Modified-CS: Modifying Compressive Sensing for Problems with Partially Known Support

Namrata Vaswani and Wei Lu
ECE Dept., Iowa State University, Ames, IA 50011, USA, Email: namrata@iastate.edu

***Abstract*— We study the problem of reconstructing a sparse signal from a limited number of its linear projections when a part of its support is known. This may be available from prior knowledge. Alternatively, in a problem of recursively reconstructing time sequences of sparse spatial signals, one may use the support estimate from the previous time instant as the "known" part of the support. The idea of our solution (modified-CS) is to solve a convex relaxation of the following problem: find the signal that satisfies the data constraint and whose support contains the smallest number of new additions to the known support. We obtain sufficient conditions for exact reconstruction using modified-CS. These turn out to be much weaker than those needed for CS, particularly when the known part of the support is large compared to the unknown part.**

I. Introduction

Consider the problem of recursively and causally reconstructing a time sequence of sparse spatial signals (or images) from a sequence of observations, when the current observation vector contains a limited (less-than-Nyquist) number of linear projections of the current signal. The observation vector is assumed to be incoherent with respect to the sparsity basis of the signal/image [1], [2]. Important applications include real-time (causal and recursive) dynamic MRI reconstruction [3], [4], real-time single-pixel video imaging [5] or real-time time-varying spatial field estimation using sensor networks [6].

Since the introduction of compressive sensing (CS) in recent work [1], [7] the static version of the above problem has been thoroughly studied. But, with the exception of [8], [9], most existing solutions for the time series problem are non-causal or batch solutions with very high complexity since they jointly reconstruct the entire video by first collecting all the observations. The alternative solution - separately doing CS at each time (simple CS) - is causal and low complexity, but requires many more measurements for accurate reconstruction.

In recent work [10], [11], we studied the causal reconstruction problem from noisy measurements and proposed a solution called Kalman filtered CS and its non-Bayesian version, least squares CS (LS-CS). Our solutions used the empirically observed fact that the *sparsity pattern (support set of the signal) changes slowly over time*. The key idea of LS-CS was to replace CS on the observation by CS on the LS observation residual computed using the previous estimate of the support. Kalman filtered CS replaced the LS residual by the Kalman filter residual. The reason LS-CS, or Kalman filtered CS, significantly outperformed simple CS was that the signal minus its LS estimate (computed using the previous support estimate) contains much fewer significantly nonzero elements than the signal itself. But note that its exact sparsity size (total number of nonzero coefficients) is larger/equal to that of the signal. Since the number of measurements required for exact reconstruction is governed by the exact sparsity size, one thing we were not able to achieve was exact reconstruction using fewer (noiseless) measurements than those needed by CS.

Exact reconstruction using fewer measurements is the focus of the current work. The idea of our solution (modified-CS) is to modify CS for problems where part of the support is known (in the time sequence case, it is the estimated support from the previous time instant). Denote the known part of the support by T. Modified-CS solves an ℓ_1 relaxation of the following problem: find the signal that satisfies the data constraint and whose support contains the smallest number of new additions to T (or in other words the support set difference from T is smallest). We derive sufficient conditions for exact reconstruction using modified-CS. These turn out to be much weaker than the sufficient conditions required for simple CS. Experimental results showing greatly improved performance of modified-CS over simple CS are also shown.

Notice that the same idea also applies to a static reconstruction problem where we know a part of the support from prior knowledge. For example, consider MR image reconstruction using the wavelet basis as the sparsifying basis. If it is known that an image has no (or very little) black background, all (or most) elements of the lowest subband of its wavelet coefficients will be nonzero. In this case, the set T is the set of indices of the lowest subband coefficients.

A. Problem Definition

We measure an n-length vector y where

$$y = Ax \tag{1}$$

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

In a static problem, the support T is available from prior knowledge, e.g. it may be the set of the lowest subband wavelet coefficients. Typically there is a small black background in an image, so that only most (not all) lowest subband wavelet coefficients will be nonzero. The indices of the lowest subband coefficients which are zero form Δ_d. For the time series problem, $y \equiv y_t$ and $x \equiv x_t$ with support, $N_t = T \cup \Delta \setminus \Delta_d$. Here $T := \hat{N}_{t-1}$ is the support estimate from $t-1$. Also, $\Delta_d := T \setminus N_t$ is the set of indices of elements that were nonzero at $t-1$, but are now zero while $\Delta := N_t \setminus T$ is the newly added coefficients at time t. Both Δ, Δ_d are typically much smaller than $|T|$. This follows from the empirical observation that sparsity patterns change slowly [11], [4].

In our proposed solution, we compute $\hat{x}$ by assuming that the support of x contains T. When n is large enough for exact reconstruction (i.e. the conditions of Theorem 1 hold), $\hat{x} = x$ and so $\hat{x}$ can be used to compute N (and Δ_d if needed).

We assume that the measurement matrix, A, is "approximately orthonormal" for sub-matrices containing $S = (|T| + 2|\Delta|)$ or less columns, i.e. it satisfies the S-RIP [2].

Notation: We use $'$ for transpose. The notation $||c||_k$ denotes the ℓ_k norm of the vector c. For a matrix, $||M||$ denotes its spectral norm (induced ℓ_2 norm). We use the notation A_T to denote the sub-matrix containing the columns of A with indices belonging to T. For a vector, the notation $(\beta)_T$ forms a sub-vector that contains elements with indices in T.

The S-restricted isometry constant [2], δ_S, for a matrix, A, is defined as the smallest real number satisfying

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

for all subsets $T \subset [1:m]$ of cardinality $|T| \leq S$ and all real vectors c of length $|T|$. S-RIP means that $\delta_S < 1$. A related quantity, the restricted orthogonality constant [2], $\theta_{S,S'}$, is defined as the smallest real number that satisfies

$$|{c_1}' {A_{T_1}}' A_{T_2} c_2| \leq \theta_{S,S'} ||c_1||_2 ||c_2||_2 \qquad (3)$$

for all disjoint sets $T_1, T_2 \subset [1:m]$ with $|T_1| \leq S$ and $|T_2| \leq S'$ and with $S + S' \leq m$, 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 (3), it is easy to see that $||{A_{T_1}}' A_{T_2}|| \leq \theta_{S,S'}$.

II. Modified Compressive Sensing (mod-CS)

Our goal is to find the sparsest possible signal estimate whose support contains T and which satisfies the data constraint (1), i.e. we would like to find a $\hat{x}$ which solves

$$\min_\beta ||(\beta)_{T^c}||_0 \text{ subject to } y = A\beta \qquad (4)$$

where $T^c := [1:m] \setminus T$ denotes the complement of T.

As is well known, minimizing the ℓ_0 norm has combinatorial complexity. We propose to use the same trick that resulted in compressive sensing. 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

$$\min_\beta ||(\beta)_{T^c}||_1 \text{ subject to } y = A\beta \qquad (5)$$

A. Recursive Reconstruction of Signal Sequences

Consider the recursive reconstruction problem where $y \equiv y_t$ and $x \equiv x_t$ with support $N \equiv N_t$. The known part of the support, $T = \hat{N}_{t-1}$. In this case, at each time, t, we solve (5) and denote its output by $\hat{x}_{t,\text{modCS}}$. The support at t, $\hat{N}_t$ is computed by thresholding $\hat{x}_{t,\text{modCS}}$, i.e.

$$\hat{N}_t = \{i \in [1:m] : (\hat{x}_{t,\text{modCS}})_i^2 > \alpha\} \qquad (6)$$

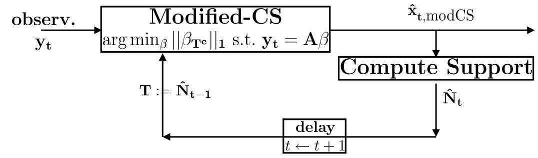


Fig. 1. Modified-CS for time sequence reconstruction

where α is a small threshold (ideally zero). With this we automatically estimate $\hat{\Delta} = \hat{N}_t \setminus T$ and $\hat{\Delta}_d = T \setminus \hat{N}_t$.

A block diagram of our proposed approach is given in Fig. 1. Note that at $t = 1$, we perform CS and use enough observations for CS to give exact reconstruction.

III. Exact Reconstruction Result

We first study the ℓ_0 version and then the actual ℓ_1 version.

A. Exact Reconstruction: ℓ_0 version of modified-CS

Consider the ℓ_0 problem, (4). Using a rank argument similar to [2, Lemma 1.2] we can show the following

Proposition 1: Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_d$, where Δ and T are disjoint and $\Delta_d \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (4). The true signal, x, is its unique minimizer if $\delta_{|T|+2|\Delta|} < 1$.

Compare this with [2, Lemma 1.2]. Since the ℓ_0 version of CS does not use the knowledge of T, it requires $\delta_{2|T|+2|\Delta|} < 1$ which is much stronger.

B. Exact Reconstruction: modified-CS

We do not solve (4) but its ℓ_1 relaxation, (5). Just like in CS, the sufficient conditions for this to give exact reconstruction will be slightly stronger. We show the following.

Theorem 1 (Exact Reconstruction): Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_d$, where Δ and T are disjoint and $\Delta_d \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (5). x is its unique minimizer if $\delta_{|T|+|\Delta|} < 1$ and if $a(2|\Delta|, |\Delta|, |T|) + a(|\Delta|, |\Delta|, |T|) < 1$, where

$$a(S, S', |T|) := \frac{\theta_{S',S} + \frac{\theta_{S',|T|}\,\theta_{S,|T|}}{1-\delta_{|T|}}}{1 - \delta_S - \frac{\theta_{S,|T|}^2}{1-\delta_{|T|}}} \qquad (7)$$

To understand the above condition better and relate it to the corresponding CS result [2, Theorem 1.3], let us simplify it. $a(2|\Delta|, |\Delta|, |T|) + a(|\Delta|, |\Delta|, |T|) \leq \frac{\theta_{|\Delta|,2|\Delta|} + \theta_{|\Delta|,|\Delta|} + \frac{\theta_{2|\Delta|,|T|}^2 + \theta_{|\Delta|,|T|}^2}{1-\delta_{|T|}}}{1-\delta_{2|\Delta|} - \frac{\theta_{2|\Delta|,|T|}^2}{1-\delta_{|T|}}}$. A sufficient condition for this is $\theta_{|\Delta|,2|\Delta|} + \theta_{|\Delta|,|\Delta|} + \frac{2\theta_{2|\Delta|,|T|}^2 + \theta_{|\Delta|,|T|}^2}{1-\delta_{|T|}} + \delta_{2|\Delta|} < 1$. Further, a sufficient condition for this is $\theta_{|\Delta|,|\Delta|} + \delta_{2|\Delta|} + \theta_{|\Delta|,2|\Delta|} + \delta_{|T|} + \theta_{|\Delta|,|T|}^2 + 2\theta_{2|\Delta|,|T|}^2 < 1$. To get a condition only in terms of δ_S's, use the fact that $\theta_{S,S'} \leq \delta_{S+S'}$. A sufficient condition is $2\delta_{2|\Delta|} + \delta_{3|\Delta|} + \delta_{|T|} + \delta_{|T|+|\Delta|}^2 + 2\delta_{|T|+2|\Delta|}^2 < 1$. Further, notice that if $|\Delta| \leq |T|$ and if $\delta_{|T|+2|\Delta|} < 1/5$, then $2\delta_{2|\Delta|} + \delta_{3|\Delta|} + \delta_{|T|} + \delta_{|T|+|\Delta|}^2 + 2\delta_{|T|+2|\Delta|}^2 < 4\delta_{|T|+2|\Delta|} + \delta_{|T|+2|\Delta|}(3\delta_{|T|+2|\Delta|}) \leq (4 + 3/5)\delta_{|T|+2|\Delta|} < 23/25 < 1$.

Corollary 1 (Exact Reconstruction): Given a sparse vector, x, whose support, $N = T \cup \Delta \setminus \Delta_d$, where Δ and T are disjoint and $\Delta_d \subseteq T$. Consider reconstructing it from $y := Ax$ by solving (5). x is its unique minimizer if $\delta_{|T|+|\Delta|} < 1$ and $\theta_{|\Delta|,|\Delta|} + \delta_{2|\Delta|} + \theta_{|\Delta|,2|\Delta|} + \delta_{|T|} + \theta^2_{|\Delta|,|T|} + 2\theta^2_{2|\Delta|,|T|} < 1$. This holds if $2\delta_{2|\Delta|} + \delta_{3|\Delta|} + \delta_{|T|} + \delta^2_{|T|+|\Delta|} + 2\delta^2_{|T|+2|\Delta|} < 1$. This, in turn, holds if $|\Delta| \le |T|$ and $\delta_{|T|+2|\Delta|} < 1/5$.

Compare the above with the requirement for CS: $2\delta_{2(|T|+|\Delta|)} + \delta_{3(|T|+|\Delta|)} < 1$ which holds if $\delta_{3(|T|+|\Delta|)} < 1/3$. It is clear that if $|\Delta|$ is small compared to $|T|$, $\delta_{|T|+2|\Delta|} < 1/5$ is a much weaker requirement.

C. Proof of Theorem 1

For the proof of Theorem 1, we use an approach similar to that used to prove [2, Theorem 1.3]. Suppose that we want to minimize a convex function $J(\beta)$ subject to $A\beta = y$ and that J is differentiable. The Lagrange multiplier optimality condition requires that there exists a Lagrange multiplier, w, s.t. $\nabla J(\beta) - A'w = 0$. Thus for x to be a solution we need $A'w = \nabla J(x)$. In our case, $J(x) = ||x_{T^c}||_1 = \sum_{j \in T^c} |x_j|$. Thus $(\nabla J(x))_j = 0$ for $j \in T$ and $(\nabla J(x))_j = \text{sgn}(x_j)$ for $j \in \Delta$. For $j \notin T \cup \Delta$, $x_j = 0$. Since J is not differentiable at 0, we require that $(A'w)_j = {A_j}'w = w'A_j$ lie in the subgradient set of $J(x_j)$ at 0, which is the set $[-1, 1]$. In summary, we need a w that satisfies $w'A_j = 0$ if $j \in T$, $w'A_j = \text{sgn}(x_j)$ if $j \in \Delta$, and, $|w'A_j| \le 1$, if $j \notin T \cup \Delta$. We show below that by using the above conditions but with $|w'A_j| \le 1$ replaced by $|w'A_j| < 1$ for $j \notin T \cup \Delta$, we get a set of sufficient conditions to ensure that x is the unique solution of (5).

Lemma 1: The sparse signal, x, with support as defined in Theorem 1, is the unique minimizer of (5) if $\delta_{|T|+|\Delta|} < 1$ and if we can find a vector w satisfying

1) $w'A_j = 0$ if $j \in T$
2) $w'A_j = \text{sgn}(x_j)$ if $j \in \Delta$
3) $|w'A_j| < 1$, if $j \notin T \cup \Delta$

Proof. Standard convex arguments give that there is at least one minimizer of (5). We need to prove that, if the conditions of the lemma hold, any minimizer, β, of (5) is equal to x. Since x also satisfies the data constraint,

$$||(\beta)_{T^c}||_1 \le ||(x)_{T^c}||_1 := \sum_{j \in \Delta} |x_j| \tag{8}$$

for any minimizer β. Take a w that satisfies the conditions of the lemma. Recall that x is zero outside of $T \cup \Delta$. Then,

$$\begin{aligned} ||(\beta)_{T^c}||_1 &= \sum_{j \in \Delta} |x_j + (\beta_j - x_j)| + \sum_{j \notin T \cup \Delta} |\beta_j| \\ &\ge \sum_{j \in \Delta} |x_j + (\beta_j - x_j)| + \sum_{j \notin T \cup \Delta} w'A_j\beta_j \\ &\ge \sum_{j \in \Delta} \text{sgn}(x_j)(x_j + (\beta_j - x_j)) + \sum_{j \notin T \cup \Delta} w'A_j\beta_j \\ &= \sum_{j \in \Delta} |x_j| + \sum_{j \in \Delta} w'A_j(\beta_j - x_j) + \sum_{j \notin T \cup \Delta} w'A_j\beta_j \\ &\quad + \sum_{j \in T} w'A_j(\beta_j - x_j) \\ &= ||x_{T^c}||_1 + w'(A\beta - Ax) = ||x_{T^c}||_1 \end{aligned} \tag{9}$$

Now, the only way (9) and (8) can hold simultaneously is if all inequalities in (9) are actually equalities. Consider the first inequality. Since $|w'A_j|$ is strictly less than 1, for all $j \notin T \cup \Delta$, the only way $\sum_{j \notin T \cup \Delta} |\beta_j| = \sum_{j \notin T \cup \Delta} w'A_j\beta_j$ is if $\beta_j = 0$ for all $j \notin T \cup \Delta$.

Since both β and x solve (5), $y = Ax = A\beta$. Since $\beta_j = 0 = x_j$ for all $j \notin T \cup \Delta$, this means that $y = A_{T\cup\Delta}(\beta)_{T\cup\Delta} = A_{T\cup\Delta}(x)_{T\cup\Delta}$ or that $A_{T\cup\Delta}((\beta)_{T\cup\Delta} - (x)_{T\cup\Delta}) = 0$. Since $\delta_{|T|+|\Delta|} < 1$, $A_{T\cup\Delta}$ is full rank and so the only way this can happen is if $(\beta)_{T\cup\Delta} = (x)_{T\cup\Delta}$. Thus any minimizer, $\beta = x$, i.e. x is the unique minimizer of (5). This proves the claim.■

Next, we begin by developing a lemma (Lemma 2) that constructs a w which satisfies ${A_T}'w = 0$ and ${A_{T_d}}'w = c$ for any given vector c and any set T_d disjoint with T of size $|T_d| \le S$. The lemma also bounds $|{A_j}'w|$ for all $j \notin T \cup T_d \cup E$ where E is called an "exceptional set". Finally, we use this lemma to find a w that satisfies the conditions of Lemma 1.

Lemma 2: Given the known part of the support, T, of size $|T|$. Let S, S' be such that $\delta_{|T|+S} < 1$ and $|T| + S + S' \le m$. Let c be a vector supported on a set T_d, that is disjoint with T, of size $|T_d| \le S$. Then there exists a vector w s.t. ${A_j}'w = c_j$, $\forall\ j \in T_d$, and ${A_j}'w = 0$, $\forall\ j \in T$. Also, there exists an exceptional set E, disjoint with $T \cup T_d$, of size $|E| < S'$ s.t.

$$\begin{aligned} |{A_j}'w| &\le \frac{a(S, S', |T|)}{\sqrt{S'}} ||c||_2 \ \forall j \notin T \cup T_d \cup E \text{ and} \\ ||{A_E}'w||_2 &\le a(S, S', |T|)||c||_2 \end{aligned} \tag{10}$$

where $a(S, S', |T|)$ is defined in (7). Also, $||w|| \le K(S, |T|)||c||_2$, where

$$K(S, |T|) := \frac{\sqrt{1 + \delta_S}}{1 - \delta_S - \frac{\theta^2_{S,|T|}}{1 - \delta_{|T|}}} \tag{11}$$

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

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

We need to find a γ s.t. ${A_{T_d}}'w = c$, i.e. ${A_{T_d}}'M\gamma = c$. Let $\gamma = M'A_{T_d}\eta$. Then $\eta = ({A_{T_d}}'MM'A_{T_d})^{-1}c = ({A_{T_d}}'MA_{T_d})^{-1}c$ (since $MM' = M^2 = M$). Thus,

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

Consider a set T'_d of size $|T'_d| \le S'$ disjoint with $T \cup T_d$. Then

$$||{A_{T'_d}}'w||_2 \le ||{A_{T'_d}}'MA_{T_d}||\ ||({A_{T_d}}'MA_{T_d})^{-1}||\ ||c||_2 \tag{14}$$

Consider the first term from the RHS of (14).

$$\begin{aligned} ||{A_{T'_d}}'MA_{T_d}|| &\le ||{A_{T'_d}}'A_{T_d}|| + ||{A_{T'_d}}'A_T({A_T}'A_T)^{-1}{A_T}'A_{T_d}|| \\ &\le \theta_{S',S} + \frac{\theta_{S',|T|}\ \theta_{S,|T|}}{1 - \delta_{|T|}} \end{aligned} \tag{15}$$

Consider the second term from the RHS of (14).

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

Now, ${A_{T_d}}'MA_{T_d} = {A_{T_d}}'A_{T_d} - {A_{T_d}}'A_T({A_T}'A_T)^{-1}{A_T}'A_{T_d}$. This is the difference of two non-negative definite matrices. It is easy to see that if B_1 and B_2 are two non-negative definite matrices, then $\lambda_{\min}(B_1 - B_2) \geq \lambda_{\min}(B_1) - \lambda_{\max}(B_2)$. Let $B_1 := A_{T_d}{}' A_{T_d}$ and $B_2 := A_{T_d}{}' A_T (A_T{}' A_T)^{-1} A_T{}' A_{T_d}$.

Then $\lambda_{\min}(B_1) \geq (1 - \delta_S)$. Also, $\lambda_{\max}(B_2) = ||B_2|| \leq \frac{||(A_{T_d}{}' A_T)||^2}{1-\delta_{|T|}} \leq \frac{\theta^2_{S,|T|}}{1-\delta_{|T|}}$. Thus,

$$||(A_{T_d}{}' M A_{T_d})^{-1}|| \leq \frac{1}{1 - \delta_S - \frac{\theta^2_{S,|T|}}{1-\delta_{|T|}}} \qquad (17)$$

as long as the denominator is positive. Using (15) and (17) to bound (14), we get

$$||A_{T_d'}{}' w||_2 \leq a(S, S', |T|) \ \ ||c||_2 \qquad (18)$$

where $a(\ )$ is defined in (7). Notice that $a(\ )$ is a non-decreasing function of all its arguments.

Define an "exceptional set" E as

$$E := \{j \in (T \cup T_d)^c : |A_j{}' w| > \frac{a(S, S', |T|)}{\sqrt{S'}} ||c||_2\} \qquad (19)$$

Notice that $|E|$ must obey $|E| < S'$ since otherwise we can contradict (18) by taking $T_d' \subseteq E$.

Since $|E| < S'$ and E is disjoint with $T \cup T_d$, (18) holds for $T_d' \equiv E$. Finally, notice that

$$\begin{aligned} ||w||_2 &\leq ||M A_{T_d} (A_{T_d}{}' M A_{T_d})^{-1}|| \ ||c||_2 \\ &\leq ||M|| \ ||A_{T_d}|| \ ||(A_{T_d}{}' M A_{T_d})^{-1}|| \ ||c||_2 \\ &\leq \frac{\sqrt{1+\delta_S}}{1 - \delta_S - \frac{\theta^2_{S,|T|}}{1-\delta_{|T|}}} ||c||_2 = K(S, |T|) ||c||_2 \qquad (20) \end{aligned}$$

This proves the lemma. ■

Proof of Theorem 1. Let us apply Lemma 2 iteratively to make the size of the exceptional set E smaller and smaller. At iteration zero, apply Lemma 2 with $T_d \equiv \Delta$ (so that $S \equiv |\Delta|$), $c_j \equiv \text{sgn}(x_j), \ \forall \ j \in \Delta$ (so that $||c||_2 = \sqrt{|\Delta|}$), and with $S' \equiv |\Delta|$. Call the exceptional set $T_{d,1}$.Thus there exists a w_1 and an exceptional set $T_{d,1}$ s.t.

$$\begin{aligned} A_j{}' w_1 &= \text{sgn}(x_j), \ \forall \ j \in \Delta \\ A_j{}' w_1 &= 0, \ \forall \ j \in T \\ |T_{d,1}| &\leq S' \equiv |\Delta| \\ ||A_{T_{d,1}}{}' w_1||_2 &\leq a(|\Delta|, |\Delta|, |T|) \sqrt{|\Delta|} \\ |A_j{}' w_1| &\leq a(|\Delta|, |\Delta|, |T|), \ \forall j \notin T \cup \Delta \cup T_{d,1} \\ ||w_1|| &\leq K(|\Delta|, |T|) \sqrt{|\Delta|} \qquad (21) \end{aligned}$$

At iteration n, apply Lemma 2 with $T_d \equiv \Delta \cup T_{d,n}$ (so that $S \equiv 2|\Delta|$), $c_j \equiv 0, \ \forall \ j \in \Delta$ and $c_j \equiv A_j{}' w_n, \ \forall \ j \in T_{d,n}$ and with $S' \equiv |\Delta|$. Call the exceptional set $T_{d,n+1}$. Thus there exists a w_{n+1} and an exceptional set $T_{d,n+1}$ that satisfy

$$\begin{aligned} A_j{}' w_{n+1} &= 0, \ \forall \ j \in \Delta \\ A_j{}' w_{n+1} &= A_j{}' w_n, \ \forall \ j \in T_{d,n} \\ A_j{}' w_{n+1} &= 0 \ \forall \ j \in T \\ |T_{d,n+1}| &\leq S' \equiv |\Delta| \\ ||A_{T_{d,n+1}}{}' w_{n+1}||_2 &\leq a(2|\Delta|, |\Delta|, |T|) ||A_{T_{d,n}}{}' w_n|| \\ |A_j{}' w_{n+1}| &\leq \frac{a(2|\Delta|, |\Delta|, |T|)}{\sqrt{|\Delta|}} ||A_{T_{d,n}}{}' w_n|| \\ & \quad \forall j \notin T \cup \Delta \cup T_{d,n} \cup T_{d,n+1} \\ ||w_{n+1}|| &\leq K(2|\Delta|, |T|) ||A_{T_{d,n}}{}' w_n|| \qquad (22) \end{aligned}$$

The last three equations above simplify to

$$\begin{aligned} & ||A_{T_{d,n+1}}{}' w_{n+1}||_2 \leq a(2|\Delta|, |\Delta|, |T|)^n a(|\Delta|, |\Delta|, |T|) \sqrt{|\Delta|} \\ & |A_j{}' w_{n+1}| \leq a(2|\Delta|, |\Delta|, |T|)^n a(|\Delta|, |\Delta|, |T|), \\ & \forall j \notin T \cup \Delta \cup T_{d,n} \cup T_{d,n+1} \qquad (23) \\ & ||w_{n+1}|| \leq K(2|\Delta|, |T|) a(2|\Delta|, |\Delta|, |T|)^{n-1} a(|\Delta|, |\Delta|, |T|) \sqrt{|\Delta|} \qquad (24) \end{aligned}$$

Now, assume that $a(2|\Delta|, |\Delta|, |T|) < 1$ and define

$$w := \sum_{n=1}^{\infty} (-1)^{n-1} w_n \qquad (25)$$

Since $a(2|\Delta|, |\Delta|, |T|) < 1$, the above summation is absolutely convergent and so w is a well defined vector. Also,

$$\begin{aligned} A_j{}' w &= \text{sgn}(x_j), \ \forall \ j \in \Delta \\ A_j{}' w &= 0, \ \forall \ j \in T \qquad (26) \end{aligned}$$

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

$$A_j{}' w = \sum_{n: j \notin T_{d,n} \cup T_{d,n-1}} (-1)^{n-1} A_j{}' w_n, \ \forall j \notin T \cup \Delta \qquad (27)$$

Consider a given n in the above summation. Since $j \notin T_{d,n} \cup T_{d,n-1} \cup T \cup \Delta$, we can use (23) to get $|A_j{}' w_n| \leq a(2|\Delta|, |\Delta|, |T|)^{n-1} a(|\Delta|, |\Delta|, |T|)$. Thus, for $j \notin T \cup \Delta$,

$$|A_j{}' w| \leq \sum_{n: j \notin T_{d,n} \cup T_{d,n-1}} a(2|\Delta|, |\Delta|, |T|)^{n-1} a(|\Delta|, |\Delta|, |T|) \qquad (28)$$

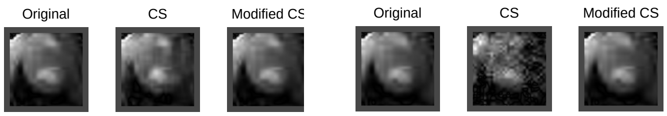
If $a(2|\Delta|, |\Delta|, |T|) < 1$, this simplifies to

$$|A_j{}' w| \leq \frac{a(|\Delta|, |\Delta|, |T|)}{1 - a(2|\Delta|, |\Delta|, |T|)}, \ \forall j \notin T \cup \Delta \qquad (29)$$

Thus, if we can assume that $a(2|\Delta|, |\Delta|, |T|) + a(|\Delta|, |\Delta|, |T|) < 1$, then we will have

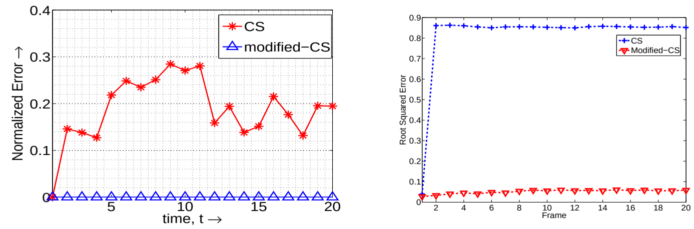
$$|A_j{}' w| < 1, \ \forall j \notin T \cup \Delta \qquad (30)$$

Thus, from (26) and (30), if $a(2|\Delta|, |\Delta|, |T|) + a(|\Delta|, |\Delta|, |T|) < 1$ then, we have found a w that satisfies the three conditions of Lemma 1. Applying Lemma 1, the exact reconstruction claim of Theorem 1 follows. ■



(a) Fourier meas. (b) Gaussian meas.

Fig. 2. (a) Reconstructing a 32×32 sparsified cardiac image ($m = 1024$) from $n = 0.19m = 195$ random Fourier measurements. Support size $|T \cup \Delta| = 107$ and $|T| = 64$. Modified-CS achieved exact reconstruction, while the CS reconstruction error (square root of normalized MSE) was 13%. (b) Reconstruction using $n = 0.29m$ random Gaussian measurements. Modified-CS achieved exact reconstruction, while the CS reconstruction error was 34%.



(a) Sparsified seq, Fourier meas (b) Real seq, Gaussian meas

Fig. 3. (a) Exact reconstruction of a sparsified cardiac sequence from only $n = 0.16m$ random Fourier measurements (MR imaging). Support size, $|N_t| \approx 0.1m$. Simple CS (referred to as CS in the figure) has very large (20-25%) error while modified-CS gives exact reconstruction. (b) Reconstructing a real cardiac sequence from $n = 0.19m$ random Gaussian measurements. We plot the square root of normalized MSE everywhere.

IV. Simulation Results

We first evaluated the static problem. The image used was a sparsified 32×32 block ($m = 1024$) of a cardiac image. This was obtained by taking a discrete wavelet transform (DWT) of the original image block, retaining the largest 107 coefficients (corresponds to 99% of image energy) while setting the rest to zero and taking the inverse DWT. A 2-level DWT served as the sparsifying basis. We used its lowest subband as the known part of the support, T. Thus, $|T| = 64$. Support size $|N| = 107$. We show reconstruction from only $n = 0.19m = 195$ random Fourier measurements in Fig. 2(a). Modified-CS achieved exact reconstruction, while CS reconstruction error (square root of normalized MSE) was 13%. Notice that $195 < 2|N| = 214$, which is the minimum n necessary for exact reconstruction using CS for a $|N|$-sparse vector. Comparison for random-Gaussian measurements is shown in Fig. 2(b).

Next, we evaluated the time sequence problem using a sparsified cardiac image sequence created the same way as above. At $t = 1$, we did simple CS and used $n = 0.5m = 256$ random Fourier measurements. For $t > 1$ we did modified-CS and used only $n = 0.16m = 164$ measurements. The size of the change in the support from $t-1$ to t, $|\Delta| \approx 0.01m = 10$ or less. The support size, $|N_t| \approx 0.1m = 103$. We show the reconstruction results in Fig. 3(a). Simple CS (referred to as CS in the figure) has very large (20-25%) error while modified-CS gives exact reconstruction.

Finally, we evaluated modified-CS for a real cardiac sequence (not sparsified). In this case, the wavelet transform is only compressible. The comparison is given in Fig. 3(b).

V. Conclusions and Future Work

We studied the problem of reconstructing a sparse signal from a limited number of its linear projections when a part of its support is known. This may be available from prior knowledge. Alternatively, in a problem of recursively reconstructing time sequences of sparse spatial signals, one may use the support estimate from the previous time instant as the "known" part of the support. We derived sufficient conditions for exact reconstruction using our proposed solution - modified-CS - and discussed why these are weaker than the sufficient conditions required by simple CS. Experiments showing greatly improved performance of modified-CS over simple CS are also given.

Future work includes (a) bounding the reconstruction error of modified-CS for compressible signals, (b) combining modified-CS with Least Squares CS [11] for the noisy measurements case, and (c) developing Bayesian extensions which also use knowledge of the previously reconstructed signal values and analyzing their performance. (d) Whenever exact reconstruction does not occur, an important question to answer is when will the algorithm be stable over time, i.e. under what conditions will the reconstruction error remain bounded. This automatically holds for modified-CS for noiseless measurements if the assumption of Theorem 1 holds at all times. It has been shown to hold with high probability for LS-CS and KF-CS for noisy measurements in [11] under strong assumptions. Our goal would be to prove it for modified-CS for noisy measurements under weaker assumptions.

Acknowledgment

This research was supported by NSF grant ECCS-0725849.

References

[1] 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.

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

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

[4] C. Qiu, W. Lu, and N. Vaswani, "Real-time dynamic mri reconstruction using kalman filtered cs," in *IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP)*, 2009.

[5] 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.

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

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

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

[9] H. Jung, K. H. 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*, To appear.

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

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