G. Alengrin, "Doppler Ambiguity Resolution using Multiple PRF," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 738-751, Jul. 1997.
- [86] J. H. Ender, "On Compressive Sensing Applied to Radar," *Signal Processing*, vol. 90, no. 5, pp. 1402-1414, 2010.
- [87] Y. Yanan, L. Chunsheng, and Y. Ze, "Parallel Frequency Radar via Compressive Sensing," *IEEE Geoscience and Remote Sensing Symposium*, pp. 2696-2699, Jul. 2011.
- [88] I. Kyriakides, "Radar Tracking Performance when Sensing and Processing Compressive Measurements," *Conference on Information Fusion (FUSION)*, pp. 1-8, Jul. 2010.
- [89] P. Maechler, N. Febler, and H. Kaeslin, "Compressive Sensing for WiFi-based Passive Bistatic Radar," *European Signal Processing Conference (EUSIPCO)*, pp. 1444-1448, Aug. 2012.
- [90] S. Kemkemian and M. Nouvel-Fiani, "Is Compressive Sensing Really Useful for Radar?," *IEEE Radar Conference*, pp. 786-791, May 2014.
- [91] D. Venkatraman and A. Makur, "A Compressive Sensing Approach to Object-Based Surveillance Video Coding," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3513-3516, 2009.
- [92] A. A. Muhit, M. R. Pickering, M. R. Frater and J. F. Arnold, "Video Coding Using Elastic Motion Model and Larger Blocks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 5, pp. 661-672, 2010.
- [93] T. T. Do, Y. Chen, D.T. Nguyen, N. Nguyen, L. Gan and T.D. Tran, "Distributed Compressed Video Sensing," *IEEE International Conference on Image Processing (ICIP)*, pp. 1393-1396, 2009.

- [94] R. Venkatesh Babu and A. Makur, "Object-based Surveillance Video Compression using Foreground Motion Compensation," *International Conference on Control, Automation, Robotics and Vision*, pp.1-6, 2006.
- [95] I. F. Gorodnitsky and B. D. Rao, "Sparse Signal Reconstruction from Limited Data using FOCUSS: A Re-Weighted Minimum Norm Algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600-616, Mar. 1997.
- [96] S. K. Ambat and K. V. S. Hari, "An Iterative Framework for Sparse Signal Reconstruction Algorithms," *Signal Processing*, vol. 108, pp. 351-364, Mar. 2015.
- [97] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J. P. Thiran, and Y. Wiaux, "Sparsity Averaging for Compressive Imaging," *IEEE Signal Processing Letters*, vol. 20, no. 6, pp. 591-594, Jun. 2013.
- [98] *Sparselab*. Available Online: <http://sparselab.stanford.edu>.
- [99] A. Y. Yang, S. S. Sastry, A. Ganesh, and Y. Ma, "Fast L1-minimization Algorithms and an Application in Robust Face Recognition: A Review," *IEEE International Conference on Image Processing (ICIP)*, pp.1849-1852, Sep. 2010.
- [100] G. B. Moody and R. G. Mark, "The Impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology*, vol. 20, no. 3, pp. 45-50, Jun. 2001.
- [101] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse Solutions to Linear Inverse Problems with Multiple Measurement Vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477-2488, Jul. 2005.
- [102] E. van den Berg and M. P. Friedlander, "SPGL1: A Solver for Large-scale Sparse Reconstruction," Jun. 2007. Available Online: <http://www.cs.ubc.ca/labs/scl/spgl1>.