Iowa State University
Digital Repository

Graduate Theses and Dissertations

Iowa State University Capstones, Theses and Dissertations

2011

Sparse Recovery with Partial Support and Signal Value Knowledge and Applications in Dynamic MRI

Wei Lu
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Electrical and Computer Engineering Commons

Recommended Citation

Lu, Wei, "Sparse Recovery with Partial Support and Signal Value Knowledge and Applications in Dynamic MRI" (2011). *Graduate Theses and Dissertations*. 10408.
https://lib.dr.iastate.edu/etd/10408

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

المنارة للاستشارات

www.manaraa.com

Sparse recovery with partial support and signal value knowledge and applications in dynamic MRI

by

Wei Lu

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
Namrata Vaswani, Major Professor
Aleksandar Dogandzic
Nicola Elia
Zhengdao Wang
Ranjan Maitra

Iowa State University

Ames, Iowa

2011

Copyright © Wei Lu, 2011. All rights reserved.

المنارة للاستشارات

www.manaraa.com

DEDICATION

I would like to dedicate this dissertation to my family without whose support I would not have been able to complete this work.

للاستشارات المنارة

www.manaraa.com

TABLE OF CONTENTS

المنارة للاستشارات

www.manaraa.com

www.manaraa.com

www.manaraa.com

LIST OF TABLES

المنارة للاستشارات

www.manaraa.com

LIST OF FIGURES

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

ACKNOWLEDGEMENTS

Abundant thanks go to my advisor, prof. Namrata Vaswani, whose guidance, teamwork and unfathomable know-how gave me a unique foundation upon which to complete this dissertation. With her help and inspiration, I can complete and enjoy my research work. May she continue to carry her insight and wisdom all across her research.

I would like to thank Prof. Aleksandar Dogandzic, Prof. Nicola Elia, Prof. Zhengdao Wang, and Prof. Ranjan Maitra for their suggestions on my research work. Their valuable comments help me to perfect this dissertation. Also, I would like to thank prof. Ian C. Atkinson from University of Illinois at Chicago who provides us the real experimental data and make our research working in the real application.

I thank my colleagues Taoran Li, Chenlu Qiu, Samarjit Das and Fardad Raisali for their suggestions and feedback for my work. Especially, I would like to thank Taoran Li for his hard working for our joint project. He gives me many useful suggestions and helps me fully understand the related knowledge in our work. In addition, I also would like to express my thanks to my friends Lei Ke and Kun Qiu in Iowa State University who study together with me and enrich me with every kind of interesting things.

Finally, I would like to thank my family, my father Mr. Qinfu Lu, my mother Mrs. Shilian Gui, my wife Zhiju Zheng and my daughter Connie Boya Lu. They keep me energetic and happy during my life. Their consistent support and encouragement wipe off all the troubles that impede me.

المنارة للاستشارات

www.manaraa.com

ABSTRACT

In this work, we study the problem of reconstructing a sparse signal from a limited number of its linear projections when the following knowledge is available. (1) We are given partial, and partly erroneous, knowledge of the signal's support, denoted by T. (2) We are also given an erroneous estimate of the signal values on T. Alternatively, in recursive reconstruction applications, like real-time dynamic MRI, one can use the support estimate and the signal value estimate from the previous time instant. We presented algorithms by modifying Compressive Sensing (CS) using the partly erroneous support and also the erroneous signal estimate for both noiseless and noisy measurements. The idea of our proposed solution is to solve a convex relaxation of the following problem: find the signal that is sparsest outside the set T, while being "close enough" to signal estimate on T and satisfying the data constraint. We obtain sufficient conditions for exact reconstruction using modified-CS and regularized modified-BP. These are much weaker than those needed for CS when the size of the unknown part of the support is small compared to the support size. We also propose solutions modified-BPDN and regularized modified-BPDN for noisy measurements using the similar idea. We obtain the computable and tighter bounds without any sufficient conditions for the reconstruction error. Simulation comparisons for both sparse and compressible signals are shown. In this work, we also study the application of CS based approaches for blood oxygenation level dependent (BOLD) contrast functional MR imaging (fMRI). In particular, we show, via exhaustive experiments on actual MR scanner data for brain fMRI, that our recently proposed approach for recursive reconstruction of sparse signal sequences, modified-CS-residual, outperforms other existing CS based approaches.

المنارة للاستشارات

www.manaraa.com

CHAPTER 1. Introduction

In traditional signal processing technology, it is required to sample the signal with Nyquist rate which is twice of the signal's bandwidth to exactly recover the signal, see Fig. 1.1. Fig. 1.1 shows the diagram of the conventional transmission scheme. The signal is first sampled at Nyquist rate so that we can obtain N samples. Then, they are compressed to only K samples where $K \ll N$. After that, the compressed data will be transmitted to the receiver and the receiver will decompress the data. Finally, the original signal will be recovered. However, we will have such a question that why we are bothering to use such a high sampling rate since we only use K sample during transmission. Therefore, our question is whether we can do sampling in a lower rate than Nyquist rate and combine the sampling and compression into one simple step. If we can recover the signal with highly undersampled measurements, we can speed up the data acquisition significantly and greatly reduce the data capturing time. Especially, in medical image reconstruction such as CT or MRI, this will greatly lower the risk of radiation and help to reduce the motion artifact which brings trouble for the reconstruction and clinical diagnosis. In addition, undersampling can allow longer scanning read-out time or increase of the radiation dose and this can increase signal-to-noise ratio (SNR) so that the reconstructed images bear good quality.

Compressive Sensing (CS) provides an answer to this question. CS theories have proved that if the signal is sparse or compressible in itself or some transform domain, we are able to recover the original signal exactly or with small loss from highly undersampled linear projections [1, 2, 3, 4, 5, 6, 7, 8, 9]. "Sparse" means the signal only has very few nonzero elements and we define the locations of nonzero elements as the support of this signal. Similarly, "compressible" means only very few elements are significantly large while others are much smaller. We also define $\beta\%$ energy support as the locations of those large coefficients containing $\beta\%$ signal energy. As is known, many medical images are sparse or

www.manaraa.com

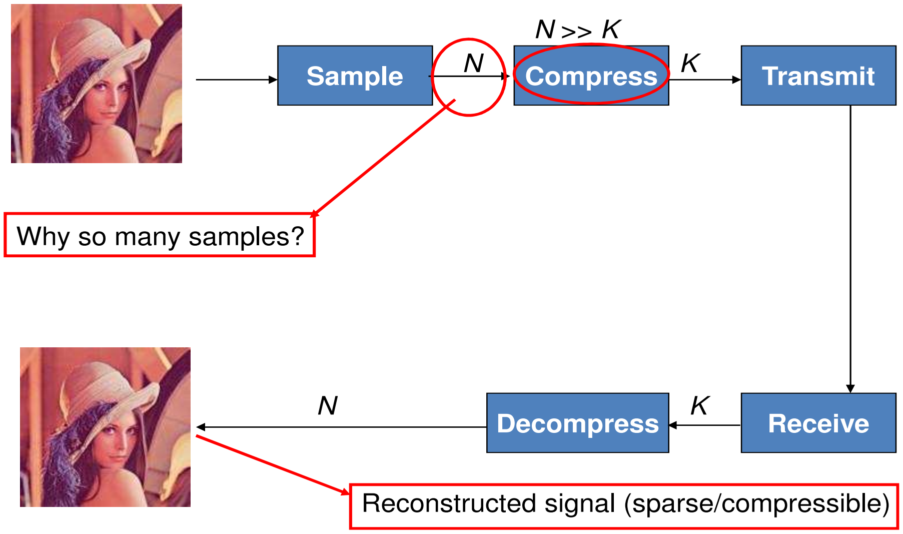


Figure 1.1   The limitation of conventional transmission scheme

compressible in wavelet domain, e.g., in the cardiac and larynx image sequence of Fig. 1.2, the sizes of their 99% energy support are only 6% or 7% of the image sizes. Many other images can be sparse in discrete cosine transform (DCT), discrete Fourier transform (DFT), total variation (TV) and other domains. To recover the original signal, the simplest way to find the sparsest solution is to exhaustively search the entire signal space in a brute force way. However, we know it is computationally expensive. CS provides practical solutions which can be solved in polynomial complexity for the sparse reconstruction. Two famous groups of CS algorithms are greedy methods and convex relaxation approaches. The greedy methods include subspace pursuit[6], Orthogonal Matching Pursuit (OMP)[7], Stagewise OMP[8], CoSAMP[9], etc. The convex relation approaches include Basis Pursuit(BP) and Basis Pursuit Denoising (BPDN)[1], Dantzig selector[10], etc. There are many other sparse reconstruction methods such as FOCUSS[11], Sparse Bayesian Learning[12] and Bayesian Compressive Sensing[13], etc.

In many real applications such as video compression or dynamic MRI reconstruction, the consecutive frames are usually correlated. Thus, when we are considering the problem of recursive reconstruction for a time sequence of sparse signals, it is easy to use the correlated information within the sequence. This gives the motivation of our work which is to causally and recursively reconstruct a time sequence of signals with slowly changing sparsity pattern. Hence, the goal of this work is to solve the

المنارة للاستشارات
www.manaraa.com

sparse recovery problem from a limited number of its linear projections by utilizing the prior information. We try to reconstruct an m-length sparse vector, x, with support, N, from an $n < m$ length noiseless measurement vector,

$$y := Ax \tag{1.1}$$

or noisy measurement vector,

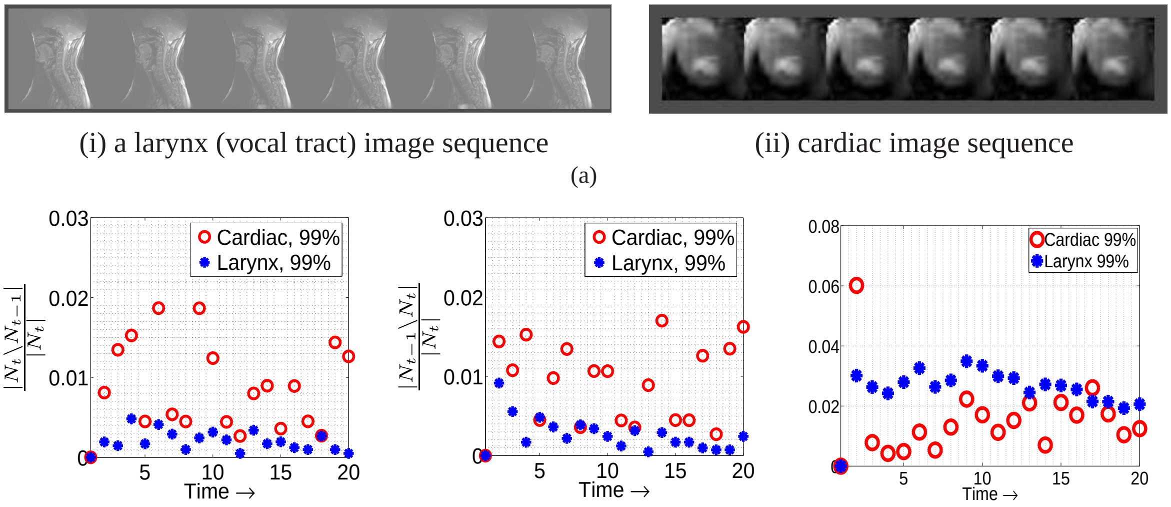
$$y := Ax + w \tag{1.2}$$

when the partial and partly erroneous knowledge of the signal's support, denoted by T, is available. Then we also study the case when an erroneous estimate of the signal values on T, denoted by $(\hat{\mu})_T$, is also available. In (1.2), w is an n-length measurement noise vector and A is an $n \times m$ measurement matrix. For simplicity, in this work, we just refer to x *as the signal* and to A *as the measurement matrix*. However, in general, x is the sparsity basis vector (which is either the signal itself or some linear transform of the signal) and $A = H\Phi$ where H is the measurement matrix and Φ is the sparsity basis matrix. If Φ is the identity matrix then x is the signal itself.

In practical applications, T and $\hat{\mu}$ may be available from prior knowledge. Alternatively, in applications requiring recursive reconstruction of (approximately) sparse signal or image sequences, with slow time-varying sparsity patterns and slow changing signal values, one can use the support estimate and the signal value estimate from the previous time instant as the "prior knowledge". A key domain where this problem occurs is in fast (recursive) dynamic MRI reconstruction from highly undersampled measurements. In MRI, we typically assume that the images are wavelet sparse. We show slow support and signal value change for two medical image sequences in Fig. 1.2. From the figure, we can see that the maximum support changes for both sequences are less than 2% of the support size and almost all signal values' changes are less than 0.16% of the signal energy. Slow signal value change also implies that a signal value is small before it gets removed from the support. Other potential applications include single-pixel camera based real-time video imaging [14]; video compression; ReProCS (recursive projected CS) based video denoising or video layering (separating video in foreground and background layers) [15, 16]; and spectral domain optical coherence tomography [17] based dynamic imaging.

Recent work on compressive sensing (CS) gives conditions for exact reconstruction [3, 4, 18] and bounds the error when this is not possible [2, 10]. In this work, we provide the exact reconstruction

www.manaraa.com



(i) a larynx (vocal tract) image sequence

(ii) cardiac image sequence

(a)

(i) support additions

(ii) support removals

(iii) signal value change

(b)

Figure 1.2 In (a), we show two medical image sequences (a cardiac and a larynx sequence). In (b), x_t is the two-level Daubechies-4 2D discrete wavelet transform (DWT) of the cardiac or the larynx image at time t and the set N_t is its 99% energy support (the smallest set containing 99% of the vector's energy). Its size, $|N_t|$ varied between 4121-4183 ($\approx 0.07m$) for larynx and between 1108-1127 ($\approx 0.06m$) for cardiac. *Notice that all support changes are less than 2% of the support size and almost all signal values changes are less than 4% of* $\|(x_t)_{N_t}\|_2$.

www.manaraa.com

conditions in noiseless case for our proposed modified compressive sensing (modified-CS) and regularized modified basis pursuit (reg-mod-BP) and also bound the reconstruction errors for our proposed modified basis pursuit denoising (mod-BPDN)and regularized modified basis pursuit denoising (reg-mod-BPDN).

1.1 Notations and Problem Definition

For any set T and vector b, b_T denotes a sub-vector containing the elements of b with indices in T. $\|b\|_k$ refers to the ℓ_k norm of the vector b. Also, $\|b\|_0$ counts the number of nonzero elements of b.

The notation T^c denotes the set complement of T, i.e., $T^c = \{i \in [1, ..., m], i \notin T\}$. $\emptyset$ is the empty set.

We use $'$ for transpose. For the matrix A, A_T denotes the sub-matrix containing the columns of A with indices in T. The matrix norm $\|A\|_p$, is defined as $\|A\|_p \triangleq \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$. I_T is an identity matrix on the set of rows and columns indexed by elements in T. $\mathbf{0}_{T,S}$ is a zero matrix on the set of rows and columns indexed by elements in T and S respectively.

$b \succeq 0$ ($b \succ 0$) means that each element of the vector b is greater than or equal to (strictly greater than) zero. Similarly $b \preceq 0$ ($b \prec 0$) means each element is less than or equal to (strictly less than) zero. We define the sign pattern, sgn(b) as: $[\text{sgn}(b)]_i = b_i/|b_i|$ if $b_i \neq 0$ and $[\text{sgn}(b)]_i = 0$ if $b_i = 0$.

The notation $\nabla L(b)$ denotes the gradient of the function $L(b)$ with respect to b.

When we say *b is supported on $T \cup S$* we mean that the support of b (set of indices where b is nonzero, denoted as supp(b)) is a subset of $T \cup S$.

The S-restricted isometry constant [18], δ_S, for a matrix, A, is defined as the smallest real number satisfying

$$(1 - \delta_S)\|c\|_2^2 \leq \|A_T c\|_2^2 \leq (1 + \delta_S)\|c\|_2^2 \qquad (1.3)$$

for all subsets $T \subset [1, n]$ of cardinality $|T| \leq S$ and all real vectors c of length $|T|$. The restricted orthogonality constant [18], θ_{S_1,S_2}, is defined as the smallest real number satisfying

$$|{c_1}' {A_{T_1}}' A_{T_2} c_2| \leq \theta_{S_1,S_2} \|c_1\|_2 \|c_2\|_2 \qquad (1.4)$$

المنارة للاستشارات
www.manaraa.com

for all disjoint sets $T_1, T_2 \subset [1, n]$ with $|T_1| \leq S_1$, $|T_2| \leq S_2$ and $S_1 + S_2 \leq n$, and for all vectors c_1, c_2 of length $|T_1|$, $|T_2|$ respectively. By setting $c_1 \equiv {A_{T_1}}' A_{T_2} c_2$ in (1.4),

$$\|{A_{T_1}}' A_{T_2}\| \leq \theta_{S_1,S_2} \tag{1.5}$$

Our goal is to reconstruct a sparse vector, x, with support, N, from the measurement vector, y satisfying (1.1) or (1.2). We assume partial knowledge of the support, denoted by T, and of the signal estimate on T, denoted by $(\hat{\mu})_T$. The support estimate may contain errors – misses Δ and extras Δ_e.

1.2 Related Work

The sparse reconstruction problem, without using any support or signal value knowledge, has been studied for a long time [18, 3, 4, 1, 2, 19, 10, 5]. It tries to find the sparsest signal among all signals that satisfy the data constraint, i.e. it solves $\min_b \|b\|_0$ s.t. $y = Ab$. This brute-force search has exponential complexity. One class of practical approaches to solve this is *basis pursuit (BP)* which replaces $\|b\|_0$ by $\|b\|_1$ [1]. The ℓ_1 norm is the closest norm to ℓ_0 that makes the problem convex. Therefore, for noiseless measurements, BP solves

$$\min_b \quad \|b\|_1 \quad \text{s.t.} \quad y = Ab \tag{1.6}$$

Exact reconstruction conditions are obtained in [18, 3, 4, 19]. For noisy measurements, the data constraint becomes an inequality constraint. However, this assumes that the noise is bounded and the noise bound is available. In practical applications where this may not be available, one can use the Lagrangian version which solves

$$\min_b \quad \gamma\|b\|_1 + \frac{1}{2}\|y - Ab\|_2^2 \tag{1.7}$$

This is called *basis pursuit denoising (BPDN) [1]*. Since this solves an unconstrained optimization problem, it is also faster. An error bound of BPDN was obtained in [2]. Error bounds for its constrained version were obtained in [19, 20].

Very recent work on causal sparse reconstruction for time sequences includes [21] (focusses on the time-invariant support case) and [22, 23] (use past estimates to only speed up the current optimization but not to improve reconstruction error). The problem of sparse reconstruction with partial support

المنارة للاستشارات
www.manaraa.com

knowledge was introduced in our work [24, 25]; and also in parallel in Khajehnejad et al [26] and in vonBorries et al [27]. In [24, 25], we proposed an approach called *modified-CS* which tries to find the signal that is sparsest outside the set T and satisfies the data constraint. We presented our solution using convex relations approaches. We obtained exact reconstruction conditions for it by using the restricted isometry approach [18]. When measurements are noisy, for the same reasons as above, one can use the Lagrangian version *modified-BPDN (mod-BPDN)*. Its error was bounded in our work [28], while the error of its constrained version was bounded in Jacques [29]. Also, some later work based on our suggested methods include [30] (which used the same idea of modified-CS but implemented using greedy algorithm OMP) and [31] (which iteratively used the support estimate from modified-CS reconstruction at each iteration).

In [26], Khajehnejad et al assumed a probabilistic support prior and proposed a weighted ℓ_1 solution. They also obtained exact reconstruction thresholds for weighted ℓ_1 by using the overall approach of Donoho [32]. It solves:

$$\min_b \ \|b_{T^c}\|_1 + \gamma \|b_T\|_1 \quad \text{s.t.} \quad y = Ab \tag{1.8}$$

for noiseless measurements or

$$\min_b \ \gamma \|b_{T^c}\|_1 + \gamma' \|b_T\|_1 + \frac{1}{2}\|y - Ab\|_2^2 \tag{1.9}$$

for noisy measurements.

Another related work is called *CS-residual or CS-diff* which computes

$$\begin{aligned} &\hat{x} = \hat{\mu} + \hat{b}, \qquad \text{where } \hat{b} \text{ solves} \\ &\min_b \ \|b\|_1 \quad \text{s.t.} \quad y = Ab \qquad\qquad \text{(noiseless)} \qquad (1.10) \\ &\min_b \ \gamma\|b\|_1 + \frac{1}{2}\|y - A\hat{\mu} - Ab\|_2^2 \qquad \text{(noisy)} \qquad (1.11) \end{aligned}$$

This has the following limitation. It does not use the fact that when T is an accurate estimate of the true support, $(x)_{T^c}$ is much more sparse compared with the full $(x - \hat{\mu})$ (the support size of x_{T^c} is $|\Delta|$ while that of $(x - \hat{\mu})$ is $|T| + |\Delta|$ which is much larger). The exception is if the signal value prior is so strong that $(x - \hat{\mu})$ is zero (or very small) on all or a part of T.

CS-residual is also related to LS-CS and KF-CS [33, 34]. LS-CS solves (1.10) or (1.11) but with $\hat{\mu}_T$ being the LS estimate computed assuming that the signal is supported on T and with $(\hat{\mu})_{T^c} = 0$.

المنارة للاستشارات
www.manaraa.com

For a static problem, KF-CS can be interpreted as computing the regularized LS estimate on T and using that as $\hat{\mu}_T$. LS-CS and KF-CS also have a limitation similar to CS-residual.

There are some other CS-based methods used in the application of MRI reconstruction. The application of CS to MRI was first developed in detail in [35]. The most straightforward application of CS to fMRI images reconstruction would be to perform CS on each slice of data independently (simple-CS). For time sequences, batch-CS [36] improves simple-CS by jointly reconstructing the entire sequence by treating it as a 3D sparse signal. Because it uses sparsity also along the time axis, it is able to achieve accurate reconstructions using much fewer measurements than simple-CS. But the reconstruction can only be performed on the entire *batch* of data after all sampling is completed. Also, for an l-frame acquisition, its computational complexity is roughly l^2 times that of simple-CS, while its memory requirement is l times that of simple-CS. In recent work, [37, 38] proposed Kt-FOCUSS, which uses the fact that a sequence of MR image data is sparse in the $y - f$ domain where f denotes temporal frequency. The key idea is to reconstruct $kY - t$ "frames" using FOCUSS[39] where kY denotes the phase encoding direction (y-axis of the 2D discrete Fourier transform (DFT) plane). Kt-FOCUSS is still a batch method, which means it is still (a) non-causal, i.e. it needs to wait to acquire the entire l frame sequence before doing the reconstruction (or one needs to re-run it in a batch fashion again at each time which is slow), and (b) its memory requirement is still l times that of simple-CS. But its reconstruction is fast because it is done on one $kY - t$ "frame" at a time and because often it only runs a a few iterations of FOCUSS starting from previous "frame" as initial guess. The same memory and non-causality issues also remain with Kt-FOCUSS with motion compensation (MC) [37].

1.3 Dissertation Organization

The dissertation is organized as follows. Exact recovery of Modified-CS and Reg-mod-BP for noiseless measurements and their sufficient conditions for exact reconstruction are introduced in Chapter 2. The error bounds for Mod-BPDN and Reg-mod-BPDN for noisy measurements are discussed in Chapter 3. The application of our algorithms in functional MRI to detect active regions is demonstrated in Chapter 4. Finally, conclusions are summarized in Chapter 5.

المنارة للاستشارات
www.manaraa.com

CHAPTER 2. Sparse Reconstruction for Noiseless Measurements with Partial Support and Signal Knowledge

In this chapter, we discuss the problem of reconstructing from noiseless measurements when partial support are signal knowledge are known.[24, 40, 25, 41] We first introduce modified-CS when only partial support is known. Then we discuss regularized modified-BP when the signal estimate is also available.

2.1 Modified-CS for problems with partially known support

We measure an m-length vector y where

$$y := Ax \tag{2.1}$$

We need to estimate x which is a sparse n-length vector with $n > m$. The support of x, denoted N, can be split as $N = T \cup \Delta \setminus \Delta_e$ where T is the "known" part of the support, $\Delta_e := T \setminus N$ is the error in the the known part and $\Delta := N \setminus T$ is the unknown part. Thus, $\Delta_e \subseteq T$, Δ, T are disjoint and $|N| = |T| + |\Delta| - |\Delta_e|$.

We use $s := |N|$ to denote the size of the (s)upport, $k := |T|$ to denote the size of the (k)nown part of the support, $e = |\Delta_e|$ to denote the size of the (e)rror in the known part and $u = |\Delta|$ to denote the size of the (u)nknown part of the support.

We assume that A satisfies the S-restricted isometry property (RIP) [18] for $S = (s + e + u) = (k + 2u)$. S-RIP means that $\delta_S < 1$ where δ_S is the RIP constant for A defined in (1.3).

In a static problem, T is available from prior knowledge. For example, in the MRI problem described in the introduction, let N be the (unknown) set of all DWT coefficients with magnitude above a certain zeroing threshold. Assume that the smaller coefficients are set to zero. Prior knowledge tells

المنارة للاستشارات
www.manaraa.com

us that most image intensities are nonzero and so the approximation coefficients are mostly nonzero. Thus we can let T be the (known) set of indices of all the approximation coefficients. The (unknown) set of indices of the approximation coefficients which are zero form Δ_e. The (unknown) set of indices of the nonzero detail coefficients form Δ.

For the time series problem, $y \equiv y_t$ and $x \equiv x_t$ with support, $N_t = T \cup \Delta \setminus \Delta_e$, and $T = \hat{N}_{t-1}$ is the support estimate from the previous time instant. If exact reconstruction occurs at $t-1$, $T = N_{t-1}$. In this case, $\Delta_e = N_{t-1} \setminus N_t$ is the set of indices of elements that were nonzero at $t-1$, but are now zero (deletions) while $\Delta = N_t \setminus N_{t-1}$ is the newly added coefficients at t (additions). Slow sparsity pattern change over time, e.g. see Fig. 1.2, then implies that $u \equiv |\Delta|$ *and* $e \equiv |\Delta_e|$ *are much smaller than* $s \equiv |N|$.

When exact reconstruction does not occur, Δ_e includes both the current deletions and the extras from $t-1$, $\hat{N}_{t-1} \setminus N_{t-1}$. Similarly, Δ includes both the current additions and the misses from $t-1$, $N_{t-1} \setminus \hat{N}_{t-1}$. In this case, slow support change, along with $\hat{N}_{t-1} \approx N_{t-1}$, still implies that $u \ll s$ and $e \ll s$.

2.1.1 Modified-CS

Our goal is to find a signal that satisfies the data constraint given in (1.1) and whose support contains the smallest number of new additions to T, although it may or may not contain all elements of T. In other words, we would like to solve

$$\min_b \quad \|(b)_{T^c}\|_0 \qquad \text{s.t.} \quad y = Ab \tag{2.2}$$

If Δ_e is empty, i.e. if $N = T \cup \Delta$, then the solution of (2.2) is also the sparsest solution whose support contains T.

As is well known, minimizing the ℓ_0 norm is a combinatorial optimization problem [42]. We propose to use the same trick that resulted in CS [1, 3, 4, 2]. We replace the ℓ_0 norm by the ℓ_1 norm, which is the closest norm to ℓ_0 that makes the optimization problem convex, i.e. we solve

$$\min_b \quad \|(b)_{T^c}\|_1 \qquad \text{s.t.} \quad y = Ab \tag{2.3}$$

www.manaraa.com

Denote its output by $\hat{x}$. If needed, the support is estimated as

$$\hat{N} := \{i \in [1, n] : (\hat{x})_i^2 > \alpha\} \tag{2.4}$$

where $\alpha \geq 0$ is a zeroing threshold. If exact reconstruction occurs, α can be zero. We discuss threshold setting for cases where exact reconstruction does not occur in Chapter 2.1.2.3.

2.1.2 Exact Reconstruction Result

We first analyze the ℓ_0 version of modified-CS in Chapter 2.1.2.1. We then give the exact reconstruction result for the actual ℓ_1 problem in Chapter 2.1.2.2.

2.1.2.1 Exact Reconstruction Result: ℓ_0 version of modified-CS

Consider the ℓ_0 problem, (2.2). Using a rank argument similar to [18, Lemma 1.2] we can show the following. The proof is given in the Appendix.

Proposition 1 *Given a sparse vector, x, with support, $N = T \cup \Delta \setminus \Delta_e$, where Δ and T are disjoint and $\Delta_e \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (2.2). x is the unique minimizer of (2.2) if $\delta_{k+2u} < 1$ (A satisfies the $(k+2u)$-RIP).*

Using $k = s + e - u$, this is equivalent to $\delta_{s+e+u} < 1$. Compare this with [18, Lemma 1.2] for the ℓ_0 version of CS. It requires $\delta_{2s} < 1$ which is much stronger when $u \ll s$ and $e \ll s$, as is true for time series problems.

2.1.2.2 Exact Reconstruction Result: modified-CS

Of course we do not solve (2.2) but its ℓ_1 relaxation, (2.3). Just like in CS, the sufficient conditions for this to give exact reconstruction will be slightly stronger. In the next few subsections, we prove the following result.

Theorem 1 (Exact Reconstruction) *Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_e$, where Δ and T are disjoint and $\Delta_e \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (2.3). x is the unique minimizer of (2.3) if*

المنارة للاستشارات

www.manaraa.com

1. *$\delta_{k+u} < 1$ and $\delta_{2u} + \delta_k + \theta^2_{k,2u} < 1$ and*

2. *$a_k(2u,u) + a_k(u,u) < 1$ where*

$$a_k(S,\check{s}) \triangleq \frac{\theta_{\check{s},S} + \frac{\theta_{\check{s},k}\ \theta_{S,k}}{1-\delta_k}}{1-\delta_S - \frac{\theta^2_{S,k}}{1-\delta_k}} \tag{2.5}$$

The above conditions can be rewritten using $k = s + e - u$.

We will not give the proof of Theorem 1 since it is a special case for reg-mod-BP and this theorem can be obtained by proving the exact reconstruction of reg-mod-BP. To understand the second condition better and relate it to the corresponding CS result, let us simplify it. $a_k(2u,u) + a_k(u,u) \leq \frac{\theta_{u,2u}+\theta_{u,u}+\frac{\theta^2_{2u,k}+\theta^2_{u,k}}{1-\delta_k}}{1-\delta_{2u}-\frac{\theta^2_{2u,k}}{1-\delta_k}}$. Simplifying further, a sufficient condition for $a_k(2u,u) + a_k(u,u) < 1$ is $\theta_{u,2u} + \theta_{u,u} + \frac{2\theta^2_{2u,k}+\theta^2_{u,k}}{1-\delta_k} + \delta_{2u} < 1$. Further, a sufficient condition for this is $\theta_{u,u} + \delta_{2u} + \theta_{u,2u} + \delta_k + \theta^2_{u,k} + 2\theta^2_{2u,k} < 1$.

To get a condition only in terms of δ_S's, use the fact that $\theta_{S,\check{s}} \leq \delta_{S+\check{s}}$ [18]. A sufficient condition is $2\delta_{2u} + \delta_{3u} + \delta_k + \delta^2_{k+u} + 2\delta^2_{k+2u} < 1$. Further, notice that if $u \leq k$ and if $\delta_{k+2u} < 1/5$, then $2\delta_{2u} + \delta_{3u} + \delta_k + \delta^2_{k+u} + 2\delta^2_{k+2u} < 4\delta_{k+2u} + \delta_{k+2u}(3\delta_{k+2u}) \leq (4+3/5)\delta_{k+2u} < 23/25 < 1$.

Corollary 1 (Exact Reconstruction) *Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_e$, where Δ and T are disjoint and $\Delta_e \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (2.3).*

- *x is the unique minimizer of (2.3) if $\delta_{k+u} < 1$ and*

$$(\delta_{2u} + \theta_{u,u} + \theta_{u,2u}) + (\delta_k + \theta^2_{k,u} + 2\theta^2_{k,2u}) < 1 \tag{2.6}$$

- *This, in turn, holds if*

$$2\delta_{2u} + \delta_{3u} + \delta_k + \delta^2_{k+u} + 2\delta^2_{k+2u} < 1.$$

- *This, in turn, holds if $u \leq k$ and*

$$\delta_{k+2u} < 1/5.$$

www.manaraa.com

These conditions can be rewritten by substituting $k = s + e - u$.

Compare (2.6) to the sufficient condition for CS given in [18]:

$$\delta_{2s} + \theta_{s,s} + \theta_{s,2s} < 1 \tag{2.7}$$

As shown in Fig. 1.2, usually $u \ll s$, $e \ll s$ and $u \approx e$ (which means that $k \approx s$). Under this assumption, compare (2.6) with (2.7). The first bracket of (2.6) will be small compared to the left hand side (LHS) of (2.7), particularly when s/m is larger. Also, if $\theta_{k,2u} < 1/2$ (requires s/m to not be too large), then each term of the second bracket will also be smaller than the LHS of (2.6). The last two terms of the second bracket are θ^2 terms, which makes them even smaller. Thus, for a certain range of values of s/m, the LHS of (2.6) will be small compared to that of (2.7). Since δ, θ are non-increasing in m, this means that, if u, e are small enough, (2.6) can hold for much smaller values of m than (2.7), i.e. *exact reconstruction with modified-CS can be guaranteed for smaller values of m than what is needed for CS.* A detailed comparison is done in Chapter 2.3.1.1.

2.1.2.3 Dynamic Modified-CS: Modified-CS for Recursive Reconstruction of Signal Sequences

The most important application of modified-CS is for recursive reconstruction of time sequences of sparse or compressible signals. To apply it to time sequences, at each time t, we solve (2.3) with $T = \hat{N}_{t-1}$ where $\hat{N}_{t-1}$ is the support estimate from $t-1$ and is computed using (2.4). At $t = 0$ we can either initialize with CS, i.e. set T to be the empty set, or with modified-CS with T being the support available from prior knowledge, e.g. for wavelet sparse images, T could be the set of indices of the approximation coefficients. The prior knowledge is usually not very accurate and thus at $t = 0$ one will usually need more measurements i.e. one will need to use $y_0 = A_0 x_0$ where A_0 is an $m_0 \times n$ measurement matrix with $m_0 > m$. The full algorithm is summarized in Algorithm 1.

Setting the support estimation threshold, α. If m is large enough for exact reconstruction, α can be zero. In case of very accurate reconstruction, if we set α to be slightly smaller than the magnitude of the smallest element of the support (if that is roughly known), it will ensure zero misses and fewest false additions. As m is reduced further (error increases), α should be increased further to prevent too many false additions.

المنارة للاستشارات
www.manaraa.com

For compressible signals, one should do the above but with support replaced by the $\beta\%$-support, i.e. α should be equal to/slightly smaller than the magnitude of the smallest element of the $\beta\%$-support. $\beta\%$-support is defined as below.

Definition 1 ($\beta\%$-energy support or $\beta\%$-support) *For sparse signals, clearly the support is* $N := \{i \in [1, n] : x_i^2 > 0\}$. *For compressible signals, we misuse notation slightly and let* N *be the* $\beta\%$-support, *i.e.* $N := \{i \in [1, n] : x_i^2 > \zeta\}$, *where* ζ *is the largest real number for which* N *contains at least* β*% of the signal energy, e.g.* $\beta = 99$ *in Fig. 1.2.*

Choose β so that, with the given m, the elements of the $\beta\%$-support are accurately reconstructed.

Alternatively, one can use the approach proposed in [43, Section II]. First, only detect additions to the support using a small threshold (or keep adding largest elements into T as long as A_T remains well-conditioned), then compute an LS estimate on that support and then use this LS estimate to perform support deletion using a larger threshold, α, selected as above. If there are few misses in the support addition step, the LS estimate will have lower error than the output of modified-CS, thus making deletion accurate even with a larger threshold.

Algorithm 1 **Dynamic Modified-CS**

```
At $t = 0$, compute $\hat{x}_0$ as the solution of $\min_b \ \|(b)_{T^c}\|_1, $ s.t. $y_0 = A_0 b$, where $T$ is either empty or is available from prior knowledge. Compute $\hat{N}_0 = \{i \in [1, n] : (\hat{x}_0)_i^2 > \alpha\}$.
For $t > 0$, do
    1. Modified-CS. Let $T = \hat{N}_{t-1}$. Compute $\hat{x}_t$ as the solution of $\min_b \ \|(b)_{T^c}\|_1, $ s.t. $y_t = Ab$.
    2. Estimate the Support. $\hat{N}_t = \{i \in [1, n] : (\hat{x}_t)_i^2 > \alpha\}$.
    3. Output the reconstruction $\hat{x}_t$.
Feedback $\hat{N}_t$, increment $t$, and go to step 1.
```

2.2 Regularized Modified-BP for Noiseless Sparse Reconstruction with Partial Erroneous Support and Signal Value Knowledge

In previous section, we discussed modified-CS which only uses the partially known support for reconstruction. In this section, we study the case when both the partial support and also the signal estimate on it are available. Our goal is to solve the sparse reconstruction problem, i.e. reconstruct an

المنارة للاستشارات
www.manaraa.com

m-length sparse vector, x, with support, N, from an $n < m$ length measurement vector,

$$y := Ax, \tag{2.8}$$

when an erroneous estimate of the signal's support, denoted by T; and an erroneous estimate of the signal values on T, denoted by $(\hat{\mu})_T$, are available. The support estimate, T, can be rewritten as $T \triangleq N \cup \Delta_e \setminus \Delta$ where Δ contains the misses while Δ_e contains the extras in the support estimate.

The signal value estimate is assumed to be zero along T^c, i.e.,

$$\hat{\mu} = \begin{bmatrix} (\hat{\mu})_T \\ \mathbf{0}_{T^c} \end{bmatrix} \tag{2.9}$$

and it satisfies

$$(\hat{\mu})_T \quad = \quad (x)_T + \nu, \text{ with } \|\nu\|_\infty \leq \rho. \tag{2.10}$$

Recall the following functions of the RIC and ROC of A in previous section:

$$a_k(s, \check{s}) \triangleq \frac{\theta_{\check{s},s} + \frac{\theta_{\check{s},k}\ \theta_{s,k}}{1-\delta_k}}{1 - \delta_s - \frac{\theta_{s,k}^2}{1-\delta_k}} \tag{2.11}$$

$$K_k(u) \triangleq \frac{\sqrt{1+\delta_u}}{1 - \delta_u - \frac{\theta_{u,k}^2}{1-\delta_k}} \tag{2.12}$$

For the matrix A, and for any set S for which $A_S{}'A_S$ is full rank, we define the matrix $M(S)$ as

$$M(S) \triangleq I - A_S(A_S{}'A_S)^{-1}A_S{}' \tag{2.13}$$

2.2.1 Regularized Modified Basis Pursuit

Mod-CS given in (2.3) puts no cost on b_T and no explicit constraint except $y = Ab$. Thus, when very few measurements are available, b_T can become larger than required in order to satisfy $y = Ab$ with the smallest $\|b_{T^c}\|_1$. A similar, though less, bias will also occur with (1.8) when $\gamma < 1$. However, if a signal value estimate on T, $(\hat{\mu})_T$, is also available, one can use that to constrain b_T. One way to do this, is to add $\lambda\|b_T - \hat{\mu}_T\|_2^2$ to the mod-CS cost. However, as we saw from simulations, while this does achieve lower reconstruction error, it cannot achieve exact recovery with fewer measurements (smaller

المنارة للاستشارات

www.manaraa.com

n) than mod-CS [25]. The reason is it puts a cost on the entire ℓ_2 distance from $\hat{\mu}_T$ and so encourages elements on the extras set, Δ_e, to be closer to $(\hat{\mu})_{\Delta_e}$ which is nonzero.

On the other hand, if we instead use the ℓ_∞ distance from $\hat{\mu}_T$, and add it as a constraint, then, at least in certain situations, we can achieve exact recovery with a smaller n than mod-CS. Thus, we study

$$\min_b \quad \|b_{T^c}\|_1, \quad \text{s.t.} \quad y = Ab \text{ and } \|b_T - \hat{\mu}_T\|_\infty \leq \rho \tag{2.14}$$

and call it *reg-mod-BP*. We see from simulations, that *whenever one or more of the inequality constraints are active, i.e.* $|b_i - \hat{\mu}_i| = \rho$ *for some* $i \in T$*, (2.14) does achieve exact recovery with fewer measurements than mod-CS*. We use this observation to derive a better exact recovery result below[1].

2.2.2 Exact Reconstruction Conditions

In this section, we obtain exact reconstruction conditions for reg-mod-BP by exploiting the above fact. We give the result and discuss its implications below in Chapter 2.2.2.1. The key lemmas leading to its proof are given in Chapter 2.2.2.2 and the proof outline in Chapter 2.2.2.3.

2.2.2.1 Exact Reconstruction Result

Let us begin by defining the two types of active sets (set of indices for which the inequality constraint is active), $T_{\text{a+}}$ and $T_{\text{a-}}$, and the inactive set, T_{in}, as follows.

$$\begin{aligned} T_{\text{a+}} &\triangleq \{i \in T : x_i - \hat{\mu}_i = \rho\} \\ T_{\text{a-}} &\triangleq \{i \in T : x_i - \hat{\mu}_i = -\rho\} \\ T_{\text{in}} &\triangleq \{i \in T : |x_i - \hat{\mu}_i| < \rho\} \end{aligned} \tag{2.15}$$

In the result below, we try to find the sets $T_{\text{a+g}} \subseteq T_{\text{a+}}$ and $T_{\text{a-g}} \subseteq T_{\text{a-}}$ so that $|T_{\text{a+g}}| + |T_{\text{a-g}}|$ is maximized while $T_{\text{a+g}}$ and $T_{\text{a-g}}$ satisfy certain constraints. We call these the "good" sets. We define the "bad" subset of T, as $T_b := T \setminus (T_{\text{a+g}} \cup T_{\text{a-g}})$. As we will see, the smaller the size of this bad set, the weaker are our exact recovery conditions.

[1] One can also try to constrain the ℓ_2 distance instead of the ℓ_∞ distance. When the ℓ_2 constraint is active, one should again need a smaller n for exact recovery. When we check this via simulations, this does happen, but since it is at most one active constraint, the reduction in n required is small compared to what is achieved by (2.14) and hence we do not study this further.

المنارة للاستشارات
www.manaraa.com

Theorem 2 (Exact Recovery Conditions) *Consider recovering a sparse vector, x, with support N, from $y := Ax$ by solving (2.14). The support estimate, T, and the misses and extras in it, Δ, Δ_e, satisfy $T \triangleq N \cup \Delta_e \setminus \Delta$. The signal estimate, $\hat{\mu}$, satisfies (2.10), i.e. $\|x_T - \hat{\mu}_T\|_\infty \leq \rho$. Recall the sizes of the sets T and Δ are defined as*

$$k := |T|, \ u := |\Delta|. \tag{2.16}$$

The true x is the unique minimizer of (2.14) if

1. $\delta_{k+u} < 1, \ \ \delta_{2u} + \delta_k + \theta_{k,2u}^2 < 1$, *and*

2. $a_k(2u, u) + a_{k_b}(u, u) < 1$ *where* $k_b := |T_b|$,

$$\begin{aligned} T_b \ &\triangleq \ T \setminus (T_{a+g} \cup T_{a-g}), \ \ \textit{and} \\ \{T_{a+g}, T_{a-g}\} \ &= \ \arg \max_{\tilde{T}_{a+g}, \tilde{T}_{a-g}} \ (|\tilde{T}_{a+g}| + |\tilde{T}_{a-g}|) \ \textit{subject to} \\ & \quad \tilde{T}_{a+g} \subseteq T_{a+}, \ \ \tilde{T}_{a-g} \subseteq T_{a-}, \\ & \quad A_i{}'w > 0 \ \ \forall \, i \in \tilde{T}_{a+g}, \ \textit{and} \ A_i{}'w < 0 \ \ \forall \, i \in \tilde{T}_{a-g} \\ & \quad \textit{where} \\ & \quad w \triangleq M(\tilde{T}_b) A_\Delta (A_\Delta{}' M(\tilde{T}_b) A_\Delta)^{-1} \textit{sgn}(x_\Delta), \\ & \quad \tilde{T}_b \triangleq T \setminus (\tilde{T}_{a+g} \cup \tilde{T}_{a-g}), \end{aligned} \tag{2.17}$$

$M(S)$ is specified in (2.13), $a_k(s, \check{s})$ is defined in (2.11), and the sets T_{a+}, T_{a-} are defined in (2.15). ■

Notice that $a_k(s, \check{s})$ is a non-decreasing function of k. Since $k_b = k - |T_{\text{a+g}}| - |T_{\text{a-g}}|$, thus, finding the largest possible sets $T_{\text{a+g}}$ and $T_{\text{a-g}}$ ensures that the condition $a_k(2u, u) + a_{k_b}(u, u) < 1$ is the weakest. The reason for defining $T_{\text{a+g}}$ and $T_{\text{a-g}}$ in the above fashion will become clear in the proof of Lemma 2.

Notice also that the first condition of the above result ensures that $\delta_k < 1$. Since $|\tilde{T}_b| \leq k$, thus, $A_{\tilde{T}_b}{}' A_{\tilde{T}_b}$ is positive definite and thus invertible. Thus $M(\tilde{T}_b)$ is always well defined. The first condition also ensures that $a_k(2u, u) > 0$. Since $k_b \leq k$, and since δ_s and θ_{s_1, s_2} are non-decreasing functions of s, s_1, s_2, it also ensures that $a_{k_b}(u, u) > 0$.

المنارة للاستشارات

www.manaraa.com

Remark 1 (Computation complexity) *Finding the best T_{a+g} and $T_{a\text{-}g}$ requires that one check all possible subsets of T_{a+} and $T_{a\text{-}}$ and find the pair with the largest sum of sizes that satisfies (2.17). To do this, one would start with $\tilde{T}_{a+g} = T_{a+}$, $\tilde{T}_{a\text{-}g} = T_{a\text{-}}$; compute $\tilde{T}_b$ and w and check if (2.17) holds; if it does not, remove one element from $\tilde{T}_{a+g}$ and then check (2.17); then remove an element from $\tilde{T}_{a\text{-}g}$ and check (2.17); keep doing this until one finds a pair for which (2.17) holds. In the worst case, one will need to check (2.17) $2^{|T_{a+}|+|T_{a\text{-}}|}$ times. However, the complexity of computing the RIC $\delta_{|T|}$ or any of the ROC's is anyway exponential in $|T|$ and $|T| \geq |T_{a+}| + |T_{a\text{-}}|$. In summary, computing the conditions of Theorem 2 has complexity that is exponential in the support size, but the same is true for all sparse recovery results that use the RIC. We should mention though that, for certain random matrices, e.g. random Gaussian, there are results that upper bound the RIC values with high probability, e.g. see [18]. However, the resulting bounds are usually quite loose.*

Remark 2 (Applicability) *A practical case where some of the inequality constraints will be active with nonzero probability is when dealing with quantized signals and quantized signal estimates. If the range of values that the signal estimate can take given the signal (or vice versa) is known, the smallest choice of ρ is easily computed. We show some examples in Chapter 2.3. In general, even if just the range of values both can take is known, we can compute ρ. The fewer the number values that $x_i - \hat{\mu}_i$ can take, the larger will be the expected size of the active set, $T_a := T_{a+} \cup T_{a\text{-}}$. Also, the condition (2.17) will hold for non-empty $T_g := T_{a+g} \cup T_{a\text{-}g}$ with positive probability, e.g. in our simulations (see Tables 2.3, 2.4), the average size of the good set T_g was about half the average size of the active set T_a. Some real applications where quantized signals and signal estimates occur are recursive CS based video compression [44, 45] (the original video itself is quantized) or in recursive projected CS (ReProCS) [15, 16] based moving or deforming foreground objects' extraction (e.g. a person moving towards a camera) from very large but correlated noise (e.g. very similar looking but slowly changing backgrounds), particularly when the videos are coarsely quantized (low bit rate). A common example where low bit rate videos occur is mobile telephony applications. In any of these applications, if we know a bound on the maximum change of the sparse signal's value from one time instant to the next, that can serve as ρ.*

Remark 3 (Comparison with BP, mod-CS, other results) *The worst case for Theorem 2 is when both*

المنارة للاستشارات
www.manaraa.com

the sets T_{a+g} and $T_{a\text{-}g}$ are empty either because no constraint is active (T_{a+} and $T_{a\text{-}}$ are both empty) or because (2.17) does not hold for any pair of subsets of T_{a+} and $T_{a\text{-}}$. In this case, we have $k_b = k$ and so the required sufficient conditions are the same as those of mod-CS (Theorem 1). A small extra requirement is that x satisfies (2.10). Thus, in the worst case, Theorem 2 holds under the same conditions on A (needs the same number of measurements) as mod-CS. In previous section, we have already argued that the mod-CS result holds under weaker conditions than the results for BP [18, 19] as long as the size of the support errors, $|\Delta|, |\Delta_e|$, are small, and hence the same can be said about Theorem 2. Small $|\Delta|, |\Delta_e|$ is a valid assumption in recursive recovery applications like recursive dynamic MRI, recursive CS based video compression, or ReProCS based foreground extraction from large but correlated background noise.

Moreover, if some inequality constraints are active and (2.17) holds, as in case of quantized signals and signal estimates, Theorem 2 holds under weaker conditions on A than the mod-CS result.

Remark 4 (Small reconstruction error) *The reconstruction error of reg-mod-BP is significantly smaller than that of mod-CS, weighted ℓ_1 or BP,* even when none of the constraints is active, *as long as ρ is small (see Table 2.5). On the other hand, the exact recovery conditions* do not *depend on the value of ρ, but only on the size of the good subsets of the active sets. This is also observed in our simulations. In Table 2.5, we show results for $\rho = 0.1$. Even when we tried $\rho = 0.5$, the exact reconstruction probability or the smallest n needed for exact reconstruction remained the same, but the reconstruction error increased.*

2.2.2.2 Proof of Theorem 2: Key Lemmas

Our overall proof strategy is similar to that of [18] for BP. We first find a set of sufficient conditions on an $n \times 1$ vector, w, that help ensure that x is the unique minimizer of (2.14). This is done in Lemma 1. Next, we find sufficient conditions that the measurement matrix A should satisfy so that one such w can be found. This is done in an iterative fashion in the theorem's proof. The proof uses Lemma 2 at the zeroth iteration, followed by applications of Lemma 3 at later iterations.

To obtain the sufficient conditions on w, as suggested in [18], we first write out the Karush-Kuhn-Tucker (KKT) conditions for x to be *a* minimizer of (2.14) [46, Chapter 5]. By strengthening these a

المنارة للاستشارات

www.manaraa.com

little, we get a set of *sufficient* conditions for x to be *the unique* minimizer. The necessary conditions for x to be a minimizer are: there exists an $n \times 1$, vector w (Lagrange multiplier for the constraints in $y = Ax$), a $|T_{\text{a+}}| \times 1$ vector, λ_1, and a $|T_{\text{a-}}| \times 1$ vector, λ_2, such that (s.t.)

1. every element of λ_1 and λ_2 is non-negative, i.e. $\lambda_1 \succeq 0$ and $\lambda_2 \succeq 0$,

2. $A_{T_{\text{in}}}{}'w = 0$, $A_{T_{\text{a+}}}{}'w = \lambda_1$, $A_{T_{\text{a-}}}{}'w = -\lambda_2$, $A_\Delta{}'w = \text{sgn}(x_\Delta)$, and $\|A_{(T\cup\Delta)^c}{}'w\|_\infty \leq 1$.

As we will see in the proof of Lemma 1, strengthening $\|A_{(T\cup\Delta)^c}{}'w\|_\infty \leq 1$ to $\|A_{(T\cup\Delta)^c}{}'w\|_\infty < 1$, keeping the other conditions the same, and requiring that $\delta_{k+u} < 1$ gives us a set of *sufficient* conditions.

Lemma 1 *Let x be as defined in Theorem 2. x is the unique minimizer of (2.14) if $\delta_{k+u} < 1$ and if we can find an $n \times 1$ vector, w, s.t.*

1. *$A_{T_{in}}{}'w = 0$, $A_{T_{a+}}{}'w \succeq 0$, $A_{T_{a-}}{}'w \preceq 0$,*

2. *$A_\Delta{}'w = sgn(x_\Delta)$,*

3. *$|A_j{}'w| < 1$ for all $j \notin T \cup \Delta$*

Recall that T_{a+}, T_{a-} and T_{in} are defined in (2.15) and k, u in Theorem 2. ■

Proof: The proof is given in Appendix A.2.

Next, we try to obtain sufficient conditions on the measurement matrix, A (on its RIC's and ROC's) to ensure that such a w can be found. This is done by using Lemmas 2 and 3 given below. Lemma 2 helps ensure that the first two conditions of Lemma 1 hold and provides the starting point for ensuring that the third condition also holds. Then, Lemma 3 applied iteratively helps ensure that the third condition also holds.

Lemma 2 *Assume that $k + u \leq m$. Let $\check{s}$ be such that $k + u + \check{s} \leq m$. If $\delta_u + \delta_{k_b} + \theta^2_{k_b,u} < 1$, then there exists an $n \times 1$ vector $\tilde{w}$ and an "exceptional" set, E, disjoint with $T \cup \Delta$, s.t.*

1. *$A_{T_b}{}'\tilde{w} = 0$, $A_{T_{a+g}}{}'\tilde{w} \succ 0$, $A_{T_{a-g}}{}'\tilde{w} \prec 0$,*

2. *$A_\Delta{}'\tilde{w} = sgn(x_\Delta)$,*

www.manaraa.com

3. *$|E| < \check{s}$, $\|A_E{}'\tilde{w}\|_2 \le a_{k_b}(u,\check{s})\sqrt{u}$, $|A_j{}'\tilde{w}| \le \frac{a_{k_b}(u,\check{s})}{\sqrt{\check{s}}}\sqrt{u}$ $\forall j \notin T \cup \Delta \cup E$,*

4. *$\|\tilde{w}\|_2 \le K_{k_b}(u)\sqrt{u}$.*

Recall that $a_k(s,\check{s})$, $K_k(s)$ are defined in (2.11), (2.12) and T_{a+g}, $T_{a\text{-}g}$, T_b, k_b, k and u in Theorem 2. ■

Notice that because we have assumed that $\delta_u + \delta_{k_b} + \theta_{k_b,u}^2 < 1$, $a_{k_b}(u,\check{s})$ and $K_{k_b}(u)$ are positive. We call the set E an "exceptional" set, because except on the set $E \subseteq (T \cup \Delta)^c$, everywhere else on $(T \cup \Delta)^c$, $|A_j{}'\tilde{w}|$ is bounded. This notion is taken from [18]. Notice that the first two conditions of the above lemma are one way to satisfy the first two conditions of Lemma 1 since $T_b = T_{\text{in}} \cup (T_{\text{a+}} \setminus T_{\text{a+g}}) \cup (T_{\text{a-}} \setminus T_{\text{a-g}})$.

Proof: The proof is given in Appendix A.3. We let $\tilde{w} = M(T_b)A_\Delta(A_\Delta{}'M(T_b)A_\Delta)^{-1}\text{sgn}(x_\Delta)$. Since the good sets $T_{\text{a+g}}$, $T_{\text{a-g}}$ are appropriately defined (see (2.17)), the first two conditions hold. The rest of the proof bounds $\|\tilde{w}\|_2$, and finds the set $E \subseteq (T \cup \Delta)^c$ of size $|E| < \check{s}$ so that $|A_j{}'\tilde{w}|$ is bounded for all $i \notin T \cup \Delta \cup E$ and also $\|A_E{}'\tilde{w}\|_2$ is bounded.

Lemma 3 *Assume that $k \le m$. Let s, $\check{s}$ be such that $k + s + \check{s} \le m$. Assume that $\delta_s + \delta_k + \theta_{k,s}^2 < 1$. Let T_d be a set that is disjoint with T, of size $|T_d| \le s$ and let c be a $|T_d| \times 1$ vector. Then there exists an $n \times 1$ vector, $\tilde{w}$, and a set, E, disjoint with $T \cup T_d$, s.t. (i) $A_T{}'\tilde{w} = 0$, (ii) $A_{T_d}{}'\tilde{w} = c$, (iii) $|E| < \check{s}$, $\|A_E{}'\tilde{w}\|_2 \le a_k(s,\check{s})\|c\|_2$, $|A_j{}'\tilde{w}| \le \frac{a_k(s,\check{s})}{\sqrt{\check{s}}}\|c\|_2$, $\forall j \notin T \cup T_d \cup E$, and (iv) $\|\tilde{w}\|_2 \le K_k(s)\|c\|_2$. Recall that $a_k(s,\check{s})$, $K_k(s)$ are defined in (2.11), (2.12), and k, u in Theorem 2.* ■

Proof: The proof of this lemma is given in Appendix A.4.

Notice that because we have assumed that $\delta_s + \delta_k + \theta_{k,s}^2 < 1$, $a_k(s,\check{s})$ and $K_k(s)$ are positive.

2.2.2.3 Proof Outline of Theorem 2

We give only the outline here and the complete proof is given in the Appendix A.5. At iteration zero, we apply Lemma 2 with $\check{s} \equiv u$, to get a w_1 and an exceptional set $T_{d,1}$, disjoint with $T \cup \Delta$, of size less than u. Lemma 2 can be applied because $k_b \le k$ and condition 1 of the theorem holds. At iteration $r > 0$, we apply Lemma 3 with $T_d \equiv \Delta \cup T_{d,r}$ (so that $s \equiv 2u$), $c_\Delta \equiv 0$, $c_{T_d} \equiv A_{T_d}{}'w_r$ and

المنارة للاستشارات

www.manaraa.com

$\check{s} \equiv u$ to get a w_{r+1} and an exceptional set $T_{d,r+1}$ disjoint with $T \cup \Delta \cup T_{d,r}$ of size less than u. Lemma 3 can be applied because condition 1 of the theorem holds. Define $w \triangleq \sum_{r=1}^{\infty}(-1)^{r-1}w_r$. We then argue that if condition 2 of the theorem holds, w is well-defined and satisfies the conditions of Lemma 1. Applying Lemma 1, the result follows.

2.2.3 Reconstruction Error Bound

When exact reconstruction cannot be achieved, we want to bound the error of $h = \hat{x} - x$. We adapt the approach of [19, 29] to bound the ℓ_2 norm of the error $\|h\|_2$. First consider modCS, i.e. (2.3). When exact reconstruction condition does not hold, the following lemma provides one way to bound the error.

Lemma 4 *Pick a $\tilde{\Delta} \subseteq \Delta$ and a $\tilde{T} \subseteq T$ such that $\delta_{|\tilde{T}|+2|\tilde{\Delta}|} < \sqrt{2}-1$. Denote $\hat{x}$ as the unique minimizer of (2.3), then*

$$\|x-\hat{x}\|_2 \leq \frac{1-\delta_{|\tilde{T}|+2|\tilde{\Delta}|}}{1-(\sqrt{2}+1)\delta_{|\tilde{T}|+2|\tilde{\Delta}|}} \cdot \frac{2\|x_{(\tilde{T}\cup\tilde{\Delta})^c}\|_1}{\sqrt{|\tilde{\Delta}|}} \tag{2.18}$$

As long as the true x is always part of the feasible set of (2.14), i.e. as long as $\|x_T - \mu_T\|_\infty \leq \rho$, the above lemma also holds for reg-mod-BP. In the next lemma we also use this prior constraint to obtain another error bound for reg-mod-BP, which is tighter than that of Lemma 4 when ρ is small enough, i.e. prior information is strong.

Lemma 5 *Let $\hat{x}$ solve (2.14) and $\|x_T - \mu_T\|_\infty \leq \rho$. If $\delta_{2u} \leq \sqrt{2}-1$ and $\delta_{k+2u} < 1$ hold, then*

$$\|x-\hat{x}\|_2 \leq \left(\frac{2\sqrt{k}\delta_{k+2u}}{1-(\sqrt{2}+1)\delta_{2u}} + 2\right)\rho \tag{2.19}$$

Combining the above two lemmas, we have the following Theorem to bound the error for reg-mod-BP.

Theorem 3 (Reconstruction Error Bound) *Let $\hat{x}$ solve (2.14). If $\|x_T - \mu_T\|_\infty \leq \rho$ and if $\delta_{2u} \leq$*

المنارة للاستشارات
www.manaraa.com

$\sqrt{2}-1$ *and* $\delta_{k+2u} < 1$, *then*

$$\|x-\hat{x}\|_2 \leq \min\{B_1, B_2\}, \text{ where}$$
$$B_1 \triangleq \left(\frac{2\sqrt{k}\delta_{k+2u}}{1-(\sqrt{2}+1)\delta_{2u}} + 2\right)\rho$$
$$B_2 \triangleq \min_{\substack{\tilde{T}\subseteq T, \tilde{\Delta}\subseteq\Delta \\ \delta_{|\tilde{T}|+2|\tilde{\Delta}|} < \sqrt{2}-1}} \frac{1-\delta_{|\tilde{T}|+2|\tilde{\Delta}|}}{1-(\sqrt{2}+1)\delta_{|\tilde{T}|+2|\tilde{\Delta}|}} \cdot \frac{2\|x_{(\tilde{T}\cup\tilde{\Delta})^c}\|_1}{\sqrt{|\tilde{\Delta}|}}$$

The complete proof is in the Appendix A.5.1. Clearly the bound for modCS is B_2 since modCS is a special case of reg-mod-BP when $\rho = \infty$ and $B_1 = \infty$ in this case. Therefore, reg-mod-BP bound, which is $\min\{B_1, B_2\}$, will never be larger than modCS bound. One particular case is when $\delta_{k+2u} < \sqrt{2}-1$ and in this case $B_2 = 0$ which implies that exact reconstruction occurs for both modCS and reg-mod-BP. However, when the number of measurements is very small, δ_{k+2u} will be much larger than $\sqrt{2}-1$. Thus, $|\tilde{T}|$ and $|\tilde{\Delta}|$ in modCS bound B_2 must be small such that $\delta_{|\tilde{T}|+2|\tilde{\Delta}|} < \sqrt{2}-1$. However, the set $(\tilde{T}\cup\tilde{\Delta})^c$ becomes larger resulting in $\frac{\|x_{(\tilde{T}\cup\tilde{\Delta})^c}\|_1}{\sqrt{|\tilde{\Delta}|}}$ to be very large. Hence, modCS bound will be very large. But for reg-mod-BP, if the signal estimate μ_T is good which allows a small ρ, then $B_1 \ll B_2$ resulting a much smaller bound than modCS.

2.2.4 Variation of Regularized Modified-BP

So far we have studied the exact recovery conditions for reg-mod-BP. As we stated in the beginning of this chapter, we study the exact reconstruction conditions of (2.14) because it can have better conditions when some constraints are active. In practice, when exact reconstruction cannot be achieved, a variant version of reg-mod-BP is to move the signal estimate constraint to the cost function which reduces the reconstruction error by solving

$$\min_b \quad \|(b)_{T^c}\|_1 + \gamma\|(b)_T - \mu_T\|_2^2 \quad \text{s.t.} \quad y = Ab \tag{2.20}$$

We call the above regularized modified-CS or reg-mod-CS. Denote its output by $\hat{x}_{reg}$. The parameter γ is easier to adjust in practical applications. However, as we claimed at the beginning, reg-mod-CS can not get better exact recovery conditions than modified-CS. We will study it through some simulations below.

المنارة للاستشارات

www.manaraa.com

2.2.4.1 Setting γ using an MAP interpretation of reg-mod-CS

One way to select γ is to interpret the solution of (2.20) as a maximum a posteriori (MAP) estimate under the following prior model and under the observation model of (1.1). Given the prior support and signal estimates, T and μ_T, assume that x_T and x_{T^c} are mutually independent and

$$\begin{aligned} p(x_T|T,\mu_T) &= \mathcal{N}(x_T;\mu_T,\sigma_p^2 I), \\ p(x_{T^c}|T,\mu_T) &= \left(\frac{1}{2\lambda_p}\right)^{|T^c|} e^{-\frac{\|x_{T^c}\|_1}{\lambda_p}}, \end{aligned} \qquad (2.21)$$

i.e. all elements of x are mutually independent; each element of T^c is zero mean Laplace distributed with parameter λ_p; and the i^{th} element of T is Gaussian with mean μ_i and variance σ_p^2. Under the above model, if $\gamma = \lambda_p/2\sigma_p^2$ in (2.20), then, clearly, its solution, $\hat{x}_{reg}$, will be an MAP solution.

Given i.i.d. training data, the maximum likelihood estimate (MLE) of λ_p, σ_p^2 can be easily computed in closed form [47].

2.2.4.2 Dynamic Regularized Modified-CS (reg-mod-CS)

To apply reg-mod-CS to time sequences, we solve (2.20) with $T = \hat{N}_{t-1}$ and $\mu_T = (\hat{x}_{t-1})_T$. Thus, we use Algorithm 1 with step 1 replaced by

$$\min_b \quad \|(b)_{\hat{N}_{t-1}^c}\|_1 + \gamma\|(b)_{\hat{N}_{t-1}} - (\hat{x}_{t-1})_{\hat{N}_{t-1}}\|_2^2 \quad \text{s.t.} \quad y_t = Ab \qquad (2.22)$$

In the last step, we feed back $\hat{x}_t$ and $\hat{N}_t$.

In Appendix A.6, we give the conditions under which the solution of (2.22) becomes a causal MAP estimate. To summarize that discussion, if we set $\gamma = \lambda_p/2\sigma_p^2$ where λ_p, σ_p^2 are the parameters of the signal model given there, and if we assume that the previous signal is perfectly estimated from $y_0, \dots y_{t-1}$ with the estimate being zero outside $\hat{N}_{t-1}$ and equal to $(\hat{x}_{t-1})_{\hat{N}_{t-1}}$ on it, then the solution of (2.22) will be the causal MAP solution under that model.

In practice, the model parameters are usually not known. But, if we have a training time sequence of signals, we can compute their MLEs using (A.44), also given in Appendix A.6.

المنارة للاستشارات
www.manaraa.com

2.3 Numerical Experiments

In this section, we did the simulations to verify all results we obtained in the above two sections. First, we show a set of experiments for modified-CS. Then, we give the other set of experiments for reg-mod-BP.

2.3.1 Experimental results of modified-CS

We first compared the sufficient conditions of modified-CS and CS using their high probability bounds and also through a detailed simulation. Then, we simulated two applications: CS-based image/video compression (or single-pixel camera imaging) and static/dynamic MRI. The measurement matrix was $A = H\Phi$ where Φ is the sparsity basis of the image and H models the measurement acquisition. All operations are explained by rewriting the image as a 1D vector. We used $\Phi = W'$ where W is an orthonormal matrix corresponding to a 2D-DWT for a 2-level Daubechies-4 wavelet. For video compression (or single-pixel imaging), *H is a random Gaussian matrix, denoted G_r*, (i.i.d. zero mean Gaussian $m \times n$ matrix with columns normalized to unit ℓ_2 norm). For MRI, *H is a partial Fourier matrix, i.e.* $H = MF$ where M is an $m \times n$ mask which contains a single 1 at a different randomly selected location in each row and all other entries are zero and F is the matrix corresponding to the 2D discrete Fourier transform (DFT).

N-RMSE, defined here as $\|x_t - \hat{x}_t\|_2 / \|x_t\|_2$, is used to compare the reconstruction performance. We first used the sparsified and then the true image and then did the same for image sequences. In all cases, the image was sparsified by computing its 2D-DWT, retaining the coefficients from the 99%-energy support while setting others to zero and taking the inverse DWT. We used the 2-level Daubechies-4 2D-DWT as the sparsifying basis. We compare modified-CS with simple CS, CS-residual or CS-diff [48] and LS-CS [43].

For solving the minimization problems given in (2.3), we used CVX, `http://www.stanford.edu/~boyd/cvx/`, for smaller sized problems ($n < 4096$). All simulations of Chapter 2.3.1.1 and all results of Table 2.2 and Figs. 2.2 used CVX. For bigger signals/images, (i) the size of the matrix A becomes too large to store on a PC (needed by most existing solvers including the ones in CVX) and (ii) direct matrix multiplications take too much time. For bigger images and structured matrices like

www.manaraa.com

DFT times DWT, we wrote our own solver for (2.3) by using a modification of the code in L1Magic [49]. We show results using this code on a 256×256 larynx image sequence ($n = 65536$) in Fig. 2.3. This code used the operator form of primal-dual interior point method. With this, one only needs to store the sampling mask which takes $O(n)$ bits of storage and one uses FFT and fast DWT to perform matrix-vector multiplications in $O(n \log n)$ time instead of $O(n^2)$ time. In fact for a $\sqrt{m} \times \sqrt{m}$ image the cost difference is $O(m \log m)$ versus $O(b^4)$. All our code, for both small and large problems, is posted online at `http://www.ece.iastate.edu/~namrata/SequentialCS.html`. This page also links to more experimental results.

2.3.1.1 Comparison of CS and Modified-CS

In Theorem 1 and Corollary 1, we derived sufficient conditions for exact reconstruction using modified-CS. We first compare the sufficient conditions for modified-CS and for CS, expressed only in terms of δ_S's. Sufficient conditions for an algorithm serve as a designer's tool to decide the number of measurements needed for it and in that sense comparing the two sufficient conditions is meaningful.

For modified-CS, from Corollary 1, the sufficient condition in terms of only δ_S's is $2\delta_{2u} + \delta_{3u} + \delta_k + \delta_{k+u}^2 + 2\delta_{k+2u}^2 < 1$. Using $k = s + e - u$, this becomes

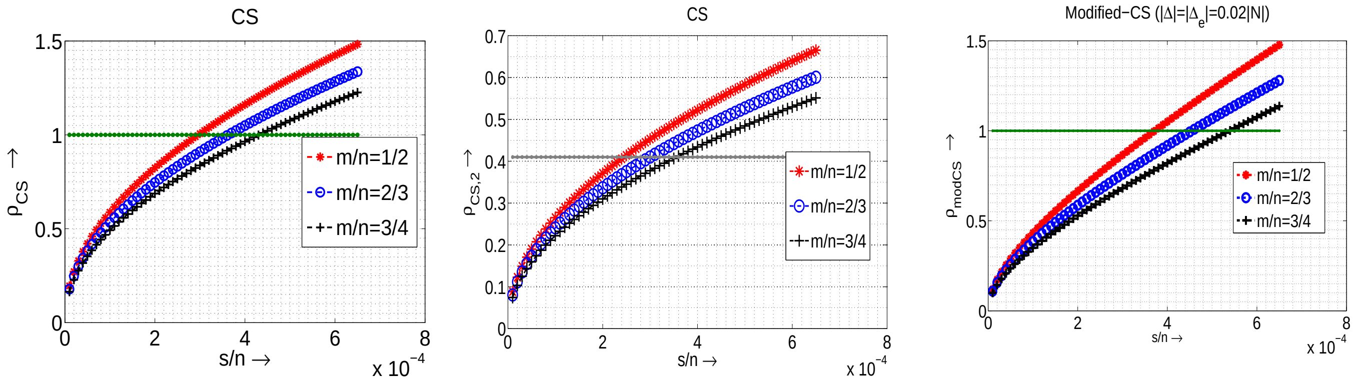
$$2\delta_{2u} + \delta_{3u} + \delta_{s+e-u} + \delta_{s+e}^2 + 2\delta_{s+e+u}^2 < 1. \tag{2.23}$$

For CS, two of the best (weakest) sufficient conditions that use only δ_S's are given in [19, 11] and [10]. Between these two, it is not obvious which one is weaker. Using [19] and [10], CS achieves exact reconstruction if either

$$\delta_{2s} < \sqrt{2} - 1 \ \text{ or } \ \delta_{2s} + \delta_{3s} < 1. \tag{2.24}$$

To compare (2.23) and (2.24), we use $u = e = 0.02s$ which is typical for time series applications (see Fig. 1.2). One way to compare them is to use $\delta_{cr} \leq c\delta_{2r}$ [9, Corollary 3.4] to get the LHS's of both in terms of a scalar multiple of δ_{2u}. Thus, (2.23) holds if $\delta_{s+e+u} < 1/2$ and $\delta_{2u} < 1/132.5$. Since $\delta_{s+e+u} = \delta_{52u} < 52\delta_{2u}$, the second condition implies the first, and so only $\delta_{2u} < 1/132.5$ is sufficient. But, (2.24) holds if $\delta_{2u} < 1/241.5$ *which is clearly stronger.*

المنارة للاستشارات
www.manaraa.com



(a) Plots of ρ_{CS} defined in (2.26) (b) Plots of $\rho_{CS,2}$ defined in (2.26) (c) Plots of ρ_{modCS} defined in (2.25)

Figure 2.1 Plots of ρ_{CS} and $\rho_{CS,2}$ (in (a) and (b)) and ρ_{modCS} (in (c)) against s/n for 3 different values of m/n. For ρ_{modCS}, we used $u = e = s/50$. Notice that, for any given m/n, the maximum allowed sparsity, s/n, for $\rho_{modCS} < 1$ is larger than that for which either $\rho_{CS} < 1$ or $\rho_{CS,2} < \sqrt{2} - 1$. Also, both are much smaller than what is observed in simulations.

Alternatively, we can compare (2.23) and (2.24) using the high probability upper bounds on δ_S as in [18]. Using [18, Eq 3.22], for an $m \times n$ random Gaussian matrix, with high probability (w.h.p.), $\delta_S < g_{n/m}(\frac{S}{n})$, where

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

and binary entropy $H(r) := -r \log r - (1-r)\log(1-r)$ for $0 \leq r \leq 1$. Thus, w.h.p., modified-CS achieves exact reconstruction from random-Gaussian measurements if

$$\rho_{modCS} := 2g_{n/m}\left(\frac{2u}{n}\right) + g_{n/m}\left(\frac{3u}{n}\right) + g_{n/m}\left(\frac{s+e-u}{n}\right) + g_{n/m}\left(\frac{s+e}{n}\right)^2 + 2g_{n/m}\left(\frac{s+e+u}{n}\right)^2 < 1. \quad (2.25)$$

Similarly, from (2.24), w.h.p., CS achieves exact reconstruction from random-Gaussian measurements if either

$$\rho_{CS} := g_{n/m}\left(\frac{2s}{n}\right) + g_{n/m}\left(\frac{3s}{n}\right) < 1 \text{ or } \rho_{CS,2} := g_{n/m}\left(\frac{2s}{n}\right) < \sqrt{2} - 1. \quad (2.26)$$

In Fig. 2.1, we plot ρ_{CS}, $\rho_{CS,2}$ and ρ_{modCS} against s/n for three different choices of m/n. For ρ_{modCS}, we use $u = e = 0.02s$ (from Fig. 1.2). As can be seen, the maximum allowed sparsity, i.e.

المنارة للاستشارات
www.manaraa.com

the maximum allowed value of s/n, for which either $\rho_{CS} < 1$ or $\rho_{CS,2} < \sqrt{2} - 1$ is smaller than that for which $\rho_{modCS} < 1$. Thus, for a given number of measurements, m, w.h.p., modified-CS will give exact reconstruction from random-Gaussian measurements, for larger sparsity sizes, s/n, than CS would. As also noted in [18], in all cases, the maximum allowed s/n is much smaller than what is observed in simulations, because of the looseness of the bounds. For the same reason, the difference between CS and modified-CS is also not as significant.

Table 2.1 Probability of exact reconstruction for modified-CS. Notice that $u = s$ and $e = 0$ corresponds to CS.

(a) $m = 0.16n$

e u	0	$0.08s$	$0.24s$	$0.40s$
$0.04s$	0.9980	0.9900	0.8680	0.4100
$0.08s$	0.8880	0.8040	0.3820	0.0580
s	**(CS)** 0.0000			

(b) $m = 0.19n$

e u	0	$0.08s$	$0.24s$	$0.40s$
$0.08s$	0.9980	0.9980	0.9540	0.7700
$0.12s$	0.9700	0.9540	0.7800	0.4360
s	**(CS)** 0.0000			

(c) $m = 0.25n$

e u	0	$0.08s$	$0.24s$	$0.40s$
$0.04s$	1	1	1	1
$0.20s$	1	1	0.9900	0.9520
$0.35s$	0.9180	0.8220	0.6320	0.3780
$0.50s$	0.4340	0.3300	0.1720	0.0600
s	**(CS)** 0.0020			

(d) $m = 0.30n$

e u	0	$0.08s$	$0.24s$	$0.40s$
$0.04s$	1	1	1	1
$0.20s$	1	1	1	1
$0.35s$	1	1	0.9940	0.9700
$0.50s$	0.9620	0.9440	0.8740	0.6920
s	**(CS)** 0.1400			

(e) $m = 0.40n$

e u	0	$0.40s$
$0.04s$	1	1
$0.20s$	1	1
$0.35s$	1	1
$0.50s$	1	1
s	**(CS)** 0.9820	

المنارة للاستشارات

www.manaraa.com

So far we only compared sufficient conditions. The actual allowed s for CS may be much larger. To actually compare exact reconstruction ability of modified-CS with that of CS, we thus need Monte Carlo. We use the following procedure to obtain a Monte Carlo estimate of the probability of exact reconstruction using CS and modified-CS, for a given A (i.e. we average over the joint distribution of x and y given A).

1. Fix signal length, $n = 256$ and its support size, $s = 0.1n = 26$. Select m, u and e.
2. Generate the $m \times n$ random-Gaussian matrix, A (generate an $m \times n$ matrix with independent identically distributed (i.i.d.) zero mean Gaussian entries and normalize each column to unit ℓ_2 norm)
3. Repeat the following tot $= 500$ times
   (a) Generate the support, N, of size s, uniformly at random from $[1, n]$.
   (b) Generate $(x)_N \sim \mathcal{N}(0, 100I)$. Set $(x)_{N^c} = 0$.
   (c) Set $y := Ax$.
   (d) Generate Δ of size u uniformly at random from the elements of N.
   (e) Generate Δ_e of size e, uniformly at random from the elements of $[1, n] \setminus N$.
   (f) Let $T = N \cup \Delta_e \setminus \Delta$. Run modified-CS, i.e. solve (2.3)). Call the output $\hat{x}_{modCS}$.
   (g) Run CS, i.e. solve (2.3) with T being the empty set. Call the output $\hat{x}_{CS}$.
4. Estimate the probability of exact reconstruction using modified-CS by counting the number of times $\hat{x}_{modCS}$ was equal to x (“equal” was defined as $\|\hat{x}_{modCS} - x\|_2 / \|x\|_2 < 10^{-5}$) and dividing by tot $= 500$.
5. Do the same for CS using $\hat{x}_{CS}$.
6. Repeat for various values of m, u and e.

We set $n = 256$ and $s = 0.1n$ and we varied m between $0.16n = 1.6s$ and $0.4n = 4s$. For each m, we varied u between $0.04s$ to s and e between 0 to $0.4s$. We tabulate our results in Table 2.1. *The*

www.manaraa.com

case $u = s$ and $e = 0$ corresponds to CS. Notice that when m is just $0.19n = 1.9s < 2s$, modified-CS achieves exact reconstruction more than 99.8% of the times if $u \leq 0.08s$ and $e \leq 0.08s$. In this case, CS has *zero* probability of exact reconstruction. With $m = 0.3n = 3s$, CS has a very small (14%) chance of exact reconstruction. On the other hand, modified-CS works almost all the time for $u \leq 0.2s$ and $e \leq 0.4s$. CS needs at least $m = 0.4n = 4s$ to work reliably.

The above simulation was done in a fashion similar to that of [18]. It does not compute the m required for Theorem 1 to hold. Theorem 1 says that if m is large enough for a given s, u, e, so that the two conditions given there hold, modified-CS will *always* work. But all we show above is that (1) for certain large enough values of m, the Monte Carlo estimate of the probability of exact reconstruction using modified-CS is 1 (probability computed by averaging over the joint distribution of x and y); *and (2) when u, e are small, this happens for much smaller values of m with modified-CS than with CS.*

This issue has been discussed in detail in [50, 51] (probability or expected chance of exact reconstruction). In [50], the authors give a greedy pursuit algorithm to find these pathological cases for CS, i.e. to find the sparsest vector x for which CS does not give exact reconstruction. The support size of this vector then gives an upper bound on the sparsity that CS can handle. Developing a similar approach for modified-CS is a useful open problem.

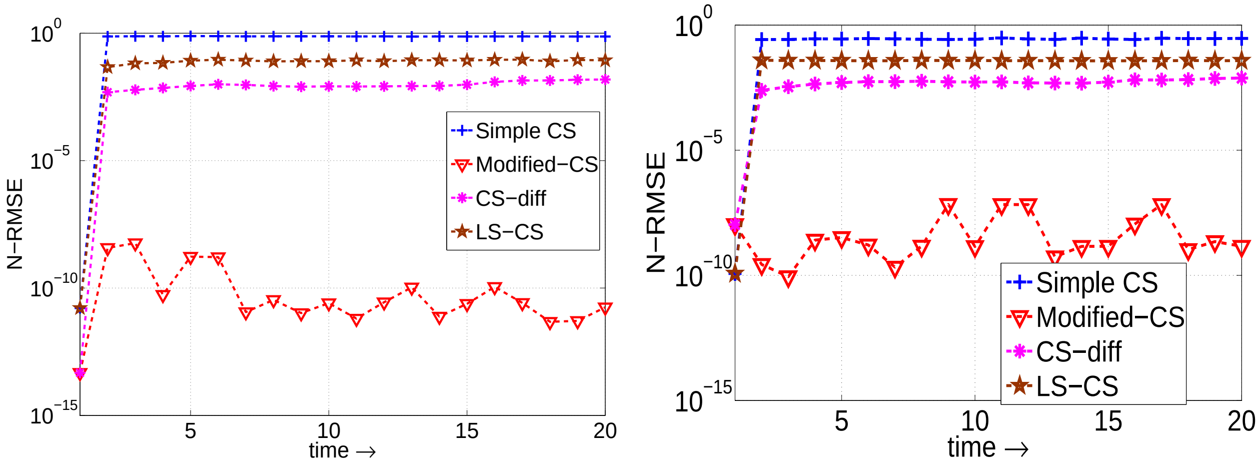
2.3.1.2 Sparsified and True (Compressible) Single Image

We first evaluated the single image reconstruction problem for a sparsified image. The image used was a 32×32 cardiac image (obtained by decimating the full 128×128 cardiac image shown in Fig. 1.2), i.e. $n = 1024$. Its support size $s = 107 \approx 0.1n$. We used the set of indices of the approximation coefficients as the known part of the support, T. Thus, $k = |T| = 64$ and so $u = |\Delta| \geq 43$ which is a significantly large fraction of s. We compare the N-RMSE in Table 2.2. Even with such a large unknown support size, modified-CS achieved exact reconstruction from 29% random Gaussian and 19% partial Fourier measurements. CS error in these cases was 34% and 13% respectively.

We also did a comparison for actual cardiac and larynx images (which are only approximately sparse). The results are tabulated in Table 2.2. Modified-CS works better than CS, though not by much since $|\Delta|$ is a large fraction of $|N|$. Here N refers to the $\beta\%$ support for any large β, e.g. $\beta = 99$.

المنارة للاستشارات

www.manaraa.com



(a) $H = G_r$, m_0=0.5n, m=0.16n

(b) $H = MF$, m_0=0.5n, m=0.16n

Figure 2.2 Reconstructing the *sparsified* 32×32 cardiac image sequence. $s \approx 0.1n$, $u \approx 0.01n$, $e \approx 0.005n$. (a) $H = G_r$, (b) $H = MF$. Similar results were also obtained for the larynx sequence.

Table 2.2 Reconstruction Error (N-RMSE)

	Sparsified Cardiac	True Cardiac	True Larynx
CS ($H = G_r$, $m = 0.29n$)	0.34	0.36	0.090
Mod-CS ($H = G_r$, $m = 0.29n$)	0	0.14	0.033
CS ($H = MF$, $m = 0.19n$)	0.13	0.12	0.097
Mod-CS ($H = MF$, $m = 0.19n$)	0	0.11	0.025

2.3.1.3 Sparsified Image Sequences

We compared modified-CS with simple CS (CS at each time instant), CS-diff(CS-residual) and LS-CS [43] for the sparsified 32×32 cardiac sequence in Fig. 2.2. Modified-CS was implemented as in Algorithm 1. At $t = 0$, the set T was empty and we used 50% measurements. For this sequence, $|N_t| \approx 0.1n = 107$, $u = |\Delta| \leq 10 \approx 0.01n$ and $e = |\Delta_e| \leq 5 \approx 0.005n$. Since $u \ll |N_t|$ and $e \ll |N_t|$, modified-CS achieves exact reconstruction with as few as 16% measurements at $t > 0$. Fig. 2.2(a) used $H = G_r$ (compression/single-pixel imaging) and Fig. 2.2(b) used $H = MF$ (MRI). As can be seen, simple CS has very large error. CS-diff and LS-CS also have significantly nonzero error since the exact sparsity size of both the signal difference and the signal residual is equal to/larger than the signal's sparsity size. *Modified-CS error is 10^{-8} or less (exact for numerical implementation)*. Similar

www.manaraa.com

conclusions were also obtained for the sparsified larynx sequence, see [40, Fig. 3].

2.3.1.4 True (Compressible) Image Sequences

Finally we did the comparison for actual image sequences which are only compressible. We show results on the larynx (vocal tract) image sequence of Fig. 1.2. For Fig. 2.3 we used the entire 256×256 image sequence with partial Fourier measurements. *At $t = 0$, modified-CS and LS-CS used T to be the set of indices of the approximation coefficients.*

Fig. 2.3 shows reconstruction of the full larynx sequence using $H = MF$, $m = 0.19n$ and three choices of m_0. In 2.3(a), we compare the reconstructed image sequence using modified-CS with that using simple CS. The error (N-RMSE) was 8-11% for CS, while it was stable at 2% or lesser for modified-CS. Since m_0 is large enough for CS to work, the N-RMSE of CS-diff (not shown) also started at a small value of 2% for the first few frames, but kept increasing slowly over time. In 2.3(b), 2.3(c), we show N-RMSE comparisons with simple CS, CS-diff and LS-CS. In the plot shown, the LS-CS error is close to that of modified-CS because we implemented LS estimation using conjugate gradient and did not allow the solution to converge (forcibly ran it with a reduced number of iterations). Without this tweeking, LS-CS error was much higher, since the computed initial LS estimate itself was inaccurate.

Notice from Fig. 2.3, that *modifiedCS significantly outperform CS and CS-diff. In most cases, both also outperform LS-CS.* In Fig. 2.3(c), CS-diff performs so poorly primarily because the initial error at $t = 0$ is very large (since we use only $m_0 = 0.19n$). As a result the difference signal at $t = 1$ is not compressible enough, making its error large and so on. But even when m_0 is larger and the initial error is small, CS-diff is still the worst, although the difference in errors is not as large, e.g. in Fig. 2.3(b).

2.3.2 Experimental results of reg-mod-BP and reg-mod-CS

2.3.2.1 Comparing reg-mod-BP with modified-CS

In this section, we show two types of numerical experiments. The first simulates quantized signals and signal estimates. This is the case where some constraints are active with nonzero probability. The good set, $T_g = T_{\text{a+g}} \cup T_{\text{a-g}}$ is also non empty with nonzero probability. Hence, for a given small

المنارة للاستشارات
www.manaraa.com

enough n, reg-mod-BP has significantly higher exact reconstruction probability, $p_{\text{exact}}(n)$, as compared to both mod-CS [25] and weighted ℓ_1 [26] and much higher than that of BP [1, 18]. Alternatively, it also requires a significantly reduced n for exact reconstruction with probability one, $n_{\text{exact}}(1)$. In computing $p_{\text{exact}}(n)$ we average over the distribution of x, T and $\hat{\mu}$, as also in [25, 18]. All numbers are computed based on 100 Monte Carlo simulations. To compute $n_{\text{exact}}(1)$, we tried various values of n for each algorithm and computed the smallest n required for exact recovery always (in all 100 simulations).

We also do a second simulation where signal estimates are not quantized.

In the following steps, the notation $z \sim$ discrete-uniform$(a_1, a_2, \ldots a_n)$ means that z is equally likely to be equal to a_1, a_2, $\ldots$ or a_n. We use $\pm a$ as short for $+a, -a$. Also, $z \sim$ uniform(a, b) generates a scalar uniform random variable in the range $[a, b]$. The notation $x_i \overset{iid}{\sim}$ P for all $i \in S$ means that, for all $i \in S$, each x_i is identically distributed according to P and is independent of all the others.

	$2K$	**BP**	**mod-CS**	**weighted ℓ_1**	**Reg-mod-BP**
$p_{\text{exact}}(0.15m)$	4	0	0.18	0.16	0.64
N-RMSE($0.15m$)	4	1.011	0.059	0.060	0.029
$n_{\text{exact}}(1)$	4	0.39m	0.21m	0.21m	0.18m
$p_{\text{exact}}(0.15m)$	10	0	0.18	0.16	0.39
N-RMSE($0.15m$)	10	1.011	0.059	0.060	0.032
$n_{\text{exact}}(1)$	10	0.4m	0.21m	0.21m	0.20m

Table 2.3 Quantized signals and signal estimates. Recall that $k = |T| = 26$. For $2K = 4$, the expected sizes of T_a, T_g and T_b are $\mathbb{E}[|T_a|] = 10.01$, $\mathbb{E}[|T_g|] = 5.27$ and $\mathbb{E}[|T_b|] = 20.73$. For $2K = 10$, $\mathbb{E}[|T_a|] = 4.28$, $\mathbb{E}[|T_g|] = 2.3$ and $\mathbb{E}[|T_b|] = 23.7$.

	BP	**mod-CS**	**weighted ℓ_1**	**Reg-mod-BP**
$p_{\text{exact}}(0.15m)$	0	0.26	0.26	0.57
N-RMSE($0.15m$)	0.967	0.152	0.152	0.082
$n_{\text{exact}}(1)$	0.4m	0.21m	0.21m	0.20m

Table 2.4 Quantized signals and signal estimates: case 2. Recall that $k = |T| = 26$. The expected sizes of T_a, T_g and T_b are $\mathbb{E}[|T_a|] = 9.02$, $\mathbb{E}[|T_g|] = 4.58$ and $\mathbb{E}[|T_b|] = 21.42$.

المنارة للاستشارات
www.manaraa.com

	BP	**mod-CS**	**weighted ℓ_1**	**Reg-mod-BP**
$p_{\text{exact}}(0.18\text{m})$	0	0.87	0.87	0.87
N-RMSE(0.18m)	0.961	0.0175	0.0177	0.0123
N-RMSE(0.11m)	1.05	0.179	0.175	0.0635
$n_{\text{exact}}(1)$	0.39m	0.21m	0.21m	0.21m

Table 2.5 The non quantized case.

For the quantized case, x was a $m = 256$ length sparse vector with support size $|N| = 0.1m = 26$ and support estimate error sizes $u = |\Delta| = |\Delta_e| = 0.1|N| = 3$. We generated the matrix A once as an $n \times m$ random Gaussian matrix (generate an $n \times m$ matrix with i.i.d zero mean Gaussian entries and normalize each column to unit ℓ_2 norm). The following steps were repeated tot $= 100$ times.

1. The support set, N, of size $|N|$, was generated uniformly at random from $[1, m]$. The support misses set, Δ, of size u, was generated uniformly at random from the elements of N. The support extras set, Δ_e, also of size u, was generated uniformly at random from the elements of N^c. The support estimate, $T = N \cup \Delta_e \setminus \Delta$ and thus $|T| = |N| = 26$.

2. We generated $x_i \overset{iid}{\sim}$ discrete-uniform(± 1) for $i \in N \cap T$; $x_i \overset{iid}{\sim}$ discrete-uniform(± 0.1) for $i \in \Delta$, and $x_i = 0$ for $i \in N^c$. $x_{N \cap T}$ and x_Δ are also independent of each other. We generated $\hat{\mu}_T = x_T + \nu$ where $\nu_i \overset{iid}{\sim}$ discrete-uniform$(0, \pm\frac{\rho}{K}, \pm 2\frac{\rho}{K}, \cdots \pm \rho)$ for $i \in T \cap N$ and $\nu_i \overset{iid}{\sim}$ discrete-uniform$(\pm\frac{\rho}{K}, \pm 2\frac{\rho}{K}, \cdots \pm \rho)$ for $i \in \Delta_e$. We used $\rho = 0.1$ and tried two choices of K. Notice that, for a given K, the number of equally likely values that $x_i - \hat{\mu}_i$ for $i \in T$ can take are roughly $2K + 1$ ($2K$ when $i \in \Delta_e$). The constraint is active when $x_i - \hat{\mu}_i$ is equal to $\pm\rho$. Thus, the expected size of the active set is roughly $\frac{2}{2K+1}|T|$.

3. We generated $y = Ax$. We solved reg-mod-BP given in (2.14) with $\rho = 0.1$; BP given in (1.6); mod-CS given in (2.3); and weighted ℓ_1 given in (1.8) with various choices of γ: $[0.1\ 0.05\ 0.01\ 0.001]$. We used the CVX optimization package, `http://www.stanford.edu/boyd/cvx/`, which uses primal-dual interior point method for solving the minimization problem.

We computed $p_{\text{exact}}(n)$ as the the number of times $\hat{x}$ was equal to x (“equal” was defined as $\|\hat{x} -$

المنارة للاستشارات
www.manaraa.com

$x\|_2/\|x\|_2 < 10^{-5}$) divided by tot $= 100$. For weighted ℓ_1, we computed $p_{\text{exact}}(n)$ for each choice of γ and recorded the largest one. This corresponded to $\gamma = 0.1$. We tabulate results in Table 2.3. In the first row, we record $p_{\text{exact}}(0.15m)$ for all the methods, when using $K = 2$. We also record the Monte Carlo average of the sizes of the active set $|T_a| = |T_{\text{a+}} \cup T_{\text{a-}}|$; of the good set, $|T_g| = |T_{\text{a+g}} \cup T_{\text{a-g}}|$ and of the bad set $|T_b| = k - |T_g|$. In the second row, we record the normalized root mean squared error (N-RMSE). In the third row, we record $n_{\text{exact}}(1)$. In the next three rows, we repeat the same things with $K = 5$.

As can be seen, $|T_g|$ is about half the size of the active set, $|T_a|$. As K is increased, $|T_a|$ and hence $|T_g|$ reduces ($|T_b|$ increases) and thus $p_{\text{exact}}(0.15m)$ decreases and $n_{\text{exact}}(1)$ increases. Also, for mod-CS and weighted ℓ_1, $p_{\text{exact}}(0.15m)$ is significantly smaller than for reg-mod-BP, while $n_{\text{exact}}(1)$ is larger.

Next, we simulated a more realistic scenario – the case of 3-bit quantized images (both x and $\hat{\mu}$ take integer values between 0 to 7). Here again $m = 256$, $|N| = 0.1m = 26$, and $u = |\Delta| = |\Delta_e| = 0.1|N| = 3$. The sets N, Δ, Δ_e and T were generated as before. We generated $x_i \overset{iid}{\sim}$ discrete-uniform$(3, 4, \dots 7)$ for $i \in N \cap T$; $x_i \sim$ discrete-uniform$(1, 2)$ for $i \in \Delta$; and $x_i = 0$ for $i \in N^c$. Also, $\hat{\mu}_T =$ clip$(x_T + \nu)$ where $\nu_i \sim$ discrete-uniform$(-2, -1, 0, 1, 2)$ for $i \in T \cap N$; and $\nu_i \sim$ discrete-uniform$(-2, -1, 1, 2)$ for $i \in \Delta_e$. Also clip(z) clips any value more than 7 to 7 and any value less than zero to zero. Clearly, in this case $\rho = 2$. We record our results in Table 2.4. Similar conclusions as before can be drawn.

Finally, we simulated the non-quantized case. We used $m = 256$, $|N| = 0.1m = 26$, and $u = |\Delta| = |\Delta_e| = 0.1|N| = 3$. We generated $x_i \overset{iid}{\sim}$ discrete-uniform(± 1) for $i \in N \cap T$; $x_i \overset{iid}{\sim}$ discrete-uniform(± 0.1) for $i \in \Delta$, and $x_i = 0$ for $i \in N^c$. The signal estimate, $\hat{\mu}_T = x_T + \nu$ where $\nu_i \overset{iid}{\sim}$ uniform$(-\rho, \rho)$ with $\rho = 0.1$. We tabulate our results in Table 2.5. Since ν is a real vector (not quantized), the probability of any constraint being active is zero. Thus, as expected, p_{exact} and n_{exact} are the same for reg-mod-BP and mod-CS and weighted ℓ_1, though significantly better than BP. However, the N-RMSE for reg-mod-BP is significantly lower than that for mod-CS and weighted ℓ_1 also, particularly when $n = 0.11m$.

www.manaraa.com

2.3.2.2 Comparing reg-mod-CS with Modified-CS

We ran a Monte Carlo simulation to compare Modified-CS with reg-mod-CS for sparse signals. We fixed $n = 256$, $s = 26 \approx 0.1n$, $u = e = 0.08s$. We used $m = 0.16n, 0.12n, 0.11n$ in three sets of simulations done in a fashion similar to that of Chapter 2.3.1.1, but with the following change. In each run of a simulation, we generated each element of $\mu_{N \setminus \Delta}$ to be i.i.d. ± 1 with probability (w.p.) 1/2 and each element of μ_{Δ} and of μ_{Δ_e} to be i.i.d. ± 0.25 w.p. 1/2. We generated $x_N \sim \mathcal{N}(\mu_N, 0.01I)$ and we set $x_{N^c} = 0$. We set $y := Ax$. We tested reg-mod-CS with various values of γ ($\gamma = 0$ corresponds to modified-CS). We used tot $= 50$. The results are tabulated in Table 2.6. We computed the exact reconstruction probability as in Chapter 2.3.1.1 by counting the number of times $\hat{x}_{reg}$ equals x and normalizing. As can be seen, reg-mod-CS does not improve the exact reconstruction probability, in fact it can reduce it. This is primarily because the elements of $(\hat{x}_{reg})_{\Delta_e}$ are often nonzero, though small[2]. But, it significantly reduces the reconstruction error, particularly when m is small.

Table 2.6 Comparing probability of exact reconstruction (prob) and reconstruction error (error) of reg-mod-CS with different γ's. $\gamma = 0$ corresponds to modified-CS.

(a) $m = 0.16n$

γ	0 **(modCS)**	0.001	0.05	0.1	0.5	1
prob	0.76	0.76	0.74	0.74	0.70	0.34
error	0.0484	0.0469	0.0421	0.0350	0.0273	0.0286

(b) $m = 0.12n$

γ	0 **(modCS)**	1
prob	0.04	0
error	0.2027	0.0791

(c) $m = 0.11n$

γ	0 **(modCS)**	1
prob	0	0
error	0.3783	0.0965

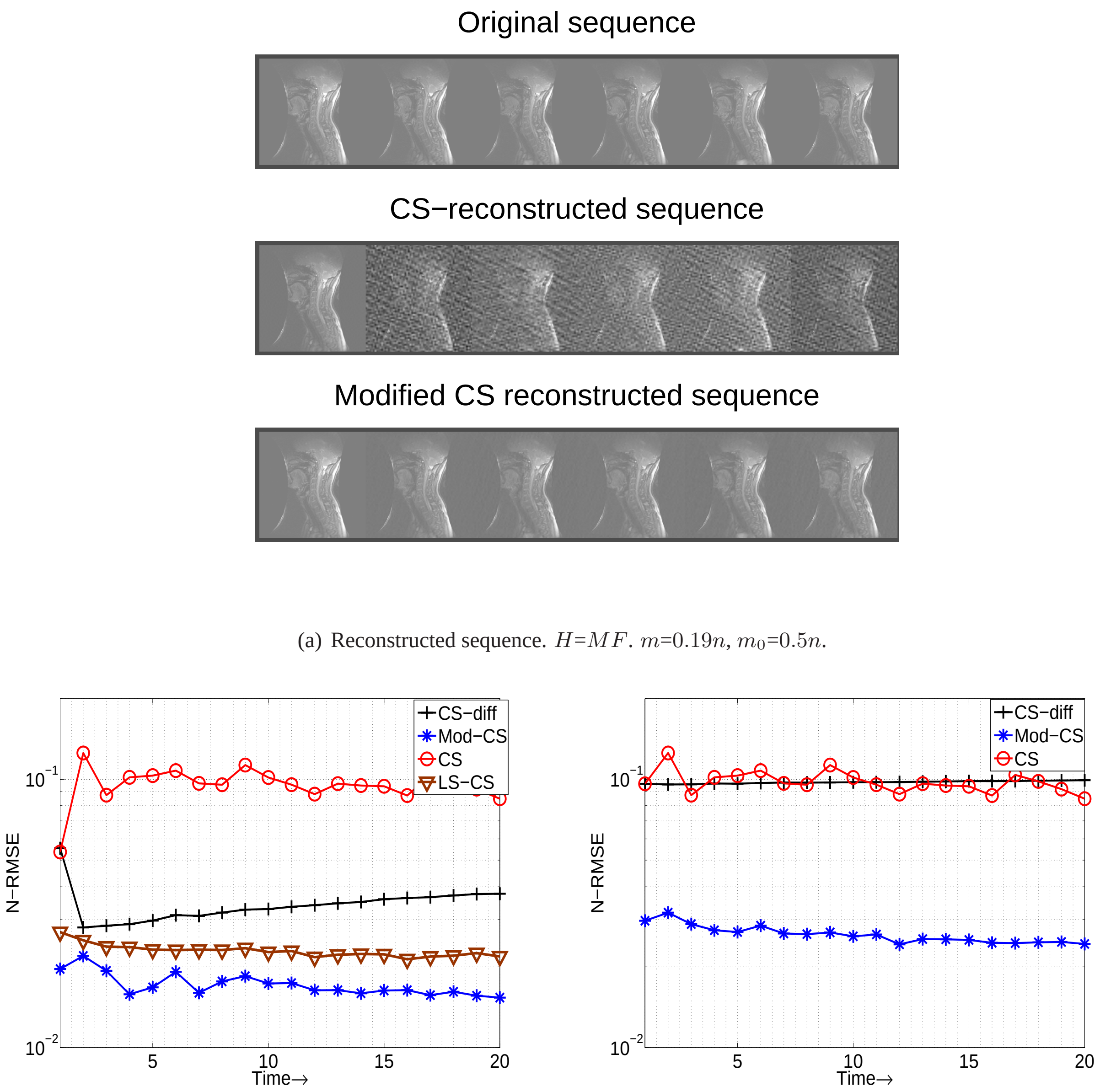
We compared reg-mod-CS with other algorithms in Fig. 2.4. We used a 32×32 block of it with random Gaussian measurements. For the subfigures in Fig. 2.4, we used $H = G_r$ (random Gaussian) and $m_0 = 0.19n$. Fig. 2.4(a) and 2.4(b) used $m = 0.19n, 0.06n$ respectively. At each t, RegModCS-MAP solved (2.22) with λ_p, σ_p^2 estimated using (A.44) from a few frames of the sequence treated as training data. The resulting $\gamma = \hat{\lambda}_p / 2\hat{\sigma_p^2}$ was 0.007. RegModCS-exp-opt solved (2.20) with

[2] But if we use $\hat{x}_{reg}$ to first estimate the support using a small threshold, α, and then estimate the signal as ${A_{\hat{N}}}^{\dagger} y$, this probability does not decrease as much and in fact it even increases when m is smaller.

المنارة للاستشارات
www.manaraa.com

$T = \hat{N}_{t-1}$, $\mu_T = (\hat{x}_{reg,t-1})_T$ and we experimented with many values of γ and chose the one which gave the smallest error. Notice from Fig. 2.4(a) that RegModCS-MAP gives MSEs which are very close to those of RegModCS-exp-opt.

المنارة للاستشارات
www.manaraa.com



(a) Reconstructed sequence. H=MF. m=0.19n, m_0=0.5n.

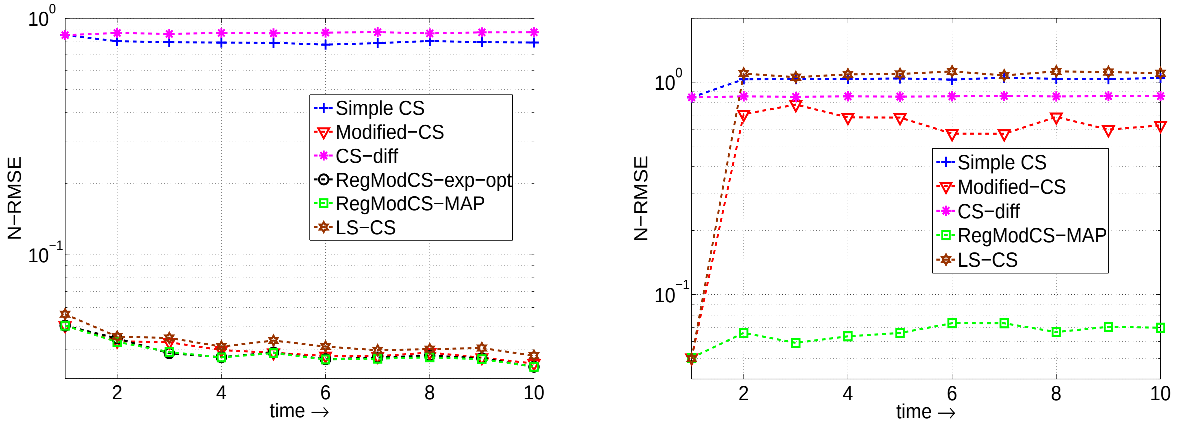
(b) H=MF, m_0=0.2n, m=0.19n

(c) H=MF, m_0=0.19n, m=0.19n

Figure 2.3 Reconstructing the 256x256 *actual (compressible)* vocal tract (larynx) image sequence from *simulated MRI* measurements, i.e. $H = MF$. All three figures used $m = 0.19n$ for $t > 0$ but used different values of m_0. Image size, $n = 256^2 = 65536$. 99% energy support, $|N_t| \approx 0.07n$; $u \approx 0.001n$. In Fig. 2.3(a), modified-CS used $\alpha = 10^2$ which is the smallest magnitude element in the 99% support.

المنارة للاستشارات

www.manaraa.com



(a) H=G_r, m_0=0.19n, m=0.19n

(b) H=G_r, m_0=0.19n, m=0.06n

Figure 2.4 Reconstructing a 32×32 block of the *actual (compressible)* larynx sequence from random Gaussian measurements. $n = 1024$, 99%-energy support size, $s \approx 0.07n$, $u \approx 0.001n$ and $e \approx 0.002n$. Modified-CS used $\alpha = 50^2$ when $m = 0.19n$ and increased it to $\alpha = 80^2$ when $m = 0.06n$.

www.manaraa.com

CHAPTER 3. Sparse Reconstruction for Noisy Measurements with Partial Support and Signal Knowledge

In Chapter 2, we introduced modified-CS and reg-mod-BP for the noiseless measurements' case. Sufficient conditions for exact reconstruction were derived and it was argued that these are much weaker than those needed for CS. In this chapter, we bound the reconstruction error of modified-BPDN and regularized modified-BPDN which are noisy cases of modified-CS and reg-mod-BP. We use a strategy similar to the results of [2] to bound the reconstruction error and hence, just like in [2], *the bounds we obtain are computable*. Then we also derive the bounds without sufficient conditions that are much tighter. Simulations are shown to compare the bounds.

3.1 Modified-BPDN

In this section, our goal is to reconstruct the m-length sparse signal x from the n-length measurement y with $m > n$

$$y := Ax + w \tag{3.1}$$

The measurement is obtained from an $n \times m$ measurement matrix A and corrupted by a n-length vector noise w. The support of x denoted as N consists of three parts: $N \triangleq T \cup \Delta \setminus \Delta_e$ where Δ and T are disjoint and $\Delta_e \subseteq T$. We use the partially known support T which the known part of support while Δ_e is the error in the known part of support and Δ is the unknown part. We also define $N_e \triangleq T \cup \Delta = N \cup \Delta_e$.

In Chapter 2, equation (2.3) gives the modified-CS algorithm under noiseless measurements. We relax the equality constraint of this equation to propose modified-BPDN algorithm using a modification

المنارة للاستشارات

www.manaraa.com

of the BPDN idea[1]. We solve

$$\min_{b} \quad \frac{1}{2}\|y - Ab\|_2^2 + \gamma \|b_{T^c}\|_1 \tag{3.2}$$

Then the solution to this convex optimization problem $\hat{x}$ will be the reconstructed signal of the problem. In [28, 52], we provided a computable bound for the reconstruction error as well as the sufficient conditions. We will not address it here since mod-BPDN is a special case of reg-mod-BPDN and the bound follows by setting $\lambda = 0$ of the bound for reg-mod-BPDN. We will compare the bounds of BPDN, mod-BPDN in the next section.

3.2 Regularized Modified-BPDN for Noisy Sparse Reconstruction with Partial Erroneous Support and Signal Value Knowledge

In previous section, we introduced modified-BPDN using partially known support to reconstruct the sparse signal from noisy measurements. In this section, we study the method to reconstruct using both the support information and the signal estimate on it in this chapter. Our goal is still to reconstruct an m-length sparse vector, x, with support, N, from an $n < m$ length noisy measurement vector, y, satisfying

$$y := Ax + w \tag{3.3}$$

when the following two things are available: (i) partial, and partly erroneous, knowledge of the signal's support, denoted by T; and (ii) an erroneous estimate of the signal values on T, denoted by $(\hat{\mu})_T$. w is the measurement noise and A is the measurement matrix. The true support of the signal, N, can be rewritten as $N = T \cup \Delta \setminus \Delta_e$ and $\Delta \triangleq N \setminus T$ and $\Delta_e \triangleq T \setminus N$ are the errors in the support estimate. The signal estimate is assumed to be zero along T^c, i.e.

$$\hat{\mu} = \begin{bmatrix} (\hat{\mu})_T \\ \mathbf{0}_{T^c} \end{bmatrix} \tag{3.4}$$

and the signal itself can be rewritten as

$$\begin{aligned} (x)_{N \cup T} &= (\hat{\mu})_{N \cup T} + \nu \\ (x)_{N^c} &= 0 \end{aligned} \tag{3.5}$$

المنارة للاستشارات
www.manaraa.com

where ν denotes the error in the prior signal estimate. It is assumed that the error energy, $\|\nu\|_2^2$, is small compared to the signal energy, $\|x\|_2^2$.

In this section, we introduce regularized modified-BPDN (reg-mod-BPDN) and obtain a computable bound on its reconstruction error using an approach motivated by [2]. Reg-mod-BPDN solves

$$\min_{b} \;\; \gamma\|b_{T^c}\|_1 + \frac{1}{2}\|y - Ab\|_2^2 + \frac{1}{2}\lambda\|b_T - \hat{\mu}_T\|_2^2 \tag{3.6}$$

i.e. it tries to find the signal that is sparsest outside the set T, while being "close enough" to $\hat{\mu}_T$ on T, and while satisfying the data constraint. Reg-mod-BPDN uses the fact that T is a good estimate of the true support, N, and that $\hat{\mu}_T$ is a good estimate of x_T. In particular, for $i \in \Delta_e$, this implies that $|\hat{\mu}_i|$ is close to zero (since $x_i = 0$ for $i \in \Delta_e$). We also show how to use the reconstruction error bound result to obtain another computable bound that holds without any sufficient conditions and is tighter. This allows easy bound comparisons of the various approaches. A similar result for mod-BPDN and BPDN follows as a direct corollary.

Before we bound the reconstruction error for reg-mod-BPDN, we will discuss some related approaches which may be confused with reg-mod-BPDN. Notice that Reg-mod-BPDN may also be interpreted as a Bayesian CS or a model-based CS approach. Recent work in this area includes [53, 54, 13, 55, 56, 57, 58].

3.2.1 Some Related Approaches

Before going further, we discuss below *a few approaches that are related to, but different from reg-mod-BPDN, and we argue when and why these will be worse than reg-mod-BPDN.* We show comparisons with all these in Fig. 3.1.

One seemingly related approach is what can be called *CS-mod-residual.* It computes

$$\hat{x}_T = \hat{\mu}_T, \;\; \hat{x}_{T^c} = \hat{b}_c, \text{ where } \hat{b}_c \text{ solves}$$

$$\min_{b_c} \;\; \frac{1}{2}\|y - A_T\hat{\mu}_T - A_{T^c}b_c\|_2^2 + \gamma\|b_c\|_1 \tag{3.7}$$

where b_c stands for $(b)_{T^c}$. This is solving a sparse recovery problem on T^c, i.e. it is implicitly assuming that x_T is either equal to $\hat{\mu}_T$ or very close to it. Thus, this also works only when the signal value prior is very strong.

www.manaraa.com

Both CS-residual and CS-mod-residual can be interpreted as extensions of BPDN, and [2, Theorem 8] can be used to bound their error. In either case, the bound will contain terms proportional to $\|(x_T - \hat{\mu}_T)\|_2$ and as a result, it will be large whenever the prior is not strong enough[1]. This is also seen from our simulation experiments shown in Fig. 3.1 where we provide comparisons for the case of good signal value prior (0.1% error in initial signal estimate) and bad signal value prior (10% error in initial signal estimate). We vary support errors from 5% to 20% misses, while keeping the extras fixed at 10%.

Reg-mod-BPDN can also be confused with *modified-CS-residual* which computes[40]

$$\begin{aligned} \hat{x} &= \hat{\mu} + \hat{b}, \text{ where } \hat{b} \text{ solves} \\ & \min_b \ \frac{1}{2}\|y - A\hat{\mu} - Ab\|_2^2 + \gamma\|b_{T^c}\|_1 \end{aligned} \tag{3.8}$$

This is indeed related to reg-mod-BPDN and in fact this inspired it. We studied this empirically in Chapter 6. However, one cannot get good error bounds for it in any easy fashion. Notice that the minimization is over the entire vector b, while the ℓ_1 cost is only on b_{T^c}.

One may also consider solving the following variant of reg-mod-BPDN (we call this *reg-mod-BPDN-var*):

$$\min_b \quad \gamma\|b_{T^c}\|_1 + \frac{1}{2}\|y - Ab\|_2^2 + \frac{1}{2}\lambda\|b - \hat{\mu}\|_2^2 \tag{3.9}$$

Since $\hat{\mu}$ is supported on T, the regularization term can be rewritten as $\lambda\|b - \hat{\mu}\|_2^2 = \lambda\|b_T - \hat{\mu}_T\|_2^2 + \lambda\|b_{T^c}\|_2^2$. Thus, in addition to the ℓ_1 norm cost on b_{T^c} imposed by the first term, this last term is also imposing an ℓ_2 norm cost on it. If λ is large enough, the ℓ_2 norm cost will encourage the energy of the solution to be spread out on T^c, thus causing it to be less sparse. Since the true x is very sparse on T^c ($|\Delta|$ is small compared to the support size also), we will end up with a larger recovery error[2]. [see Fig. 3.1(a)]. However, if we compare the two approaches for compressible signal sequences, e.g. the larynx sequence, it is difficult to say which will be better [see Fig. 3.3].

Finally, one may solve the following (*we can call it reg-BPDN*)

$$\min_b \quad \gamma\|b\|_1 + \frac{1}{2}\|y - Ab\|_2^2 + \frac{1}{2}\lambda\|b - \hat{\mu}\|_2^2 \tag{3.10}$$

[1] In either case, one can assume that $(x - \hat{\mu})$ is supported on Δ and the "noise" is $w + A_T(x_T - \hat{\mu}_T)$. Thus, CS-residual error can be bounded by $C(A,\Delta)(\|w\|_2 + \|A_T(x_T - \hat{\mu}_T)\|_2)$ while CS-mod-residual error can be bounded by $\|x_T - \hat{\mu}_T\|_2 + C(A_{T^c},\Delta)(\|w\|_2 + \|A_T(x_T - \hat{\mu}_T)\|_2)$.

[2] In the limit if $\sqrt{\lambda/2}$ is much larger than γ, we may get a completely non-sparse solution.

المنارة للاستشارات
www.manaraa.com

This has two limitations. (1) Like CS-residual, this also does not use the fact that when T is an accurate estimate of the true support, $(x)_{T^c}$ is much more sparse compared with the full $(x - \hat{\mu})$. (2) Its last term is the same as that of reg-mod-BPDN-var which also causes the same problem as above.

3.3 Regularized Modified-BPDN (Reg-mod-BPDN)

Consider the sparse recovery problem when partial support knowledge is available. As explained earlier, one can use mod-BPDN given in (3.2). When the support estimate is accurate, i.e. $|\Delta|$ and $|\Delta_e|$ are small, mod-BPDN provides accurate recovery with fewer measurements than what BPDN needs. However, it puts no cost on b_T except the cost imposed by the data term. Thus, when very few measurements are available or when the noise is large, b_T can become larger than required (in order to reduce the data term). A similar, though lesser, bias will occur with weighted ℓ_1 also when $\gamma' < \gamma$. To address this, when reliable prior signal value knowledge is available, we can instead solve

$$\min_b \quad L(b) \triangleq \gamma \|b_{T^c}\|_1 + \frac{1}{2}\|y - Ab\|_2^2 + \frac{1}{2}\lambda \|b_T - \hat{\mu}_T\|_2^2 \tag{3.11}$$

which we call *reg-mod-BPDN*. Its solution, denoted by $\hat{x}$, serves as the reconstruction of the unknown signal, x. Notice that the first term helps to find the solution that is sparsest outside T, the second term imposes the data constraint while the third term imposes closeness to $\hat{\mu}$ along T.

Mod-BPDN is the special case of (3.11) when $\lambda = 0$. BPDN is also a special case with $\lambda = 0$ and $T = \emptyset$ (so that $\Delta = N$).

3.3.1 Limitations and Assumptions

A limitation of adding the regularizing term, $\lambda \|b_T - \hat{\mu}_T\|_2^2$ is as follows. It encourages the solution to be close to $(\hat{\mu})_{\Delta_e}$ which is not zero. As a result, $(\hat{x})_{\Delta_e}$ will also not be zero (except if λ is very small) even though $(x)_{\Delta_e} = 0$. Thus, even in the noise-free case, reg-mod-BPDN will not achieve exact reconstruction. In both noise-free and noisy cases, if $(\hat{\mu})_{\Delta_e}$ is large, $(\hat{x})_{\Delta_e}$ being close to $(\hat{\mu})_{\Delta_e}$ can result in large error. Thus, we need the assumption that $(\hat{\mu})_{\Delta_e}$ is small.

For the reason above, when we estimate the support of $\hat{x}$, we need to use a nonzero threshold, i.e.

المنارة للاستشارات

www.manaraa.com

compute

$$\hat{N} = \{i : |\hat{x}_i| > \rho\} \tag{3.12}$$

with a $\rho > 0$. We note that thresholding as above is done *only* for support estimation and not for improving the actual reconstruction. Support estimation is required in dynamic reg-mod-BPDN (described below) where we use the support estimate from the previous time instant as the support knowledge, T, for the current time.

In summary, to get a small error reconstruction, reg-mod-BPDN requires the following (this can also be seen from the result of Theorem 4):

1. T is a good estimate of the true signal's support, N, i.e. $|\Delta|$ and $|\Delta_e|$ are small compared to $|N|$; and
2. $\hat{\mu}_T$ is a good estimate of x_T. For $i \in \Delta_e$, this implies that $|\hat{\mu}_i|$ is close to zero (since $x_i = 0$ for $i \in \Delta_e$).
3. For accurate support estimation, we also need that most nonzero elements of x are larger than $\max_{i \in \Delta_e} |\hat{\mu}_i|$ (for exact support estimation, we need this to hold for all nonzero elements of x).

The smallest nonzero elements of x are usually on the set Δ. In this case, the third assumption is equivalent to requiring that most elements of x_Δ are larger than $\max_{i \in \Delta_e} |\hat{\mu}_i|$.

3.3.2 Dynamic Reg-Mod-BPDN for Recursive Recovery

An important application of reg-mod-BPDN is for recursively reconstructing a time sequence of sparse signals from undersampled measurements, e.g. for dynamic MRI. To do this, at time t we solve (3.11) with $T = \hat{N}_{t-1}$, $(\hat{\mu})_T = (\hat{x}_{t-1})_T$ and $(\hat{\mu})_{T^c} = \mathbf{0}$. Here $\hat{N}_{t-1}$ is the support estimate of the previous reconstruction, $\hat{x}_{t-1}$. At the initial time, $t = 0$, we can either initialize with BPDN, or with mod-BPDN using T from prior knowledge, e.g. for wavelet sparse images, T could be the set of indices of the approximation coefficients. We summarize the stepwise dynamic reg-mod-BPDN approach in Algorithm 2. Notice that at $t = 0$, one may need more measurements since the prior knowledge of T may not be very accurate. Hence, we use $y_0 = A_0 x_0 + w_0$ where A_0 is an $n_0 \times m$ measurement matrix with $n_0 > n$.

المنارة للاستشارات
www.manaraa.com

In Algorithm 2, we should reiterate that for support estimation, we need to use a threshold $\rho > 0$. The threshold should be large enough so that most elements of $\Delta_{e,t} := T \setminus N_t = \hat{N}_{t-1} \setminus N_t$ do not get detected into the support.

We briefly discuss here the stability of dynamic reg-mod-BPDN (reconstruction error and support estimation errors bounded by a time-invariant and small value at all times). Using an approach similar to that of [59], it should be possible to show the following. If (i) ρ is large enough (so that $\hat{N}_t$ does not falsely detect any element that got removed from N_t); (ii) the newly added elements to the current support, N_t, either get added at a large enough value to get detected immediately, or within a finite delay their magnitude becomes large enough to get detected; and (iii) the matrix A satisfies certain conditions (for a given support size and support change size); reg-mod-BPDN will be stable.

Algorithm 2 **Dynamic Reg-mod-BPDN**

```
At $t = 0$, compute $\hat{x}_0$ as the solution of $\min_b \quad \gamma \|(b)_{T^c}\|_1 + \frac{1}{2}\|y_0 - Ab\|_2^2$, where $T$ is either empty or is available from prior knowledge. Compute $\hat{N}_0 = \{i \in [1, ..., m] : |(\hat{x}_0)_i| > \rho\}$. Set $T \leftarrow \hat{N}_0$ and $(\hat{\mu})_T \leftarrow (\hat{x}_0)_T$
For $t > 0$, do
    1. Reg-Mod-BPDN. Let $T = \hat{N}_{t-1}$ and let $\hat{\mu}_T = (\hat{x}_{t-1})_T$. Compute $\hat{x}_t$ as the solution of (3.11).
    2. Estimate Support. $\hat{N}_t = \{i \in [1, ..., m] : |(\hat{x}_t)|_i > \rho\}$.
    3. Output the reconstruction $\hat{x}_t$.
Feedback $\hat{N}_t$ and $\hat{x}_t$; increment $t$, and go to step 1.
```

3.4 Bounding the Reconstruction Error

In this section, we bound the reconstruction error of reg-mod-BPDN. Since mod-BPDN and BPDN are special cases, their results follow as direct corollaries. The result for BPDN is the same as [2, Theorem 8]. In Chapter 5.2.1, we define the terms needed to state our result. In 5.2.2 we state our result and discuss its implications. In 5.2.3, we give the proof outline.

المنارة للاستشارات
www.manaraa.com

3.4.1 Definitions

We begin by defining the function that we want to minimize as

$$L(b) \triangleq L_1(b) + \gamma \|b_{T^c}\|_1 \tag{3.13}$$

where

$$L_1(b) \triangleq \frac{1}{2}\|y - Ab\|_2^2 + \frac{1}{2}\lambda \|b_T - \hat{\mu}_T\|_2^2 \tag{3.14}$$

contains the two ℓ_2 norm terms (data fidelity term and the regularization term). If we constrain b to be supported on $T \cup S$ for some $S \subset T^c$, then the minimizer of $L_1(b)$ will be the regularized least squares (LS) estimator obtained when we put a weight λ on $\|b_T - \hat{\mu}_T\|_2^2$ and a weight zero on $\|b_S - \hat{\mu}_S\|_2^2$.

Let S be a given subset of Δ. Next, we define three matrices which will be frequently used in our results. Let

$$\begin{aligned} Q_{T,\lambda}(S) &\triangleq A_{T\cup S}' A_{T\cup S} + \lambda \begin{bmatrix} I_T & \mathbf{0}_{T,S} \\ \mathbf{0}_{S,T} & \mathbf{0}_{S,S} \end{bmatrix} && (3.15) \\ M_{T,\lambda} &\triangleq I - A_T (A_T' A_T + \lambda I_T)^{-1} A_T' && (3.16) \\ P_{T,\lambda}(S) &\triangleq (A_S' M_{T,\lambda} A_S)^{-1} && (3.17) \end{aligned}$$

where I_T is a $|T| \times |T|$ identity matrix and $\mathbf{0}_{T,S}$, $\mathbf{0}_{S,T}$, $\mathbf{0}_{S,S}$ are all zeros matrices with sizes $|T| \times |S|$, $|S| \times |T|$ and $|S| \times |S|$.

Assumption 1 *We assume that $Q_{T,\lambda}(\Delta)$ is invertible. This implies that, for any $S \subseteq \Delta$, the functions $L(b)$ and $L_1(b)$ are strictly convex over the set of all vectors supported on $T \cup S$.*

Proposition 2 *When $\lambda > 0$, $Q_{T,\lambda}(S)$ is invertible if A_S has full rank. When $\lambda = 0$ (mod-BPDN), this will hold if $A_{T\cup S}$ has full rank.*

The proof is easy and is given in Appendix B.1.

Let $S \subseteq \Delta$. Consider minimizing $L(b)$ over b supported on $T \cup S$. When $b_{(T\cup S)^c} = 0$ and Assumption 1 holds, $L(b_{T\cup S})$ is strictly convex and thus has a unique minimizer. The same holds for $L_1(b_{T\cup S})$. Define their respective unique minimizers as

$$\begin{aligned} d_{T,\lambda}(S) &\triangleq \arg\min_b \ L(b) \quad \text{subject to} \quad b_{(T\cup S)^c} = \mathbf{0} && (3.18) \\ c_{T,\lambda}(S) &\triangleq \arg\min_b \ L_1(b) \quad \text{subject to} \quad b_{(T\cup S)^c} = \mathbf{0} && (3.19) \end{aligned}$$

المنارة للاستشارات

www.manaraa.com

As explained earlier, $c_{T,\lambda}(S)$ is the regularized LS estimate of x when assuming that x is supported on $T \cup S$ and with the weights mentioned earlier. It is easy to see that

$$\begin{aligned} [c_{T,\lambda}(S)]_{T\cup S} &= Q_{T,\lambda}(S)^{-1} \left(A_{T\cup S}' y + \begin{bmatrix} \lambda \hat{\mu}_T \\ \mathbf{0}_S \end{bmatrix} \right) \\ [c_{T,\lambda}(S)]_{(T\cup S)^c} &= \mathbf{0} \end{aligned} \tag{3.20}$$

In a fashion similar to [2], define

$$ERC_{T,\lambda}(S) \triangleq 1 - \max_{\omega \notin T\cup S} \|P_{T,\lambda}(S) A_S' M_{T,\lambda} A_\omega\|_1 \tag{3.21}$$

This is different from the ERC of [2] but simplifies to it when $T = \emptyset$, $S = N$ and $\lambda = 0$. In [2], the ERC, which in our notation is $ERC_{\emptyset,0}(N)$, being strictly positive, along with γ approaching zero, ensured exact recovery of BPDN in the noise-free case. Hence, in [2], ERC was an acronym for *Exact Recovery Coefficient*. In this work, the same holds for mod-BPDN. If $ERC_{T,0}(\Delta) > 0$, the solution of mod-BPDN approaches the true x as γ approaches zero. We explain this further in Remark 6 below. However, no similar claim can be made for reg-mod-BPDN. On the other hand, for the reconstruction error bounds, ERC serves the exact same purpose for reg-mod-BPDN as it does for BPDN in [2]: $ERC_{T,\lambda}(\Delta) > 0$ and γ greater than a certain lower bound ensures that the reg-mod-BPDN (or mod-BPDN) error can be bounded by modifying the approach of [2].

3.4.2 Reconstruction error bound

The reconstruction error can be bounded as follows.

Theorem 4 *If $Q_{T,\lambda}(\Delta)$ is invertible, $ERC_{T,\lambda}(\Delta) > 0$ and*

$$\gamma \geq \gamma^*_{T,\lambda}(\Delta) \triangleq \frac{\|A_{(T\cup\Delta)^c}'(y - Ac_{T,\lambda}(\Delta))\|_\infty}{ERC_{T,\lambda}(\Delta)} \tag{3.22}$$

then,

1. *$L(b)$ has a unique minimizer, $\hat{x}$.*

2. *The minimizer, $\hat{x}$, is equal to $d_{T,\lambda}(\Delta)$, and thus is supported on $T \cup \Delta$.*

www.manaraa.com

3. *Its error can be bounded as*

$$\|x-\hat{x}\|_2 \leq \gamma\sqrt{|\Delta|} f_1(\Delta) + \lambda f_2(\Delta)\|x_T - \hat{\mu}_T\|_2 + f_3(\Delta)\|w\|_2$$

where

$$\begin{aligned} f_1(\Delta) &\triangleq \sqrt{\|({A_T}'A_T + \lambda I_T)^{-1}{A_T}'A_\Delta P_{T,\lambda}(\Delta)\|_2^2 + \|P_{T,\lambda}(\Delta)\|_2^2}, \\ f_2(\Delta) &\triangleq \|Q_{T,\lambda}(\Delta)^{-1}\|_2, \\ f_3(\Delta) &\triangleq \|Q_{T,\lambda}(\Delta)^{-1}{A_{T\cup\Delta}}'\|_2, \end{aligned} \tag{3.23}$$

$P_{T,\lambda}(\Delta)$ *is defined in (3.17) and* $Q_{T,\lambda}(\Delta)$ *in (3.15).*

Corollary 2 (corollaries for mod-BPDN and BPDN) *The result for mod-BPDN follows by setting* $\lambda = 0$ *in Theorem 4. The result for BPDN follows by setting* $\lambda = 0$*,* $T = \emptyset$ *(and so* $\Delta = N$*). This result is the same as [2, Theorem 8].*

Remark 5 (smallest γ) *Notice that the error bound above is an increasing function of* γ*. Thus* $\gamma = \gamma^*_{T,\lambda}(\Delta)$ *gives the smallest bound.*

In words, Theorem 4 says that, if $Q_{T,\lambda}(\Delta)$ is invertible, $ERC_{T,\lambda}(\Delta)$ is positive, and γ is large enough (larger than γ^*), then $L(b)$ has a unique minimizer, $\hat{x}$, and $\hat{x}$ is supported on $T \cup \Delta = N \cup \Delta_e$. This means that the only wrong elements that can possibly be part of the support of $\hat{x}$ are elements of Δ_e. Moreover, the error between $\hat{x}$ and the true x is bounded by a value that is small as long as the noise, $\|w\|_2$, is small, the prior term, $\|x_T - \hat{\mu}_T\|_2$, is small and $\gamma^*_{T,\lambda}(\Delta)$ is small. By rewriting $y - Ac_{T,\lambda}(\Delta) = A(x - c_{T,\lambda}(\Delta)) + w$ and using Lemma 7 (given in the Appendix B.2) one can upper bound γ^* by terms that are increasing functions of $\|w\|_2$ and $\|x_T - \hat{\mu}_T\|_2$. Thus, as long as these are small, the bound is small.

As shown in Proposition 2, $Q_{T,\lambda}(\Delta)$ is invertible if $\lambda > 0$ and A_Δ is full rank or if $A_{T\cup\Delta}$ is full rank.

Next, we use the idea of [2, Corollary 10] to show that $ERC_{T,0}(\Delta)$ is an *Exact Recovery Coefficient* for mod-BPDN.

Remark 6 (ERC and exact recovery of mod-BPDN) *For mod-BPDN,* $c_{T,0}(\Delta)$ *is the LS estimate when* x *is supported on* $T\cup\Delta$*. Using (3.20), (1.2), and the fact that* x *is supported on* $N \subseteq T\cup\Delta$*, it is easy to*

المنارة للاستشارات
www.manaraa.com

see that in the noise-free ($w = 0$) case, $c_{T,0}(\Delta) = x_{T\cup\Delta}$. Hence the numerator of $\gamma^*_{T,0}(\Delta)$ will be zero. Thus, using Theorem 4, if $ERC_{T,0}(\Delta) > 0$, the mod-BPDN error satisfies $\|x - \hat{x}\|_2 \leq \gamma\sqrt{|\Delta|}f_1(\Delta)$. Thus the mod-BPDN solution, $\hat{x}$, will approach the true x as γ approaches zero. Moreover, as long as $\gamma < \frac{\min_{i\in N}|x_i|}{\sqrt{|\Delta|}f_1(\Delta)}$, at least the support of $\hat{x}$ will equal the true support, N [3].

We show a numerical comparison of the results of reg-mod-BPDN, mod-BPDN and BPDN in Table 3.1 (simulation details given in Chapter 3.4). Notice that BPDN needs 90% of the measurements for its sufficient conditions to start holding (ERC to become positive) whereas mod-BPDN only needs 19%. Moreover, even with 90% of the measurements, the ERC of BPDN is just positive and very small. As a result, its error bound is large (27% normalized mean squared error (NMSE)). Similarly, notice that mod-BPDN needs $n \geq 19\% m$ for its sufficient conditions to start holding ($A_{T\cup\Delta}$ to become full rank which is needed for $Q_{T,0}(\Delta)$ to be invertible). For reg-mod-BPDN which only needs A_Δ to be full rank, $n = 13\% m$ suffices.

Remark 7 *A sufficient conditions comparison only provides a comparison of when a given result can be applied to provide a bound on the reconstruction error. In other words, it tells us under what conditions we can guarantee that the reconstruction error of a given approach will be small (below a bound). Of course this does not mean that we cannot get small error even when the sufficient condition does not hold, e.g., in simulations, BPDN provides a good reconstruction using much less than 90% of the measurements. However, when $n < 90\% m$ we cannot bound its reconstruction error using Theorem 4 above.*

3.4.3 Proof Outline

To prove Theorem 4, we use the following approach motivated by that of [2].

1. We first bound $\|d_{T,\lambda}(\Delta) - c_{T,\lambda}(\Delta)\|_2$ by simplifying the necessary and sufficient condition for it to be the minimizer of $L(b)$ when b is supported on $T \cup \Delta$. This is done in Lemma 6 in Appendix B.2.

[3] *If we bounded the ℓ_∞ norm of the error as done in [2] we would get a looser upper bound on the allowed γ's for this.*

المنارة للاستشارات
www.manaraa.com

2. We bound $\|c_{T,\lambda}(\Delta) - x\|_2$ using the expression for $c_{T,\lambda}(\Delta)$ in (3.20) and substituting $y = A_{T\cup\Delta}x_{T\cup\Delta} + w$ in it (recall that x is zero outside $T \cup \Delta$). This is done in Lemma 7 in Appendix B.2.

3. We can bound $\|d_{T,\lambda}(\Delta) - x\|_2$ using the above two bounds and the triangle inequality.

4. We use an approach similar to [2, Lemma 6] to find the sufficient conditions under which $d_{T,\lambda}(\Delta)$ is also the unconstrained unique minimizer of $L(b)$, i.e. $\hat{x} = d_{T,\lambda}(\Delta)$. This is done in Lemma 8 in Appendix B.2.

The last step (Lemma 8) helps prove the first two parts of Theorem 4. Combining the above four steps, we get the third part (error bound). We give the lemmas in Appendix B.2. They are proved in Appendix B.4.1, B.4.2 and B.4.3.

Two key differences in the above approach with respect to the result of [2] are

- $c_{T,\lambda}(\Delta)$ is the regularized LS estimate instead of the LS estimate in [2]. This helps obtain a better and simpler error bound of reg-mod-BPDN than when using the LS estimate. Of course, when $\lambda = 0$ (mod-BPDN or BPDN), $c_{T,0}(\Delta)$ is just the LS estimate again.

- For reg-mod-BPDN (and also for mod-BPDN), the subgradient set of the ℓ_1 term is $\partial\|b_{T^c}\|_1|_{b=d_{T,\lambda}(\Delta)}$ and so any ϕ in this set is zero on T, and only has $\|\phi_\Delta\|_\infty \leq 1$. Since $|\Delta| \ll |N|$, this helps to get a tighter bound on $\|c_{T,\lambda}(\Delta) - d_{T,\lambda}(\Delta)\|_2$ in step 1 above as compared to that for BPDN [2] (see proof of Lemma 6 for details).

3.5 Tighter Bounds without Sufficient Conditions

The problem with the error bounds for reg-mod-BPDN, mod-BPDN, BPDN or LS-CS [60] is that they all hold under different sufficient conditions. This makes it difficult to compare them. Moreover, the bound is particularly loose when n is such that the sufficient conditions just get satisfied. This is because the ERC is just positive but very small (resulting in a very large γ^* and hence a very large bound). To address this issue, in this section, we obtain a bound that holds without any sufficient conditions and that is also tighter, while still being computable.

The key idea that we use is as follows:

المنارة للاستشارات

www.manaraa.com

- we modify Theorem 4 to hold for "sparse-compressible" signals [60], i.e. for sparse signals, x, in which some nonzero coefficients out of the set Δ are small ("compressible") compared to the rest; and then
- we minimize the resulting bound over all allowed split-ups of x into non-compressible and compressible parts.

Let $\tilde{\Delta} \subseteq \Delta$ be such that the conditions of Theorem 4 hold for it. Then the first step involves modifying Theorem 4 to bound the error for reconstructing x when we treat $x_{\Delta \setminus \tilde{\Delta}}$ as the "compressible" part. The main difference here is in bounding $\|c_{T,\lambda}(\tilde{\Delta}) - x\|_2$ which now has a larger bound because of $x_{\Delta \setminus \tilde{\Delta}}$. We do this in Lemma 9 in the Appendix B.3. Notice from the proofs of Lemma 6 and Lemma 8 in Appendix B.4.1 and B.4.3 that nothing in their result changes if we replace Δ by a $\tilde{\Delta} \subseteq \Delta$. Combining Lemma 9 with Lemmas 6 and 8 applied for $\tilde{\Delta}$ instead of Δ leads to the following corollary.

Corollary 3 *Consider a $\tilde{\Delta} \subseteq \Delta$. If $Q_{T,\lambda}(\tilde{\Delta})$ is invertible, $ERC_{T,\lambda}(\tilde{\Delta}) > 0$, and $\gamma = \gamma^*_{T,\lambda}(\tilde{\Delta})$, then*

$$\|x - \hat{x}\|_2 \leq f(T, \lambda, \Delta, \tilde{\Delta}, \gamma^*_{T,\lambda}(\tilde{\Delta})) \tag{3.24}$$

where

$$f(T, \lambda, \Delta, \tilde{\Delta}, \gamma) \triangleq \gamma \sqrt{|\tilde{\Delta}|} f_1(\tilde{\Delta}) + \lambda f_2(\tilde{\Delta}) \|x_T - \hat{\mu}_T\|_2 + f_3(\tilde{\Delta}) \|w\|_2 + f_4(\tilde{\Delta}) \|x_{\Delta \setminus \tilde{\Delta}}\|_2, \tag{3.25}$$

$$f_4(\tilde{\Delta}) \triangleq \sqrt{\|Q_{T,\lambda}(\tilde{\Delta})^{-1} A_{T \cup \tilde{\Delta}}{}' A_{\Delta \setminus \tilde{\Delta}}\|_2^2 + 1}, \tag{3.26}$$

$f_1(\cdot)$,$f_2(\cdot)$, $f_3(\cdot)$ are defined in (3.23) and $\gamma^*_{T,\lambda}(\tilde{\Delta})$ in (3.22).

Proof: The proof is given in Appendix B.3.1.

In order to get a bound that depends only on $\|x_T - \hat{\mu}_T\|_2$, $\|x_{\Delta \setminus \tilde{\Delta}}\|_2$, the noise, w, and the sets T, Δ, Δ_e, we can further bound $\gamma^*_{T,\lambda}(\tilde{\Delta})$ by rewriting $y - Ac_{T,\lambda}(\tilde{\Delta}) = A(x - c_{T,\lambda}(\tilde{\Delta})) + w$ and then bounding $\|x - (c_{T,\lambda}(\tilde{\Delta}))\|_2$ using Lemma 9. Doing this gives the following corollary.

Corollary 4 *If $Q_{T,\lambda}(\tilde{\Delta})$ is invertible, $ERC_{T,\lambda}(\tilde{\Delta}) > 0$, and $\gamma = \gamma^*_{T,\lambda}(\tilde{\Delta})$, then*

$$\|x - \hat{x}\|_2 \leq g(\tilde{\Delta}) \tag{3.27}$$

المنارة للاستشارات

www.manaraa.com

where

$$\begin{aligned}
g(\tilde{\Delta}) &\triangleq g_1\|x_T - \hat{\mu}_T\|_2 + g_2\|w\|_2 + g_3\|x_{\Delta\setminus\tilde{\Delta}}\|_2 + g_4 \qquad (3.28)\\
g_1 &\triangleq \lambda f_2(\tilde{\Delta})(\frac{\sqrt{|\tilde{\Delta}|}f_1(\tilde{\Delta})\mathit{maxcor}(\tilde{\Delta})}{ERC_{T,\lambda}(\tilde{\Delta})} + 1),\\
g_2 &\triangleq \frac{\sqrt{|\tilde{\Delta}|}f_1(\tilde{\Delta})f_3(\tilde{\Delta})\mathit{maxcor}(\tilde{\Delta})}{ERC_{T,\lambda}(\tilde{\Delta})} + f_3(\tilde{\Delta}),\\
g_3 &\triangleq \frac{\sqrt{|\tilde{\Delta}|}f_1(\tilde{\Delta})f_4(\tilde{\Delta})\mathit{maxcor}(\tilde{\Delta})}{ERC_{T,\lambda}(\tilde{\Delta})} + f_4(\tilde{\Delta}),\\
g_4 &\triangleq \frac{\sqrt{|\tilde{\Delta}|}\|A_{(T\cup\tilde{\Delta})^c}{}'w\|_\infty f_1(\tilde{\Delta})}{ERC_{T,\lambda}(\tilde{\Delta})},\\
\mathit{maxcor}(\tilde{\Delta}) &\triangleq \max_{i\notin(T\cup\tilde{\Delta})^c}\|A_i{}'A_{T\cup\Delta}\|_2,
\end{aligned}$$

$f_1(\cdot)$,$f_2(\cdot)$, $f_3(\cdot)$ *and* $f_4(\cdot)$ *are defined in (3.23) and (3.26), and* $\gamma^*_{T,\lambda}(\tilde{\Delta})$ *in (3.22).*

Proof: The proof is given in Appendix B.3.2.

Using the above corollary and minimizing over all allowed $\tilde{\Delta}$'s, we get the following result.

Theorem 5 *Let*

$$\tilde{\Delta}^* \triangleq \arg\min_{\tilde{\Delta}\in\mathcal{G}} g(\tilde{\Delta}) \qquad (3.29)$$

where

$$\mathcal{G} \triangleq \{\tilde{\Delta} : \tilde{\Delta} \subseteq \Delta, ERC_{T,\lambda}(\tilde{\Delta}) > 0, Q_{T,\lambda}(\tilde{\Delta}) \textit{ is invertible}\} \qquad (3.30)$$

If $\gamma = \gamma^*_{T,\lambda}(\tilde{\Delta}^*)$*, then*

1. $L(b)$ *has a unique minimizer,* $\hat{x}$*, supported on* $T \cup \tilde{\Delta}^*$.

2. *The error bound is*

$$\|x - \hat{x}\|_2 \leq g(\tilde{\Delta}^*) \qquad (3.31)$$

($\gamma^*_{T,\lambda}(\tilde{\Delta})$ *is defined in (3.22)).*

Proof: This result follows by minimizing over all allowed $\tilde{\Delta}$'s from Corollary 4.

Compare Theorem 5 with Theorem 4. Theorem 4 holds only when the complete set Δ belongs to $\mathcal{G}$, whereas Theorem 5 holds always (we only need to set γ appropriately). Moreover, even when Δ does

المنارة للاستشارات
www.manaraa.com

belong to $\mathcal{G}$, Theorem 4 gives the error bound by choosing $\tilde{\Delta}^* = \Delta$. However, Theorem 5 minimizes over all allowed $\tilde{\Delta}$'s, thus giving a tighter bound, especially for the case when the sufficient conditions of Theorem 4 just get satisfied and $ERC_{T,\lambda}(\Delta)$ is positive but very small. A similar comparison also holds for the mod-BPDN and BPDN results.

The problem with Theorem 5 is that its bound is not computable (the computational cost is exponential in $|\Delta|$). Notice that $g(\tilde{\Delta}^*) := \min_{\tilde{\Delta}\in\mathcal{G}} g(\tilde{\Delta})$ can be rewritten as

$$\begin{aligned} & g(\tilde{\Delta}^*) \triangleq \min_{\tilde{\Delta}\in\mathcal{G}} g(\tilde{\Delta}) = \min_{0\le k\le|\Delta|} \min_{\mathcal{G}_k} g(\tilde{\Delta}) \ \text{ where} \\ & \mathcal{G}_k \triangleq \mathcal{G} \cap \{\tilde{\Delta} \subseteq \Delta : |\tilde{\Delta}| = k\} \end{aligned} \tag{3.32}$$

Let $d := |\Delta|$. The minimization over $\mathcal{G}_k$ is expensive since it requires searching over all $\binom{d}{k}$ size k subsets of Δ to first find which ones belong to $\mathcal{G}_k$ and then find the minimum over all $\tilde{\Delta} \subseteq \mathcal{G}_k$. The total computation cost to do the former for all sets $\mathcal{G}_0, \mathcal{G}_1, \dots \mathcal{G}_d$ is $O(\sum_{k=0}^{d} \binom{d}{k}) = O(2^d)$, i.e. it is *exponential in* d. This makes the bound computation intractable for large problems.

3.5.1 Obtaining a Computable Bound

In most cases of practical interest, the term that has the maximum variability over different sets in $\mathcal{G}_k$ is $\|x_{\Delta\setminus\tilde{\Delta}}\|_2$. The multipliers g_1, g_2, g_3 and g_4 vary very slightly for different sets in a given $\mathcal{G}_k$. Using this fact, we can obtain the following upper bound on $\min_{\mathcal{G}_k} g(\tilde{\Delta})$ which is only slightly looser and also holds without sufficient conditions, but is computable in polynomial time.

Define $\tilde{\Delta}^{**}(k)$ and B_k as follows

$$\begin{aligned} \tilde{\Delta}^{**}(k) &\triangleq \arg \min_{\{\tilde{\Delta}\subseteq\Delta, |\tilde{\Delta}|=k\}} \|x_{\Delta\setminus\tilde{\Delta}}\|_2 \\ B_k &\triangleq \begin{cases} g(\tilde{\Delta}^{**}(k)) & \text{if } \tilde{\Delta}^{**}(k) \in \mathcal{G}_k \\ \infty & \text{otherwise} \end{cases} \end{aligned} \tag{3.33}$$

Then, clearly

$$\min_{\mathcal{G}_k} g(\tilde{\Delta}) \le B_k \tag{3.34}$$

since $\min_{\mathcal{G}_k} g(\tilde{\Delta}) \le g(\tilde{\Delta})$ for any $\tilde{\Delta} \in \mathcal{G}_k$ and it is also less than infinity. For any k, the set $\tilde{\Delta}^{**}(k)$ can be obtained by sorting the elements of x_Δ in decreasing order of magnitude and letting $\tilde{\Delta}^{**}(k)$ contain

www.manaraa.com

the indices of the k largest elements. Doing this takes $O(d \log d)$ time since sorting takes $O(d \log d)$ time. Computation of B_k requires matrix multiplications and inversions which are $O(k^3)$. Thus, the total cost of doing this is at most $O(d^4)$ which is still polynomial in d.

Therefore, we get the following bound that is *computable in polynomial time and that still holds without sufficient conditions and is much tighter than Theorem 4.*

Theorem 6 *Let*

$$\begin{aligned} k_{\min} &\triangleq \arg \min_{0 \le k \le |\Delta|} B_k \quad \textit{and} \\ \tilde{\Delta}^{**} &\triangleq \tilde{\Delta}^{**}(k_{\min}) \end{aligned} \tag{3.35}$$

where B_k and $\tilde{\Delta}^{**}(k)$ are defined in (3.33). If $\gamma = \gamma^*_{T,\lambda}(\tilde{\Delta}^{**})$,

1. *$L(b)$ has a unique minimizer, $\hat{x}$, supported on $T \cup \tilde{\Delta}^{**}$.*

2. *The error bound is*

$$\|x - \hat{x}\|_2 \le g(\tilde{\Delta}^{**}) \tag{3.36}$$

($\gamma^*_{T,\lambda}(\tilde{\Delta})$ is defined in (3.22)).

Corollary 5 (corollaries for mod-BPDN and BPDN) *The result for mod-BPDN follows by setting $\lambda = 0$ in Theorem 6. The result for BPDN follows by setting $\lambda = 0$, $T = \emptyset$ (and so $\Delta = N$) in Theorem 6.*

When n and $s \triangleq |N|$ are large enough, the above bound is either only slightly larger, or often actually equal, to that of Theorem 5 (e.g. in Fig. 3.4(a), $m = 256$, $n = 0.13m = 33$, $s = 0.1m = 26$). The reason for the equality is that the minimizing value of k is the one that is small enough to ensure that g_1, g_2, g_3, g_4 are small. When k is small, g_1, g_2, g_3, g_4, ERC and $Q(\tilde{\Delta})$ have very similar values for all sets $\tilde{\Delta}$ of the same size k. In (3.28), the only term with significant variability for different sets $\tilde{\Delta}$ of the same size k is $\|x_{\Delta \setminus \tilde{\Delta}}\|_2$. Thus, (a) $\arg\min_{\mathcal{G}_k} g(\tilde{\Delta}) = \arg\min_{\mathcal{G}_k} \|x_{\Delta \setminus \tilde{\Delta}}\|_2$ and (b) $\mathcal{G}_k$ is equal to $\{\tilde{\Delta} \subseteq \Delta, |\tilde{\Delta}| = k\}$. Thus, (3.34) holds with equality and so the bounds from Theorems 6 and 5 are equal. As n and $s \triangleq |N|$ approach infinity, *it is possible to use a law of large numbers (LLN) argument to prove that both bounds will be equal with high probability (w.h.p.)*. The key idea will be the same as

www.manaraa.com

above: show that as n, s go to infinity, w.h.p., g_1, g_2, g_3, g_4, Q and ERC are equal for all sets $\tilde{\Delta}$ of any given size k. We will develop this result in future work.

3.6 Numerical Experiments

In this section, we show both upper bound comparisons and actual reconstruction error comparisons. The upper bound comparison only tells us that the performance guarantees of reg-mod-BPDN are better than those for the other methods. To actually demonstrate that reg-mod-BPDN outperforms the others, we need to compare the actual reconstruction errors. This section is organized as follows. After giving the simulation model in 5.4.1, we show the reconstruction error comparisons for recovering simulated sparse signals from random Gaussian measurements in 5.4.2. In 5.4.3, we show comparisons for recursive dynamic MRI reconstruction of a larynx image sequence. In this comparison, we also show the usefulness of the Theorem 6 in helping us select a good value of γ. In the last three subsections, we show numerical comparisons of the results of the various theorems. The upper bound comparisons of Theorem 6 and the comparison of the corresponding reconstruction errors suggests that the bounds for reg-mod-BPDN and BPDN are tight under the scenarios evaluated. Hence, they can be used as a proxy to decide which algorithm to use when. We show this for both random Gaussian and partial Fourier measurements.

3.6.1 Simulation Model

The notation $z = \pm a$ means that we generate each element of the vector z independently and each is either $+a$ or $-a$ with probability 1/2. The notation $\nu \sim \mathcal{N}(0, \Sigma)$ means that ν is generated from a Gaussian distribution with mean 0 and covariance matrix Σ. We use $\lfloor a \rfloor$ to denote the largest integer less than or equal to a. Independent and identically distributed is abbreviated as iid. Also, N-RMSE refers to the normalized root mean squared error.

We use the recursive reconstruction application [33, 25] to motivate the simulation model. In this case, assuming that slow support and slow signal value change hold [see Fig. 1.2], we can use the reconstructed value of the signal at the previous time as $\hat{\mu}$ and its support as T. To simulate the effect of slow signal value change, we let $x_N = \mu_N + \nu$ where ν is a small iid Gaussian deviation and we let

المنارة للاستشارات

www.manaraa.com

$\hat{\mu}_{T\cap N} = \mu_{T\cap N}$ (and so $x_{T\cap N} = \hat{\mu}_{T\cap N} + \nu_{T\cap N}$).

The extras set, $\Delta_e = T \setminus N$, contains elements that got removed from the support at the current time or at a few previous times (but so far did not get removed from the support estimate). In most practical applications, only small valued elements at the previous time get removed from the support and hence the magnitude of $\hat{\mu}$ on Δ_e will be small. We use β_s to denote this small magnitude, i.e. we simulate $(\hat{\mu})_{\Delta_e} = \pm\beta_s$.

The misses' set at time t, Δ, definitely includes the elements that just got added to the support at t or the ones that previously got added but did not get detected into the support estimate so far. The new elements typically get added at a small value and their value slowly increases to a large one. Thus, elements in Δ will either have small magnitude (corresponding to the current newly added ones), or will have larger magnitude but still smaller than that of elements already in $N \cap T$. To simulate this, we do the following. (a) We simulate the elements on $N \cap T$ to have large magnitude, β_l, i.e. we let $(\mu)_{N\cap T} = \pm\beta_l$. (b) We split the set Δ into two disjoint parts, Δ_1 and $\Delta_2 = \Delta \setminus \Delta_1$. The set Δ_1 contains the small (e.g. newly added) elements, i.e. $(\mu)_{\Delta_1} = \pm\beta_s$. The set Δ_2 contains the larger elements, though still with magnitudes smaller than those in $N \cap T$, i.e. $(\mu)_{\Delta_2} = \pm\beta_m$, where $\beta_l \geq \beta_m \geq \beta_s$.

In summary, we use the following simulation model.

$$\begin{aligned}
(x)_N &= (\mu)_N + \nu, \;\; \nu \sim \mathcal{N}(0, \sigma_p^2 I) \\
(x)_{N^c} &= 0 \qquad (3.37) \\
\text{where } \; (\mu)_{N\cap T} &= \pm\beta_l \\
(\mu)_{\Delta_1} &= \pm\beta_s, \;\; (\mu)_{\Delta_2} = \pm\beta_m \\
(\mu)_{N^c} &= 0 \qquad (3.38)
\end{aligned}$$

and

$$\begin{aligned}
(\hat{\mu})_{T\cap N} &= (\mu)_{T\cap N} = \pm\beta_l \\
(\hat{\mu})_{\Delta_e} &= \pm\beta_s \\
(\hat{\mu})_{T^c} &= 0 \qquad (3.39)
\end{aligned}$$

المنارة للاستشارات

www.manaraa.com

We generate the support of x, N, of size $|N|$, uniformly at random from $[1, ..., m]$. We generate Δ with size $|\Delta|$ and Δ_e with size $|\Delta_e|$ uniformly at random from N and from N^c respectively. The set Δ_1 of size $|\Delta_1| = \lfloor|\Delta|/2\rfloor$ is generated uniformly at random from Δ. The set $\Delta_2 = \Delta \setminus \Delta_1$. We let $T = N \cup \Delta_e \setminus \Delta$. We generate μ and then x using (3.38) and (3.37). We generate $\hat{\mu}$ using (3.39).

In some simulations, we simulated the more difficult case where $\beta_m = \beta_s$. In this case, all elements on Δ were identically generated and hence we did not need Δ_1.

3.6.2 Reconstruction Error Comparisons

In Fig. 3.1, we compare the Monte Carlo average of the reconstruction error of reg-mod-BPDN with that of mod-BPDN, BPDN, weighted ℓ_1 [26] given in (1.9), CS-residual given in (1.11), CS-mod-residual given in (3.7) and modified-CS-residual[40] given in (3.8). Simulation was done according to the model specified above. We used random Gaussian measurements in this simulation, i.e. we generated A as an $n \times m$ matrix with iid zero mean Gaussian entries and normalized each column to unit ℓ_2 norm.

We experimented with two choices of n, $n = 0.13m$ (where reg-mod-BPDN outperforms mod-BPDN) and $n = 0.3m$ (where both are similar) and two values of σ_p^2, $\sigma_p^2 = 0.001$ (good prior) and $\sigma_p^2 = 0.1$ (bad prior). For the cases of Fig 3.1(a) ($n = 0.13m$, $\sigma_p^2 = 0.001$) and Fig 3.1(b) ($n = 0.13m$, $\sigma_p^2 = 0.1$), we used signal length $m = 256$, support size $|N| = 0.1m = 26$ and support extras size, $|\Delta_e| = 0.1|N| = 3$. The misses' size, $|\Delta|$, was varied between 0 and $0.2|N|$ (these numbers were motivated by the medical imaging application, we used larger numbers than what are shown in Fig. 1.2). We used $\beta_l = 1$, $\beta_m = 0.4$ and $\beta_s = 0.2$. The noise variance was $\sigma_w^2 = 10^{-5}$. For the last two figures, Fig 3.1(c) ($n = 0.3m$, $\sigma_p^2 = 0.001$) and Fig 3.1(d) ($n = 0.3m$, $\sigma_p^2 = 0.1$), for which n was larger, we used $\beta_m = \beta_s = 0.25$ which is a more difficult case for reg-mod-BPDN. For Fig. 3.1(c), we also used a larger noise variance $\sigma_w^2 = 10^{-4}$. All other parameters were the same.

In Fig. 3.2, we show a plot of reg-mod-BPDN and BPDN from Fig 3.1(a) extended all the way to $|\Delta|/|N| = 1$ (which is the same as $\Delta = N$). Notice that if $|\Delta_e| = 0$, then the point $|\Delta|/|N| = 1$ of reg-mod-BPDN (or of mod-BPDN) is the same as BPDN. But in this plot, $|\Delta_e| = 3$ and hence the two points are different, even though the errors are quite similar.

المنارة للاستشارات

www.manaraa.com

For applications where some training data is available, γ and λ for reg-mod-BPDN can be chosen by interpreting the reg-mod-BPDN solution as the maximum a posteriori (MAP) estimate under a certain prior signal model (assume x_T is Gaussian with mean $\hat{\mu}_T$ and variance σ_p^2 and x_{T^c} is independent of x_T and is iid Laplacian with parameter λ_p). This idea is explained in detail in [25]. However, there is no easy way to do this for the other methods. Alternatively, choosing γ and λ according to Theorem 6 gives another good start point. We can do this for mod-BPDN and BPDN, but we cannot do this for the other methods (we show examples using this approach later). For a fair error comparison, for each algorithm, we selected γ from a set of values $[0.00001\ 0.00005\ 0.0001\ 0.0005\ 0.001\ 0.005\ 0.01\ 0.1]$. We tried all these values for a small number of simulations (10 simulations) and then picked the best one (one with the smallest N-RMSE) for each algorithm. For weighted ℓ_1 reconstruction, we also pick the best γ' in (1.9) from the same set in the same way[4]. For reg-mod-BPDN, λ should be larger when the signal estimate is good and should be decreased when the signal estimate is not so good. We can use $\lambda = \alpha\sigma_w^2/\sigma_p^2$ to adaptively determine its value for different choices of σ_w^2 and σ_p^2. In our simulations, we used $\alpha = 0.2$ for Fig. 3.1 (a), (b) and (d) and $\alpha = 0.05$ for Fig. 3.1(c).

We fixed the chosen γ, γ' and λ and did Monte Carlo averaging over 100 simulations. We conclude the following. (1) When the signal estimate is not good (Fig. 3.1(b),(d)) or when n is small (Fig. 3.1(a),(b)), CS-residual and CS-mod-residual have significantly larger error than reg-mod-BPDN. (2) In case of Fig. 3.1(d) ($n = 0.3m$), they also have larger error than mod-BPDN. (3) In all four cases, weighed ℓ_1 and mod-BPDN have similar performance. This is also similar to that of reg-mod-BPDN in case of $n = 0.3m$, but is much worse in case of $n = 0.13m$. (4) We also show a comparison with regmodBPDN-var in Fig. 3.1(a). Notice that it has larger errors than reg-mod-BPDN for reasons explained in the beginning of this chapter.

3.6.3 Dynamic MRI application using γ from Theorem 6

In Fig. 3.3, we show comparisons for simulated dynamic MR imaging of an actual larynx image sequence (Fig. 1.2 (a)(i)). The larynx image is not exactly sparse but is only compressible in the

[4] To give an example, our finally selected numbers for Fig. 3.1(d) were $\gamma = 0.01, 0.001, 0.001, 0.001, 0.001, 0.001, 0.01, 0.01$ for BPDN, mod-BPDN, reg-mod-BPDN, weighted ℓ_1, LS-CS, CS-residual, CS-mod-residual, mod-CS-residual respectively and $\gamma' = 0.0001$

المنارة للاستشارات
www.manaraa.com

wavelet domain. We used a two-level Daubechies-4 2D discrete wavelet transform (DWT). The 99%-energy support size of its wavelet transform vector, $|N_t| \approx 0.07m$. Also, $|\Delta_t| \approx 0.001m$ and $|\Delta_{e,t}| \approx 0.002m$. We used a 32×32 block of this sequence and at each time and simulated undersampled MRI, i.e. we selected n 2D discrete Fourier transform (DFT) coefficients using the variable density sampling scheme of [35], and added iid Gaussian noise with zero mean and variance $\sigma_w^2 = 10$ to each of them. Using a small 32×32 block allows easy implementation using CVX (for full sized image sequences, one needs specialized code). We used $n_0 = 0.18m$ at $t = 0$ and $n = 0.06m$ at $t > 0$.

We implemented dynamic reg-mod-BPDN as described in Algorithm 2. In this problem, the matrix $A = F_u \cdot W^{-1}$ where F_u contains the selected rows of the 2D-DFT matrix and W is the inverse 2D-DWT matrix (for a two-level Daubechies-4 wavelet). Reg-mod-BPDN was compared with similarly implemented reg-mod-BPDN-var and CS-residual algorithms (CS-residual only solved simple BPDN at $t = 0$). We also compared with simple BPDN (BPDN done for each frame separately). For reg-mod-BPDN and reg-mod-BPDN-var, the support estimation threshold, ρ, was chosen as suggested in [25]: we used $\rho = 20$ which is slightly larger than the smallest magnitude element in the 99%-energy support which is 15. At $t = 0$, we used T_0 to be the set of indices of the wavelet approximation coefficients. To choose γ and λ we tried two different things. (a) We used λ and γ from the set $[0.00001 \; 0.00005 \; 0.0001 \; 0.0005 \; 0.001 \; 0.005 \; 0.01 \; 0.1]$ to do the reconstruction for a short training sequence (5 frames), and used the average error to pick the best λ and γ. We call the resulting reconstruction error plot reg-mod-BPDN-opt. (b) We computed the average of the γ^* obtained from Theorem 6 for the 5-frame training sequence and used this as γ for the test sequence. We selected λ from the above set by choosing the one that minimizes the average of the bound of Theorem 6 for the 5 frames. We call the resulting error plot reg-mod-BPDN-γ^*. The same two things were also done for BPDN and CS-residual as well. For reg-mod-BPDN-var, we only did (a).

From Fig. 3.3, we can conclude the following. (1) Reg-mod-BPDN significantly outperforms the other methods when using so few measurements. (2) Reg-mod-BPDN-var and reg-mod-BPDN have similar performance in this case. (3) The reconstruction performance of reg-mod-BPDN using γ^* from Theorem 6 is close to that of reg-mod-BPDN using the best γ chosen from a large set. This indicates that Theorem 6 provides a good way to select γ in practice.

المنارة للاستشارات

www.manaraa.com

3.6.4 Comparing the result of Theorem 4

In Table 3.1, we compare the result of Theorem 4 for reg-mod-BPDN, mod-BPDN and BPDN. We used $m = 256$, $|N| = 26 = 0.1m$, $|\Delta| = 0.04|N| = |\Delta_e|$, $\sigma_p^2 = 10^{-3}$, $\beta_l = 1$ and $\beta_m = \beta_s = 0.25$. Also, $\sigma_w^2 = 10^{-5}$ and we varied n. For each experiment with a given n, we did the following. We did 100 Monte Carlo simulations. Each time, we evaluated the sufficient conditions for the bound of reg-mod-BPDN to hold. We say the bound *holds* if all the sufficient conditions hold for at least 98 realizations. If this did not happen, we record *not hold* in Table 3.1. If this did happen, then we recorded $\sqrt{\frac{\mathbb{E}[\text{bound}^2]}{\mathbb{E}[\|x\|_2^2]}}$ where $\mathbb{E}[\cdot]$ denotes the Monte Carlo average computed over those realizations for which the sufficient conditions do hold. Here, "bound" refers to the right hand side of (3.23) computed with $\gamma = \gamma^*_{T,\lambda}(\Delta)$ given in (3.22). An analogous procedure was followed for both mod-BPDN and BPDN.

The comparisons are summarized in Table 3.1. For reg-mod-BPDN, we selected λ from the set $[0.00001\ 0.00005\ 0.0001\ 0.0005\ 0.001\ 0.005\ 0.01\ 0.1]$ by picking the one that gave the smallest bound. Clearly the reg-mod-BPDN result holds with the smallest n, while the BPDN result needs a very large n ($n \geq 90\%$). Also even with $n = 90\%$, the BPDN error bound is very large.

n	Reg-mod-BPDN	Mod-BPDN	BPDN
$0.13m$	0.885	not hold	not hold
$0.19m$	0.161	0.303	not hold
$0.5m$	0.0199	0.0199	not hold
$0.9m$	0.014	0.014	0.27

Table 3.1 Sufficient conditions and normalized bounds comparison of reg-mod--BPDN, mod-BPDN and BPDN. Signal length $m = 256$, support size $|N| = 0.1m$, $|\Delta| = 4\%|N|$, $|\Delta_e| = 4\%|N|$, $\sigma_w^2 = 10^{-5}$ and $\sigma_p^2 = 10^{-3}$. "not hold" means the one or all of the sufficient conditions does not hold.

3.6.5 Comparing Theorems 4, 5, 6

In Fig. 3.4 (a), we compare the results from Theorems 4, 5 and 6 for one simulation. We plot $\frac{\text{bound}}{\|x\|_2}$ for $|\Delta|/|N|$ ranging from 0 to 0.2. Also, we used $m = 256$, $|N| = 26$, $|\Delta_e| = 0.1|N|$, $\sigma_p^2 = 10^{-3}$,

المنارة للاستشارات
www.manaraa.com

$\beta_l = 1$ and $\beta_m = \beta_s = 0.25$. Also, $n = 0.13m$ and $\sigma_w^2 = 10^{-5}$. We used $\gamma = \gamma^*$ given in the respective theorems, and we set $\lambda = 10\sigma_w^2/\sigma_p^2$. We notice the following. (1) The bound of Theorem 4 is much larger than that of Theorem 5 or 6, even for $|\Delta| = 0.04|N|$. (2) For larger values of $|\Delta|$, the sufficient conditions of Theorem 4 do not hold and hence it does not provide a bound at all. (3) For reasons explained in Chapter 3.3, in this case, the bound of Theorem 6 is equal to that of Theorem 5. Recall that the computational complexity of the bound from Theorem 5 is exponential in $|\Delta|$. However if $|\Delta|$ is small, e.g. in our simulations $|\Delta| \leq 5$, this is still doable.

3.6.6 Upper bound comparisons using Theorem 6

In Fig. 3.4(b), we do two things. (1) We compare the reconstruction error bounds from Theorem 6 for reg-mod-BPDN, mod-BPDN and BPDN and compare them with the bounds for LS-CS error given in [60, Corollary 1]. All bounds hold without any sufficient conditions which is what makes this comparison possible. (2) We also use the γ^* given by Theorem 6 to obtain the reconstructions and compute the Monte Carlo averaged N-RMSE. Comparing this with the Monte Carlo averaged upper bound on the N-RMSE, $\sqrt{\frac{\mathbb{E}[\text{bound}^2]}{\mathbb{E}[\|x\|_2^2]}}$, allows us to evaluate the tightness of a bound. Here $\mathbb{E}[\cdot]$ denotes the mean computed over 100 Monte Carlo simulations and "bound" refers to the right hand side of (3.36). We used $m = 256$, $|N| = 26$, $|\Delta_e| = 0.1|N|$, $\sigma_p^2 = 10^{-3}$, $\beta_l = 1$, $\beta_m = \beta_s = 0.25$, and $|\Delta|$ was varied from 0 to $0.2|N|$. Also, $n = 0.13m$ and $\sigma_w^2 = 10^{-5}$.

From the figure, we can observe the following. (1) Reg-mod-BPDN has much smaller bounds than those of mod-BPDN, BPDN and LS-CS. The differences between reg-mod-BPDN and mod-BPDN bounds is minor when $|\Delta|$ is small but increases as $|\Delta|$ increases. (2) The conclusions from the reconstruction error comparisons are similar to those seen from the bound comparisons, indicating that the bound can serve as a useful proxy to decide which algorithm to use when (notice bound computation is much faster than computing the reconstruction error). (3) Also, reg-mod-BPDN and mod-BPDN bounds are quite tight as compared to the LS-CS bound. BPDN bound and error are both 100%. 100% error is seen because the reconstruction is the all zeros' vector.

In Fig. 3.4(c), we did a similar set of experiments for the case where A corresponds to a simulated MRI experiment, i.e. $A = F_u \cdot W^{-1}$ where F_u contains randomly selected rows of the 2D-DFT matrix

المنارة للاستشارات

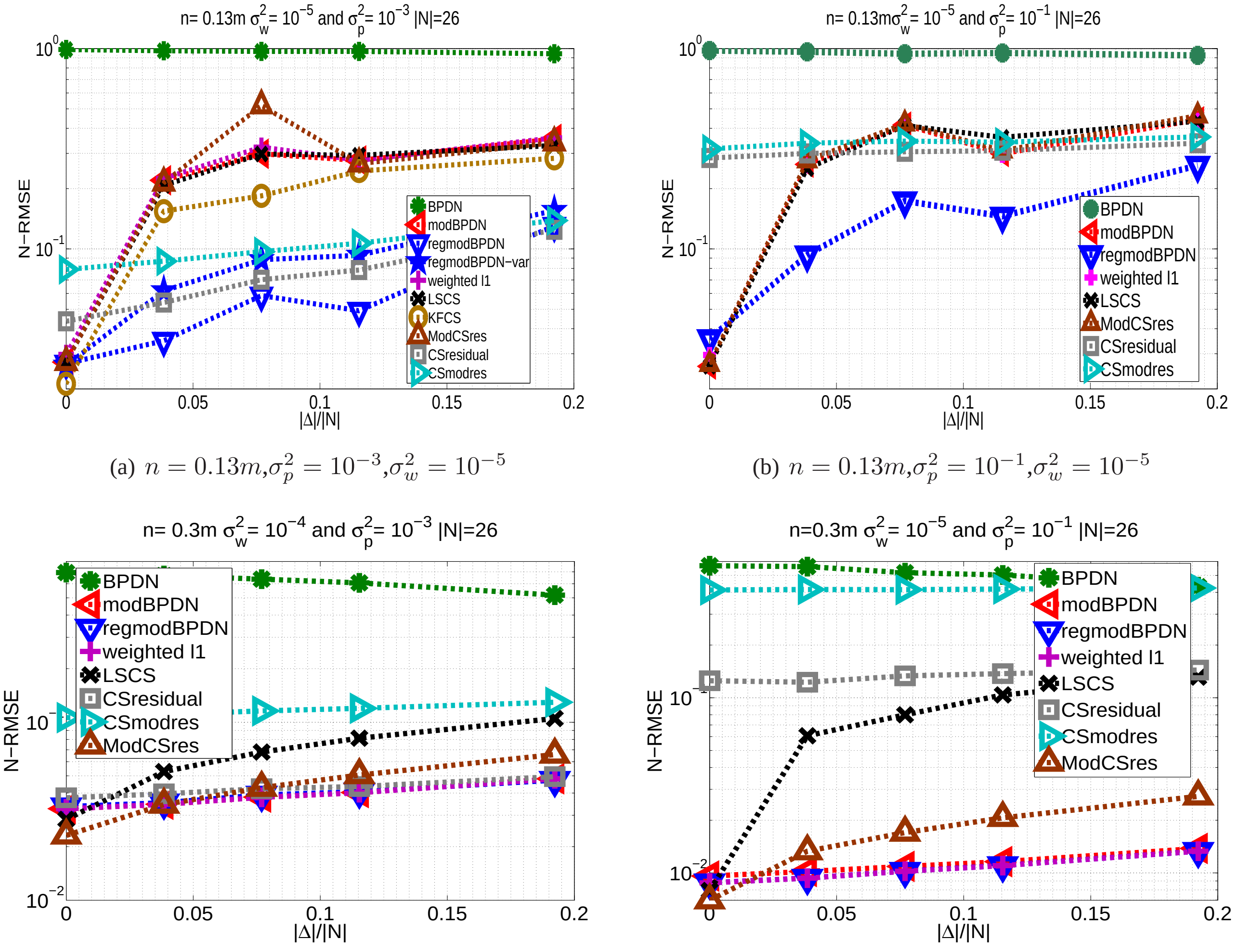
www.manaraa.com

and W is the inverse 2D-DWT matrix (for a two-level Daubechies-4 wavelet). We used $n = 0.17m$ and $\sigma_w^2 = 10^{-3}$. All other parameters were the same as in Fig. 3.4(b). Our conclusions are also the same.

The complexity for Theorem 6 is polynomial in $|\Delta|$ whereas that of the LS-CS bound [60, Corollary 1] is exponential in $|\Delta|$. To also show comparison with the LS-CS bound, we had to choose a small value of $m = 256$ so that the maximum value of $|\Delta| = 0.2|N| = 5$ was small enough. In terms of MATLAB time, computation of the Theorem 6 bound for reg-mod-BPDN took 0.2 seconds while computing the LS-CS bound took 1.2 seconds. For all methods except LS-CS, we were able to do the same thing fairly quickly even for $m = 4096$, or even larger. It took only 8 seconds to compute the bound of Theorem 6 when $m = 4096$, $n = 0.13m$, $|N| = 410 = 0.1m$ and $|\Delta| = |\Delta_e| = 0.1|N| = 41$.

المنارة للاستشارات

www.manaraa.com



(a) $n = 0.13m$,$\sigma_p^2 = 10^{-3}$,$\sigma_w^2 = 10^{-5}$

(b) $n = 0.13m$,$\sigma_p^2 = 10^{-1}$,$\sigma_w^2 = 10^{-5}$

(c) $n = 0.3m$,$\sigma_p^2 = 10^{-3}$,$\sigma_w^2 = 10^{-4}$

(d) $n = 0.3m$,$\sigma_p^2 = 10^{-1}$,$\sigma_w^2 = 10^{-5}$

Figure 3.1 The N-RMSE for reg-mod-BPDN, mod-BPDN, BPDN, LS-CS, KF-CS, weighted ℓ_1, CS-residual, CS-mod-residual and modified-CS-residual are plotted. For $n = 0.13m$, reg-mod-BPDN has smaller errors than those of mod-BPDN and the gap is larger when the signal estimate is good. For $n = 0.3m$, the errors of reg-mod-BPDN, mod-BPDN and weighted ℓ_1 are close and all small.

www.manaraa.com

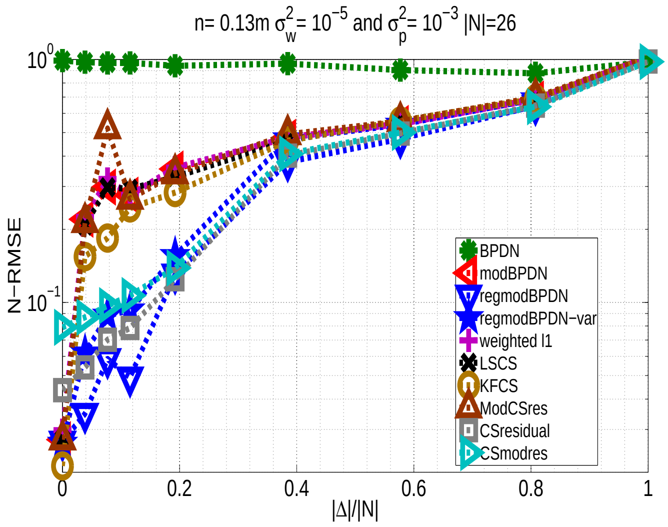


Figure 3.2 Plot of Fig 3.1(a) extended all the way to $|\Delta|/|N| = 1$ (which is the same as $\Delta = N$). Notice that if $|\Delta_e| = 0$, then the point $|\Delta|/|N| = 1$ of reg-mod-BPDN (or of mod-BPDN) is the same as BPDN. But in our plot, $|\Delta_e| = 3$ and hence the two points are different, even though the errors are quite similar.

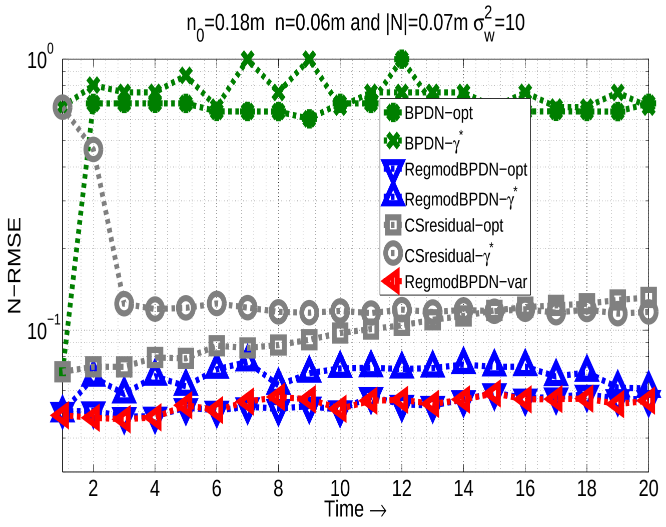
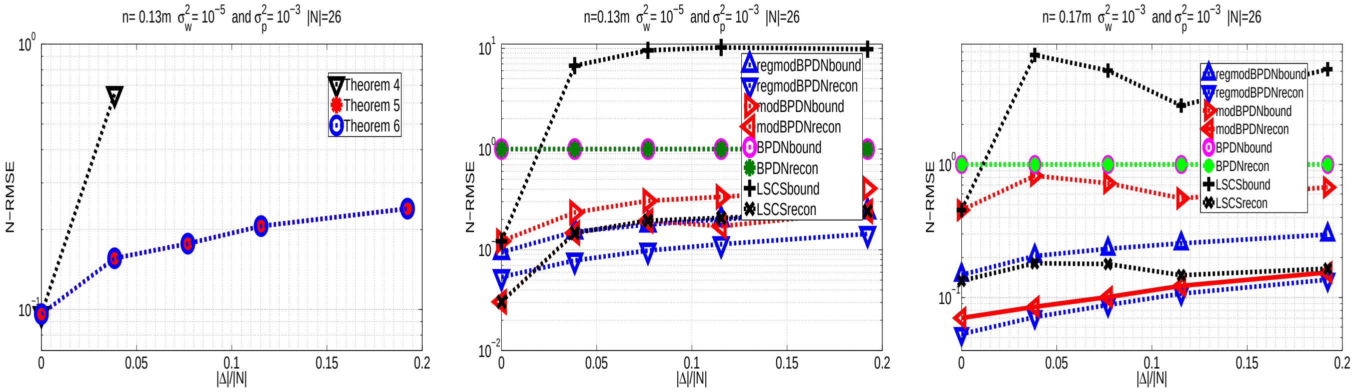


Figure 3.3 Reconstructing a 32×32 block of the actual (compressible) larynx sequence from partial Fourier measurements. Measurements $n = 0.18m$ for $t = 0$ and $n = 0.06m$ for $t > 0$. Reg-mod-BPDN has the smallest reconstruction error among all methods.

المنارة للاستشارات

www.manaraa.com



(a) $n = 0.13m$,$\sigma_p^2 = 10^{-3}$,$\sigma_w^2 = 10^{-5}$ (b) $n = 0.13m$,$\sigma_p^2 = 10^{-3}$,$\sigma_w^2 = 10^{-5}$ (c) $n = 0.17m$,$\sigma_p^2 = 10^{-3}$,$\sigma_w^2 = 10^{-3}$

Figure 3.4 In (a), we compare the three bounds from Theorem 4, 5 and 6 for one realization of x. In (b) and (c), we compare the normalized average bounds from Theorem 6 and reconstruction errors with random Gaussian and partial Fourier measurements respectively.

المنارة للاستشارات
www.manaraa.com

CHAPTER 4. Modified-CS-residual for Recursive Reconstruction of Highly Undersampled Functional MRI Sequences

In previous four chapters, we have discussed algorithms and analyzed the exact recovery conditions or bounded the reconstruction errors. In this chapter, we study the application of modified-CS based approaches for blood oxygenation level dependent (BOLD) contrast functional MR imaging (fMRI). In particular, we show, via exhaustive experiments on actual MR scanner data for brain fMRI, that our recently proposed approach for recursive reconstruction of sparse signal sequences, modified-CS-residual, outperforms other existing CS based approaches. Modified-CS-residual exploits the fact that the sparsity pattern of brain fMRI sequences and their signal values change slowly over time. It provides a fast, yet accurate, reconstruction approach that is able to accurately track the changes of the active pixels, while using only about 30% measurements per frame. Significantly improved performance over existing work is shown in terms of practically relevant metrics such as active pixel time courses, activation maps and receiver operating characteristic (ROC) curves.

In BOLD contrast fMRI, a time-series of T_2^*-weighted images are collected as the subject is presented a controlled stimulus. To achieve whole-brain coverage fMRI is typically performed at a low spatial (e.g., $3 \times 3 \times 3\ mm^3$ voxels) and temporal (e.g., volume repetition time of $2 - 3$ seconds) resolution. This provides a sufficient signal-to-noise ratio for robust detection of BOLD contrast by statistical testing. However, if CS based approaches can be applied to fMRI it may ultimately enable higher spatial and temporal resolution functional brain imaging, which potentially provides a new view of human brain function [61].

The application of CS to MRI was first developed in detail in [35]. The most straightforward application of CS to fMRI images reconstruction would be to perform CS on each slice of data independently (simple-CS). For time sequences, batch-CS [36] improves simple-CS by jointly reconstructing the en-

المنارة للاستشارات
www.manaraa.com

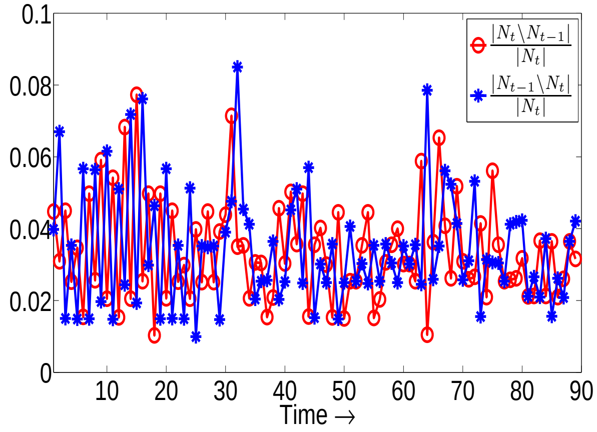


Figure 4.1 Slow support change plots for a simulated brain fMRI sequence (details are given in Chapter 4.3). N_t refers to the 99% energy support of the two-level Daubechies-4 2D discrete wavelet transform (DWT) of the image at time t. $|N_t| \approx 0.05m$. We the plot support changes, additions and deletions, with respect to the previous frame

tire sequence by treating it as a 3D sparse signal. Because it uses sparsity also along the time axis, it is able to achieve accurate reconstructions using much fewer measurements than simple-CS. But the reconstruction can only be performed on the entire *batch* of data after all sampling is completed. Also, for an l-frame acquisition, its computational complexity is roughly l^2 times that of simple-CS, while its memory requirement is l times that of simple-CS. In recent work, [37, 38] proposed Kt-FOCUSS, which uses the fact that a sequence of MR image data is sparse in the $y - f$ domain where f denotes temporal frequency. The key idea is to reconstruct $kY - t$ "frames" using FOCUSS[39] where kY denotes the phase encoding direction (y-axis of the 2D discrete Fourier transform (DFT) plane). Kt-FOCUSS is still a batch method, which means it is still (a) non-causal, i.e. it needs to wait to acquire the entire l frame sequence before doing the reconstruction (or one needs to re-run it in a batch fashion again at each time which is slow), and (b) its memory requirement is still l times that of simple-CS. But its reconstruction is fast because it is done on one $kY - t$ "frame" at a time and because often it only runs a a few iterations of FOCUSS starting from previous "frame" as initial guess. The same memory and non-causality issues also remain with Kt-FOCUSS with motion compensation (MC) [37]. Moreover, as we demonstrate in our experiments, for the fMRI based BOLD contrast detection application

المنارة للاستشارات
www.manaraa.com

that we study here, its performance is, in fact, slightly worse than our proposed recursive approach (modified-CS-residual) because of its assumption of Fourier sparsity along the time axis – it tries to recover the sparsest sum of sinusoids to represent the time sequence for a given pixel.

In recent work, we studied the problem of recursively reconstructing a time sequence of (approximately) sparse signals from highly undersampled measurements and proposed two sets of approaches – LS-CS and KF-CS [33] and later modified-CS and modified-CS-residual [25, 40]. By "recursive", we mean that we use only the previous reconstruction and the current measurements' vector to recover the current signal. As a result, these are (a) causal approaches, i.e. they can recover the current frame as soon as its MR data gets acquired; and (b) they have the same storage (memory) and computational complexity as that of simple-CS (and hence much lower than that of batch methods), but they can achieve significantly lower reconstruction errors than simple-CS when the available number of measurements is too few for simple-CS.

In all the above works, we have done experiments only on either fully simulated data or simulated MRI data, i.e. real medical image sequences, but random-sampled MRI is simulated by taking the 2D discrete Fourier transform (DFT) of the image and randomly sampling it. Moreover, only the mean squared error (MSE) has been used as the performance evaluation metric. But we know that when using actual MR scanner data, (a) there are multiple sources of noise and modeling error so that the resulting 2D-DFT of the image is no longer conjugate symmetric (its inverse DFT is not fully real); and (b) randomly sampling the 2D-DFT plane is not a practical scanning approach. In practice, one can only random sample in one direction e.g. one can only random sample rows or columns of the 2D-DFT plane. (c) Moreover, it is well known to the image processing and medical imaging communities that MSE over the entire image is not a useful performance metric since it does not capture errors in individual pixels very well. But often errors in even a few pixels can be quite problematic, e.g. they can indicate incorrect active regions.

In this chapter, we perform a detailed experimental evaluation of modified-CS-residual for

1. a real functional MRI application (that of detecting the active region in the brain as a stimulus is provided to the subject);

2. with actual MR scanner data that is acquired in a practically sensible fashion (randomly sample

المنارة للاستشارات
www.manaraa.com

the ky axis); and

3. using practically relevant performance metrics – activation maps and receiver operating characteristic (ROC) curves.

Modified-CS relies on a key assumption that the sparsity pattern (support change in the sparsity basis) changes slowly with time for most practical image sequences. We demonstrate this for brain fMRI sequences in Fig. 4.1. Notice that the maximum support change is less than 7% of the support size in most cases and in the worst case it is less than 10%. Denote the support estimate from the previous time by T. The key idea of modified-CS is to find the solution that is sparsest outside of T while satisfying the data constraint.

Some other related approaches include Dynamic-LASSO [62] which is a causal but batch approach (with very high computational and storage cost) and it assumes that the sparsity pattern of the image sequence *does not* change with time; or [48] which recovers the difference image by doing CS on the measurement differences(CS-diff). Both CS-diff and our earlier work on LS-CS and KF-CS have already been demonstrated to have worse performance than modified-CS [25, 40]. Approaches related to modified-CS for a static problem but with partial support knowledge include [27, 26].

4.1 Problem Formulation

We formulate the problem for a single slice of fMRI acquired over time. Let $(I_t)_{m_1 \times m_1}$ denote the image at time t and let $m := m_1^2$ be its dimension. The full sampling measurement model is

$$Y_{full,t} = S_t + Z_t \tag{4.1}$$

where $Y_{full,t}$ is the measured k-space data at time t. S_t is the ideal k-space data and Z_t is the measurement noise, which is modeled as a complex Gaussian noise. The image reconstructed from the full Fourier samples, I_t, can be rewritten as

$$I_t = F' Y_{full,t} F' = I_{true,t} + \eta_t \tag{4.2}$$

where F is the DFT matrix and $I_{true,t}$ is the ideal image reconstructed from noise-free k-space data. $\eta_t = F' Z_t F'$ is the degrading noise in image domain, which is complex and zero mean Gaussian with

www.manaraa.com

variance σ_η^2. We further model the complex image I_t as follows. Each pixel is made up of the baseline MR signal, the functional signal of interest, nuissance signals[63], and the degrading noise signal. Then we model a slice in an fMRI time-sequence as [64].

$$I_t(i,j) = I_b(i,j) + \nu_t(i,j) + \alpha(i,j) \cdot b_t(i,j) + \eta_t(i,j) \tag{4.3}$$

Here, i, j are the pixel indices with $i, j \in \{1, \ldots, m\}$. I_b is the baseline MR signal which does not change over time. $b_t(i,j)$ denotes the unit-amplitude BOLD signal shape in pixel (i,j), the exact form of which depends on the hemodynamic response function (HDR) corresponding to the pixel. $\alpha(i,j)$ is the non-negative amplitude of the BOLD signal in pixel (i,j) that will be equal to zero in inactive pixels. ν_t is the nuissance signal, which are modeled only for completeness since we aim to faithfully reconstruct I_t from highly undersampled data. From these definitions, the contrast-to-noise ratio (CNR) of the BOLD signal in each pixel can be expressed as $CNR(i,j) = \frac{\alpha(i,j)}{\sigma_\eta}$. MR images, especially MR brain images are known to be compressible in the wavelet transform domain[35]. Hence, we set up the measurement model of CS as follows. Let X_t denote the 2D discrete wavelet transform (DWT) of the image representation from ideal k-space, i.e. $X_t := WI_{true,t}W'$, where W is the DWT matrix. Then $Y_{full,t} = FW'X_tWF + Z_t$. We capture a smaller number, $n < m$, of Fourier coefficients of the images. Since we only sample in kY direction, this can be modeled by applying an $\frac{n}{m_1} \times m_1$ sampling mask, M_{2D} (which contains a single 1 at a different location in each row and all other entries are zero) to $Y_{full,t}$ to obtain the measurements Y_t,i.e. $Y_t = M_{2D}Y_{full,t} = M_{2D}(FW'X_tWF + Z_t)$. The above can also be transformed to a 1D problem by using Kronecker product, denoted by $\bigotimes$. Let $y_{full,t} := vec(Y_{full,t})$, $x_t := vec(X_t)$ and $z_t := vec(Z_t)$. Here, $vec(X_t)$ denotes the vectorization of the matrix X_t formed by stacking the columns of X_t into a single column vector. Then $y_{full,t} = F_{1D}W'_{1D}x_t + z_t$ where $F_{1D} = F \bigotimes F$, $W'_{1D} = W' \bigotimes W'$. An $n \times m$ mask $M_{1D} = Id_{m_1} \bigotimes M_{2D}$ is applied to $y_{full,t}$ to undersample the Fourier coefficients to obtain y_t where Id_{m_1} is an $m_1 \times m_1$ identity matrix. The above can be rewritten as

$$y_t = Ax_t + z_t,\ where\ A := H\Phi, \tag{4.4}$$

where $H := M_{1D}F_{1D}$ and $\Phi := W'_{1D}$. For our algorithm, we require A be satisfying $S = (|T| + 2|\Delta|)$ RIP property[18].

المنارة للاستشارات
www.manaraa.com

Our final goal is to detect the active pixels' region from the reconstructed sequence, i.e. detect the region where $b_t(i, j) > 0$.

4.2 Modified-CS-residual

BPDN[1] is the most commonly used method in noisy CS. Modified-BPDN[28] tries to find the signal sparsest outside of the set T while satisfying the data constraint. For signal sequences with slow changing support, we can use $T = \hat{N}_{t-1}$. When the measurements are few(smaller than what CS needs), modified-BPDN is known to have much smaller reconstruction error than that of CS(as long as $|\Delta|$ and $|\Delta_e|$ are small) [28].

Furthermore, by using this fact that signal/image also changes slowly over time, we can apply modified-BPDN on the observation residual computed using the previous signal estimate (or using the first signal estimate), i.e. we can solve

$$\arg\min_b \ \|y_t - Ax_{t,temp} - Ab\|_2^2 + \gamma \|b_{T^c}\|_1 \tag{4.5}$$

with $\hat{x}_{t,temp} = \hat{x}_{t-1}$ or $\hat{x}_{t,temp} = \hat{x}_1$. The reconstructed signal $\hat{x}_t$ is then given by

$$\hat{x}_t = \hat{b} + \hat{x}_{t,temp} \tag{4.6}$$

We refer the above as modified-CS-residual. If n is small and γ is not large enough, modified-BPDN will not have a unique minimizer. Modified-CS-residual in (4.5) ensures that the chosen minimizer is the one closest to $\hat{x}_{t,temp}$. Assuming that $\hat{x}_{t,temp}$ is a good initial estimate of x_t, this would be the correct one. In our experiments, we used $\hat{x}_{t,temp} = \hat{x}_1$, the baseline signal at the first frame. The entire algorithm is summarized in Algorithm 3.

4.3 Experimental Results

In this section, we show experiments on real fMRI sequences. We evaluate the performance of detection using 'activation map', 'Receiver operating characteristic(ROC)' and 'time course'. Two-level Daubechies-4 2D discrete wavelet transform(DWT) is used as the sparsifying basis. N_t refers to the 99% energy support of the wavelet coefficients of each frame. Variable density undersampling

المنارة للاستشارات
www.manaraa.com

```
**Algorithm 3** Modified-CS-residual
Initialization: Do inverse DFT for $x_1$ and set $\hat{N}_1 = \{k : |(\hat{x}_1)_k| \geq \tau\}$. For $t > 0$, do,
    1. **Modified-CS-residual**
        (a) Set $\hat{x}_{t,temp} = \hat{x}_1$.
        (b) **Do Modified-CS-residual.** Compute $\hat{b} = \arg\min_b \|y_t - A\hat{x}_{t,temp} - Ab\|_2^2 + \gamma\|(b)_{\hat{N}^c_{t-1}}\|_1$.
        (c) **Compute the support.** Set $\hat{x}_t = \hat{x}_{t,temp} + \hat{b}$ and compute $\hat{N}_t = \{k : |(\hat{x}_t)_k| \geq \tau\}$.
    2. **Output** $\hat{N}_t$ and $\hat{x}_t$. Increment $t$ and go to step 1.
```

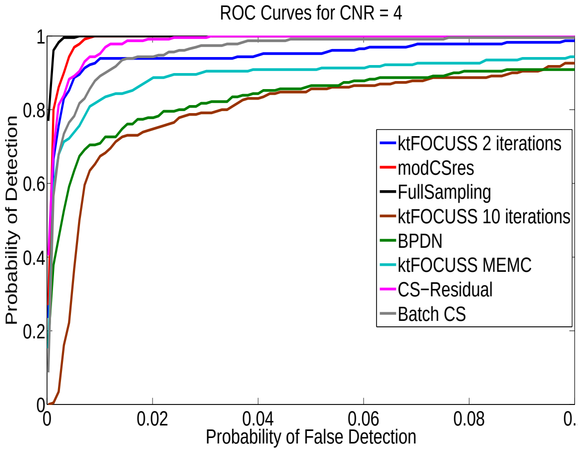


Figure 4.2 Comparing modified-CS-residual, Kt-FOCUSS with different iterations and ME/MC, BPDN, CS-residual,and Batch CS with full sampling. At $t = 1$, $n = 100\%m$ measurements are used. For $t > 1$, $n = 0.3m$ measurements are used.

scheme(which samples from a distribution that has more weight on the low frequencies) [35] is used in our experiments. The sampling mask, M_{2D}, is varying for each t. In our experiments, the reconstruction of the whole sequences takes 4 seconds for all BPDN, modified-CS-residual, CS-residual, Kt-FOCUSS with 2 iterations.

4.3.1 Real Brain Sequence(Simulated Activation)

To quantify detection performance using ROC curves, we need to know the ground truth for active regions. Hence in the first experiment, we captured a rest brain sequence (brain fMRI when no stimulus was provided to the subject) using a real MR scanner, but we added the activation later in software.

www.manaraa.com

Rest fMRI (TR/TE $= 2500/24.3$ ms, 90 degree flip angle, 3 mm slick thickness, 22 cm FOV, 64×64 matrix, 90 volumes) was performed using a $3T$ whole-body MR scanner and a gradient-echo echo-planar imaging(EPI) acquisition sequence. We added synthetic BOLD contrast at an average CNR of 4 to pixels corresponding to motor activation on one slice. The 64×64 slice image has 23 active pixels. The BOLD signal was created by convolving a bi-Gamma HDR model (6-s onset delay, 4-s FWHM) with binary-valued function representing a block stimulus (30 s active, 30 s rest; start/end in rest condition). 10 separate observations were generated by resampling with the wavestrapping technique[65] the original rest fMRI data and adding activation to the appropriate pixels to compute descriptive statistics and compute meaningful performance curves.

We compare modified-CS-residual, Kt-FOCUSS, BPDN, batch-CS, CS-residual with IDFT using full sampling. CS-residual, an improved version of CS-diff, refers to doing BPDN on the observation residual computed using the first reconstructed frame. Fig. 4.2 shows the ROC curves of all methods. From the figure, it is clear that modified-CS-residual has the best performance since the its ROC curve is strictly higher than those of other methods and closest to full sampling. We do not show N-RMSE plot since it can not show the detection performance. But modified-CS-residual has similar N-RMSE as those of Kt-FOCUSS and CS-residual and they are much smaller than other methods. For Kt-FOCUSS, increasing the number of iterations will not help improve the detection performance even if it can reduce N-RMSE. With more iterations, the temporal DC component of Kt-FOCUSS reconstruction becomes better while many other nonzero frequency components are eliminated. Hence, the reconstructed signal is more 'flat' with more iterations which worsens the detection for active pixels but reduces N-RMSE. Similarly, Kt-FOCUSS with ME/MC also has smaller N-RMSE but worse detection performance. CS-residual does not use the slow support change, therefore it has worse detection than modified-CS-residual.

Time courses for one active pixel are shown in Fig. 4.3. It is also observed that modified-CS-residual does best to track the time course of true(fully sampled) signal, thus providing good reconstruction and detection.

المنارة للاستشارات
www.manaraa.com

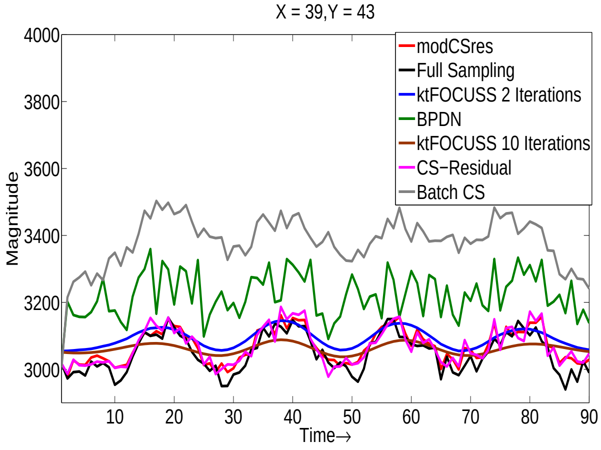
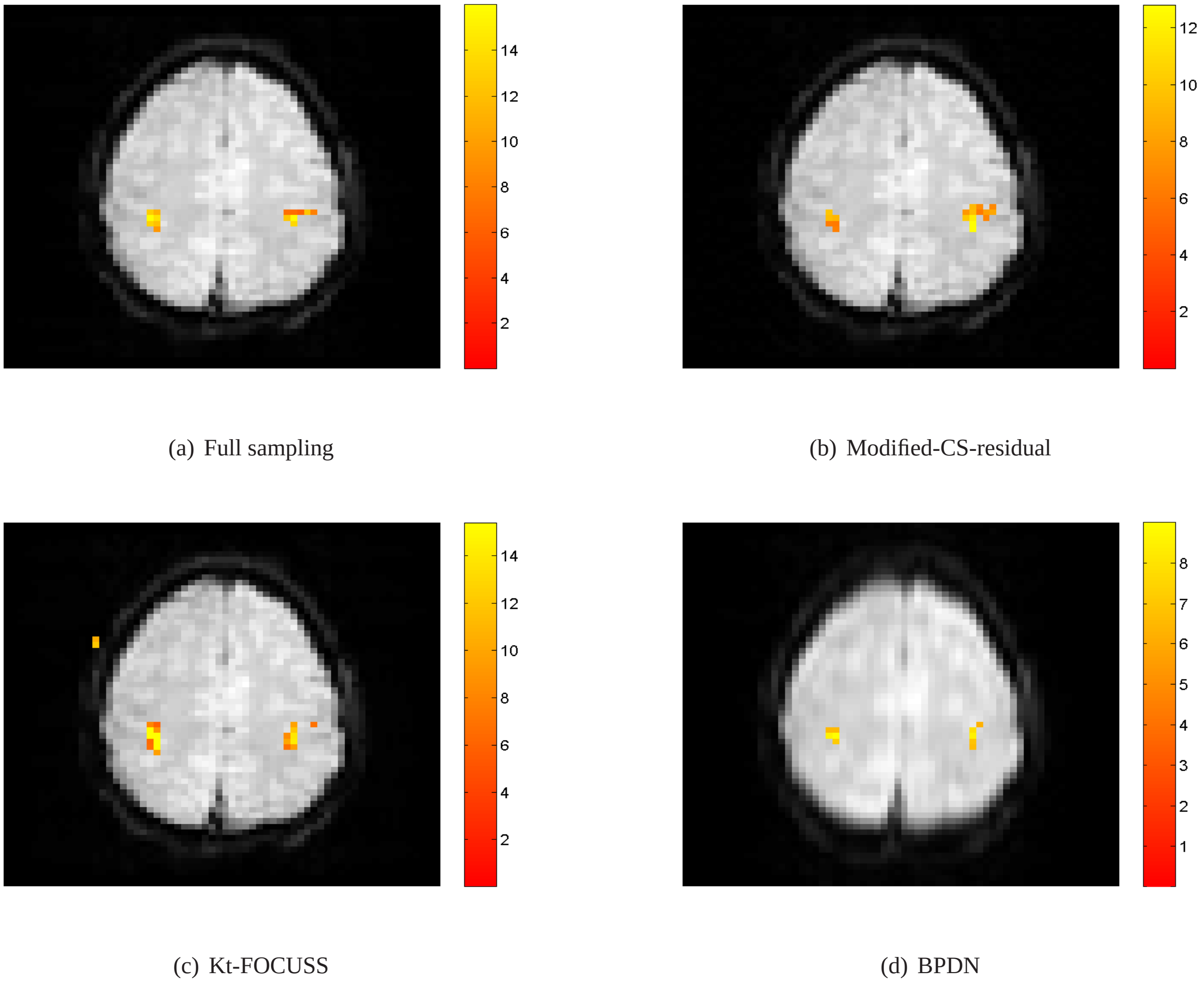


Figure 4.3 Time courses of one active pixel.

4.3.2 Real Brain Sequence(Real Activation)

For real data sequences, we cannot use ROC curves to compare the performances of different methods since no ground truth is available. Our comparison is based on how the detected activation can approximate the activation of IDFT using full Fourier samples. Activation maps for a given threshold in t-test are used to study the detected activation. Different from the simulated sequence, the activations of the real data are not so ideal. For active brain imaging, we used the same experimental setup as the one in 4.3.1 except using $n = 0.33m$ measurements for $t > 1$. The activation maps are shown in Fig. 4.4 for the reconstructions using modified-CS-residual, Kt-FOCUSS and BPDN compared with full sampling when threshold for t-test is set the same for all algorithms. The Bonferroni-corrected threshold is chosen as 5 computed from the dataset. We easily observe that modified-CS-residual has most active pixels detected and few false detection while both Kt-FOCUSS and BPDN has many missing detection.

المنارة للاستشارات
www.manaraa.com



(a) Full sampling

(b) Modified-CS-residual

(c) Kt-FOCUSS

(d) BPDN

Figure 4.4 Comparing activation maps for modified-CS-residual, Kt-FOCUSS, and BPDN with full sampling for each reconstruction. We can see modified-CS-residual has the closest detected regions to full sampling. Modified-CS-residual only has 1 missing active pixel and 5 false ones while Kt-FOCUSS has 4 missing and 11 false ones. BPDN has 7 missing active pixels and 2 false ones.

www.manaraa.com

CHAPTER 5. Conclusions and Future Directions

In this work we studied the problem of sparse reconstruction from noiseless or noisy undersampled measurements when partial and partly erroneous, knowledge of the signal's support and an erroneous estimate of the signal values on the "partly known support" is also available. Denote the support knowledge by T and the signal value estimate on T by $\hat{\mu}_T$. We proposed and studied the solutions modified-CS and regularized modified-BP for noiseless measurements as well as modified-BPDN and regularized modified-BPDN for noisy measurements.

Modified-CS for noiseless measurements solves an ℓ_1 relaxation of the following problem: find the signal that is sparsest outside of T and that satisfies the data constraint. We derived sufficient conditions for exact reconstruction using mod-CS. These are much weaker than those for CS when the sizes of the unknown part of the support and of errors in the known part are small compared to the support size. Simulation results showing greatly improved performance of mod-CS using both random Gaussian and partial Fourier measurements are shown on both sparse and compressible signals and image sequences. An important extension of mod-CS, Regularized modified-BP, was developed that also uses prior signal estimate knowledge. We obtained the exact reconstruction conditions for reg-mod-BP and argued that if some of the inequality constraints are active and if even a subset of the set of active constraints satisfies certain conditions, then reg-mod-BP achieves exact recovery under weaker conditions than what mod-CS needs. A practical situation where this would happen is when both the signal and its estimate are quantized. In other cases, the conditions are only as weak as those for mod-CS. In either case they are much weaker than those for BP as long as T is a good support estimate. We also provided the reconstruction error bound when the exact recovery can not happen. Similarly, the error bound is smaller or at least as large as that for mod-CS. From simulations, we see that even without any active constraints, the reg-mod-BP reconstruction error is much lower than that of mod-CS.

المنارة للاستشارات

www.manaraa.com

We also proposed a modification of the BPDN idea, modified-BPDN, for sparse reconstruction from noisy measurements when a part of the support is known, and bounded its reconstruction error. A key feature of our work is that the bounds that we obtain are computable. Hence, we used Monte Carlo to show that the average value of the bound increases as the unknown support size or the size of the error in the known support increases and mod-BPDN requires weaker conditions than BPDN needs. Also, Regularized Modified-BPDN, the extension of mod-BPDN, was proposed when signal estimate is also available. We bounded its reconstruction error and introduced the tightest bounds for regularized modified-BPDN and modified-BPDN. We showed how to obtain computable error bounds that hold without any sufficient conditions. This made it easy to compare bounds for the various approaches (corresponding results for mod-BPDN and BPDN follow as direct corollaries). Empirical error comparisons with these and many other existing approaches are also provided.

In this work, we also studied the problem of recursively and causally reconstructing a sequence of fMRI sequences from a reduced number of Fourier measurements. We demonstrated improved reconstruction and activation pattern detection performance of our proposed solution, modified-CS-residual on the real fMRI sequences, compared to existing work.

In ongoing work, we want to evaluate the utility of reg-mod-BPDN for recursive functional MR imaging to detect brain activation patterns in response to stimuli [66]. On the other end, we are also working on obtaining conditions under which it will remain "stable" (its error will be bounded by a time-invariant and small value) for a recursive recovery problem. In [59], this has been done for the constrained version of reg-mod-BPDN. That result uses the restricted isometry constants (RIC) and the restricted orthogonality constants (ROC) [18, 19] in its sufficient conditions and bounds. However, this means that the conditions and bounds are not computable. Also, since the stability holds under a different set of sufficient conditions and has a different error bound than that for mod-CS [67] or LS-CS [33] or CS [19], comparison of the various results is difficult. An open question is how to extend the results of the current work (which are computable) to show the stability of unconstrained reg-mod-BPDN. In future, we also want to do joint real-time detection and reconstruction to further improve performance. Also, higher spatial and temporal resolution sequences will be experimented.

المنارة للاستشارات
www.manaraa.com

APPENDIX A. Appendix for the Proof of Exact Reconstruction Conditions of Regularized Modified-BP

Recall that $k = |T|$, $u = |\Delta|$, $e = |\Delta_e|$ and $s = |N|$.

A.1 Proof of Proposition 1

The proof follows by contradiction. Suppose that we can find two different solutions b_1 and b_2 that satisfy $y = Ab_1 = Ab_2$ and have the same ℓ_0 norm, u, along T^c. Thus b_1 is nonzero along T (or a subset of it) and some set Δ_1 of size u while b_2 is nonzero along T (or a subset of it) and some set Δ_2 also of size u. The sets Δ_1 and Δ_2 may or may not overlap. Thus $A(b_1 - b_2) = 0$. Since $(b_1 - b_2)$ is supported on $T \cup \Delta_1 \cup \Delta_2$, this is equivalent to $A_{T\cup\Delta_1\cup\Delta_2}(b_1 - b_2)_{T\cup\Delta_1\cup\Delta_2} = 0$. But if $\delta_{k+2u} < 1$, $A_{T\cup\Delta_1\cup\Delta_2}$ is full rank and so the only way this can happen is if $b_1 - b_2 = 0$, i.e $b_1 = b_2$.

Therefore there can be only one solution with ℓ_0 norm u along T^c that satisfies that data constraint. Since x is one such solution, any other solution has to be equal to x.

A.2 Proof of Lemma 1

Denote a minimizer of (2.14) by b. Since $y = Ax$ and x satisfies (2.10), x is feasible for (2.14). Thus,

$$\|b_{T^c}\|_1 \leq \|x_{T^c}\|_1 = \|x_\Delta\|_1. \tag{A.1}$$

المنارة للاستشارات

www.manaraa.com

Next, we use the conditions on w given in Lemma 1 and the fact that x is supported on $N \subseteq T \cup \Delta$ to show that $\|b_{T^c}\|_1 \geq \|x_{T^c}\|_1$ and hence $\|x_{T^c}\|_1 = \|b_{T^c}\|_1$. Notice that

$$\begin{aligned}
\|b_{T^c}\|_1 &= \sum_{j\in\Delta} |x_j + b_j - x_j| + \sum_{j\notin T\cup\Delta} |b_j| \geq \sum_{j\in\Delta} |x_j + b_j - x_j| + \sum_{j\notin T\cup\Delta} w' A_j b_j \qquad (A.2)\\
&\geq \sum_{j\in\Delta} \text{sgn}(x_j)(x_j + (b_j - x_j)) + \sum_{j\notin T\cup\Delta} w' A_j (b_j - x_j) \qquad (A.3)\\
&= \|x_\Delta\|_1 + \sum_{j\notin T} w' A_j (b_j - x_j) = \|x_\Delta\|_1 + w'(Ab - Ax) - \sum_{j\in T} w' A_j (b_j - x_j) \qquad (A.4)\\
&= \|x_\Delta\|_1 - \sum_{j\in T} w' A_j (b_j - \hat{\mu}_j + \hat{\mu}_j - x_j) \qquad (A.5)\\
&= \|x_\Delta\|_1 - \sum_{j\in T_{\text{a+}}} w' A_j (b_j - \hat{\mu}_j - \rho) - \sum_{j\in T_{\text{a-}}} w' A_j (b_j - \hat{\mu}_j + \rho) \qquad (A.6)\\
&\geq \|x_\Delta\|_1 = \|x_{T^c}\|_1 \qquad (A.7)
\end{aligned}$$

In the above, the inequality in (A.2) follows because $w' A_j \leq |w' A_j| < 1$ for $j \notin T \cup \Delta$ and because $|b_j| \geq b_j$. Inequality (A.3) uses the fact that $|z| \geq \text{sgn}(b) z$ for any two scalars z and b and that $x_j = 0$ for $j \notin T \cup \Delta$. In (A.4), the first equality uses $\text{sgn}(x_j) x_j = |x_j|$ and $w' A_j = \text{sgn}(x_j)$ for $j \in \Delta$. The second equality just rewrites the second term in a different form. In (A.5), we use the fact that $Ab = Ax = y$ (since both b and x are feasible) to eliminate $w'(Ab - Ax)$. Equation (A.6) uses $w' A_j = 0$ for $j \in T_{\text{in}}$ and the definitions of $T_{\text{a+}}$ and $T_{\text{a-}}$ given in (2.15). Finally, (A.7) follows because $-\sum_{j\in T_{\text{a+}}} w' A_j (b_j - \hat{\mu}_j - \rho) - \sum_{j\in T_{\text{a-}}} w' A_j (b_j - \hat{\mu}_j + \rho) \geq 0$. This holds since $-\rho \leq b_j - \hat{\mu}_j \leq \rho$ for all $j \in T$; $w' A_j \geq 0$ for $j \in T_{\text{a+}}$; and $w' A_j \leq 0$ for $j \in T_{\text{a-}}$.

Both inequalities (A.1) and (A.2)-(A.7) can hold only when $\|b_{T^c}\|_1 = \|x_{T^c}\|_1$, i.e. all the inequalities in (A.2)-(A.7) hold with equality. Consider the inequality in (A.2). Since $|w' A_j| < 1$ for $j \notin T \cup \Delta$, this holds with equality only if $b_j = 0$ for all $j \notin T \cup \Delta$. Since $Ab = y = Ax$ and since both b and x are supported on $T \cup \Delta$ (or on its subset), $A_{T\cup\Delta}(b_{T\cup\Delta} - x_{T\cup\Delta}) = 0$. Since $\delta_{k+u} < 1$, $A_{T\cup\Delta}$ has full rank. Therefore, this means that $b_{T\cup\Delta} = x_{T\cup\Delta}$. Thus, we can conclude that $b = x$, i.e., x is the unique minimizer.

A.3 Proof of Lemma 2

This proof uses the following simple facts. Let $\lambda_{\min}(M)$, $\lambda_{\max}(M)$ denote the minimum and maximum eigenvalues of a matrix M. (i) For positive semi-definite matrices, M, Q, $\|M\| = \lambda_{\max}(M)$;

المنارة للاستشارات

www.manaraa.com

$\|MQ\| \le \|M\|\|Q\|$; $\lambda_{\min}(M-Q) \ge \lambda_{\min}(M) - \lambda_{\max}(Q)$; and for a positive definite matrix, M, $\|M^{-1}\| = 1/\lambda_{\min}(M)$; (ii) for any matrices, B, C, $\|B - C\| \le \|B\| + \|C\|$; (iii) for disjoint sets T_1, T_2, $\|{A_{T_1}}'A_{T_2}\| \le \theta_{|T_1|,|T_2|}$ [25, equation (3)]; (iv) $1-\delta_{|T_1|} \le \lambda_{\min}({A_{T_1}}'A_{T_1}) \le \lambda_{\max}({A_{T_1}}'A_{T_1}) \le 1+\delta_{|T_1|}$ [18]; (v) $M(T_b)$ is a projection matrix and so $M(T_b)M(T_b)' = M(T_b)$ and $\|M(T_b)\| = 1$; (vi) $\|\text{sgn}(x_\Delta)\|_2 = \sqrt{u}$.

The lemma assumes that $\delta_u + \delta_{k_b} + \theta_{k_b,u}^2 < 1$. This implies that (a) $\delta_u < 1$ and so ${A_\Delta}'A_\Delta$ is positive definite and so $u \le n$; (b) $\delta_{k_b} < 1$ and so ${A_{T_b}}'A_{T_b}$ is positive definite and $M(T_b)$ is well-defined; and (c) as we show next, ${A_\Delta}'M(T_b)A_\Delta$ is positive definite and hence full rank. Since ${A_\Delta}'M(T_b)A_\Delta = {A_\Delta}'A_\Delta - {A_\Delta}'A_{T_b}({A_{T_b}}'A_{T_b})^{-1}{A_{T_b}}'A_\Delta$ is a difference of two positive semi-definite matrices, thus,

$$\lambda_{\min}({A_\Delta}'M(T_b)A_\Delta) \ge \lambda_{\min}({A_\Delta}'A_\Delta) - \lambda_{\max}({A_\Delta}'A_{T_b}({A_{T_b}}'A_{T_b})^{-1}{A_{T_b}}'A_\Delta) \ge (1-\delta_u) - \frac{\theta_{k_b,u}^2}{1-\delta_{k_b}} > 0 \quad \text{(A.8)}$$

Thus, ${A_\Delta}'M(T_b)A_\Delta$ is positive definite. The first inequality in (A.8) follows from fact (i). The second one follows because $\lambda_{\min}({A_\Delta}'A_\Delta) \ge (1-\delta_u)$ (using fact (iv)); $\lambda_{\max}({A_\Delta}'A_{T_b}({A_{T_b}}'A_{T_b})^{-1}{A_{T_b}}'A_\Delta) = \|{A_\Delta}'A_{T_b}({A_{T_b}}'A_{T_b})^{-1}{A_{T_b}}'A_\Delta\| \le \|{A_\Delta}'A_{T_b}\| \, \|({A_{T_b}}'A_{T_b})^{-1}\| \, \|{A_{T_b}}'A_\Delta\|$ (using fact (i)); $\|{A_\Delta}'A_{T_b}\| = \|{A_{T_b}}'A_\Delta\| \le \theta_{k_b,u}$ (using fact (iii)); and $\|({A_{T_b}}'A_{T_b})^{-1}\| = \frac{1}{\lambda_{\min}({A_{T_b}}'A_{T_b})} \le \frac{1}{1-\delta_{k_b}}$ (since ${A_{T_b}}'A_{T_b}$ is positive definite, this follows using fact (i) and fact (iv)). The third inequality of (A.8) follows because $(1-\delta_u) - \frac{\theta_{k_b,u}^2}{1-\delta_{k_b}} = \frac{1-\delta_u-\delta_{k_b}+\delta_u\delta_{k_b}-\theta_{k_b,u}^2}{1-\delta_{k_b}} > 0$. Both the numerator and the denominator are positive because we have assumed that $\delta_u + \delta_{k_b} + \theta_{k_b,u}^2 < 1$.

Using fact (v), ${A_\Delta}'M(T_b)A_\Delta = {A_\Delta}'M(T_b)M(T_b)'A_\Delta$. Thus, using the above, ${A_\Delta}'M(T_b)M(T_b)'A_\Delta$ is positive definite and hence has full rank u. Thus, the $u \times n$ fat matrix, ${A_\Delta}'M(T_b)$ has full rank, u.

To prove the lemma, we first try to construct an $n \times 1$ vector, $\tilde{w}$, that satisfies the first two conditions of the lemma. Then, we show that we can find an exceptional set E so that the constructed $\tilde{w}$ and E satisfy all the required conditions. Any $\tilde{w}$ that satisfies ${A_{T_b}}'\tilde{w} = 0$ lies in the null space of ${A_{T_b}}'$ and hence is of the form $\tilde{w} = M(T_b)\gamma$. To satisfy the second condition, we need a γ that satisfies ${A_\Delta}'M(T_b)\gamma = \text{sgn}(x_\Delta)$. As shown above, ${A_\Delta}'M(T_b)$ is full rank and so this system of equations has a solution (in fact has infinitely many solutions). We can compute the minimum ℓ_2 norm solution in closed form as $\gamma = M(T_b)'A_\Delta({A_\Delta}'M(T_b)M(T_b)'A_\Delta)^{-1}\text{sgn}(x_\Delta)$. Since $M(T_b)M(T_b)' = M(T_b)$,

المنارة للاستشارات

www.manaraa.com

$\tilde{w} = M(T_b)\gamma$ can be rewritten as

$$\tilde{w} = M(T_b)A_\Delta(A_\Delta{}'M(T_b)A_\Delta)^{-1}\text{sgn}(x_\Delta) \tag{A.9}$$

Using the definition of $T_{\text{a+g}}$, $T_{\text{a-g}}$ given in (2.17) in Theorem 2, we can see that $\tilde{w}$ satisfies the first two conditions of the lemma. Recall that $A_i{}'w > 0$ for all $i \in T_{\text{a+g}}$ is equivalent to $A_{T_{\text{a+g}}}{}'w \succ 0$, and similarly, $A_i{}'w < 0$ for all $i \in T_{\text{a-g}}$ is equivalent to $A_{T_{\text{a-g}}}{}'w \prec 0$.

Consider any set $\check{T}_d$ disjoint with $T \cup \Delta$ of size $|\check{T}_d| \leq \check{s}$. Then,

$$\begin{aligned} \|A_{\check{T}_d}{}'\tilde{w}\|_2 &\leq \|A_{\check{T}_d}{}'M(T_b)A_\Delta\| \, \|(A_\Delta{}'M(T_b)A_\Delta)^{-1}\| \, \|\text{sgn}(x_\Delta)\|_2 \\ &\leq (\theta_{\check{s},u} + \frac{\theta_{\check{s},k_b}\theta_{u,k_b}}{1-\delta_{k_b}})\frac{1}{1-\delta_u - \frac{\theta_{u,k_b}^2}{1-\delta_{k_b}}}\sqrt{u} = a_{k_b}(u,\check{s})\sqrt{u} \end{aligned} \tag{A.10}$$

Notice that $a_{k_b}(u,\check{s})$ is positive because we have assumed that $\delta_u + \delta_{k_b} + \theta_{k_b,u}^2 < 1$. The bound in (A.10) follows using the simple facts given in the beginning. We obtain (A.10) as follows. Consider the first term $\|A_{\check{T}_d}{}'M(T_b)A_\Delta\|$. Using the definition of $M(T_b)$ and fact (ii), $\|A_{\check{T}_d}{}'M(T_b)A_\Delta\| \leq \|A_{\check{T}_d}{}'A_\Delta\| + \|A_{\check{T}_d}{}'A_{T_b}(A_{T_b}{}'A_{T_b})^{-1}A_{T_b}{}'A_\Delta\|$. Using fact (iii), $\|A_{\check{T}_d}{}'A_\Delta\| \leq \theta_{\check{s},u}$, $\|A_{\check{T}_d}{}'A_{T_b}\| \leq \theta_{\check{s},k_b}$ and $\|A_{T_b}{}'A_\Delta\| \leq \theta_{u,k_b}$. Since $A_{T_b}{}'A_{T_b}$ is positive definite, using fact (i) and fact (iv), $\|(A_{T_b}{}'A_{T_b})^{-1}\| = \frac{1}{\lambda_{\min}(A_{T_b}{}'A_{T_b})} \leq \frac{1}{1-\delta_{k_b}}$. Thus, we get $\|A_{\check{T}_d}{}'M(T_b)A_\Delta\| \leq (\theta_{\check{s},u} + \frac{\theta_{\check{s},k_b}\theta_{u,k_b}}{1-\delta_{k_b}})$. Consider the second term $\|(A_\Delta{}'M(T_b)A_\Delta)^{-1}\|$. Since $A_\Delta{}'M(T_b)A_\Delta$ is positive definite, using fact (i) and (A.8), $\|(A_\Delta{}'M(T_b)A_\Delta)^{-1}\| = \frac{1}{\lambda_{\min}(A_\Delta{}'M(T_b)A_\Delta)} \leq \frac{1}{(1-\delta_u)-\frac{\theta_{u,k_b}^2}{1-\delta_{k_b}}}$. Using fact (vi), the third term, $\|\text{sgn}(x_\Delta)\|_2 = \sqrt{u}$.

Define the set, E, as $E := \{j \in (T \cup \Delta)^c : |A_j{}'\tilde{w}| > \frac{a_{k_b}(u,\check{s})\sqrt{u}}{\sqrt{\check{s}}}\}$. Notice that $|E|$ must obey $|E| < \check{s}$ since otherwise we can contradict (A.10) by taking $\check{T}_d \subseteq E$. Since $|E| < \check{s}$ and E is disjoint with $T \cup \Delta$, (A.10) holds for $\check{T}_d \equiv E$, i.e., $\|A_E{}'\tilde{w}\|_2 \leq a_{k_b}(u,\check{s})\sqrt{u}$. Also, by definition of E, $|A_j{}'\tilde{w}| \leq \frac{a_{k_b}(u,\check{s})\sqrt{u}}{\sqrt{\check{s}}}$, for all $j \notin T \cup \Delta \cup E$. Thus $\tilde{w}$ satisfies the third condition of the lemma.

Finally, $\|\tilde{w}\|_2 \leq \|M(T_b)\| \, \|A_\Delta\| \, \|(A_\Delta{}'M(T_b)A_\Delta)^{-1}\| \, \sqrt{u} \leq K_{k_b}(u)\sqrt{u}$. This follows using fact (v); $\|A_\Delta\| \leq \sqrt{1+\delta_u}$; and fact (i) and (A.8). Thus, we have found a $\tilde{w}$ and E that satisfy all required conditions.

المنارة للاستشارات

www.manaraa.com

A.4 Proof of Lemma 3

Let $M = M(T)$.

The lemma assumes that $\delta_s + \delta_k + \theta_{k,s}^2 < 1$. This means that (a) $\delta_k < 1$ and so $A_T{}'A_T$ is positive definite; (b) $\delta_s < 1$ and so for any set T_d of size $|T_d| \leq s$, $A_{T_d}{}'A_{T_d}$ is positive definite; and (c) as we show next, for any set T_d of size $|T_d| \leq s$, $A_{T_d}{}'MA_{T_d}$ is also positive definite. Notice that $A_{T_d}{}'MA_{T_d} = A_{T_d}{}'A_{T_d} - A_{T_d}{}'A_T(A_T{}'A_T)^{-1}A_T{}'A_{T_d}$ which is the difference of two symmetric non-negative definite matrices. Let B_1 denote the first matrix and B_2 the second one. Use the fact that $\lambda_{\min}(B_1 - B_2) \geq \lambda_{\min}(B_1) + \lambda_{\min}(-B_2) = \lambda_{\min}(B_1) - \lambda_{\max}(B_2)$ where $\lambda_{\min}(.), \lambda_{\max}(.)$ denote the minimum, maximum eigenvalue. Since $\lambda_{\min}(B_1) \geq (1 - \delta_s)$ and $\lambda_{\max}(B_2) = \|B_2\| \leq \frac{\|(A_{T_d}{}'A_T)\|^2}{1-\delta_k} \leq \frac{\theta_{s,k}^2}{1-\delta_k}$, thus

$$\lambda_{\min}(A_{T_d}{}'MA_{T_d}) \geq 1 - \delta_s - \frac{\theta_{s,k}^2}{1 - \delta_k} > 0 \tag{A.11}$$

(the last inequality holds because $\delta_s + \delta_k + \theta_{k,s}^2 < 1$). Thus, $A_{T_d}{}'MA_{T_d}$ is positive definite.

Since M is a projection matrix, $MM' = M$, and so $A_{T_d}{}'MA_{T_d} = A_{T_d}{}'MM'A_{T_d}$. Thus, from above, $A_{T_d}{}'MM'A_{T_d}$ is also positive definite. Thus, $A_{T_d}{}'M$ is full rank.

Any $\tilde{w}$ that satisfies $A_T{}'\tilde{w} = 0$ will be of the form

$$\tilde{w} = [I - A_T(A_T{}'A_T)^{-1}A_T{}']\gamma := M\gamma \tag{A.12}$$

We need to find a γ s.t. $A_{T_d}{}'\tilde{w} = c$, i.e. $A_{T_d}{}'M\gamma = c$. Since $A_{T_d}{}'M$ is full rank, this system of equations has a solution (in fact, it has infinitely many solutions). Let $\gamma = M'A_{T_d}\eta$. Then $\eta = (A_{T_d}{}'MM'A_{T_d})^{-1}c = (A_{T_d}{}'MA_{T_d})^{-1}c$. This follows because $MM' = M^2 = M$ since M is a projection matrix. Thus,

$$\tilde{w} = MM'A_{T_d}(A_{T_d}{}'MA_{T_d})^{-1}c = MA_{T_d}(A_{T_d}{}'MA_{T_d})^{-1}c \tag{A.13}$$

Consider any set $\check{T}_d$ with $|\check{T}_d| \leq \check{s}$ disjoint with $T \cup T_d$. Then

$$\|A_{\check{T}_d}{}'\tilde{w}\|_2 \quad \leq \quad \|A_{\check{T}_d}{}'MA_{T_d}\| \, \|(A_{T_d}{}'MA_{T_d})^{-1}\| \, \|c\|_2 \tag{A.14}$$

www.manaraa.com

Consider the first term from the right hand side (RHS) of (A.14).

$$\begin{aligned} \|A_{\check{T}_d}{}'MA_{T_d}\| &\le \|A_{\check{T}_d}{}'A_{T_d}\| + \|A_{\check{T}_d}{}'A_T(A_T{}'A_T)^{-1}A_T{}'A_{T_d}\| \\ &\le \theta_{\check{s},s} + \frac{\theta_{\check{s},k}\,\theta_{s,k}}{1-\delta_k} \end{aligned} \tag{A.15}$$

This follows in a fashion exactly analogous to the derivation of the upper bound on the first term of (A.10) in the proof of Lemma 2. Consider the second term from the RHS of (A.14). Since $A_{T_d}{}'MA_{T_d}$ is positive definite,

$$\|(A_{T_d}{}'MA_{T_d})^{-1}\| = \frac{1}{\lambda_{\min}(A_{T_d}{}'MA_{T_d})} \tag{A.16}$$

Using (A.11),

$$\|(A_{T_d}{}'MA_{T_d})^{-1}\| \le \frac{1}{1-\delta_s-\frac{\theta_{s,k}^2}{1-\delta_k}} \tag{A.17}$$

Recall that the denominator is positive because we have assumed that $\delta_s+\delta_k+\theta_{k,s}^2<1$. Using (A.15) and (A.17) to bound (A.14), we get that for any set $\check{T}_d$ with $|\check{T}_d| \le \check{s}$,

$$\|A_{\check{T}_d}{}'\tilde{w}\|_2 \le \frac{\theta_{\check{s},s}+\frac{\theta_{\check{s},k}\,\theta_{s,k}}{1-\delta_k}}{1-\delta_s-\frac{\theta_{s,k}^2}{1-\delta_k}}\|c\|_2 = a_k(s,\check{s})\|c\|_2 \tag{A.18}$$

Notice that $a_k(s,\check{s})$ is non-decreasing in k, s, $\check{s}$. Define an exceptional set, E, as

$$E := \{j \in (T\cup T_d)^c : |A_j{}'\tilde{w}| > \frac{a_k(s,\check{s})}{\sqrt{\check{s}}}\|c\|_2\} \tag{A.19}$$

Notice that $|E|$ must obey $|E| < \check{s}$ since otherwise we can contradict (A.18) by taking $\check{T}_d \subseteq E$.

Since $|E| < \check{s}$ and E is disjoint with $T\cup T_d$, (A.18) holds for $\check{T}_d \equiv E$, i.e. $\|A_E{}'\tilde{w}\|_2 \le a_k(s,\check{s})\|c\|_2$. Also, by definition of E, $|A_j{}'\tilde{w}| \le \frac{a_k(s,\check{s})}{\sqrt{\check{s}}}\|c\|_2$, for all $j \notin T\cup T_d\cup E$. Finally,

$$\begin{aligned} \|\tilde{w}\|_2 &\le \|MA_{T_d}(A_{T_d}{}'MA_{T_d})^{-1}\|\ \|c\|_2 \\ &\le \|M\|\ \|A_{T_d}\|\ \|(A_{T_d}{}'MA_{T_d})^{-1}\|\ \|c\|_2 \\ &\le \frac{\sqrt{1+\delta_s}}{1-\delta_s-\frac{\theta_{s,k}^2}{1-\delta_k}}\|c\|_2 = K_k(s)\|c\|_2 \end{aligned} \tag{A.20}$$

since $\|M\| = 1$ (holds because M is a projection matrix). Thus we have found a $\tilde{w}$ and a set E that satisfy all conditions of the lemma.

المنارة للاستشارات
www.manaraa.com

A.5 Proof of Theorem 2

We construct a w that satisfies the conditions of Lemma 1 by first applying Lemma 2 and then applying Lemma 3 iteratively as explained below. Finally we define w using (A.25) below. At iteration zero, we apply Lemma 2 with $\check{s} \equiv u$. Lemma 2 can be applied because $k_b \leq k$ and $\delta_u + \delta_k + \theta_{k,u}^2 < 1$ (holds because condition 1 of the theorem holds). Thus, there exists a w_1 and an exceptional set $T_{d,1}$, disjoint with $T \cup \Delta$, of size less than $\check{s} = u$, s.t.

$$\begin{aligned}
{A_j}'w_1 &> 0,\ \forall\ j \in T_{\text{a+g}} \\
{A_j}'w_1 &< 0,\ \forall\ j \in T_{\text{a-g}} \\
{A_j}'w_1 &= 0,\ \forall\ j \in T_b \\
{A_j}'w_1 &= \text{sgn}(x_j),\ \forall\ j \in \Delta \\
|T_{d,1}| &< u \\
\|{A_{T_{d,1}}}'w_1\|_2 &\leq a_{k_b}(u,u)\sqrt{u} \\
|{A_j}'w_1| &\leq a_{k_b}(u,u),\ \forall j \notin T \cup \Delta \cup T_{d,1} \\
\|w_1\|_2 &\leq K_{k_b}(u)\sqrt{u}
\end{aligned} \qquad \text{(A.21)}$$

At iteration $r > 0$, apply Lemma 3 with $T_d \equiv \Delta \cup T_{d,r}$ (so that $s \equiv 2u$), $c_j \equiv 0\ \forall\ j \in \Delta$, $c_j \equiv {A_j}'w_r\ \forall\ j \in T_{d,r}$ and $\check{s} \equiv u$. Call the exceptional set $T_{d,r+1}$. Lemma 3 can be applied because $\delta_{2u} + \delta_k + \theta_{k,2u}^2 < 1$ (condition 1 of the theorem). From Lemma 3, there exists a w_{r+1} and an

المنارة للاستشارات

www.manaraa.com

exceptional set $T_{d,r+1}$, disjoint with $T \cup \Delta \cup T_{d,r}$, of size less than $\check{s} = u$, s.t.

$$\begin{aligned}
{A_j}'w_{r+1} &= 0 \, \forall \, j \in T \\
{A_j}'w_{r+1} &= 0, \, \forall \, j \in \Delta \\
{A_j}'w_{r+1} &= {A_j}'w_r, \, \forall \, j \in T_{d,r} \\
|T_{d,r+1}| &< u \\
\|{A_{T_{d,r+1}}}'w_{r+1}\|_2 &\leq a_k(2u,u)\|{A_{T_{d,r}}}'w_r\|_2 \\
|{A_j}'w_{r+1}| &\leq \frac{a_k(2u,u)}{\sqrt{u}}\|{A_{T_{d,r}}}'w_r\|_2 \\
& \quad \forall j \notin T \cup \Delta \cup T_{d,r} \cup T_{d,r+1} \\
\|w_{r+1}\|_2 &\leq K_k(2u)\|{A_{T_{d,r}}}'w_r\|_2
\end{aligned} \tag{A.22}$$

Notice that $|T_{d,1}| < u$ (at iteration zero) and $|T_{d,r+1}| < u$ (at iteration r) ensures that $|\Delta \cup T_{d,r}| < s = 2u$ for all $r \geq 1$.

The last three equations of (A.22), combined with the sixth equation of (A.21), simplify to

$$\begin{aligned}
\|{A_{T_{d,r+1}}}'w_{r+1}\|_2 &\leq a_k(2u,u)^r a_{k_b}(u,u)\sqrt{u} \\
|{A_j}'w_{r+1}| &\leq a_k(2u,u)^r a_{k_b}(u,u), \\
& \quad \forall j \notin T \cup \Delta \cup T_{d,r} \cup T_{d,r+1}
\end{aligned} \tag{A.23}$$

$$\|w_{r+1}\|_2 \leq K_k(2u)a_k(2u,u)^{r-1}a_{k_b}(u,u)\sqrt{u} \tag{A.24}$$

We can define

$$w \triangleq \sum_{r=1}^{\infty}(-1)^{r-1}w_r \tag{A.25}$$

Since $a_k(2u,u) < 1$, $\|w_r\|_2$ approaches zero with r, and so the above summation is absolutely convergent, i.e. w is well-defined.

www.manaraa.com

From the first four equations of (A.21) and first two equations of (A.22),

$$\begin{aligned} A_j{}'w &> 0,\ \forall\ j \in T_{\text{a+g}} \\ A_j{}'w &< 0,\ \forall\ j \in T_{\text{a-g}} \\ A_j{}'w &= 0,\ \forall\ j \in T_b \\ A_j{}'w &= A_j{}'w_1 = \text{sgn}(x_j),\ \forall\ j \in \Delta \end{aligned} \tag{A.26}$$

Consider $A_j{}'w = A_j{}'\sum_{r=1}^{\infty}(-1)^{r-1}w_r$ for some $j \notin T \cup \Delta$. If for a given r, $j \in T_{d,r}$, then $A_j{}'w_r = A_j{}'w_{r+1}$ (gets canceled by the $r+1^{th}$ term). If $j \in T_{d,r-1}$, then $A_j{}'w_r = A_j{}'w_{r-1}$ (gets canceled by the $r-1^{th}$ term). Since $T_{d,r}$ and $T_{d,r-1}$ are disjoint, j cannot belong to both of them. Thus,

$$A_j{}'w = \sum_{r:j \notin T_{d,r} \cup T_{d,r-1}} (-1)^{r-1} A_j{}'w_r,\ \forall j \notin T \cup \Delta \tag{A.27}$$

Consider a given r in the above summation. Since $j \notin T_{d,r} \cup T_{d,r-1} \cup T \cup \Delta$, we can use (A.23) to get $|A_j{}'w_r| \leq a_k(2u,u)^{r-1} a_{k_b}(u,u)$. Thus, for all $j \notin T \cup \Delta$,

$$\begin{aligned} |A_j{}'w| &\leq \sum_{r:j \notin T_{d,r} \cup T_{d,r-1}} a_k(2u,u)^{r-1} a_{k_b}(u,u) \\ &\leq \frac{a_{k_b}(u,u)}{1 - a_k(2u,u)} \end{aligned} \tag{A.28}$$

Since $a_k(2u,u) + a_{k_b}(u,u) < 1$ (condition 2 of the theorem),

$$|A_j{}'w| < 1,\ \forall j \notin T \cup \Delta \tag{A.29}$$

Thus, from (A.26) and (A.29), we have found a w that satisfies the conditions of Lemma 1. From condition 1 of the theorem, $\delta_{k+u} < 1$. Applying Lemma 1, the claim follows.

A.5.1 Proof of Lemma 5

Let Δ_1 denote the set of indices of h with the $|\Delta|$ largest values outside of $T \cup \Delta$ and Δ_2 denote the indices of the next $|\Delta|$ largest values and so on. We bound the error in 3 parts: h_T, $h_{\Delta \cup \Delta_1}$ and $h_{(T \cup \Delta \cup \Delta_1)^c}$ and we can obtain the following theorem. First, we bound $\|h_T\|_2$ by using our second constraint. Since x and $\hat{x}$ are both feasible, so

$$\|h_T\|_2 \leq \|x_T - \mu_T\|_2 + \|\hat{x}_T - \mu_T\|_2 \leq 2\rho\sqrt{k} \tag{A.30}$$

المنارة للاستشارات
www.manaraa.com

Next, we bound $\|h_{(T\cup\Delta\cup\Delta_1)^c}\|_2$.

$$\|h_{(T\cup\Delta\cup\Delta_1)^c}\|_2 \leq \sum_{j\geq 2}\|h_{\Delta_j}\|_2 \leq \frac{1}{\sqrt{u}}\|h_{(T\cup\Delta)^c}\|_1 \tag{A.31}$$

Since $\hat{x} = x + h$ is the minimizer of (2.14) and since both x and $\hat{x}$ are feasible,

$$\|x_{T^c}\|_1 \geq \|(x+h)_{T^c}\|_1 \geq \|x_\Delta\|_1 - \|h_\Delta\|_1 + \|h_{(T\cup\Delta)^c}\|_1 - \|x_{(T\cup\Delta)^c}\|_1 \tag{A.32}$$

and since $x_{(T\cup\Delta)^c} = 0$ then

$$\|h_{(T\cup\Delta)^c}\|_1 \leq \|h_\Delta\|_1 \tag{A.33}$$

Combining this with (A.31), and using $\frac{\|h_\Delta\|_1}{\sqrt{u}} \leq \|h_\Delta\|_2$, we get

$$\|h_{(T\cup\Delta\cup\Delta_1)^c}\|_2 \quad\leq\quad \sum_{j\geq 2}\|h_{\Delta_j}\|_2 \leq \|h_\Delta\|_2$$

Next, since both x and $\hat{x}$ are feasible,

$$Ah = A(\hat{x} - x) = 0 \tag{A.34}$$

To upper bound $\|h_{\Delta\cup\Delta_1}\|_2$, use RIP to get

$$(1-\delta_{2u})\|h_{\Delta\cup\Delta_1}\|_2^2 \leq \|Ah_{\Delta\cup\Delta_1}\|_2^2 \tag{A.35}$$

To bound the right hand side of the above, notice that $Ah_{\Delta\cup\Delta_1} = Ah - \sum_{j\geq 2} Ah_{\Delta_j} - Ah_T$ and thus

$$\|Ah_{\Delta\cup\Delta_1}\|_2^2 =< Ah_{\Delta\cup\Delta_1}, Ah > - \sum_{j\geq 2} < Ah_{\Delta\cup\Delta_1}, Ah_{\Delta_j} > - < Ah_{\Delta\cup\Delta_1}, Ah_T > \tag{A.36}$$

Using (A.34),

$$| < Ah_{\Delta\cup\Delta_1}, Ah > | = 0 \tag{A.37}$$

Using RIP and (A.34),

$$\begin{aligned} |\sum_{j\geq 2} < Ah_{\Delta\cup\Delta_1}, Ah_{\Delta_j} > | &\leq |\sum_{j\geq 2} < Ah_\Delta, Ah_{\Delta_j} > | + |\sum_{j\geq 2} < Ah_{\Delta_1}, Ah_{\Delta_j} > | \\ &\leq \sqrt{2}\delta_{2u}\|h_{\Delta\cup\Delta_1}\|_2\|h_\Delta\|_2 \end{aligned} \tag{A.38}$$

المنارة للاستشارات

www.manaraa.com

Finally, using RIP and (A.30),

$$\begin{aligned} | < Ah_{\Delta\cup\Delta_1}, Ah_T > | &\leq \delta_{k+2u} \|h_{\Delta\cup\Delta_1}\|_2 \|h_T\|_2 && \text{(A.39)} \\ &\leq \delta_{k+2u} \|h_{\Delta\cup\Delta_1}\|_2 2\rho\sqrt{k} && \text{(A.40)} \end{aligned}$$

Combining the above 5 equations, we get

$$(1-\delta_{2u})\|h_{\Delta\cup\Delta_1}\|_2 \leq 2\delta_{k+2u}\rho\sqrt{k} + \sqrt{2}\delta_{2u}\|h_\Delta\|_2 \tag{A.41}$$

Using $\|h_\Delta\|_2 \leq \|h_{\Delta\cup\Delta_1}\|_2$ and simplifying,

$$\|h_{\Delta\cup\Delta_1}\|_2 \leq \frac{2\sqrt{k}\delta_{k+2u}}{1-(\sqrt{2}+1)\delta_{2u}}\rho \tag{A.42}$$

Combining with (A.34) and (A.30), we get

$$\|h\|_2 \leq \|h_{\Delta\cup\Delta_1}\|_2 + \|h_{(T\cup\Delta\cup\Delta_1)^c}\|_2 + \|h_T\|_2 \leq 2\|h_{\Delta\cup\Delta_1}\|_2 + 2\rho \leq B_1 \tag{A.43}$$

A.6 Causal MAP Interpretation of Dynamic RegModCS

The solution of (2.22) becomes a causal MAP estimate under the following assumptions. Let $p(X|Y)$ denote the conditional PDF of X of given Y and let $\delta(X)$ denote the Dirac delta function. Assume that

1. the random processes $\{x_t\}, \{y_t\}$ satisfy the hidden Markov model property; $p(y_t|x_t) = \delta(y_t - Ax_t)$ (re-statement of the observation model); and

   $$\begin{aligned} & p(x_t|x_{t-1}) = p((x_t)_{N_{t-1}}|x_{t-1})p((x_t)_{N^c_{t-1}}|x_{t-1}), \text{where} \\ & p((x_t)_{N_{t-1}}|x_{t-1}) = \mathcal{N}((x_t)_{N_{t-1}}; (x_{t-1})_{N_{t-1}}, \sigma_p^2 I) \\ & p((x_t)_{N^c_{t-1}}|x_{t-1}) = \left(\frac{1}{2\lambda_p}\right)^{|N^c_{t-1}|} \exp\left(-\frac{\|(x_t)_{N^c_{t-1}}\|_1}{\lambda_p}\right) \end{aligned}$$

   i.e. given x_{t-1} (and hence given N_{t-1}), $(x_t)_{N_{t-1}}$ and $(x_t)_{N^c_{t-1}}$ are conditionally independent; $(x_t)_{N_{t-1}}$ is Gaussian with mean $(x_{t-1})_{N_{t-1}}$ while $(x_t)_{N^c_{t-1}}$ is zero mean Laplace.

2. x_{t-1} is perfectly estimated from $y_0, y_1, \ldots y_{t-1}$, and

   $$p(x_{t-1}|y_0, \ldots y_{t-1}) = \delta\left(x_{t-1} - \begin{bmatrix} (\hat{x}_{t-1})_{\hat{N}_{t-1}} \\ 0_{\hat{N}^c_{t-1}} \end{bmatrix}\right)$$

www.manaraa.com

3. $\hat{x}_t$ is the solution of (2.22) with $\gamma = \frac{\lambda_p}{2\sigma_p^2}$.

If the first two assumptions above hold, it is easy to see that the "causal posterior" at time t, $p(x_t|y_1, \ldots y_t)$, satisfies

$$p(x_t|y_1, \ldots y_t) = C\delta(y_t - Ax_t)e^{-\frac{\|(x_t)_T - (\hat{x}_{t-1})_T\|_2^2}{2\sigma_p^2}} e^{-\frac{\|(x_t)_{T^c}\|_1}{\lambda_p}}$$

where $T := \hat{N}_{t-1}$ and C is the normalizing constant. If the last assumption also holds, then clearly the solution of (2.22) is a maximizer of $p(x_t|y_1, \ldots y_t)$, i.e. it is a causal MAP solution.

The MLE of λ_p, σ_p^2 can be computed from a training time sequence of signals, $\tilde{x}_0, \tilde{x}_1, \tilde{x}_2, \ldots \tilde{x}_{t_{\max}}$ as follows. Denote their supports ($\beta\%$-energy supports in case of compressible signal sequences) by $\tilde{N}_0, \tilde{N}_1, \ldots \tilde{N}_{t_{\max}}$. Then the MLE is

$$\begin{aligned} \hat{\lambda}_p &= \frac{\sum_{t=1}^{t_{\max}} \|(\tilde{x}_t)_{\tilde{N}_{t-1}^c}\|_1}{\sum_{t=1}^{t_{\max}} |\tilde{N}_{t-1}^c|}, \\ \hat{\sigma_p^2} &= \frac{\sum_{t=1}^{t_{\max}} \|(\tilde{x}_t - \tilde{x}_{t-1})_{\tilde{N}_{t-1}}\|_2^2}{\sum_{t=1}^{t_{\max}} |\tilde{N}_{t-1}|} \end{aligned} \tag{A.44}$$

المنارة للاستشارات

www.manaraa.com

APPENDIX B. Appendix for the Proof of Reconstruction Error Bound of Regularized Modified-BPDN

B.1 Proof of Proposition 1

When $\lambda = 0$, $Q_{T,0}(S) = A_{T\cup S}{}'A_{T\cup S}$. Thus, $Q_{T,\lambda}(S)$ is invertible iff $A_{T\cup S}$ is full rank. When $\lambda > 0$, $Q_{T,\lambda}(S)$ is as defined in (3.15). Apply block matrix inversion lemma

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} (\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & \mathbf{D}^{-1}+\mathbf{D}^{-1}\mathbf{C}(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}\mathbf{B}\mathbf{D}^{-1} \end{bmatrix}$$

with $\mathbf{A} = A_T{}'A_T + \lambda I_T$, $\mathbf{B} = A_T{}'A_S$, $\mathbf{C} = A_S{}'A_T$ and $\mathbf{D} = A_S{}'A_S$, clearly $Q_{T,\lambda}(S)$ is invertible iff $A_S{}'A_S$ and $A_T{}'RA_T + \lambda I_T$ are invertible where $R := [I - A_S(A_S{}'A_S)^{-1}A'_S]$. When A_S is full rank, (i) $A_S{}'A_S$ is full rank; and (ii) R is a projection matrix. Thus $R = R'R$ and so $A_T{}'RA_T = (RA_T)'(RA_T)$ is positive semi-definite. As a result, $A_T{}'RA_T + \lambda I_T$ is positive definite and thus invertible. Hence, when A_S is invertible, $Q_{T,\lambda}(S)$ is also invertible.

B.2 Proof of Theorem 4

In this subsection, we give the three lemmas for the proof of Theorem 4. *To keep notation simple we remove the subscripts $_{T,\lambda}$ from $Q(\Delta)$, M, $P(\Delta)$, $d(\Delta)$, $c(\Delta)$, $ERC(\Delta)$ in this and other Appendices.*

Lemma 6 *Suppose that $Q(\Delta)$ is invertible, then*

$$\|d(\Delta) - c(\Delta)\|_2 \leq \gamma\sqrt{|\Delta|} \cdot f_1(\Delta) \tag{B.1}$$

Lemma 6 can be obtained by setting $\nabla L(b) = 0$ and then using block matrix inversion on $Q(\Delta)$. The proof of Lemma 6 is in Appendix B.4.1. Next, $\|c(\Delta) - x\|_2$ can be bounded using the following lemma.

المنارة للاستشارات
www.manaraa.com

Lemma 7 *Suppose that $Q(\Delta)$ is invertible. Then*

$$\|c(\Delta) - x\|_2 \leq \lambda f_2(\Delta)\|x_T - \hat{\mu}_T\|_2 + f_3(\Delta)\|w\|_2 \tag{B.2}$$

The proof of Lemma 7 is in Appendix B.4.2.

Lemma 8 *If $Q(\Delta)$ is invertible, $ERC(\Delta) > 0$, and $\gamma \geq \gamma^*(\Delta)$, then $L(b)$ has a unique minimizer which is equal to $d(\Delta)$.*

Lemma 8 can be obtained in a fashion similar to [2, 28]. Its proof is given in Appendix B.4.3.

Combining Lemmas 6, 7 and 8, and using the fact $\|d(\Delta) - x\|_2 \leq \|d(\Delta) - c(\Delta)\|_2 + \|c(\Delta) - x\|_2$, we get Theorem 4.

B.3 Proof of Theorem 5

The following lemma is needed for the proof of the corollaries leading to Theorem 5.

Lemma 9 *Suppose that $Q(\tilde{\Delta})$ is invertible. Then*

$$\|c(\tilde{\Delta}) - x\|_2 \leq \lambda f_2(\tilde{\Delta})\|x_T - \hat{\mu}_T\|_2 + f_3(\tilde{\Delta})\|w\|_2 + f_4(\tilde{\Delta})\|x_{\Delta\setminus\tilde{\Delta}}\|_2 \tag{B.3}$$

Since $c(\tilde{\Delta})$ is only supported on $T \cup \tilde{\Delta}$ and $y = A_{T\cup\tilde{\Delta}} x_{T\cup\tilde{\Delta}} + A_{\Delta\setminus\tilde{\Delta}} x_{\Delta\setminus\tilde{\Delta}} + w$, the last term of (B.3) can be obtained by separating $x_{\Delta\setminus\tilde{\Delta}}$ out. The proof of Lemma 9 is given in Appendix B.4.4.

Using Lemma 9, we can obtain Corollary 3 and then Corollary 4. Then minimize over all allowed $\tilde{\Delta}$'s in Corollary 3, we get Theorem 5. The proof of Corollary 3 and 4 are given as follows.

B.3.1 Proof of Corollary 3

Notice from the proof of Lemma 6 and Lemma 8 that nothing in the result changes if we replace Δ by a $\tilde{\Delta} \subseteq \Delta$. By Lemma 6 for $\tilde{\Delta}$, we are able to bound $\|d(\tilde{\Delta}) - c(\tilde{\Delta})\|_2$. Hence, we get the first term of (3.25). Next, invoke Lemma 9 to bound $\|c(\tilde{\Delta}) - x\|_2$ and we can obtain the rest three terms of (3.25). Lemma 8 for $\tilde{\Delta}$ gives the sufficient conditions under which $d(\tilde{\Delta})$ is the unique unconstrained minimizer of $L(b)$.

المنارة للاستشارات

www.manaraa.com

B.3.2 Proof of Corollary 4

Corollary 4 is obtained by bounding $\gamma^*(\tilde\Delta)$. $\gamma^*(\tilde\Delta) = \|A_{(T\cup\tilde\Delta)^c}{}'(y - Ac(\tilde\Delta))\|_\infty / ERC(\tilde\Delta)$ can be bounded by rewriting $y - Ac(\tilde\Delta) = A_{T\cup\Delta}(x_{T\cup\Delta} - (c(\tilde\Delta))_{T\cup\Delta}) + w$ and then bounding $\|x_{T\cup\Delta} - (c(\tilde\Delta))_{T\cup\Delta}\|_2 = \|x - c(\tilde\Delta)\|_2$ using Lemma 9. Doing this, we get

$$\begin{aligned}
&\|A_{(T\cup\tilde\Delta)^c}{}'(y - Ac(\tilde\Delta))\|_\infty \\
&\le \max_{i\notin T\cup\tilde\Delta} \ |A_i{}'A_{T\cup\Delta}(x_{T\cup\Delta} - (c(\tilde\Delta))_{T\cup\Delta})| + |A_i{}'w| \\
&\le \max_{i\notin T\cup\tilde\Delta} \ \|A_i{}'A_{T\cup\Delta}\|_2 \|x_{T\cup\Delta} - (c(\tilde\Delta))_{T\cup\Delta})\|_2 + |A_i{}'w| \\
&\le \text{maxcor}(\tilde\Delta)\lambda f_2(\tilde\Delta)\|x_T - \mu_T\|_2 + \text{maxcor}(\tilde\Delta) f_3(\tilde\Delta)\|w\|_2 \\
&+ \text{maxcor}(\tilde\Delta) f_4(\tilde\Delta)\|x_{\Delta\setminus\tilde\Delta}\|_2 + \|A_{(T\cup\tilde\Delta)^c}{}'w\|_\infty
\end{aligned}$$

Using the above inequality to bound $\gamma^*(\tilde\Delta)$ and replacing γ in $f(T, \lambda, \Delta, \tilde\Delta, \gamma)$, given in (3.25), by this bound, we can get (3.27).

B.4 Proof of Lemmas 6, 7, 8, 9

B.4.1 Proof of Lemma 6

We use the approach of [2, Lemma 3]. We can minimize the function $L(b)$ over all vectors supported on set $T \cup \Delta$ by minimizing:

$$F(b) = \frac{1}{2}\|y - A_{T\cup\Delta} b_{T\cup\Delta}\|_2^2 + \frac{1}{2}\lambda\|b_T - \hat\mu_T\|_2^2 + \gamma\|b_\Delta\|_1 \tag{B.4}$$

Since $Q(\Delta)$ is invertible, $F(b)$ is strictly convex as a function of $b_{T\cup\Delta}$. Then at the unique minimizer, $d(\Delta)$, $0 \in \nabla F(b)|_{b=d(\Delta)}$. Let $\partial\|b_{T^c}\|_1|_{b=d(\Delta)}$ denote the subgradient set of $\|b_{T^c}\|_1$ at $b = d(\Delta)$. Then clearly any ϕ in this set satisfies

$$\phi_T \ = \ 0 \tag{B.5}$$

$$\|\phi_{T^c}\|_\infty \ \le \ 1 \tag{B.6}$$

Now, $0 \in \nabla F(b)|_{b=d(\Delta)}$ implies that

$$(A_{T\cup\Delta}{}'A_{T\cup\Delta})[d(\Delta)]_{T\cup\Delta} - A_{T\cup\Delta}{}'y + \lambda \begin{bmatrix} [d(\Delta)]_T - \hat\mu_T \\ \mathbf{0}_\Delta \end{bmatrix} + \gamma\phi_{T\cup\Delta} = 0 \tag{B.7}$$

المنارة للاستشارات

www.manaraa.com

Simplifying the above equation, we get

$$[d(\Delta)]_{T\cup\Delta} = Q(\Delta)^{-1}(A_{T\cup\Delta}{}'y + \lambda \begin{bmatrix} \hat{\mu}_T \\ \mathbf{0}_\Delta \end{bmatrix} - \gamma\phi_{T\cup\Delta}) \tag{B.8}$$

Therefore, using (B.5) and (3.20), we have

$$[c(\Delta)]_{T\cup\Delta} - [d(\Delta)]_{T\cup\Delta} = Q(\Delta)^{-1} \begin{bmatrix} \mathbf{0}_T \\ \gamma\phi_\Delta \end{bmatrix} \tag{B.9}$$

Since

$$Q(\Delta) = \begin{bmatrix} A_T{}'A_T + \lambda I_T & A_T{}'A_\Delta \\ A_\Delta{}'A_T & A_\Delta{}'A_\Delta \end{bmatrix}, \tag{B.10}$$

using the block matrix inversion lemma

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{bmatrix}$$

with $\mathbf{A} = A_T{}'A_T + \lambda I_T$, $\mathbf{B} = A_T{}'A_\Delta$, $\mathbf{C} = A_\Delta{}'A_T$ and $\mathbf{D} = A_\Delta{}'A_\Delta$ and using $\phi_T = 0$, we obtain

$$[c(\Delta)]_{T\cup\Delta} - [d(\Delta)]_{T\cup\Delta} = \begin{bmatrix} -\gamma(A_T{}'A_T + \lambda I_{|T|})^{-1}A_T A_\Delta(A_\Delta{}'MA_\Delta)^{-1}\phi_\Delta \\ \gamma(A_\Delta{}'MA_\Delta)^{-1}\phi_\Delta \end{bmatrix}$$

Since $\|\phi_\Delta\|_\infty \leq 1$, the bound of (B.1) follows.

B.4.2 Proof of Lemma 7

Recall $c(\Delta)$ is given in (3.20). Since both x and $c(\Delta)$ are zero outside $T \cup \Delta$, then $\|c(\Delta) - x\|_2 = \|[c(\Delta)]_{T\cup\Delta} - x_{T\cup\Delta}\|_2$. With $y = Ax + w$ and $Ax = A_{T\cup\Delta}x_{T\cup\Delta}$, we have

$$A_{T\cup\Delta}{}'y = A_{T\cup\Delta}{}'(A_{T\cup\Delta}x_{T\cup\Delta} + w) \tag{B.11}$$

Notice $A'_{T\cup\Delta}A_{T\cup\Delta} = Q(\Delta) - \lambda \begin{bmatrix} I_T & \mathbf{0}_{T,S} \\ \mathbf{0}_{S,T} & \mathbf{0}_{S,S} \end{bmatrix}$. Using (B.11), we obtain the following equation

$$A_{T\cup\Delta}{}'y = Q(\Delta)x_{T\cup\Delta} - \lambda \begin{bmatrix} x_T \\ \mathbf{0}_\Delta \end{bmatrix} + A_{T\cup\Delta}{}'w \tag{B.12}$$

المنارة للاستشارات

www.manaraa.com

Then, using (3.20) we can obtain

$$[c(\Delta)]_{T\cup\Delta} - x_{T\cup\Delta} = \lambda Q(\Delta)^{-1}\begin{bmatrix} \hat{\mu}_T - x_T \\ \mathbf{0}_{\Delta} \end{bmatrix} + Q(\Delta)^{-1}A_{T\cup\Delta}{}'w$$

Finally, this gives (B.2).

B.4.3 Proof of Lemma 8

The proof is similar to that in [2] and [28]. Recall that $d(\Delta)$ minimizes the function $L(b)$ over all b supported on $T \cup \Delta$. We need to show that if $\gamma \geq \gamma^*(\Delta)$, then $d(\Delta)$ is the unique global minimizer of $L(b)$.

The idea is to prove under the given condition, any small perturbation h on $d(\Delta)$ will increase function $L(d(\Delta))$,i.e. $L(d(\Delta)+h) - L(d(\Delta)) > 0, \forall \|h\|_\infty \leq \epsilon$ for ϵ small enough. Then since $L(b)$ is a convex function, $d(\Delta)$ will be the unique global minimizer[2].

Similar to [28], we first split the perturbation into two parts $h = u + v$ where u is supported on $T \cup \Delta$ and v is supported on $(T \cup \Delta)^c$. Clearly $\|u\|_\infty \leq \|h\|_\infty \leq \epsilon$. We consider the case $v \neq 0$ since the case $v = 0$ is already covered in Lemma 1. Then

$$L(d(\Delta)+h) = \frac{1}{2}\|y - A(d(\Delta)+u) - Av\|_2^2 +$$
$$\frac{1}{2}\lambda\|[d(\Delta)]_T + u_T + v_T - \hat{\mu}_T\|_2^2 + \gamma\|(d(\Delta)+u)_{T^c} + v_{T^c}\|_1$$

Then, we can obtain

$$L(d(\Delta)+h) - L(d(\Delta)) = L(d(\Delta)+u) - L(d(\Delta))$$
$$+\frac{1}{2}\|Av\|_2^2 - \langle y - Ad(\Delta), Av\rangle + \langle Au, Av\rangle + \gamma\|v_{T^c}\|_1$$

Since $d(\Delta)$ minimizes $L(b)$ over all vectors supported on $T \cup \Delta$, $L(d(\Delta)+u) - L(d(\Delta)) \geq 0$. Then since $L(d(\Delta)+u) - L(d(\Delta)) \geq 0$ and $\|Av\|_2^2 \geq 0$, we need to prove that the rest are positive,i.e.,$\gamma\|v_{T^c}\|_1 - \langle y - Ad(\Delta), Av\rangle + \langle Au, Av\rangle \geq 0$. Instead, we can prove this by proving a stronger condition $\gamma\|v_{T^c}\|_1 - |\langle y - Ad(\Delta), Av\rangle| - |\langle Au, Av\rangle| \geq 0$. Since $\langle y - Ad(\Delta), Av\rangle = v'A'(y - Ad(\Delta))$ and v is supported on $(T \cup \Delta)^c$,

$$\begin{aligned} |\langle y - Ad(\Delta), Av\rangle| &= |v_{(T\cup\Delta)^c}{}'A_{(T\cup\Delta)^c}{}'(y - Ad(\Delta))| \\ &\leq \|v\|_1\|A_{(T\cup\Delta)^c}{}'(y - Ad(\Delta))\|_\infty \end{aligned}$$

www.manaraa.com

Thus,

$$|\langle y - Ad(\Delta), Av\rangle| \le \max_{\omega \notin T\cup\Delta} |\langle y - Ad(\Delta), A_\omega\rangle| \|v\|_1$$

Meanwhile,

$$|\langle Au, Av\rangle| \le \|A'Au\|_\infty \|v\|_1 \le \epsilon \|A'A\|_\infty \|v\|_1 \tag{B.13}$$

And $\|v\|_1 = \|v_{T^c}\|_1$ since v is supported on $(T \cup \Delta)^c \subseteq T^c$. Then what we need to prove is

$$\big[\gamma - \max_{\omega \notin T\cup\Delta} |\langle y - Ad(\Delta), A_\omega\rangle| - \epsilon \|A'A\|_\infty\big] \|v\|_1 > 0 \tag{B.14}$$

Since we can select $\epsilon > 0$ as small as possible, then we just need to show

$$\gamma - \max_{\omega \notin T\cup\Delta} |\langle y - Ad(\Delta), A_\omega\rangle| > 0 \tag{B.15}$$

Since $y - Ad(\Delta) = (y - Ac(\Delta)) + A(c(\Delta) - d(\Delta))$, and by Lemma 1 we know $A(c(\Delta) - d(\Delta)) = \gamma M A_\Delta (A_\Delta{}' M A_\Delta)^{-1} \phi_\Delta$ and since $\|\phi_\Delta\|_\infty \le 1$, we conclude that $d(\Delta)$ is the unique global minimizer if

$$\|A_{(T\cup\Delta)^c}{}'(y - Ac(\Delta))\|_\infty < \gamma\big[1 - \max_{\omega \notin T\cup\Delta} \|P(\Delta) A_\Delta{}' M A_\omega\|_1\big] \tag{B.16}$$

Next, we will show that $d(\Delta)$ is also the unique global minimizer under the following condition

$$\|A_{(T\cup\Delta)^c}{}'(y - Ac_{T,\lambda}(\Delta))\|_\infty = \gamma\big[1 - \max_{\omega \notin T\cup\Delta} \|P(\Delta) A_\Delta{}' M A_\omega\|_1\big] \tag{B.17}$$

Since the perturbation $h \neq 0$, then $u \neq 0$ or $v \neq 0$. Therefore, we will discuss the following three cases.

1. $u \neq 0$. In this case, we know $L(d(\Delta) + u) - L(d(\Delta)) > 0$ since $d(\Delta)$ is the unique minimizer over all vectors supported on $T \cup \Delta$. Therefore, $L(d(\Delta) + h) - L(d(\Delta)) > 0$ if (B.17) holds.

2. $u = 0$, $v \neq 0$ and v is not in the null space of A, i.e., $Av \neq 0$. In this case, we know $\|Av\|_2^2 > 0$. Hence, $L(d(\Delta) + h) - L(d(\Delta)) > 0$ when (B.17) holds.

3. $u = 0$, $v \neq 0$ and $Av = 0$. In this case, $L(d(\Delta) + h) - L(d(\Delta)) = \gamma \|v_{T^c}\|_1$. Thus, $L(d(\Delta) + h) - L(d(\Delta)) > 0$ if $\gamma > 0$. Clearly, $L(d(\Delta) + h) - L(d(\Delta)) > 0$ when (B.17) holds.

Finally, combining (B.16) and (B.17), we can conclude that $d(\Delta)$ is the unique global minimizer if the following condition holds

$$\|A_{(T\cup\Delta)^c}{}'(y - Ac(\Delta))\|_\infty \le \gamma \mathrm{ERC}(\Delta) \tag{B.18}$$

المنارة للاستشارات

www.manaraa.com

B.4.4 Proof of Lemma 9

Consider a $\tilde{\Delta} \subseteq \Delta$ such that $A_{\tilde{\Delta}}$ has full rank. Since $A_{T\cup\tilde{\Delta}}'y = A_{T\cup\tilde{\Delta}}'(A_{T\cup\tilde{\Delta}}x_{T\cup\tilde{\Delta}} + w + A_{\Delta\setminus\tilde{\Delta}}x_{\Delta\setminus\tilde{\Delta}})$, expanding these terms we have

$$A_{T\cup\tilde{\Delta}}'y = Q(\Delta)x_{T\cup\tilde{\Delta}} - \lambda \begin{bmatrix} x_T \\ \mathbf{0}_{\tilde{\Delta}} \end{bmatrix} + A_{T\cup\tilde{\Delta}}'w + A_{T\cup\tilde{\Delta}}'A_{\Delta\setminus\tilde{\Delta}}x_{\Delta\setminus\tilde{\Delta}} \tag{B.19}$$

Then, using this in the expression for $c(\tilde{\Delta})$ from (3.20), we get

$$\begin{aligned}[c(\tilde{\Delta})]_{T\cup\Delta} - x_{T\cup\Delta} = &\begin{bmatrix} \lambda Q(\tilde{\Delta})^{-1} \begin{bmatrix} \hat{\mu}_T - x_T \\ \mathbf{0}_{\tilde{\Delta}} \end{bmatrix} \\ \mathbf{0}_{\Delta\setminus\tilde{\Delta}} \end{bmatrix} \\ &+ \begin{bmatrix} Q(\tilde{\Delta})^{-1}A_{T\cup\tilde{\Delta}}'w \\ \mathbf{0}_{\Delta\setminus\tilde{\Delta}} \end{bmatrix} + \begin{bmatrix} Q(\tilde{\Delta})^{-1}A_{T\cup\tilde{\Delta}}'A_{\Delta\setminus\tilde{\Delta}}x_{\Delta\setminus\tilde{\Delta}} \\ -x_{\Delta\setminus\tilde{\Delta}} \end{bmatrix}\end{aligned} \tag{B.20}$$

Therefore, we get (B.3).

B.5 Sufficient Conditions' Comparison using RIC and ROC

We briefly compare the results for reg-mod-BPDN, mod-BPDN and BPDN, primarily by comparing the sufficient conditions required for them to hold. The comparison of the bounds is not easy since each holds under a different set of sufficient conditions. This will be done later using the results of Section IV which hold without any sufficient conditions. For the comparison of sufficient conditions, we use the restricted isometry constant (RIC), δ_S and restricted orthogonality constant (ROC), $\theta_{S,S'}$ [18] defined next. These depend only on the sizes of the sets T, Δ and N and hence make a theoretical comparison easier. However the comparison can only be qualitative. The RIC and ROC are not computable (computation complexity is exponential in the set size) and hence cannot be used for numerical comparisons. On the other hand, the ERC and the bounds obtained based on the ERC approach are computable and can be used for a quantitative numerical comparison.

Consider mod-BPDN versus BPDN first. Let us compare their ERC's. Using the facts that $\|A_T'A_\Delta\|_2 \leq \theta_{|T|,|\Delta|}$, $\|(A_T'A_T + \lambda I_T)^{-1}\|_2 \leq 1/(1 - \delta_{|T|} + \lambda)$ and the fact that for a vector z of length l,

المنارة للاستشارات
www.manaraa.com

$\|z\|_1 \leq \sqrt{l}\|z\|_2$,

$$\begin{aligned} ERC_{T,\lambda}(\Delta) &\geq 1 - \sqrt{|\Delta|}\|P_{T,\lambda}(\Delta)\|_2\|A_\Delta{}'M_{T,\lambda}A_\omega\|_2 \\ &\geq 1 - \sqrt{|\Delta|}\frac{(\theta_{|\Delta|,1} + \frac{\theta_{|\Delta|,|T|}\theta_{|\Delta|,1}}{1-\delta_{|T|}+\lambda})}{1 - \delta_{|\Delta|} - \frac{\theta^2_{|T|,|\Delta|}}{1-\delta_{|T|}+\lambda}} \end{aligned} \tag{B.21}$$

where the numerator of the second term comes from bounding $\|A_\Delta{}'M_{T,\lambda}A_\omega\|_2$ and the denominator of the second term comes from bounding $\|P_{T,\lambda}(\Delta)\|_2$. In practice, for example in recursive reconstruction applications like real-time dynamic MRI, usually $|\Delta| \approx |\Delta_e| \ll |N|$ and $|N| \approx |T| \approx |T \cup \Delta|$ [40]. Under this assumption, when fewer measurements are available (but still enough to ensure that $\delta_{|N|} < 1$), the denominator for the second term of $ERC_{\emptyset,0}(N)$ (BPDN), $1 - \delta_{|N|}$, will be smaller than that of $ERC_{T,0}(\Delta)$ (mod-BPDN), $1 - \delta_{|\Delta|} - \frac{\theta^2_{|T|,|\Delta|}}{1-\delta_{|T|}}$. Also, $\sqrt{|N|}$ in its numerator will be larger than $\sqrt{|\Delta|}$ for mod-BPDN, while the other numerator terms will be similar in both cases. This can result in a smaller (and possibly negative) lower bound on the ERC for BPDN.

To compare reg-mod-BPDN and mod-BPDN, notice that mod-BPDN needs $A_{T\cup\Delta}$ to be full rank where as reg-mod-BPDN only needs A_Δ to be full rank which is much weaker.

We show a numerical comparison in Table 3.1 (simulation details given in Chapter 5.4). Notice that BPDN needs 90% measurements for its ERC to become positive where as mod-BPDN only needs 19%. Moreover even with 90% measurements, its ERC is just positive and very small. As a result its error bound is large (27% normalized mean squared error (NMSE)). Similarly notice that mod-BPDN needs $n \geq 19\%$ while for reg-mod-BPDN $n = 13\%$ also suffices.

Remark 8 *A sufficient conditions' comparison only provides a comparison of when a given result can be applied to provide a bound on the reconstruction error. For example, in simulations, of course BPDN provides a good reconstruction using much lesser than 90% measurements. However, when* $n < 90\%$ *we cannot bound its reconstruction error using Theorem 4 above (for BPDN this is the same as the result of [2]). We address this issue in the next section.*

B.5.1 Equivalence between Theorem 5 and Theorem 6 bounds

We can use the weak law of large numbers (WLLN) to argue that as n, $s \triangleq |N|$ approach to infinity the bound from Theorem 6 converges to that of Theorem 5 in probability. We give the basic idea here.

المنارة للاستشارات
www.manaraa.com

The complete proof will be in future work. The WLLN argument applies when

- Each element of A is iid with zero mean and variance $1/n$, i.e. $A = \frac{1}{\sqrt{n}} Z$ where each element of Z is iid with zero mean and unit variance.
- The noise w is bounded in ℓ_2 norm, i.e. $\|w\|_2 \leq \eta$ and
- $n, s \to \infty$

WLLN can be used to argue that as $n, s \to \infty$, with high probability (w.h.p.), $ERC(\tilde{\Delta})$ and the multipliers g_1, g_2, g_3, g_4 depend only on the size, k, of the set $\tilde{\Delta}$, i.e. they are the same for all sets $\tilde{\Delta}$ of a given size. Thus, the only term in $g(\tilde{\Delta})$ that varies for different sets $\tilde{\Delta} \in \mathcal{G}_k$ is $\|x_{\Delta \setminus \tilde{\Delta}}\|_2$. Thus $\arg\min_{\mathcal{G}_k} g(\tilde{\Delta}) = \arg\min_{\mathcal{G}_k} \|x_{\Delta \setminus \tilde{\Delta}}\|_2$. Since ERC also only depends on k, for a given k, either $ERC(k) > 0$ or $ERC(k) < 0$. When $ERC(k) > 0$, $\mathcal{G}_k = \{\tilde{\Delta} \subseteq \Delta, |\tilde{\Delta}| = k\}$, where as when $ERC(k) < 0$, $\mathcal{G}_k$ is empty. The minimum value over an empty set is infinity. Thus, $\min_{\mathcal{G}_k} \|x_{\Delta \setminus \tilde{\Delta}}\|_2 = B_k$. Using (3.36), (3.32) and (3.35), this means that $g(\tilde{\Delta}^*) = g(\tilde{\Delta}^{**})$, i.e. the bounds from Theorems 5 and 6 are equal.

The WLLN argument is as follows. Note that all terms in g_1, g_2, g_3, g_4 and ERC that depend on $\tilde{\Delta}$ are functions of either $A_{\tilde{\Delta}}{}' A_{\tilde{\Delta}}$ or $A_T{}' A_{\tilde{\Delta}}$ or $A_{\tilde{\Delta}}{}' M_{T,\lambda} A_{\tilde{\Delta}}$ or $A'_{T \cup \tilde{\Delta}} w$ Consider $A_{\tilde{\Delta}}{}' A_{\tilde{\Delta}}$.

$$(A_{\tilde{\Delta}}{}' A_{\tilde{\Delta}})_{i,j} = \begin{cases} \sum_{r=1}^n A_{i,r}^2 = \frac{1}{n} \sum_{r=1}^n Z_{i,r}^2 & \text{if } i = j \\ \sum_{r=1}^n A_{i,r} A_{j,r} = \frac{1}{n} \sum_{r=1}^n Z_{i,r} Z_{j,r} & \text{if } i \neq j \end{cases}$$

Clearly $\mathbb{E}[Z_{i,r}^2] = 1$ and its variance, $Var[Z_{i,r}^2] = 3$ where as $\mathbb{E}[Z_{i,r} Z_{j,r}] = 0$ while $Var[Z_{i,r} Z_{j,r}] = 1$. Here, $\mathbb{E}[\cdot]$ and $Var[\cdot]$ denote the expectation and variance computed over the distribution of A. Thus by WLLN, as $n \to \infty$, $A_{\tilde{\Delta}}{}' A_{\tilde{\Delta}}$ approaches the identity matrix, I_k w.h.p.. A similar argument can be made for each element of $A_T{}' A_{\tilde{\Delta}}$ to show that this approaches the zero matrix as $n \to \infty$. A similar argument can also be made for $M_{T,\lambda}$ when $s := |N|$ (and hence $|T|$) goes to infinity to show that all its diagonal elements converge to one value and all the non-diagonal ones converge to another value. This fact can then be used to make a WLLN argument for each element of $A_{\tilde{\Delta}}{}' M_{T,\lambda} A_{\tilde{\Delta}}$. Now consider g_4 which contains the term $\|A_{(T \cup \Delta)^c}{}' w\|_\infty$. Notice that $(A_{(T \cup \Delta)^c}{}' w)_i = \sum_{j=1}^n w_j A_{j,i}$. Taking expectations only over the elements of A, $\mathbb{E}[(A_{(T \cup \Delta)^c}{}' w)_i] = 0$ and $Var[(A_{(T \cup \Delta)^c}{}' w)_i] = \sum_{j=1}^n w_j^2 \frac{1}{n} \leq \frac{\eta^2}{n}$. Thus

www.manaraa.com

by WLLN, each element of the vector $A_{(T\cup\Delta)^c}{}'w$ approaches zero, and hence its infinity norm also approaches zero w.h.p.. Thus, w.h.p., for a given size k, all these three matrices and $\|A_{(T\cup\Delta)^c}{}'w\|_\infty$, and as a result all of ERC, g_1, g_2, g_3, g_4, converge to a value that does not depend on the set $\tilde{\Delta}$.

المنارة للاستشارات

www.manaraa.com

BIBLIOGRAPHY

[1] S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Journal of Scientific Computation*, vol. 20, pp. 33 – 61, 1998.

[2] J. A. Tropp, “Just relax: Convex programming methods for identifying sparse signals in noise,” *IEEE Trans. on Information Theory*, vol. 52(3), pp. 1030 – 1051, March 2006.

[3] E. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Info. Th.*, vol. 52(2), pp. 489–509, February 2006.

[4] D. Donoho, “Compressed sensing,” *IEEE Trans. on Information Theory*, vol. 52(4), pp. 1289 – 1306, April 2006.

[5] “Rice compressive sensing resources [online],” in *Available: http://www.dsp.rice.edu/cs*.

[6] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” May 2009.

[7] J. Tropp and A. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. on Information Theory*, vol. 53(12), pp. 4655–4666, December 2007.

[8] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, “Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit,” *IEEE Trans. on Information Theory*, (Submitted) 2007.

[9] D. Needell and J. Tropp., “Cosamp: Iterative signal recovery from incomplete and inaccurate samples,” *Appl. Comp. Harmonic Anal.*, 2008.

المنارة للاستشارات

www.manaraa.com

[10] E. Candes and T. Tao, “The dantzig selector: statistical estimation when p is much larger than n,” vol. 35(6), pp. 2313 – 2351, September 2006.

[11] S. Foucart and M. J. Lai, “Sparsest solutions of underdetermined linear systems via ell-q-minimization for 0 ¡= q ¡= 1,” *Applied and Computational Harmonic Analysis*, vol. 26, pp. 395–407, 2009.

[12] D. Wipf and B. Rao, “Sparse bayesian learning for basis selection,” *IEEE Trans. on Signal Processing*, vol. 52, pp. 2153–2164, Aug 2004.

[13] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Trans. Signal Process.*, vol. 56(6), pp. 2346–2356, June 2008.

[14] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, “Compressive imaging for video representation and coding,” in *Proc. Picture Coding Symposium (PCS), Beijing, China*, April 2006.

[15] C. Qiu and N. Vaswani, “Reprocs: A missing link between recursive robust pca and recursive sparse recovery in large but correlated noise,” *arXiv: 1106.3286*, 2011.

[16] ——, “Support-predicted modified-cs for principal components’ pursuit,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2011.

[17] X. Liu and J. U. Kang, “Compressive sd-oct: the application of compressed sensing in spectral domain optical coherence tomography,” *Optics Express*, vol. 18, pp. 22 010 – 22 019, 2010.

[18] E. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. on Infomation Theory*, vol. 51(12), pp. 4203 – 4215, 2005.

[19] E. Candes, “The restricted isometry property and its implications for compressed sensing,” *Compte Rendus de lAcademie des Sciences, Paris, Serie I*, 2008.

[20] E. Candes, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59(8), pp. 1207–1223, August 2006.

المنارة للاستشارات

www.manaraa.com

[21] D. Angelosante and G. Giannakis, “Rls-weighted lasso for adaptive estimation of sparse signals,” in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2009.

[22] C. Rozell, D. Johnson, R. Baraniuk, and B. Olshausen, “Locally competitive algorithms for sparse approximation,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2007.

[23] M. Asif and J. Romberg, “Dynamic updating for sparse time varying signals,” in *Conference on Information Sciences and Systems (CISS)*, 2009.

[24] N. Vaswani and W. Lu, “Modified-cs: Modifying compressive sensing for problems with partially known support,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2009.

[25] ——, “Modified-cs: Modifying compressive sensing for problems with partially known support,” *IEEE Trans. on Sig. Proc.*, vol. 58(9), pp. 4595 – 4607, September 2010.

[26] A. Khajehnejad, W. Xu, A. Avestimehr, and B. Hassibi, “Weighted l1 minimization for sparse recovery with prior information,” in *IEEE Intl. Symp. Info. Theory(ISIT)*, 2009.

[27] R. von Borries, C. J. Miosso, and C. Potes, “Compressive sensing reconstruction with prior information by iteratively reweighted least-squares,” *IEEE Transactions on Signal Processing*, vol. 57(6), pp. 2424 – 2431, June 2009.

[28] W. Lu and N. Vaswani, “Modified basis pursuit denoising (modified-bpdn) for noisy compressive sensing with partially known support,” in *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, 2010.

[29] L. Jacques, “A short note on compressed sensing with partially known signal support,” *Signal Processing*, vol. 90(12), pp. 3308–3312, December 2010.

[30] V. Stankovic, L. Stankovic, and S. Cheng, “Compressive image sampling with side information,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2009.

[31] L. F. P. Rafael E. Carrillo and K. E. Barner, “Iterative algorithms for compressed sensing with partially known support,” in *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, 2010.

المنارة للاستشارات

www.manaraa.com

[32] D. Donoho and J. Tanner, “High-dimensional centrally symmetric polytopes with neighborliness proportional to dimension,” vol. 102(27), pp. 617–652, 2006.

[33] N. Vaswani, “Ls-cs-residual (ls-cs): Compressive sensing on the least squares residual,” *IEEE Trans. Signal Process.*, vol. 58(8), pp. 4108 – 4120, August 2010.

[34] ——, “Kalman filtered compressed sensing,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2008.

[35] M. Lustig, D. Donoho, and J. M. Pauly, “Sparse mri: The application of compressed sensing for rapid mr imaging,” *Magnetic Resonance in Medicine*, vol. 58(6), pp. 1182–1195, December 2007.

[36] U. Gamper, P. Boesiger, and S. Kozerke, “Compressed sensing in dynamic mri,” *Magnetic Resonance in Medicine*, vol. 59(2), pp. 365 – 373, January 2008.

[37] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, “k-t focuss: A general compressed sensing framework for high resolution dynamic mri,” *Magnetic Resonance in Medicine*, vol. 61, pp. 103 – 116, January 2009.

[38] J. C. Y. Hong Jung and E. Y. Kim, “Improved k-t blask and k-t sense using focuss,” *Phys. Med. Biol.*, vol. 52, pp. 3201 – 3226, 2007.

[39] I. Gorodnitsky and B. Rao, “Sparse signal reconstruction from limited data using focuss: A reweighted norm minimization algorithm,” *IEEE Trans. on Signal Processing*, vol. 45, pp. 600 – 616, March 1997.

[40] W. Lu and N. Vaswani, “Modified compressive sensing for real-time dynamic mr imaging,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2009.

[41] ——, “Exact reconstruction conditions and error bounds for regularized modified basis pursuit (reg-modified-bp),” in *44th Asilomar Conference on Signals, Systems and Computers*, 2010.

[42] B. K. Natarajan, “Sparse approximate solutions to linear systems,” *SIAM J. Comput.*, vol. 24, pp. 227–234, 1995.

[43] N. Vaswani, “Analyzing least squares and kalman filtered compressed sensing,” in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2009.

المنارة للاستشارات

www.manaraa.com

[44] V. Stankovic, L. Stankovic, and S. Cheng, “Compressive video sampling,” in *European Signal Processing Conf. (EUSIPCO), Lausanne, Switzerland,*, August 2008.

[45] J. Y. Park and M. B. Wakin, “A multiscale framework for compressive sensing of video,” in *Proc. Picture Coding Symposium (PCS), Chicago, Illinois*, May 2009.

[46] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[47] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer.

[48] V. Cevher, A. Sankaranarayanan, M. Duarte, D. Reddy, R. Baraniuk, and R. Chellappa, “Compressive sensing for background subtraction,” in *Eur. Conf. on Comp. Vis. (ECCV)*, 2008.

[49] E. Candes and J. Romberg, “L1 Magic Users Guide,” October 2005.

[50] C. Dossal, G. Peyre, and J. Fadili, “A numerical exploration of compressed sampling recovery,” in *Signal Processing with Adaptive Sparse Structured Representations (SPARS)*.

[51] C. Dossal, “A necessary and sufficient condition for exact recovery by l1 minimization,” in *Preprint*, 2007.

[52] W. Lu and N. Vaswani, “Regularized Modified BPDN for Noisy Sparse Reconstruction with Partial Erroneous Support and Signal Value Knowledge,” *to appear in IEEE Trans. Sig. Proc 2012, arXiv: 1002.0019, 2011*.

[53] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. on Information Theory*, vol. 56, pp. 1982–2001, April 2010.

[54] Y. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. on Information Theory*, vol. 55(11), pp. 5302 – 5316, November 2009.

[55] P. Schniter, L. Potter, and J. Ziniel, “Fast bayesian matching pursuit: Model uncertainty and parameter estimation for sparse linear models,” in *Information Theory and Applications (ITA)*, 2008.

المنارة للاستشارات

www.manaraa.com

[56] S. Som, L. C. Potter, and P. Schniter, “Compressive imaging using approximate message passing and a markov-tree prior,” in *Asilomar Conf. on Sig. Sys. Comp.*, 2010.

[57] C. La and M. Do, “Signal reconstruction using sparse tree representations,” in *SPIE Wavelets XI, San Diego, California*, September 2005.

[58] M. Duarte, M. Wakin, and R. Baraniuk, “Fast reconstruction of piecewise smooth signals from random projections,” in *SPARS Workshop*, November 2005.

[59] F. Raisali, “Stability (over time) of regularized modified-cs (noisy) for recursive causal sparse reconstruction,” in *45th Annual Conference on Information Sciences and Systems (CISS)*, 2011.

[60] C. Qiu and N. Vaswani, “Ls-cs-residual (ls-cs): Compressive sensing on the least squares residual,” in *http://home.engineering.iastate.edu/~chenlu/KFLS_v3*.

[61] A. I. Nemani A and T. KR, “Investigating the consistency of brain activation using individual trial analysis of high resolution fmri in the human primary visual cortex,” *NeuroImage*, vol. 42, pp. 1417–1424, 2009.

[62] D. Angelosante, E. Grossi, and G. B. Giannakis, “Compressed sensing of time-varying signals,” in *DSP*, 2009.

[63] C. R. Genovese, N. A. Lazar, and T. E. Nichols, “Thresholding of statistical maps in functional neurimaging using the false discovery rate,” *NeuroImage*, vol. 15, p. 870C878, 2002.

[64] I. C. Atkinson, D. L. J. Farzad Kamalabadi, and Thulborn, “Blind estimation for localized low contrast-to-noise ratio bold signals,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 350–364, 2008.

[65] E. Bullmore, C. Long, J. Suckling, J. Fadili, G. Calvert, F. Zelaya, T. A. Carpenter, and M. Brammer, “Colored noise and computational inference in neurophysiological (fmri) time series analysis: Resampling methods in time and wavelet domains,” *Hum. Brain Mapp.*, 61-78.

[66] W. Lu, T. Li, I. C. Atkinson, and N. Vaswani, “Modified-cs-residual for recursive reconstruction of highly undersampled functional mri sequences,” in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2011.

المنارة للاستشارات

www.manaraa.com

[67] N. Vaswani, “Stability (over time) of Modified-CS for Recursive Causal Sparse Reconstruction,” in *Allerton Conf. Communication, Control, and Computing*, 2010.

www.manaraa.com