Recursive Reconstruction of Sparse Signal Sequences

Namrata Vaswani, Wei Lu

Dept. of Electrical and Computer Engineering

Iowa State University, Ames, IA 50011, USA

Email: namrata@iastate.edu

Abstract

In this chapter, we describe our recent work on the design and analysis of recursive algorithms for causally reconstructing a time sequence of (approximately) sparse signals from a greatly reduced number of linear projection measurements. The signals are sparse in some transform domain referred to as the sparsity basis and their sparsity patterns (support set of the sparsity basis coefficients) can change with time. By "recursive", we mean use only the previous signal's estimate and the current measurements to get the current signal's estimate. We also briefly summarize our exact reconstruction results for the noise-free case and our error bounds and error stability results (conditions under which a time-invariant and small bound on the reconstruction error holds at all times) for the noisy case. Connections with related work are also discussed.

A key example application where the above problem occurs is dynamic magnetic resonance imaging (MRI) for real-time medical applications such as interventional radiology and MRI-guided surgery, or in functional MRI to track brain activation changes. Cross-sectional images of the brain, heart, larynx or other human organ images are piecewise smooth, and thus approximately sparse in the wavelet domain. In a time sequence, their sparsity pattern changes with time, but quite slowly. The same is also often true for the nonzero signal values. This simple fact, which was first observed in our work, is the key reason that our proposed recursive algorithms can achieve provably exact or accurate reconstruction from very few measurements.

I. Introduction

In this chapter, we describe our recent work on the design and analysis of recursive algorithms for causally reconstructing a time sequence of (approximately) sparse signals from a greatly reduced number of linear projection measurements. The signals are sparse in some transform domain referred to as the sparsity basis and their sparsity patterns (support set of the sparsity basis coefficients) can change with time. The most important example of the above problem occurs in dynamic magnetic resonance imaging (MRI) for real-time medical applications such as interventional radiology, MR image guided surgery, or functional MRI to track brain activation changes. MRI is a technique for cross-sectional imaging that sequentially captures the 2D Fourier projections of the cross-section to be reconstructed. Cross-sectional images of the brain, heart, larynx or other human organ images are usually piecewise smooth, e.g. see Fig. 1, and thus approximately sparse in the wavelet domain. In a time sequence, the sparsity pattern changes with time, but slowly. Often, the signal values also change gradually over time. We demonstrate this for a larynx and a cardiac MRI sequence in Fig. 1.

Since MR data acquisition is sequential, the ability to accurately reconstruct with fewer measurements directly translates to reduced scan times. Shorter scan times along with online (causal) and fast (recursive) reconstruction allow the possibility of real-time imaging of fast changing physiological phenomena. Other example applications where real-time imaging is needed include real-time single-pixel video imaging [1], real-time video compression/decompression, real-time sensor network based sensing of time-varying fields [2], or real-time extraction of the foreground image sequence (sparse image) from a slow changing background image sequence (well modeled as lying in a low-dimensional space [3]) using recursive projected compressive sensing (CS) [4], [5]. For other potential applications, see [6], [7].

Since the recent introduction of compressive sensing (CS) [8], [9], [10], the static sparse reconstruction problem has been thoroughly studied. However most existing algorithms for the dynamic problem just use CS to jointly reconstruct the entire time sequence in one go [11], [12], [13]. This is an offline and batch solution with very high complexity. The alternative - doing CS at each time separately (simple CS) - is online and fast but requires many more measurements. The question then is: *for a time sequence of sparse signals, how can we obtain a recursive solution that improves the accuracy of simple CS by using past observations, and does this while keeping the computational complexity only as much as that of simple CS (and thus much lower than that of the batch methods)? In particular, how can we use slow or correlated sparsity pattern change, and in certain cases also slow signal value change, to do this?* By *"recursive"*, we mean a solution that uses only the previous signal estimate and the current observation vector at the current time.

This problem was first studied in [14] which proposed a solution called Kalman Filtered Compressed Sensing (KF-CS). In later work, a simpler special case of KF-CS, called Least Squares CS-residual (LS-CS) was analyzed in detail [15]; and more powerful approaches such as Modified-CS [16], [17], Modified-CS-residual [18], [19] and regularized modified-CS [20], [21]

were introduced. Performance guarantees – exact recovery conditions in the noise-free case [16], [17], [21] and time-invariant error bounds (stability) in the noisy case [15], [22] – were also obtained. We describe all of these ideas in the next few sections. We first begin by providing a short background on sparse recovery and compressed sensing, followed by giving a formal problem definition for our problem and discussing related work.

II. NOTATION AND SPARSE RECOVERY BACKGROUND

A. Notation

We use T^c to denote the complement of T w.r.t. $[1, m] := [1, 2, \dots m]$, i.e. $T^c := \{i \in [1, m] : i \notin T\}$. The notation $|T|$ denotes the size (cardinality) of the set T. The set operations $\cup$, $\cap$, and $\setminus$ have the usual meanings.

For a vector, v, and a set, T, v_T denotes the $|T|$ length sub-vector containing the elements of v corresponding to the indices in the set T. Also, $\|v\|_k$ denotes the ℓ_k norm of a vector v. When $k = 0$, $\|v\|_0$ counts the number of nonzero elements in the vector v. If just $\|v\|$ is used, it refers to $\|v\|_2$.

For a matrix M, $\|M\|_k$ denotes its induced k-norm, while just $\|M\|$ refers to $\|M\|_2$. M' denotes the transpose of M and $M^\dagger$ denotes its Moore-Penrose pseudo-inverse. For a tall matrix, M, $M^\dagger := (M'M)^{-1}M'$. For a fat matrix A, A_T denotes the sub-matrix obtained by extracting the columns of A corresponding to the indices in T.

The restricted isometry constant (RIC) [10], δ_S, for a matrix A, is the smallest real number satisfying

$$(1 - \delta_S)\|c\|^2 \leq \|A_T c\|^2 \leq (1 + \delta_S)\|c\|^2 \tag{1}$$

for all subsets $T \subseteq [1, m]$ of cardinality $|T| \leq S$ and all real vectors c of length $|T|$. It is easy to see that $\|A_T'A_T\| \leq (1 + \delta_S)$, $\|(A_T'A_T)^{-1}\| \leq 1/(1 - \delta_S)$ and $\|A_T^\dagger\| \leq 1/\sqrt{(1 - \delta_S)}$.

The restricted orthogonality constant (ROC) [10], $\theta_{S,S'}$, for a matrix A, is the smallest real number satisfying

$$|c_1'A_{T_1}'A_{T_2}c_2| \leq \theta_{S,S'} \|c_1\| \|c_2\| \tag{2}$$

for all disjoint sets $T_1, T_2 \subseteq [1, m]$ with $|T_1| \leq S$, $|T_2| \leq S'$, $S + S' \leq m$, and for all vectors c_1, c_2 of length $|T_1|$, $|T_2|$.

B. Background on Sparse Recovery

The sparse recovery problem has been studied for a very long time, e.g. see [23], [24], [25]. The goal in sparse recovery, or what is now interchangeably referred to as compressive sensing (CS), is to recover a sparse signal from a reduced number of its linear projection measurements. To be precise, we would like to recover an m length sparse vector, x, with support size, s, from $y := Ax$, or, in the noisy case, from $y := Ax + w$, when A is a fat matrix (a matrix with more columns than rows). Consider the noise-free case. The sparse recovery problem is solved if we can find the sparsest vector b among all vectors satisfying $y = Ab$, i.e. if we can solve

$$\min_b \|b\|_0 \text{ s.t. } y = Ab$$

and if A is such that every set of $2s$ columns of A are linearly independent [24], [10]. Finding the sparsest vector requires a combinatorial search and thus has complexity of the order of m^s [10]. The exponential complexity in s makes it impractical to directly solve this for any reasonable sized problem. Practical (polynomial complexity) approaches to this problem include (i) ℓ_1 minimization methods (replace the ℓ_0 norm by the ℓ_1 norm which is the closest norm to ℓ_0 that makes the problem convex) such as basis pursuit [24] and its noisy relaxations – basis pursuit denoising (BPDN) [24], [26], [27], Dantzig selector [28] and others; (ii) greedy methods such as matching pursuit [23], orthogonal matching pursuit [29] and many other recent works [30], [31]; and (iii) various other more recent approaches. While these approaches have been proposed and used since the 1990s, the recent work on compressed sensing provided strong performance guarantees for them: exact recovery conditions [8], [9], [10] and bounds on reconstruction error when exact recovery is not possible [26], [28], [27].

III. PROBLEM DEFINITION, CHAPTER ORGANIZATION AND RELATED WORK

A. Problem Definition and Chapter Organization

The recursive reconstruction problem explained here was first introduced in the ICIP 2008 paper on Kalman filtered compressed sensing (KF-CS) [14]. Let $(z_t)_{m \times 1}$ denote the spatial signal at time t and $(y_t)_{n \times 1}$, with $n < m$, denote its noise-corrupted measurements' vector at t, i.e. $y_t = Hz_t + w_t$ where w_t is measurement noise and H is the measurement matrix. The signal, z_t, is sparse in a given sparsity basis (e.g. wavelet) with orthonormal basis matrix, $\Phi_{m \times m}$, i.e. $x_t := \Phi'z_t$ is a sparse vector. Thus the observation model can be written as

$$y_t = Ax_t + w_t, \ A := H\Phi \tag{3}$$

We assume that A has unit norm columns. We study both the noise-free case, i.e. $w_t = 0$, and the bounded noise case, i.e. $\|w_t\|_2 \leq \epsilon$. We use N_t to denote the support of x_t, i.e.

$$N_t := \text{supp}(x_t) = \{i : (x_t)_i \neq 0\}.$$

The goal is to recursively estimate x_t (or equivalently the signal, $z_t = \Phi x_t$) using $y_1, \dots y_t$. By *recursively*, we mean, use only y_t and the estimate from $t-1$, $\hat{x}_{t-1}$, to compute the estimate at t. This is done under one or both of the following two assumptions.

1) *Slow Support Change.* The support additions, $|N_t \setminus N_{t-1}| \leq S_a \ll |N_t|$ and the removals, $|N_{t-1} \setminus N_t| \leq S_a \ll |N_t|$ at all times t. This assumption is verified for MRI sequences in Fig. 1.
2) *Slow Signal Value Change.* The magnitude of the nonzero signal values also changes slowly with time, i.e. $\|(x_t - x_{t-1})_{N_t}\|_2 \ll \|(x_t)_{N_t}\|_2$. This assumption is also verified in Fig. 1.

Consider first the class of problems for which only the first assumption holds. Under this assumption, the above problem can be reformulated as one of sparse recovery in the presence of partial support knowledge. We can use the support estimate obtained from the previous time instant, $\hat{N}_{t-1}$, as the "partial support knowledge". We describe this problem and the proposed solutions for it in Sec. IV. If both assumptions hold, the above problem can be reformulated as one of sparse recovery with partial support and signal value knowledge. This is discussed in Sec. V. Performance guarantees (exact reconstruction results, error bounds, and conditions for time-invariant error bounds) are briefly discussed in Sec. VI and some interesting experimental results are shown in Sec. VII. Conclusions are given in Sec. VIII.

B. Related Work

The recursive reconstruction problem was first studied in [14]. Before this, the only works that dealt with time sequences of sparse signals were batch methods [11], [12], [13].

A modification of KF-CS was introduced in [32]. Recent work on Bayesian or other model-based approaches to recursive sparse estimation with time-varying supports includes [33], [34], [35], [36], [37]. The work of [13] gives an approximate batch solution for dynamic MRI that is quite fast, but is offline. Related work on model-based and Bayesian approaches for a single signal includes [38], [39], [40], [41], [42], [43], [44].

The problem of sparse reconstruction with partial knowledge of the support was simultaneously addressed in [16], [17] and in [45], [46], [47]. The work of [46] obtains exact recovery thresholds for weighted ℓ_1, similar to those in [48], for the case when a probabilistic prior on the signal support is available. Some later work motivated by modified-CS includes modified OMP [49], modified CoSaMP [50], modified block CS [51], error bounds on modified BPDN [52], [22], [53], [20], better conditions for modified-CS based exact recovery [54], and exact support recovery conditions for multiple measurement vectors (MMV) based recursive recovery [33].

There is other recent work that may also be referred to as recursive sparse reconstruction, but whose goals are quite different from the problem that we discuss in this chapter. This includes (i) homotopy methods, e.g. [55], [56], whose goal is to only speed up the optimization algorithm using homotopy or warm starts and the previous reconstructed signal, but not to reduce the number of measurements required; (ii) [57], [55], [58], [59] which reconstruct a single signal from sequentially arriving measurements; and (iii) [60], [61], [62], which iteratively improve support estimation for a single sparse signal. Another recent work [63] proposes causal but batch methods. This can be interpreted as a causal approach to solve the multiple measurements' vector (MMV) problem.

We should note that none of the above works obtain conditions under which a time-invariant bound on the reconstruction error (i.e. stability) holds. Except [46] and [62], none of these obtain exact reconstruction conditions either.

IV. SPARSE RECOVERY WITH PARTIAL SUPPORT KNOWLEDGE

This problem was first formulated in [16], [17]. The goal is to recover a sparse vector, x, with support set N, either from noise-free undersampled measurements, $y := Ax$, or from noisy measurements, $y := Ax + w$, when partial and possibly erroneous support knowledge, T, is available. The true support N can be rewritten as

$$N = T \cup \Delta \setminus \Delta_e \quad \text{where} \quad \Delta := N \setminus T, \ \Delta_e := T \setminus N$$

It is easy to see that

$$|N| = |T| + |\Delta| - |\Delta_e|$$

Here we refer to the set Δ as the *misses* in the support knowledge and the set Δ_e is the *extras* in it. We say the *support knowledge is accurate if* $|\Delta| \ll |N|$ and $|\Delta_e| \ll |N|$.

Least Squares CS-residual (LS-CS) introduced in [64], [15] can be interpreted as the first solution to the above problem. We describe this next. The first solution that gives exact recovery under weaker conditions (using fewer measurements) than what simple CS needs was Modified-CS [16], [17]. We explain this in Sec. IV-B. For using either LS-CS or modified-CS for recursive reconstruction, we use the support estimate from the previous time instant as the partial knowledge set T. Support estimation approaches are discussed in Sec. IV-C and LS-CS or modified-CS for recursive reconstruction is given in Sec. IV-D.

(i) a larynx (vocal tract) image sequence

(ii) cardiac image sequence

(a)

(i) support additions, $\frac{|N_t \setminus N_{t-1}|}{|N_t|}$

(ii) support removals, $\frac{|N_{t-1} \setminus N_t|}{|N_t|}$

(iii) signal value change, $\frac{\|(x_t - x_{t-1})_{N_t \cup N_{t-1}}\|_2}{\|(x_t)_{N_t}\|_2}$

(b)
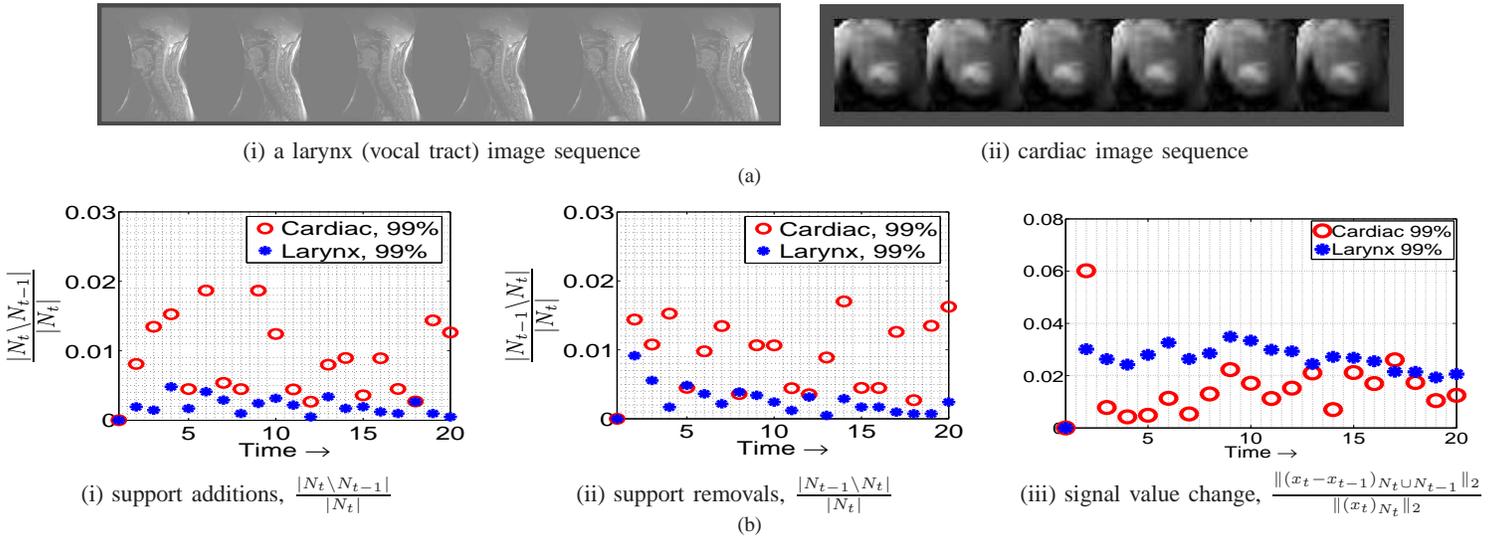
Fig. 1. In (a), we show two MRI image sequences (a cardiac and a larynx sequence). In (b), x_t is the two-level Daubechies-4 2D discrete wavelet transform (DWT) of the cardiac or the larynx image at time t and the set N_t is its 99% energy support (the smallest set containing 99% of the vector's energy). The support size was between 6-7% of the image size. In (c) we plot signal value change. As can be seen from the plots, all support changes (both additions and removals) are less than 2% of the support size. Also, almost all signal value changes are are less than 4% of $\|(x_t)_{N_t}\|_2$.

Algorithm 1 Dynamic LS-CS: LS-CS for recursive reconstruction

---

Simple CS. At $t = 0$, set $T = \emptyset$ and compute $\hat{x}_0$ as the solution of $\min_b \|b\|_1$ s.t. $\|y - Ab\|_2 \leq \epsilon$

For $t > 0$, do

1) Set $T = \hat{N}_{t-1}$.

2) *Initial LS.*

    a) Compute the initial LS estimate $(\hat{x}_{t,\text{init}})_T = (A_T'A_T)^{-1}A_T'y_t$, $\quad(\hat{x}_{t,\text{init}})_{T^c} = 0$

3) *CS-residual.*

    a) Compute the observation residual, $\tilde{y}_t = y_t - A\hat{x}_{t,\text{init}}$

    b) Solve the ℓ_1 problem for the residual, i.e. compute $\hat{\beta}_t$ as the solution of $\min_b \|b\|_1$ s.t. $\|\tilde{y}_t - Ab\|_2 \leq \epsilon$

    c) Compute $\hat{x}_t = \hat{x}_{t,\text{init}} + \hat{\beta}_t$

4) *Support Estimation via Add-LS-Del.*

$$T_{\text{add}} = T \cup \{i \in T^c : |(\hat{x}_t)_i| > \alpha_{\text{add}}\}$$
$$(\hat{x}_{\text{add}})_{T_{\text{add}}} = A_{T_{\text{add}}}{}^\dagger y_t, \quad (\hat{x}_{\text{add}})_{T_{\text{add}}^c} = 0$$
$$\hat{N}_t = T_{\text{add}} \setminus \{i \in T : |(\hat{x}_{\text{add}})_i| \leq \alpha_{\text{del}}\} \quad (4)$$

5) *Final LS Estimate.*

$$(\hat{x}_{t,\text{final}})_{\hat{N}_t} = A_{\hat{N}_t}{}^\dagger y_t, \quad (\hat{x}_{t,\text{final}})_{\hat{N}_t^c} = 0 \quad (5)$$

A. Least Squares CS-residual (LS-CS)

The key idea of LS-CS is as follows [64], [15]. Using T as the support set, compute an initial estimate of x by computing an LS estimate on T and setting all other elements to zero, i.e. compute

$$(\hat{x}_{\text{init}})_T = (A_T'A_T)^{-1}A_T'y_t, \quad (\hat{x}_{\text{init}})_{T^c} = 0 \quad (6)$$

Compute the observation residual, $\tilde{y}$,

$$\tilde{y} = y - A\hat{x}_{\text{init}} \quad (7)$$

followed by solving the ℓ_1 minimization problem for this residual, i.e. compute $\hat{\beta}$ as the solution of

$$\arg\min_b \|b\|_1 \text{ s.t. } \|\tilde{y} - Ab\|_2 \leq \epsilon \quad (8)$$

Then, the estimate of x is computed as

$$\hat{x} = \hat{x}_{\text{init}} + \hat{\beta}. \quad (9)$$

Algorithm 2 Dynamic Modified-CS: Modified-CS for recursive reconstruction

---

1) In Algorithm 1, replace steps 2 and 8 by the following *Modified-CS* step.

   a) Compute $\hat{x}_t$ as the solution of $\min_b \|b_{T^c}\|_1$ s.t. $\|y_t - Ab\|_2 \leq \epsilon$.

This is followed by support estimation and computing a final LS estimate on the estimated support as described in Sec IV-C.

Notice that, the signal residual, $\beta := x - \hat{x}_{\text{init}}$, is supported on $T \cup \Delta$ and satisfies

$$\beta_T = (A_T{}'A_T)^{-1}A_T{}'(A_\Delta x_\Delta + w), \quad \|\beta_T\|_2 \leq \frac{\theta_{|T|,|\Delta|}}{1 - \delta_{|T|}}\|x_\Delta\|_2 + \frac{1}{\sqrt{1 - \delta_{|T|}}}\epsilon$$

$$\beta_{T^c} = x_\Delta$$

If $|\Delta|$ is small enough, θ is small. If $|\Delta_e|$ is small enough, $|T| \leq |N| + |\Delta_e|$ is not too large and so $1/(1 - \delta_{|T|})$ is only a little more than one. Finally if the noise is also small, the above implies that $\|\beta_T\|_2 \ll \|x_\Delta\|_2$. Thus, if T is a good estimate of the true support, N; the measurement matrix A is incoherent enough; and the noise is small enough; then β is small on the set T. Or, in other words, β is approximately supported only on Δ. Since T is a good estimate of the true support, $|\Delta| \ll |N|$ and so the ℓ_1 problem that we need to solve in this case is much easier than in case of simple CS. As a result, it is possible to show that LS-CS results in small reconstruction error using much fewer measurements than what simple CS needs [15, Theorem 1]. We summarize LS-CS for recursive reconstruction in Algorithm 1.

However, notice that the exact sparsity size (total number of nonzero components) of the signal residual, β, is $|T| + |\Delta|$ and this is equal to or larger than that of the signal, $|N|$. Since the number of measurements required for exact reconstruction is governed by the exact sparsity size, LS-CS is not able to achieve exact reconstruction using fewer noiseless measurements than those needed by simple CS. The search for such a solution led us to our next and more powerful idea called Modified-CS.

B. Modified-CS

The key idea of Modified-CS is as follows [16], [17]. Suppose first that Δ_e is empty, i.e. $N = T \cup \Delta$. Thus, the sparse recovery problem now becomes one of trying to find the sparsest vector whose support contains T among all vectors that satisfy the data constraint. Or in other words, we would like to find the vector that is sparsest outside the set T among all vectors that satisfy the data constraint. In the noise-free case, this can be written as

$$\min_b \|b_{T^c}\|_0 \text{ s.t. } y = Ab$$

The above also works if Δ_e is not empty. It is easy to show that it can exactly recover x if $w = 0$ (noise-free case) and if every set of $|T| + 2|\Delta| = |N| + |\Delta| + |\Delta_e|$ columns of A are linearly independent [17, Proposition 1]. In comparison, the original ℓ_0 problem given in Sec. II-B requires every set of $2|N|$ columns of A to be linearly independent [10]. This is much stronger when $|\Delta| \approx |\Delta_e| \ll |N|$.

The above ℓ_0 problem also has exponential complexity, and hence as in case of CS, we replace it by the ℓ_1 problem (ℓ_1 norm is the closest norm to ℓ_0 that makes the optimization problem convex). Thus, *modified-CS* solves

$$\min_b \|b_{T^c}\|_1 \text{ s.t. } y = Ab \tag{10}$$

and we denote its solution by $\hat{x}$. Once again, this works, and can provably achieve exact recovery, even when Δ_e is not empty. We give the exact recovery conditions in Sec VI-A. For noisy measurements, one can relax the data constraint as follows.

$$\min_b \|b_{T^c}\|_1 \text{ s.t. } \|y - Ab\|_2 \leq \epsilon \tag{11}$$

We summarize modified-CS for recursive reconstruction in Algorithm 2. In practice, for large scale problems, one always adds the data term as a soft constraint and solves the following unconstrained problem (which is less expensive to solve and does not require knowledge of the noise bound). We refer to this as modified-BPDN [53], [20].

$$\min_b \gamma\|b_{T^c}\|_1 + 0.5\|y - Ab\|_2^2 \tag{12}$$

C. Support Estimation: thresholding and add-LS-del

In order to use either LS-CS or modified-CS for recursive reconstruction, we use the support estimate from the previous time as the set T. Thus, we need to estimate the support of the signal at each time. The simplest way to do this is by thresholding, i.e. we compute

$$\hat{N} = \{i : |(\hat{x})_i| > \alpha\}$$

where $\alpha \geq 0$ is the zeroing threshold. In cases of exact reconstruction, i.e. if $\hat{x} = x$, we can use $\alpha = 0$. In other situations, we need a nonzero value. In cases of very accurate reconstruction, we can set α to be slightly smaller than the magnitude of

the smallest nonzero element of x (assuming its rough estimate is available) [17]. This will ensure close to zero misses and few false additions. In general, α should depend on both the noise level and the magnitude of the smallest nonzero element of x. For compressible signals, one should do the above but with "support" replaced by the $b\%$-energy support. For a given number of measurements, b can be chosen to be the largest value so that all elements of the $b\%$-energy support can be exactly reconstructed [17].

Single step thresholding as above means that the threshold, α, needs to be large enough to ensure that most missed elements from T are correctly deleted while ensuring that there are few false detections. However, notice that, in both modified-CS and LS-CS, $\hat{x}$ is a biased estimate of x. Consider modified-CS. Along $\Delta \subset T^c$, the values of $\hat{x}$ are biased towards zero (because we minimize $\|(\beta)_{T^c}\|_1$), while, along $\Delta_e \subset T$, they may be biased away from zero (since there is no constraint on $(\beta)_T$). The same also happens for LS-CS although the reasoning is a bit different [15, Sec II-A]. Since the estimates along Δ are biased towards zero, one needs a smaller threshold to detect them, whereas, since those along Δ_e may be biased away from zero, one may need a higher threshold to delete them. One partial solution to this problem is to use the following three step Add-LS-Del approach:

$$T_{\text{add}} = T \cup \{i : |(\hat{x})_i| > \alpha_{\text{add}}\} \tag{13}$$
$$(\hat{x}_{\text{add}})_{T_{\text{add}}} = A_{T_{\text{add}}}{}^\dagger y, \quad (\hat{x}_{\text{add}})_{T_{\text{add}}^c} = 0 \tag{14}$$
$$\hat{N} = T_{\text{add}} \setminus \{i : |(\hat{x}_{\text{add}})_i| \leq \alpha_{\text{del}}\} \tag{15}$$

The above add-LS-del procedure involves a support addition step, that uses a smaller threshold, α_{add}, as in (13); followed by LS estimation on the new support estimate, T_{add}, as in (14); and then a deletion step that thresholds the LS estimate, as in (15). The addition step threshold, α_{add}, needs to be just large enough to ensure that the matrix used for LS estimation, $A_{T_{\text{add}}}$ is well-conditioned. If α_{add} is chosen properly and if the number of measurements, n, is large enough, the LS estimate on T_{add} will have smaller error, and will be less biased, than $\hat{x}$ (modified-CS or LS-CS output). As a result, deletion will be more accurate when done using this estimate. This also means that one can use a larger deletion threshold, α_{del}, which will ensure deletion of more extras.

A similar issue for noisy CS, and a possible solution (Gauss-Dantzig selector), was first discussed in [28]. The add-LS-del idea was first introduced in the KF-CS and LS-CS papers [14], [15], [22] for recursive reconstruction and simultaneously also in [30], [31] for greedy algorithms for static sparse reconstruction.

Support estimation is usually followed by LS estimation on the final support estimate, in order to get a solution with reduced bias (Gauss-Dantzig selector idea), i.e. one computes

$$(\hat{x}_{\text{final}})_{\hat{N}} = A_{\hat{N}}{}^\dagger y, \quad (\hat{x}_{\text{final}})_{\hat{N}^c} = 0 \tag{16}$$

D. Recursive Recovery

For recursive recovery, in case of slow support change, one can use $T = \hat{N}_{t-1}$. We summarize the complete algorithm for LS-CS in Algorithm 1 and that for Modified-CS in Algorithm 2. Recent work [4] has introduced solutions for the more general case where the support change may not be slow, but is still highly correlated over time.

V. SPARSE RECOVERY WITH PARTIAL SUPPORT AND SIGNAL VALUE KNOWLEDGE

So far we only talked about the case where prior support information is available. In certain applications, one may also have partial signal value knowledge. In recursive recovery problems, it often happens that signal values also change slowly over time. In this case the problem can be formulated as follows. The goal is to recover a sparse vector x, with support set N, either from noise-free undersampled measurements, $y := Ax$, or from noisy measurements, $y := Ax + w$, when partial erroneous support knowledge, T, is available and partial erroneous signal value knowledge on T, $\hat{\mu}_T$, is available. The true support N can be written as

$$N = T \cup \Delta \setminus \Delta_e \quad \text{where} \quad \Delta := N \setminus T, \ \Delta_e := T \setminus N$$

and the true signal x can be written as

$$(x)_{N \cup T} = (\hat{\mu})_{N \cup T} + e$$
$$(x)_{N^c} = 0, (\hat{\mu})_{T^c} = 0 \tag{17}$$

The error e in the prior signal estimate is assumed to be small, i.e. $\|e\| \ll \|x\|$.

A. Regularized Modified-CS

Regularized modified-CS adds the slow signal value change constraint to modified-CS and solves the following [21], [20].

$$\min_b \|b_{T^c}\|_1 \text{ s.t. } \|y - Ab\|_2 \leq \epsilon, \text{ and } \|b_T - \hat{\mu}_T\|_\infty \leq \rho \tag{18}$$

As before, the following Lagrangian version (constraints added as weighted costs to get an unconstrained problem) is more useful in practice

$$\min_b \gamma\|b_{T^c}\|_1 + 0.5\|y - Ab\|_2^2 + 0.5\lambda\|b_T - \hat{\mu}_T\|_2^2 \tag{19}$$

Regularized modified-CS is analyzed in detail in [21] and [20].

B. Modified-CS-residual

The idea of modified-CS-residual is to combine the modified-CS idea with the CS-residual idea. One solves modified-CS on the observation residual computed using $\hat{x}_{\text{init}} = \hat{\mu}$. Once again the following unconstrained version is most useful:

$$\min_b \|b_{T^c}\|_1 + 0.5\alpha\|(y - A\hat{\mu} - Ab\|_2^2 \tag{20}$$

For recursive reconstruction, one again uses $T = \hat{N}_{t-1}$. For $\hat{\mu}$, one can either use $\hat{\mu} = \hat{x}_{t-1}$, or, in certain situations where the signal values do not change much w.r.t. the first frame, using $\hat{\mu} = \hat{x}_0$ is a better idea. *For practical problems, e.g. real functional MRI sequences [19], modified-CS-residual with $\hat{\mu} = \hat{x}_0$ turns out to be the most promising approach to use.*

As we explain next, in recursive reconstruction problems, if a model on signal value change is available, one can also obtain $\hat{\mu}$ by using a Kalman filter.

C. Kalman Filtered CS-residual (KF-CS) and Kalman Filtered Modified-CS-residual (KalMoCS)

Kalman Filtered CS (KF-CS) was introduced in the context of recursive reconstruction in [14]. The key idea is to replace the initial LS step of LS-CS by a regularized LS step. One then computes the observation residual, followed by solving the ℓ_1 problem on this residual, exactly as in LS-CS. In KalMoCS, one replaces the ℓ_1 problem on this residual by the modified-ℓ_1.

Regularized LS becomes the KF in case of recursive recovery. The extra piece of information needed for KF-CS or KalMoCS is a model on signal value change. Typically, in most cases, one can assume a simple random walk model with equal change variance in all directions [14]. We summarize KF-CS and KalMoCS in Algorithm 3. This will outperform LS-CS and modified-CS when support changes occur every so often (allows the KF to stabilize to a small error before the next support change).

VI. THEORETICAL RESULTS

We first summarize the exact reconstruction results for modified-CS and regularized modified-CS and their implications. Next, we briefly discuss the error bounds for the noisy case. Finally, we address the most important question for recursive recovery: when is the algorithm "stable" over time, i.e. when can we get time-invariant bounds on its error over time?

A. Exact Reconstruction in Noise-free case

As explained earlier, LS-CS and KF-CS cannot achieve exact recovery under weaker conditions than what is needed for simple CS. However, modified-CS [17] and regularized modified-CS can [21]. We give below the RIC based exact recovery conditions for modified-CS [17]:

Theorem 1 (Exact Recovery Conditions – Modified-CS): [17, Theorem 1] Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_e$, where $\Delta = N \setminus T$ and $\Delta_e = T \setminus N$, consider reconstructing it from $y := Ax$ by solving (10). Let $k := |T|$, $u := |\Delta|$, $e := |\Delta_e|$ and $s := |N|$. Then, x is the unique minimizer of (10) if

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

2) $a_k(2u, u) + a_k(u, u) < 1$ where $a_k(S, \check{S}) := \dfrac{\theta_{\check{S},S} + \frac{\theta_{\check{S},k} \theta_{S,k}}{1-\delta_k}}{1-\delta_S - \frac{\theta_{\check{S},k}^2}{1-\delta_k}}$

The above conditions can also be rewritten in terms of s, e, u by substituting $k = s + e - u$.

A simpler sufficient condition for modified-CS that uses only the RIC is [17, Corollary 1]:

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

Compare this with simple CS which requires [27], [65], [28]

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

Algorithm 3 **Kalman Filtered Modified-CS-residual (KalMoCS) and KF-CS**

For $t > 0$ do,

1) Set $T = \hat{N}_{t-1}$.
2) *Initial KF.*

$$
\begin{aligned}
P_{t|t-1} &= P_{t-1} + \hat{Q}_t, \text{ where } \hat{Q}_t := \sigma_{sys}^2 I_T \\
K_t &= P_{t|t-1} A' (A P_{t|t-1} A' + \sigma^2 I)^{-1} \\
P_t &= (I - K_t A) P_{t|t-1} \\
\hat{x}_{t,\text{init}} &= (I - K_t A) \hat{x}_{t-1} + K_t y_t
\end{aligned}
\tag{21}
$$

3) *CS-residual or Modified-CS-residual.*
   a) Compute the KF residual, $\tilde{y}_t$, using $\tilde{y}_t = y_t - A\hat{x}_{t,\text{init}}$
   b) KalMoCS: Solve modified-ℓ_1 on the residual: compute $\hat{\beta}_t$ as the solution of

   $$
   \min_b \|b_{T^c}\|_1 \text{ s.t. } \|\tilde{y}_t - Ab\|_2 \le \epsilon
   $$

   - In case of KF-CS: replace $\|b_{T^c}\|_1$ by $\|b\|_1$.
   c) Compute $\hat{x}_t = \hat{x}_{t,\text{init}} + \hat{\beta}_t$

4) *Support Estimation via Add-LS-Del.*

$$
\begin{aligned}
T_{\text{add}} &= T \cup \{i \in T^c : |(\hat{x}_t)_i| > \alpha_{\text{add}}\} \\
(\hat{x}_{\text{add}})_{T_{\text{add}}} &= A_{T_{\text{add}}}{}^\dagger y, \quad (\hat{x}_{\text{add}})_{T_{\text{add}}^c} = 0 \\
\hat{N}_t &= T_{\text{add}} \setminus \{i \in T : |(\hat{x}_{\text{add}})_i| \le \alpha_{\text{del}}\}
\end{aligned}
\tag{22}
$$

5) *Final Estimate:* If $\hat{N}_t$ is equal to T, set

$$
\hat{x}_{t,\text{final}} = \hat{x}_{t,\text{init}}
$$

   else, compute an LS estimate using $\hat{N}_t$ and update P_t as follows.

$$
\begin{aligned}
(\hat{x}_{t,\text{final}})_{\hat{N}_t} &= A_{\hat{N}_t}{}^\dagger y_t, \quad (\hat{x}_{t,\text{final}})_{\hat{N}_t^c} = 0 \\
(P_t)_{\hat{N}_t, \hat{N}_t} &= (A_{\hat{N}_t}{}' A_{\hat{N}_t})^{-1} \sigma^2, \quad (P_t)_{\hat{N}_t^c,[1,m]} = 0, \quad (P_t)_{[1,m],\hat{N}_t^c} = 0
\end{aligned}
\tag{23}
$$

To compare these conditions numerically, we can use $u = e = 0.02s$ which is typical for time series applications (see Fig. 1). Using $\delta_{cr} \le c\delta_{2r}$ [31, Corollary 3.4], it can be show that modified-CS only requires $\delta_{2u} < 0.004$. On the other hand, simple CS requires $\delta_{2u} < 0.008$ which is clearly stronger.

Exact recovery conditions for regularized modified-CS in the noise-free case, i.e. for (18) with $\epsilon = 0$ are obtained in [21, Theorem 1]. These are weaker than those for modified-CS if $x_i - \hat{\mu}_i = \pm\rho$ for some $i \in T$ (some of the constraints $\|b_T - \hat{\mu}_T\|_\infty \le \rho$ are active for the true signal, x) and some elements of this active set satisfy the condition given in [21, Theorem 1]. One set of practical applications where $x_i - \hat{\mu}_i = \pm\rho$ with nonzero probability is when dealing with quantized signals and quantized signal estimates.

B. Error Bounds for the Noisy case

When measurements are noisy, one cannot get exact recovery, but can only bound the reconstruction error. We give here the error bounds for both LS-CS [15] and modified-CS [22]. The LS-CS-residual step error can be bounded as follows. The proof follows in exactly the same way as that given in [15] where CS is done using the Dantzig selector instead of constrained BPDN as in (8).

Theorem 2 (LS-CS-residual error bound): [15, Lemma 1] Let x be a sparse vector with support N and let $y := Ax + w$ with $\|w\| \le \epsilon$. Also, let $\Delta := N \setminus T$ and $\Delta_e := T \setminus N$. Let $\hat{x}$ be computed as in (9). If $\delta_{2|\Delta|} < (\sqrt{2}-1)/2$ and $\delta_{|T|} < 1/2$,

$$
\|x - \hat{x}\| \le C'(|T|,|\Delta|)\epsilon + \theta_{|T|,|\Delta|} C''(|T|,|\Delta|)\|x_\Delta\|
\tag{24}
$$

where $C'(|T|,|\Delta|) := C_1(2|\Delta|) + \sqrt{2}C_2(2|\Delta|)\sqrt{\frac{|T|}{|\Delta|}}$, $C''(|T|,|\Delta|) := 2C_2(2|\Delta|)\sqrt{|T|}$, $C_1(S) := \frac{4\sqrt{1+\delta_S}}{1-(\sqrt{2}+1)\delta_S}$, and $C_2(S) := 2\frac{1+(\sqrt{2}-1)\delta_S}{1-(\sqrt{2}+1)\delta_S}$.

By adapting the approach of [27], the error of modified-CS can be bounded as a function of $|T| = |N| + |\Delta_e| - |\Delta|$ and $|\Delta|$. This was done by Jacques in [66] and by us in [22].

Theorem 3 (modified-CS error bound): [22, Lemma 1] Let x be a sparse vector with support N and let $y := Ax + w$ with $\|w\| \le \epsilon$. Also, let $\Delta := N \setminus T$ and $\Delta_e := T \setminus N$. Let $\hat{x}$ denote the solution of (11). If $\delta_{|T|+3|\Delta|} < (\sqrt{2} - 1)/2$, then

$$\|x - \hat{x}\| \le C_1(|T| + 3|\Delta|)\epsilon \le 9.8\epsilon, \text{ where } C_1(S) := \frac{4\sqrt{1 + \delta_S}}{1 - (\sqrt{2} + 1)\delta_S} \tag{25}$$

For both LS-CS and modified-CS, the error after the final LS step can be bounded in terms of $\tilde{T} := \hat{N}$ and $\tilde{\Delta} := \hat{N} \setminus N$ as follows.

$$\|x - \hat{x}_{\text{final}}\| \le (1 + \frac{\theta_{|\tilde{T}|,|\tilde{\Delta}|}}{1 - \delta_{|\tilde{T}|}})\|x_{\tilde{\Delta}}\|_2 + \frac{1}{\sqrt{1 - \delta_{|\tilde{T}|}}}\epsilon \tag{26}$$

C. Recursive Reconstruction: Time-Invariant Error Bounds (Stability)

Let $\tilde{T} := \hat{N}$, $\tilde{\Delta} := \hat{N} \setminus N$ and $\tilde{\Delta}_e := N \setminus \hat{N}$. So far we bounded the LS-CS-residual error or the modified-CS error as a function of $|T|, |\Delta|$. The bound is small as long as $|\Delta_e|$ and $|\Delta|$ are small. Similarly the bound on the error of the final LS estimate, given in (26), is small if $|\tilde{\Delta}|$ and $|\tilde{\Delta}_e|$ are small. However for recursive reconstruction, what we need is conditions under which we can get a time invariant bound on $|\Delta_e|$ and $|\Delta|$ as well as on $|\tilde{\Delta}|$ and $|\tilde{\Delta}_e|$. Otherwise, it can happen that the support errors keep adding up and become large and the same will happen to the reconstruction errors.

The study of error stability over time requires a signal change model. We assume the following simple deterministic model [15, Signal Model 1]. (a) There is nonzero delay, d, between new coefficient addition and removal times; (b) at most S_a additions and removals occur at every change time; (c) new coefficients' magnitudes increase gradually from zero for sometime and finally reach a constant value; and (d) coefficients' magnitudes decrease gradually before becoming zero (getting removed from support). Under this model, one can show the following. The actual conditions in the final result are somewhat messy and so we skip them. We only state a qualitative version here.

Theorem 4 (Time-invariant error bounds): [15, Theorem 2] Assume the above signal change model. If

1) the initial simple CS step is accurate enough,
2) the noise is bounded and the number of measurements, n, is large enough so that certain conditions on the RIC and ROC hold,
3) the addition and deletion thresholds are appropriately set,
4) for a given n and noise bound,
   a) the smallest constant coefficient magnitude is large enough,
   b) the rates of coefficient magnitude increase and decrease are large enough, and
   c) the delay between addition times, d, is larger than the "worst case detection delay" plus coefficient decrease time,

then,

1) the number of final misses $|\tilde{\Delta}_t|$ and extras $|\tilde{\Delta}_{e,t}|$ as well as the initial misses $|\Delta_t|$ and extras $|\Delta_{e,t}|$ are bounded by S_a or by a quantity slightly larger than S_a,
2) within a finite delay, all new additions get detected and not falsely deleted, i.e. $|\tilde{\Delta}_t| = 0$, and all the extras get deleted, i.e. $|\tilde{\Delta}_{e,t}| = 0$,
3) and the reconstruction error is bounded by a time-invariant and small values at all times.

As long as the number of new additions or removals, $S_a \ll |N_t|$ (slow support change), the above result shows that the worst case number of misses or extras is also small compared to the support size. This makes it a meaningful result. Similarly, we can argue that the reconstruction error bound is small compared to the signal energy.

The above result was proved for LS-CS in [15, Theorem 2]. It is possible to prove an exactly analogous result for modified-CS as well. The key ideas in obtaining this result are as follows. (i) One needs to ensure that within a finite delay of a new addition time, all new additions definitely get detected and not false deleted (this delay is the "worst case detection delay"). (ii) This needs to be done while ensuring that there are no false deletions of the constant coefficients. (iii) Also, the deletion threshold needs to be high enough to definitely delete all the extras every-so-often (ensure $|\tilde{T}_t|$ is bounded). (iv) Finally, the "worst case detection delay" plus the coefficient decrease time need to be smaller than the delay between two addition times.

The above result assumes support change every d frames. One can also show stability under a more general signal model that allows support changes at every time. This has been done for both modified-CS and LS-CS in [22].

VII. EXPERIMENTS

We briefly describe three sets of experiments here. The first set consists of simulation experiments that demonstrate that modified-CS achieves exact reconstruction using significantly fewer measurements than what simple CS needs. The second set consists of simulation experiments that compare the reconstruction errors of LS-CS, KF-CS, modified-CS (actually modified-BPDN) and regularized modified-BPDN with each other and with other existing work in literature (CS-diff and weighted ℓ_1). The third set of experiments studies recursive recovery for a simulated dynamic MRI experiment. Here we took actual (not

(a) $n = 0.3m$, $\sigma_p^2 = 10^{-1}$, $\sigma_w^2 = 10^{-5}$ (b) $n = 0.13m$, $\sigma_p^2 = 10^{-3}$, $\sigma_w^2 = 10^{-5}$ (c) $n = 0.13m$, $\sigma_p^2 = 10^{-1}$, $\sigma_w^2 = 10^{-5}$
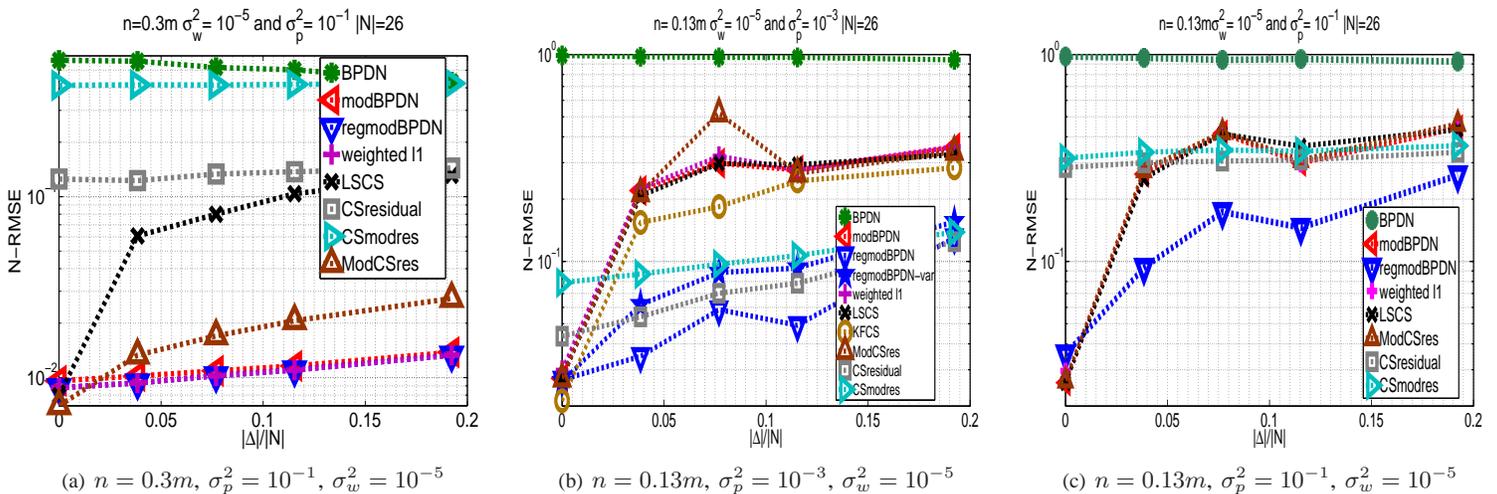
Fig. 2. The N-RMSE for reg-mod-BPDN, mod-BPDN, BPDN, LS-CS, KF-CS, weighted ℓ_1, CS-residual, CS-mod-residual and modified-CS-residual are plotted against $|\Delta|/|N|$.

sparsified) larynx or cardiac image sequences and simulated MRI by taking a randomly selected set of their partial Fourier measurements. In this case, we did not add measurement noise, however since the signal sequence is not exact sparse, one could think of the compressible coefficients as "noise" (this noise is correlated with the signal, but none of our analysis uses any probability model, so the correlation does not matter). We demonstrate error stability over time of modified-CS and LS-CS and we also show that modified-CS has lower error than LS-CS.

A. Exact Reconstruction Probability Computation via Monte Carlo

In Sec VI-A, we only compared sufficient conditions for CS and modified-CS. However, this does not mean that the required number of measurements, n, for CS is definitely smaller than what modified-CS needs. To actually compare this, we need to use Monte Carlo. We obtained a Monte Carlo estimate of the probability of exact reconstruction for CS and for modified-CS, for a given A (i.e. we averaged over the joint distribution of x and y given A) as follows [17]. Fix signal length, $m = 256$ and its support size, $s = 0.1m = 26$. In the experiment we describe here we also fixed $u = e = 0.08m$. We varied n. For each n, we generated a $n \times m$ random-Gaussian matrix, A once. We then repeated the following 500 times. (i) Generate the support, N, of size s, uniformly at random from $[1, m]$ and generate $(x)_N \sim \mathcal{N}(0, 100I)$. Set $(x)_{N^c} = 0$. (ii) Set $y := Ax$. (iii) Generate Δ of size u uniformly at random from the elements of N. (iv) Generate Δ_e of size e, uniformly at random from the elements of $[1, m] \setminus N$. (v) Set $T = N \cup \Delta_e \setminus \Delta$. (vi) Solve modified-CS, i.e. solve (10). Call the solution $\hat{x}_{modCS}$. (vii) Solve simple CS, i.e. solve (10) with T being the empty set. Call the solution $\hat{x}_{CS}$.

At the end, 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 500. Do the same for CS using $\hat{x}_{CS}$. In this experiment, we observed the following.

1) With 19% measurements, modified-CS gives exact recovery with probability (w.p.) 99.8%, while CS does this w.p. zero.
2) With 25% measurements, modified-CS gives exact recovery with probability (w.p.) 100%, while CS does this w.p. 0.2%.
3) CS requires 40% measurements to work "reliably", i.e. to give exact recovery w.p. at least 98%.

More detailed simulation results for various choices of u and e are summarized in [17, Table 1].

B. Reconstruction Error Comparisons

In Fig. 2, we compare the Monte Carlo average of the reconstruction error of reg-mod-BPDN given in (19) with that of modified-BPDN given in (12), modified-CS-residual given in (20), BPDN [24], weighted ℓ_1 [46], CS-residual [67] and CS-mod-residual. Weighted ℓ_1 solves $\min_b \gamma \|b_{T^c}\|_1 + \gamma' \|b_T\|_1 + \frac{1}{2}\|y - Ab\|_2^2$. CS-residual is an improved version of CS-diff [67]. It computes $\hat{x} = \hat{\mu} + \hat{b}$ where $\hat{b}$ solves $\min_b \gamma \|b\|_1 + \frac{1}{2}\|y - A\hat{\mu} - Ab\|_2^2$

The simulation model used is as specified in [20]. The measurements were random-Gaussian projections corrupted by zero mean i.i.d. Gaussian noise with variance σ_w^2. We used $m = 256$, support size $|N| = 0.1m = 26$ and support extras size, $|\Delta_e| = 0.1|N| = 3$. We plot the errors against $|\Delta|/|N|$. The parameters, e.g. γ, λ, γ', used in each of the minimizations were selected as explained in [20]. Notice that with $n = 30\%$ measurements and a bad signal prior (large σ_p^2), reg-mod-BPDN, mod-BPDN and weighted ℓ_1 have similar performance. LS-CS is worse than either of these, but better than simple CS and CS-residual. With $n = 13\%$ in (b) and (c), reg-mod-BPDN significantly outperforms all the others. In (b), the signal prior is good (small σ_p^2) and so CS-residual is better than modified-CS or weighted ℓ_1 (which do not use signal value knowledge at all) whereas all three of them have similar performance in (c) when the signal prior is bad.

(a) $n_0=0.2m$, $n=0.19m$ (b) $n_0=0.19m$, $n=0.19m$
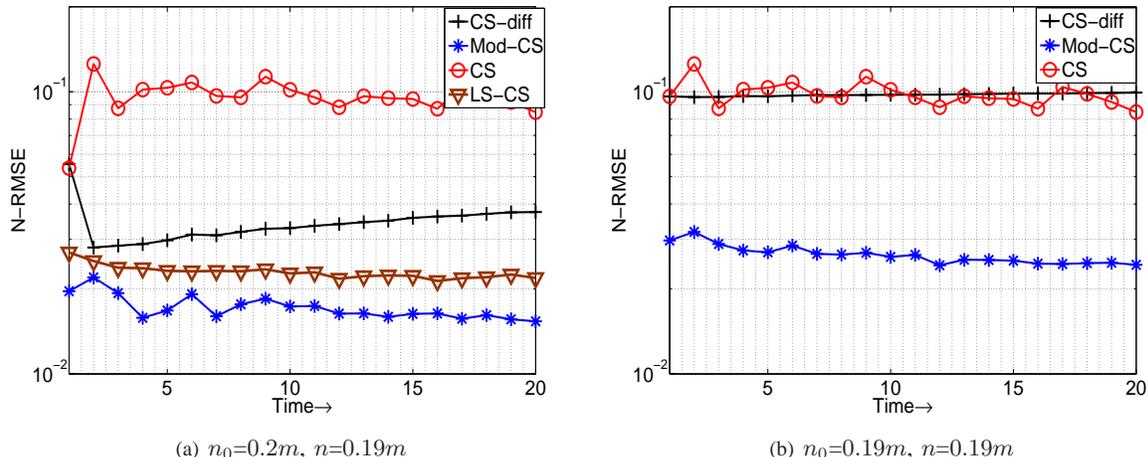
Fig. 3. Reconstructing a 256x256 *actual (compressible)* vocal tract (larynx) image sequence from *simulated MRI* measurements. Both figures used $n = 0.19m$ for $t > 0$ but used different values of n_0. Image size, $m = 256^2 = 65536$. 99% energy support size, $|N_t| \approx 0.07m$; support change size $|N_t \setminus N_{t-1}| \approx 0.001m$.

C. Recursive reconstruction: simulated dynamic MRI

We now show comparisons for recursively reconstructing an actual (compressible) vocal tract image sequence from simulated dynamic MRI measurements [17]. The original image sequence is shown in Fig. 1. In Fig. 3, we show normalized root mean squared error (N-RMSE) comparisons of modified-CS and LS-CS with simple CS [24], [10] and CS-diff [67]. 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, LS-CS error was much higher, since the computed initial LS estimate itself was inaccurate. Notice from the figure that modified-CS and LS-CS significantly outperform CS and CS-diff. Also, modified-CS has smaller error than LS-CS. In Fig 3(b), CS-diff performs so poorly, because the initial error at $t = 0$ is itself very large (since we use only $n_0 = 0.19m$). As a result the difference signal at $t = 1$ is not compressible enough, making its error large and so on. But even when n_0 is larger and the initial error is small, as in Fig. 3(a), the CS-diff error is still unstable, i.e. it increases over time.

VIII. CONCLUSIONS

In this chapter, we summarized our recent work on algorithms for recursive reconstruction of sparse signal sequences. The key ideas we used are that in many such sequences, the sparsity pattern changes slowly over time, and, in certain cases, the same is true also for signal value change. Using just the first assumption, the recursive recovery problem can be reformulated as one of sparse recovery in the presence of partial support knowledge. We discussed two solutions to this problem, the first is called least squares CS-residual (LS-CS), and the second and more powerful one is called Modified-CS. Modified-CS achieves provably exact recovery under weaker conditions (using fewer measurements) than what simple CS needs whenever the support knowledge is accurate enough. When measurements are noisy, the errors are provably bounded. For recursive recovery with noisy measurements, the most important question is, when can we obtain time-invariant bounds on the reconstruction errors, i.e. when can we show error stability over time? We showed that this can be done under fairly mild assumptions for both LS-CS and modified-CS. For problems where both slow support and signal value change hold, we introduced Kalman filtered CS-residual (KF-CS) or its improved versions, Kalman filtered Modified-CS-residual (KalMoCS). Their performance analysis is still mostly a part of ongoing work. *Among all the ideas introduced in this chapter, we think Modified-CS, explained in Sec IV-B, and Modified-CS-residual, explained in Sec V-B, are the most promising approaches.*

Ongoing work is looking at how to utilize correlated, but not necessarily slow, support change to design recursive reconstruction algorithms [4]. Another line of work is exploring the problem of recursive reconstruction in the presence of (potentially) very large but correlated noise [68], [5].

REFERENCES

[1] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, "An architecture for compressive imaging," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2006.

[2] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Info. Th.*, Sept. 2006.

[3] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *Journal of the ACM*, vol. 58 (1), pp. 1–37, 2009.

[4] C. Qiu and N. Vaswani, "Support predicted modified-cs for recursive robust principal components' pursuit," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2011.

[5] C. Qiu and N. Vaswani, "Recursive sparse recovery in large but correlated noise," in *Allerton Conf. on Communication, Control, and Computing*, 2011.

[6] I. Carron, "Nuit blanche," in *http://nuit-blanche.blogspot.com/*.

[7] "Rice compressive sensing resources," in *http://www-dsp.rice.edu/cs*.

[8] 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.

[9] D. Donoho, "Compressed sensing," *IEEE Trans. Info. Th.*, vol. 52(4), pp. 1289–1306, April 2006.

[10] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Th.*, vol. 51(12), pp. 4203 – 4215, Dec. 2005.

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

[12] 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.

[13] H. Jung, K. H. Sung, E. Y. Kim, K. S. Nayak, and J. C. Ye, "k-t focuss: a general compressed sensing framework for high resolution dynamic mri," *Magnetic Resonance in Medicine*, 2009.

[14] N. Vaswani, "Kalman filtered compressed sensing," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2008.

[15] N. Vaswani, "LS-CS-residual (LS-CS): Compressive Sensing on Least Squares residual," *IEEE Trans. Sig. Proc.*, vol. 58(8), pp. 4108–4120, August 2010.

[16] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," in *IEEE Intl. Symp. Info. Th. (ISIT)*, 2009.

[17] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Sig. Proc.*, vol. 58(9), pp. 4595–4607, September 2010.

[18] W. Lu and N. Vaswani, "Modified Compressive Sensing for Real-time Dynamic MR Imaging," in *IEEE Intl. Conf. Image Proc. (ICIP)*, 2009.

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

[20] W. Lu and N. Vaswani, "Regularized modified bpdn for noisy sparse reconstruction with partial erroneous support and signal value knowledge," *IEEE Trans. Sig. Proc.*, January 2012.

[21] W. Lu and N. Vaswani, "Exact reconstruction conditions for regularized modified basis pursuit," *IEEE Trans. Sig. Proc.*, May 2012.

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

[23] S.G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Sig. Proc.*, vol. 41(12), pp. 3397 – 3415, Dec 1993.

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

[25] D.P. Wipf and B.D. Rao, "Sparse bayesian learning for basis selection," *IEEE Trans. Sig. Proc.*, vol. 52, pp. 2153–2164, Aug 2004.

[26] J. A. Tropp, "Just relax: Convex programming methods for identifying sparse signals," *IEEE Trans. Info. Th.*, pp. 1030–1051, March 2006.

[27] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Compte Rendus de l'Academie des Sciences, Paris, Serie I*, pp. 589–592, 2008.

[28] E. Candes and T. Tao, "The dantzig selector: statistical estimation when p is much larger than n," *Annals of Statistics*, vol. 35 (6), pp. 2313–2351, 2007.

[29] 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.

[30] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Info. Th.*, vol. 55(5), pp. 2230 – 2249, May 2009.

[31] D. Needell and J.A. Tropp., "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmonic Anal.*, vol. 26(3), pp. 301–321, May 2009.

[32] A. Carmi, P. Gurfil, and D. Kanevsky, "Methods for sparse signal recovery using kalman filtering with embedded pseudo-measurement norms and quasi-norms," *IEEE Trans. Sig. Proc.*, pp. 2405–2409, April 2010.

[33] J. Kim, O. K. Lee, and J. C. Ye, "Dynamic sparse support tracking with multiple measurement vectors using compressive music," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2012, pp. 2717–2720.

[34] D. Sejdinovic, C. Andrieu, and R. Piechocki, "Bayesian sequential compressed sensing in sparse dynamical systems," in *Allerton Conf. Communication, Control, and Computing*, 2010.

[35] J. Ziniel, L. C. Potter, and P. Schniter, "Tracking and smoothing of time-varying sparse signals via approximate belief propagation," in *Asilomar Conf. on Sig. Sys. Comp.*, 2010.

[36] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse bayesian learning," *IEEE J. Sel. Topics Sig. Proc., Special Issue on Adaptive Sparse Representation of Data and Applications in Signal and Image Processing*, vol. 5, no. 5, pp. 912–926, Sept 2011.

[37] A. Charles, M. S. Asif, J. Romberg, and C. Rozell, "Sparsity penalties in dynamical system estimation," in *Conf. Info. Sciences and Systems*, 2011.

[38] J. Garcia-Frias and I. Esnaola, "Exploiting prior knowledge in the recovery of signals from noisy random projections," in *Data Comp. Conf.*, 2007.

[39] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Sig. Proc.*, to appear.

[40] 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.

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

[42] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Info. Th.*, March 2010.

[43] Yonina C. Eldar and Moshe Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inform. Theory*, vol. 55, no. 11, pp. 5302–5316, November 2009.

[44] 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.

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

[46] A. Khajehnejad, W. Xu, A. Avestimehr, and B. Hassibi, "Weighted ℓ_1 minimization for sparse recovery with prior information," *IEEE Trans. Sig. Proc.*, 2011.

[47] C. J. Miosso, R. von Borries, M. Argez, L. Valazquez, C. Quintero, and C. Potes, "Compressive sensing reconstruction with prior information by iteratively reweighted least-squares," *IEEE Trans. Sig. Proc.*, vol. 57 (6), pp. 2424–2431, June 2009.

[48] D. Donoho, "For most large underdetermined systems of linear equations, the minimal ell-1 norm solution is also the sparsest solution," *Comm. Pure and App. Math.*, vol. 59(6), pp. 797–829, June 2006.

[49] V. Stankovic, L. Stankovic, and S. Cheng, "Compressive image sampling with side information," in *ICIP*, 2009.

[50] R. Carrillo, L.F. Polania, and K. Barner, "Iterative algorithms for compressed sensing with patially known support," in *ICASSP*, 2010.

[51] M. Stojnic, "Block-length dependent thresholds for ℓ_2/ℓ_1-optimization in block-sparse compressed sensing," in *ICASSP*, 2010.

[52] L. Jacques, "A short note on compressed sensing with partially known signal support," *Signal Processing, YEAR =*.

[53] W. Lu and N. Vaswani, "Modified bpdn for noisy compressive sensing with partially known support," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2010.

[54] M.P. Friedlander, H. Mansour, R. Saab, and O. Yilmaz, "Recovering compressively sampled signals using partial support information," vol. 58, no. 2, pp. 1122–1134, 2012.

[55] M. S. Asif and J. Romberg, "Dynamic updating for sparse time varying signals," in *Conf. Info. Sciences and Systems*, 2009.

[56] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for l1-minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1(1), pp. 143–168, 2008.

[57] D.M. Malioutov, S. Sanghavi, and A. S. Willsky, "Compressed sensing with sequential observations," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2008.

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

[59] Pierre J. Garrigues and Laurent El Ghaoui, "An homotopy algorithm for the lasso with online observations," in *Adv. Neural Info. Proc. Sys. (NIPS)*, 2008.

[60] E.J. Candes, M.B Wakin, and S.P. Boyd, "Enhancing sparsity by reweighted l(1) minimization," *Journal of Fourier Analysis and Applications*, vol. 14 (5-6), pp. 877–905, 2008.

[61] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2008.

[62] Y. Wang and W. Yin, "Sparse signal reconstruction via iterative support detection," *SIAM J. on Imag. Sc.*, vol. 3(3), pp. 462–491, 2010.

[63] D. Angelosante, G.B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *Dig. Sig. Proc. Workshop*, 2009.

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

[65] 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.

[66] L. Jacques, "A short note on compressed sensing with partially known signal support," *ArXiv preprint 0908.0660*, 2009.

[67] 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.

[68] C. Qiu and N. Vaswani, "Real-time robust principal components' pursuit," in *Allerton Conf. on Communications, Control and Computing*, 2010.