

STABILITY (OVER TIME) OF REGULARIZED MODIFIED CS(NOISY) FOR RECURSIVE CAUSAL SPARSE RECONSTRUCTION

Fardad Raisali, Namrata Vaswani

Iowa State University, ECE Department
fardadr@iastate.edu, namrata@iastate.edu

ABSTRACT

In this work, we study the “stability” of Regularized modified CS(noisy) for recursive reconstruction of sparse signal sequences from noisy measurements. By “stability” we mean that the number of misses from the current support estimate; the number of extras in it; and the ℓ_2 norm of the reconstruction error remain bounded by a *time-invariant* value at all times. The concept is meaningful only if the support error bounds are small compared to the signal support size.

Regularized modified CS(noisy) is the noisy relaxation of regularized modified CS. The key assumption that reg-mod-CSN uses is that both the sparse signal’s support and its nonzero signal values change slowly over time. Denote the support estimate from the previous time by T . Modified-CS tries to find a signal that is sparsest outside of T and satisfies the data constraint. Denote the signal estimate from the previous time by μ_T . Reg-mod-CSN augments mod-CS by also putting the ℓ_2 distance of the current solution from μ_T as a constraint.

Index Terms— CS, mod-BPDN, reg-mod-CSN

1. INTRODUCTION

In this work, we study the “stability” of Regularized modified CS(noisy) for recursive reconstruction of sparse signal sequences from noisy measurements. By “stability” we mean that the number of misses from the current support estimate; the number of extras in it; and the ℓ_2 norm of the reconstruction error remain bounded by a *time-invariant* value at all times. The concept is meaningful only if the support error bounds are small compared to the signal support size.

Regularized modified CS(noisy)(reg-mod-CSN) is the noisy relaxation of regularized modified CS (reg-mod-cs) proposed in [1]. The key assumption that reg-mod-CSN uses is that both the sparse signal’s support and its nonzero signal values change slowly over time. This assumption has been empirically verified in earlier work [2] for medical image sequences. Denote the support estimate from the previous time by T . Modified-CS(mod-CS) [3] tries to find a signal that is sparsest outside of T and satisfies the data constraint. Denote the signal estimate from the previous time by μ_T . Reg-mod-CSN tries to find a signal that is sparsest outside of T ; is “close” enough to μ_T on T ; and satisfies the data constraint.

Other algorithms for recursive reconstruction include our older work on Least Squares CS-residual (LS-CS) and Kalman filtered CS-residual (KF-CS) [4, 5, 6]; modified-CS [1]; homo-

topy methods [7] (use past reconstructions to speed up current optimization but not to improve reconstruction error with fewer measurements); and [8] (a recent modification of KF-CS). Another recent work on CS for time-varying signals [9] proposed a series of causal but batch approaches that assume a *time-invariant* support. Two other algorithms that are also designed for static CS with partial knowledge of support include [10] and [11]. The work of [10] proposed an approach similar to modified-CS but did not analyze it and also did not show real experiments either. The work of [11], which appeared in parallel with modified-CS, assumed a probabilistic prior on the support.

To the best of our knowledge, stability of recursive sparse reconstruction algorithms has not been studied in any other work except in our older works [6, 3] for LS-CS and modified-CS respectively. The limitation of the result of [6] was that it assumed a signal model where support changes are only allowed every-so-often. But this assumption often does not hold in practice, e.g. for dynamic MRI sequences, support changes occur at every time. This limitation was removed in [3] where we used a signal model that allows support changes at every time t . In this work, we use the same signal model and our overall approach is also motivated by that of [3] for modified-CS. But there are significant differences since for reg-mod-CSN, the current reconstruction also depends on the previously reconstructed signal values (not just its support estimate), which makes its stability analysis more difficult.

The paper is organized as follows. We begin by giving notation and problem definition in Section 2. In Section 3, Regularized Modified CSN is developed and we obtain the reconstruction error bound under certain conditions. In Section 4 we give the stepwise algorithm which implements the Regularized Modified CSN over time. In Section 5 we show error stability over time under mild assumptions, and for a fairly realistic signal change model. The performance of reg-mod-CSN, modified BPDN [12], modified CS [1] and CS is compared using simulation in Section 6.

2. NOTATION AND PROBLEM DEFINITION

The set operations \cup, \cap, \setminus have their usual meanings. \emptyset denotes the empty set. We use T^c to denote the complement of a set T w.r.t. $[1, m] := [1, 2, \dots, m]$, i.e. $T^c := [1, m] \setminus T$. $|T|$ denotes the cardinality of T . 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 . $\|v\|_k$ denotes the ℓ_k norm of a vector v . If just $\|v\|$ is used, it refers to $\|v\|_2$. Similarly, for a matrix M , $\|M\|_k$ denotes its induced k -norm, while just $\|M\|$ refers to $\|M\|_2$. M' denotes the transpose of 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 . We obtain an n -length measurement vector y_t by:

$$y_t = Ax_t + w_t$$

A is an $n \times m$ ($n < m$) matrix that we call the measurement matrix. x_t is an m length sparse vector with support N_t , y_t is the n length observation vector and w_t is the n length noise observation vector with $\|w_t\| \leq \rho$. We assume partial knowledge of support and denote it by T_t . Also we assume partial knowledge of the signal estimate on T_t , and denote it by $(\mu_t)_{T_t}$. The signal estimate is assumed to be zero along T_t^c .

Our goal is to recursively estimate 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 time t . Recursive recovery ensures both computational and storage complexity remains the same as that of simple CS (CS done for each time instant separately). The S -restricted isometry constant δ_S , for a matrix, A , proposed in [13], is defined as the smallest positive number satisfying

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

for all subsets of T with cardinality $|T| \leq S$ and all real vectors c of length $|T|$. The S, S' restricted orthogonality constant, $\theta_{S, S'}$, proposed in [13], is defined as the smallest real number satisfying

$$\langle A_{T_1} c_1, A_{T_2} c_2 \rangle \leq \theta_{S, S'} \|c_1\| \|c_2\|$$

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

Definition 1 ($T_t, \Delta_t, \Delta_{e,t}$). We use T_t to denote the support estimate at time t from the previous time $t-1$. We use $\Delta_t := N_t \setminus T_t$ to denote the unknown part of the support estimate and $\Delta_{e,t} := T_t \setminus N_t$ to denote the "erroneous" part of T_t .

Definition 2 ($\tilde{T}_t, \tilde{\Delta}_t, \tilde{\Delta}_{e,t}$). We use \tilde{T}_t to denote the final support estimate at current time t . We use $\tilde{\Delta}_t := N_t \setminus \tilde{T}_t$ and $\tilde{\Delta}_{e,t} := \tilde{T}_t \setminus N_t$.

3. ERROR BOUND FOR REGULARIZED MODIFIED CSN

In this section we introduce regularized Modified CSN and derive the bound for its reconstruction error. In this section we consider the case where there is one measurement vector, y , and a signal vector x .

$$y := Ax + w, \quad \text{where } \|w\| \leq \epsilon$$

Let N denote the support of x , i.e $N := \{i : x_i > 0\}$. Assume that we know partial part of support denoted by T . We define $\Delta = N \setminus T$. In addition to the measurements and partial knowledge of signal support, T , we know that signal x satisfies

$$\|x_T - \mu_T\|_2 \leq \gamma$$

where μ_T is the partial knowledge of the signal estimate on T . Regularized Modified CSN solves the following problem.

$$\min \|\beta_{T^c}\|_1 \text{ s.t } \|y - A\beta\|_2 \leq \epsilon \text{ and } \|\beta_T - \mu_T\|_2 \leq \gamma \quad (1)$$

Theorem 1. Let $u := |T|$ and $k := |\Delta|$. Assume that $\delta_u < 1$, $1 - \delta_{2k} - \theta_{k, 2k} > 0$ and $\|x_T - \mu_T\| \leq \gamma$. Then the solution \hat{x} to (1) obeys

$$\|x - \hat{x}\| \leq C_{u,k}\epsilon + D_{u,k}\gamma$$

where

$$C_{u,k} = 2 \frac{\sqrt{1 + \delta_u}}{1 - \delta_u} + 2 \frac{(2 + \frac{(\sqrt{2}+1)\theta_{u,k}}{1-\delta_u})\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k} - \theta_{k, 2k}}$$

$$D_{u,k} = \frac{2(2 + \frac{(\sqrt{2}+1)\theta_{u,k}}{1-\delta_u})\theta_{u, 2k}}{1 - \delta_{2k} - \theta_{k, 2k}}$$

Proof: Proof is given in Appendix .

4. ALGORITHM FOR REGULARIZED MODIFIED CSN OVER TIME

Regularized Modified CSN was introduced in the previous section as the solution to the problem of (1). In other word, Regularized Modified CSN is the solution to the problem of sparse reconstruction (3) with partial knowledge of the support and signal value on the known support. For recursively reconstruction a time sequence of sparse signals, we use the support estimate from the previous time, \tilde{T}_{t-1} as the set T and use the signal estimate from the previous time on this support, $(\hat{x}_{t-1})_T$ as the $(\mu)_T$. At the initial time, $t = 0$, we let T be the empty set, i.e we do simple CS. Therefore at $t = 0$ we need more measurements, $n_0 > n$. Denote the $n_0 \times m$ measurement matrix used at $t = 0$ by A_0 .

We summarize the Regularized modified CSN algorithm in Algorithm 1. Here α denote the support estimation threshold. Consider that in step 3 of algorithm we update our support estimation as \tilde{T}_t at time t .

Algorithm 1 Regularized Modified CSN over time

For $t \geq 0$, do

1. *Simple CS.* If $t = 0$, set $T_0 = \emptyset$ and compute \hat{x}_0 as the solution of

$$\min \|\beta\|_1 \text{ s.t } \|y_0 - A\beta\|_2 \leq \epsilon \quad (2)$$

2. *Regularized Modified CSN.* If $t > 0$, set $T_t = \tilde{T}_{t-1}$ and compute \hat{x}_t as the solution of

$$\min \|\beta_{T_t^c}\|_1 \text{ s.t } \|y - A\beta\|_2 \leq \epsilon \text{ and } \|\beta_{T_t} - \mu_{T_t}\|_2 \leq \gamma \quad (3)$$

3. *Estimate the Support.* Compute \tilde{T}_t as

$$\tilde{T}_t = \{i \in [1, m] : |(\hat{x}_t)_i| > \alpha\} \quad (4)$$

4. Set $\mu = \hat{x}_t$. Output \hat{x}_t . Feedback μ and \tilde{T}_t .
-

5. STABILITY RESULT FOR REGULARIZED MODIFIED CSN OVER TIME

5.1. Signal Model

The proposed algorithm *does not* assume any signal model. But to prove its stability, we need certain assumptions on the signal changes over time. We use the Signal Model introduced in [3] as our signal sequence over time.

Assume the following

1. (addition) At each $t > 0$, S_a new coefficients get added to the support at magnitude r . Denote this set by A_t .
2. (increase) At each $t > 0$, the magnitude of S_a coefficients which had magnitude $(j - 1)r$ at $t - 1$ increases to jr . This occurs for all $2 \leq j \leq d$. Thus the maximum magnitude reached by any coefficient is $M := dr$.
3. (decrease) At each $t > 0$, the magnitude of S_a coefficients which had magnitude $(j + 1)r$ at $t - 1$ decreases to jr . This occurs for all $1 \leq j \leq (d - 1)$.
4. (removal) At each $t > 0$, S_a coefficients which had magnitude r at $t - 1$ get removed from the support (magnitude becomes zero). Denote this set by R_t .
5. (initial time) At $t = 0$, the support size is S_0 and it contains $2S_a$ elements each with magnitude $r, 2r, \dots, (d - 1)r$, and $(S_0 - (2d - 2)S_a)$ elements with magnitude M .

Notice that, in the above model, the size and composition of the support at any t is the same as that at $t = 0$. Also, at each t , there are S_a new additions and S_a removals. The new coefficient magnitudes increase gradually at rate r and do not increase beyond a maximum value $M := dr$. Similarly for decrease. The support size is always S_0 and the signal power is $(S_0 - (2d - 2)S_a)M^2 + 2S_a \sum_{j=1}^{d-1} j^2 r^2$.

5.2. Stability Result

In this part we are finding the conditions under which the error bound for proposed algorithm remains bounded. For this purpose, we should develop the conditions for a certain set of large coefficients to definitely get detected and the elements of Δ_e to definitely get deleted.

In the following lemma we bring some simple facts that we use through the proof of Theorem 2.

Proposition 1. *In the third step of Algorithm 1 we have the following facts*

1. An $i \in N_t$ will definitely get detected if $|x_i| > \alpha + \|x_t - \hat{x}_t\|$. This follows since $\|x_t - \hat{x}_t\| \geq \|x_t - \hat{x}_t\|_\infty \geq |x_t - \hat{x}_t|_i$
2. Similarly, all $i \in \tilde{\Delta}_{e,t}$ (the zero elements of \tilde{T}_t) will definitely not get detected if $\alpha \geq \|x_t - \hat{x}_t\|$. This is true since if $(x_t)_i = 0$ and $(\hat{x}_t)_i$ get detected as nonzero value $(\hat{x}_t)_i$, then $\alpha \leq \|(\hat{x}_t)_i - (x_t)_i\| \leq \|x_t - \hat{x}_t\|$ which is a contradiction with the assumption $\alpha \geq \|x_t - \hat{x}_t\|$.

Proposition 2. *Under proposed Signal Model we have*

$$\|(x_t)_{T_t} - (\hat{x}_{t-1})_{T_t}\|_2 \leq \|x_{t-1} - \hat{x}_{t-1}\|_2 + \sqrt{2dS_a}r$$

Proof: Proof is straightforward from the fact that by Signal Model we have $\|x_t - x_{t-1}\|_2 \leq \sqrt{2dS_a}r$.

In the following theorem we bring the conditions that makes the error bounded by a time independent value.

Theorem 2 (Stability of Regularized Modified CSN over time). *Assume the Signal Model given above. For a d_0 such that $1 \leq d_0 \leq d$, set $S_1 = (2d_0 - 2)S_a$. If the following conditions hold*

1. $\min(1 - \delta_{S_0}, 1 - \delta_{2S_1} - \theta_{S_1, 2S_1} - 2(2 + \frac{(\sqrt{2}+1)\theta_{S_0, S_1}}{1 - \delta_{S_0}})\theta_{S_0, 2S_1}) > 0$
2. $\gamma = \frac{C_{S_0, S_1}\epsilon + \sqrt{2dS_a}r}{1 - D_{S_0, S_1}}$
3. $\alpha = C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma$
4. r satisfy

$$r \geq \frac{2C_{S_0, S_1}\epsilon}{d_0(1 - D_{S_0, S_1}) - 2D_{S_0, S_1}\sqrt{2dS_a}}$$

(it ensures that $d_0 r \geq 2 \times (C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma)$)

5. n_0 is large enough so that $\|x_0 - \hat{x}_0\| \leq C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma$

Then we can conclude that

1. $|T_t| \leq S_0, |\Delta_t| \leq S_1$
2. $\|x_t - \hat{x}_t\| \leq C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma$

Proof: Our approach for the proof is based on induction. Assume that the results hold at $t - 1$. Using condition 2 and Proposition 2, we can show that $\|x_t - \hat{x}_{t-1}\| \leq \gamma$. Condition 2 is meaningful when $D_{S_0, S_1} < 1$ which is equivalent to the second term of condition 1.

Next, we try to show that $|T_t| \leq S_0$ and $|\Delta_t| \leq S_1$. Finally, this, along with conditions 1 and 2 allows us to apply Theorem 1 to get the bound on $\|x_t - \hat{x}_{t-1}\|$. To show $|T_t| \leq S_0$ and $|\Delta_t| \leq S_1$, we first use the induction assumption, conditions 3 and 4 and Proposition 1 to bound $|\tilde{T}_{t-1}|$ and $|\tilde{\Delta}_{t-1}|$; and then use the signal model to bound $|T_t|$ and $|\Delta_t|$. The complete proof is given in the Appendix.

5.3. Discussion of Theorem

We can observe some results from Theorem 2. As we can see in the first condition of Theorem 2, reg-mod-CSN needs two requirements to hold, $\delta_{S_0} < 1$ and $1 - \delta_{2S_1} - \theta_{S_1, 2S_1} - \alpha > 0$ where $\alpha = 2(2 + \frac{(\sqrt{2}+1)\theta_{S_0, S_1}}{1 - \delta_{S_0}})\theta_{S_0, 2S_1}$. Consider the case where $\delta_{S_0} = \frac{3}{4}$ and $\theta_{S_0, 2S_1} = \frac{1}{8}$ then it can be concluded that $\alpha = \frac{7}{8}$. So the second requirement of condition 1 is simplified to $\delta_{2S_1} + \theta_{S_1, 2S_1} \leq \frac{1}{8}$. Since in practise S_1 is small in compare with S_0 , we can see that the condition $\delta_{2S_1} + \theta_{S_1, 2S_1} \leq \frac{1}{8}$ will be satisfied easily.

We showed that if $\delta_{S_0} = \frac{3}{4}$ and $\theta_{S_0, 2S_1} = \frac{1}{8}$ then Theorem 2. Comparing these with the results for modified-cs [3], $\delta_{S_0 + S_1} \leq \frac{\sqrt{2}-1}{2}$, we observe that reg-mod-CSN remain stable under weaker conditions.

Also recall that CS results [14] needs $\delta_{2S_0} \leq \sqrt{2} - 1$ that is an stronger condition in compare with $\delta_{S_0} = \frac{3}{4}$ which we obtained for reg-mod-CSN.

6. SIMULATION RESULT

We compared regularized modified CSN, modified BPDN [12], modified CS [1] and simple CS for different values of $\frac{S_0}{m}$. In Figure 1 we used Signal Model with $m = 100$, $n = 50$, $S_0 = 20, 30, 40$, $S_a = 1$, $r = \frac{1}{6}$ and $w_t \sim^{iid} \text{uniform}(-c, c)$ with $c = .05$. γ_1 is the value of γ in minimization problem (1) regularized modified CSN and γ_2 is the value of γ for problem (2) in [12] for modified BPDN respectively. The measurement matrix was random Gaussian. The simulation results have been obtained by averaging over 100 samples. We set the α to some value in the noise level ($\alpha = .1$). By this value it gives a fairly accurate estimate of nonzero elements with a low number of falsely detections. In Figure 1 we showed a set of plots. Normalized MSE (NMSE), average number of extras (mean of the $|T_t \setminus N_t|$ over the 100 simulations) and average number of misses (mean of $|N_t \setminus T_t|$) are plotted in parts (b) and (c). Since in (b) and (c) the error was over 0.2 for CS, we just showed CS in (a). As it can be seen in (a) ($\frac{S_0}{m} = .2$) reg-mod-CSN, mod-BPDN and mod-CS are stable and works almost the same (the errors are under 0.02) while the CS has a large error. In (b) as $\frac{S_0}{m}$ is increased ($\frac{S_0}{m} = .3$) modified BPDN and modified CS starts to become unstable (The NMSE is increased gradually over time) while reg mod CSN is still stable (The NMSE remains under 0.02 over time). In the case where $\frac{S_0}{m} = .4$ all three methods become unstable.

7. CONCLUSION

An algorithm was proposed for using Regularized modified CSN over time in a way that we utilize the estimated signal from the previous time as an prior estimate for current time signal. It was shown that this algorithm remains stable under a proposed Signal Model and certain conditions. It was demonstrated that reg-mod-CS remain stable under weaker conditions in compare with mod-CS and simple CS.

8. APPENDIX

8.1. Proof of Theorem 1

Here we consider the general case where the signal is not sparse. Let $V = [1..m]$. Assume that we know partial part of support denoted by T . We redefine set N such that $T \subset N \subset V$. To prove the Theorem, first let us get the following relation by using the fact that both x and \hat{x} are feasible

$$\|A(\hat{x} - x)\|_2 \leq \|A\hat{x} - y\|_2 + \|y - Ax\|_2 \leq 2\epsilon \quad (5)$$

Basically our approach is a modification of the proof [14]. Let us write $\hat{x} = x + h$. Our aim in the rest of the proof is to make an upper bound for $\|h\|_2$. We decompose the vector h into a sum of vectors. We define $\Delta_0 = N \setminus T$ and Δ_j for $j \geq 1$ as the support of k largest coefficient of $h_{S_j^c}$ with $S_j = T \cup \bigcup_{i=0}^{j-1} \Delta_i$. The plan of the proof is to bound $\|h_T\|_2$, $\|h_{\Delta_0 \cup \Delta_1}\|_2$ and $\|h_{(T \cup \Delta_0 \cup \Delta_1)^c}\|_2$. Using the triangular inequality, we have $\|h_{(T \cup \Delta_0 \cup \Delta_1)^c}\|_2 \leq \sum_{j=2} \|h_{\Delta_j}\|_2$. For $j \geq 1$, $\|h_{\Delta_j}\|_2 \leq k^{\frac{1}{2}} \|h_{\Delta_j}\|_\infty \leq k^{-\frac{1}{2}} \|h_{\Delta_{j-1}}\|_1$ this leads to

$$\|h_{(T \cup \Delta_0 \cup \Delta_1)^c}\|_2 \leq \sum_{j=2} \|h_{\Delta_j}\|_2 \leq \frac{1}{\sqrt{k}} \|h_{(T \cup \Delta_0)^c}\|_1. \quad (6)$$

Since $\hat{x} = x + h$ is the solution to (1) and both \hat{x} and x are feasible, we have

$$\begin{aligned} \|x_{T^c}\|_1 &\geq \|(x + h)_{T^c}\|_1 \\ &= \|x_{\Delta_0} + h_{\Delta_0}\|_1 + \|x_{(T \cup \Delta_0)^c} + h_{(T \cup \Delta_0)^c}\|_1 \\ &\geq \|x_{\Delta_0}\|_1 - \|h_{\Delta_0}\|_1 + \|h_{(T \cup \Delta_0)^c}\|_1 - \|x_{(T \cup \Delta_0)^c}\|_1 \end{aligned}$$

So then we have

$$\|h_{(T \cup \Delta_0)^c}\|_1 \leq \|h_{\Delta_0}\|_1 + 2\|x_{(T \cup \Delta_0)^c}\|_1 \quad (7)$$

First we bound $\|h_T\|_2$. To do that, observe that $Ah_T = Ah - \sum_{j=0} Ah_{\Delta_j}$ and, therefore,

$$\|Ah_T\|^2 = \langle Ah_T, Ah \rangle - \langle Ah_T, \sum_{j=0} Ah_{\Delta_j} \rangle$$

Applying Cauchy-Schwartz, it follows from (5) and the restricted isometry and orthogonality property that

$$(1 - \delta_u) \|h_T\|_2^2 \leq \sqrt{1 + \delta_u} \|h_T\|_2 (2\epsilon) + \theta_{u,k} \|h_T\|_2 (\sum_{j=0} \|h_{\Delta_j}\|_2)$$

$$\|h_T\|_2 \leq \frac{\sqrt{1 + \delta_u}}{1 - \delta_u} (2\epsilon) + \frac{\theta_{u,k}}{1 - \delta_u} (\sum_{j=0} \|h_{\Delta_j}\|_2) \quad (8)$$

We can break the term $\sum_{j=0} \|h_{\Delta_j}\|_2 = \|h_{\Delta_0}\|_2 + \|h_{\Delta_1}\|_2 + \sum_{j=2} \|h_{\Delta_j}\|_2$. Since $\|h_{\Delta_0}\|_2 + \|h_{\Delta_1}\|_2 \leq \sqrt{2} \|h_{\Delta_0 \cup \Delta_1}\|_2$, Using (6) we can conclude

$$\sum_{j=0} \|h_{\Delta_j}\|_2 \leq \sqrt{2} \|h_{\Delta_0 \cup \Delta_1}\|_2 + \frac{1}{\sqrt{k}} \|h_{(T \cup \Delta_0)^c}\|_1 \quad (9)$$

Using (9) we can rewrite inequality (8) as,

$$\|h_T\|_2 \leq \frac{\sqrt{1 + \delta_u}}{1 - \delta_u} (2\epsilon) + \frac{\theta_{u,k}}{1 - \delta_u} (\sqrt{2} \|h_{\Delta_0 \cup \Delta_1}\|_2 + \frac{1}{\sqrt{k}} \|h_{(T \cup \Delta_0)^c}\|_1) \quad (10)$$

In the next step we bound $\|h_{\Delta_0 \cup \Delta_1}\|_2$. To do that we first make a bound for $\|h_T\|_2$. Since both x and \hat{x} are feasible and by using the second constraint of problem (1) we have

$$\|h_T\|_2 = \|x_T - \hat{x}_T\|_2 \leq \|x_T - \mu_T\|_2 + \|\hat{x}_T - \mu_T\|_2 \leq 2\gamma \quad (11)$$

Same as previous step we can write $\|Ah_{\Delta_0 \cup \Delta_1}\|_2$ as

$$\begin{aligned} \|Ah_{\Delta_0 \cup \Delta_1}\|_2^2 &= \langle Ah_{\Delta_0 \cup \Delta_1}, Ah \rangle - \langle Ah_{\Delta_0 \cup \Delta_1}, Ah_T \rangle - \\ &\langle Ah_{\Delta_0 \cup \Delta_1}, A(\sum_{j=2} h_{\Delta_j}) \rangle \end{aligned}$$

Using Cauchy-Schwartz, (5) and the restricted isometry property we have

$$\langle Ah_{\Delta_0 \cup \Delta_1}, Ah \rangle \leq 2\sqrt{1 + \delta_{2k}} \epsilon \|h_{\Delta_0 \cup \Delta_1}\|_2 \quad (12)$$

Employing the restricted orthogonality property and (11) we get

$$\langle Ah_{\Delta_0 \cup \Delta_1}, Ah_T \rangle \leq 2\theta_{u,2k} \gamma \|h_{\Delta_0 \cup \Delta_1}\|_2 \quad (13)$$

Using the restricted orthogonality property we get

$$\langle Ah_{\Delta_0 \cup \Delta_1}, A(\sum_{j=2} h_{\Delta_j}) \rangle \leq \theta_{k,2k} \|h_{\Delta_0 \cup \Delta_1}\|_2 (\sum_{j=2} \|h_{\Delta_j}\|_2) \quad (14)$$

Using (6) we can rewrite the above inequality as

$$\langle Ah_{\Delta_0 \cup \Delta_1}, A(\sum_{j=2} h_{\Delta_j}) \rangle \leq \frac{\theta_{k,2k}}{\sqrt{k}} \|h_{\Delta_0 \cup \Delta_1}\|_2 \|h_{(T \cup \Delta_0)^c}\|_1 \quad (15)$$

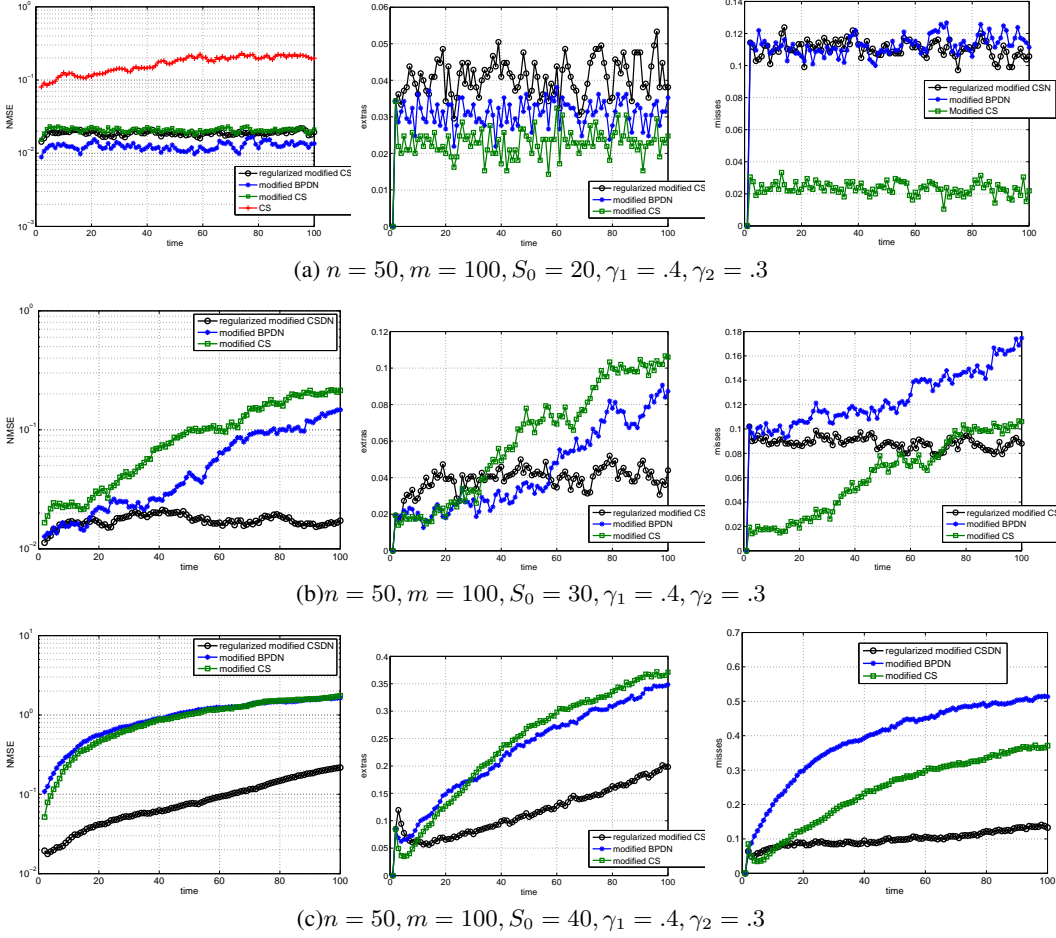


Fig. 1. Normalized MSE (NMSE), number of extras and number of misses over time for CS, modified CS, modified BPDN, and regularized modified CSN. In part (b) and (c), NMSE for CS was more than 20%. (plotted only in (a))

Combining (12), (13) and (15), we get

$$(1 - \delta_{2k}) \|h_{\Delta_0 \cup \Delta_1}\|_2^2 \leq \sqrt{1 + \delta_{2k}} (2\epsilon) \|h_{\Delta_0 \cup \Delta_1}\|_2 + \theta_{u,2k} \|h_T\|_2 \|h_{\Delta_0 \cup \Delta_1}\|_2 + \frac{\theta_{k,2k}}{\sqrt{k}} \|h_{\Delta_0 \cup \Delta_1}\|_2 \|h_{(T \cup \Delta_0)^c}\|_1 \quad (16)$$

By simplifying the above inequality we have

$$\|h_{\Delta_0 \cup \Delta_1}\|_2 \leq 2 \frac{\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k}} \epsilon + 2 \frac{\theta_{u,2k}}{1 - \delta_{2k}} \gamma + \frac{\theta_{k,2k}}{\sqrt{k}(1 - \delta_{2k})} \|h_{(T \cup \Delta_0)^c}\|_1 \quad (17)$$

By inequality (7) we can conclude

$$\|h_{(T \cup \Delta_0)^c}\|_1 \leq \sqrt{k} \|h_{\Delta_0}\|_2 + 2 \|x_{(T \cup \Delta_0)^c}\|_1 \leq \sqrt{k} \|h_{\Delta_0 \cup \Delta_1}\|_2 + 2 \|x_{(T \cup \Delta_0)^c}\|_1 \quad (18)$$

We use this inequality to replace it with $\|h_{(T \cup \Delta_0)^c}\|_1$ in inequality (17).

$$\|h_{\Delta_0 \cup \Delta_1}\|_2 \leq 2 \frac{\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k}} \epsilon + 2 \frac{\theta_{u,2k}}{1 - \delta_{2k}} \gamma + \frac{\theta_{k,2k}}{1 - \delta_{2k}} \|h_{\Delta_0 \cup \Delta_1}\|_2 + 2 \frac{\theta_{k,2k}}{\sqrt{k}(1 - \delta_{2k})} \|x_{(T \cup \Delta_0)^c}\|_1$$

Simplifying the above inequality lead to

$$\|h_{\Delta_0 \cup \Delta_1}\|_2 \leq \tilde{F}_1 \epsilon + \tilde{F}_2 \gamma + \tilde{F}_3 e_0(T, \Delta_0) \quad (19)$$

$$e_0(T, \Delta) = 2 \frac{\|x_{(T \cup \Delta)^c}\|_1}{\sqrt{|\Delta|}}$$

$$\tilde{F}_1 = \frac{2\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k} - \theta_{k,2k}}, \quad \tilde{F}_2 = \frac{2\theta_{u,2k}}{1 - \delta_{2k} - \theta_{k,2k}}$$

$$\tilde{F}_3 = \frac{\theta_{k,2k}}{1 - \delta_{2k} - \theta_{k,2k}}$$

Here we use the previous bounds on $\|h_T\|_2$, $\|h_{\Delta_0 \cup \Delta_1}\|_2$ and $\|h_{(T \cup \Delta_0 \cup \Delta_1)^c}\|_2$ to bound $\|h\|_2$. Using (10), (17) and (6) we have

$$\begin{aligned} \|h\| &\leq \|h_T\|_2 + \|h_{\Delta_0 \cup \Delta_1}\|_2 + \|h_{(T \cup \Delta_0 \cup \Delta_1)^c}\|_2 \leq \\ &\frac{\sqrt{1+\delta_u}}{1-\delta_u}(2\epsilon) + \frac{\theta_{u,k}}{1-\delta_u}(\sqrt{2}\|h_{\Delta_0 \cup \Delta_1}\|_2 \\ &+ \frac{1}{\sqrt{k}}\|h_{(T \cup \Delta_0)^c}\|_1) + \|h_{\Delta_0 \cup \Delta_1}\|_2 + \frac{1}{\sqrt{k}}\|h_{(T \cup \Delta_0)^c}\|_1 \end{aligned}$$

Using (18) and reordering the terms lead to

$$\begin{aligned} \|h\| &\leq 2\frac{\sqrt{1+\delta_u}}{1-\delta_u}\epsilon + (2 + \frac{(\sqrt{2}+1)\theta_{u,k}}{1-\delta_u})\|h_{\Delta_0 \cup \Delta_1}\|_2 \\ &+ 2(1 + \frac{\theta_{u,k}}{1-\delta_u})\frac{\|x_{(T \cup \Delta)^c}\|_1}{\sqrt{k}} \end{aligned}$$

By substitution of (19) in above inequality we get

$$\|h\| \leq C_{u,k}\epsilon + D_{u,k}\gamma + E_{u,k}e_0(T, \Delta_0)$$

where

$$E_{u,k} = \left(\frac{\theta_{k,2k}}{1-\delta_{2k}-\theta_{k,2k}} \left(2 + \frac{(\sqrt{2}+1)\theta_{u,k}}{1-\delta_u} \right) + 2 \left(1 + \frac{\theta_{u,k}}{1-\delta_u} \right) \right)$$

8.2. Proof of Theorem 2

First, recall that $u_t := |T_t|$ and $k_t := |\Delta_t|$. The proof follows using induction. Using condition 5 of the theorem, the claim holds for $t = 0$. This proves the base case. For the induction step, assume that the claim holds at $t - 1$, i.e. $|T_{t-1}| \leq S_0, |\Delta_{t-1}| \leq S_1$ and $\|x_{t-1} - \hat{x}_{t-1}\| \leq C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma$. Using these assumptions we prove that the claim holds at t .

First, notice that condition 2 of theorem states that $\gamma = C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma + \sqrt{2d}S_a r$. We claim that under conditions 1 and 2 of theorem we have $\|x_t - \hat{x}_{t-1}\| \leq \gamma$. This is true since by Proposition 2 and the assumption of induction, we have $\|x_t - \hat{x}_{t-1}\| \leq \|x_{t-1} - \hat{x}_{t-1}\| + \sqrt{2d}S_a r \leq C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma + \sqrt{2d}S_a r = \gamma$. Note that $D_{S_0, S_1} < 1$ is necessary for having condition 2. It can be shown that $D_{S_0, S_1} < 1$ is equivalent to $1 - \delta_{2S_1} - \theta_{S_1, 2S_1} - 2 \left(2 + \frac{(\sqrt{2}+1)\theta_{S_0, S_1}}{1-\delta_{S_0}} \right) \theta_{S_0, 2S_1} > 0$. This holds since condition 1 holds.

Now for the rest of proof if we show that $u_t \leq S_0$ and $k_t \leq S_1$, then Theorem 1 can be applied and we are done. That is because by condition 1, we can show that $\delta_{u_t} < 1$, $1 - \delta_{2k_t} - \theta_{k_t, 2k_t} > 0$ which are the first two requirements for applying Theorem 1. Condition 2 and the above discussion ensures that the third requirement of Theorem 1 holds.

First, notice that employing Proposition 1 and condition 3 and 4 of the theorem and by the assumption that $\|x_{t-1} - \hat{x}_{t-1}\| \leq C_{S_0, S_1}\epsilon + D_{S_0, S_1}\gamma$, we can conclude that at time $t - 1$, all the elements greater or equal to $d_0 r$ will be get detected in the support update step, i.e. when computing \tilde{T}_{t-1} . Thus, the missed set, $|\tilde{\Delta}_{t-1}| \leq (2d_0 - 2)S_a$. Also, notice that by Proposition 2 and the assumption of induction we can conclude that at time $t - 1$ no zero value element of x_{t-1} will be get detected as an element of \tilde{T}_{t-1} , in other word

$\tilde{\Delta}_{e,t-1} = 0$. From Algorithm 1 we remember that $T_t = \tilde{T}_{t-1}$ and $u_t = |\tilde{T}_{t-1}|$. Since $\tilde{T}_{t-1} = N_{t-1} \cup \tilde{\Delta}_{e,t-1} \setminus \tilde{\Delta}_{t-1}$, we have $u_t = |\tilde{T}_{t-1}| \leq |N_{t-1}| + |\tilde{\Delta}_{e,t-1}|$. We know that $|N_{t-1}| = S_0$ and $|\Delta_{e,t}| = 0$. Hence, it implies that $u_t \leq S_0$. Also, $\Delta_t = N_t \cap \tilde{T}_{t-1}^c = (N_{t-1} \cup A_t) \cap R_t^c \cap \tilde{T}_{t-1}^c \subseteq (\tilde{\Delta}_{t-1} \cup A_t) \cap R_t^c$. Here we have used the facts that $N_t = (N_{t-1} \cup A_t) \cap R_t^c$ and $\tilde{\Delta}_{t-1} = N_{t-1} \cap \tilde{T}_{t-1}^c$. So we have $\Delta_t \subseteq (\tilde{\Delta}_{t-1} \cup A_t) \cap R_t^c \subseteq (S_{t-1}(d_0) \cup A_t) \cap R_t^c$. Therefore, $k_t = |\Delta_t| \leq |S_{t-1}(d_0)| + |A_t| - |R_t| = (2d_0 - 2)S_a = S_1$.

9. REFERENCES

- [1] N. Vaswani and W. Lu, "Modified-cs: Modifying compressive sensing for problems with partially known support," *IEEE Trans. Signal Processing*, September 2010, to appear (shorter version in ISIT 2009).
- [2] Namrata Vaswani and Wei L, "Modied-cs: Modifying compressive sensing for problems with partially known support," SEPTEMBER.
- [3] N.Vaswani, "Stability (over time) of modified-cs and ls-cs for recursive causal sparse reconstruction," in *Allerton*, 2010.
- [4] N. Vaswani, "Kalman filtered compressed sensing," in *ICIP*, 2008.
- [5] N. Vaswani, "Analyzing least squares and kalman filtered compressed sensing," in *ICASSP*, 2009.
- [6] N. Vaswani, "Least Squares CS-residual (LS-CS): Compressive Sensing on Least Squares residual," *IEEE Trans. Sig. Proc.*, August 2010.
- [7] M. S. Asif and J. Romberg, "Dynamic updating for sparse time varying signals," in *CISS*, 2009.
- [8] A. Carmi, P. Gurfil, and D. Kanevsky, "A simple method for sparse signal recovery from noisy observations using kalman filtering," in *IBM Technical report*, December 2008.
- [9] D. Angelosante, G.B. Giannakis, and E. Grossi, "Compressed sensing of time-varying signals," in *DSP*, 2009.
- [10] R. von Borries, C. J. Miosso, and C. Potes, "Compressed sensing using prior information," in *IEEE Intl. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2007.
- [11] A. Khajehnejad, W. Xu, A. Avestimehr, and B. Hassibi, "Weighted l1 minimization for sparse recovery with prior information," in *ISIT*, 2009.
- [12] W. Lu and N. Vaswani, "Modified bpdn for noisy compressive sensing with partially known support," in *ICASSP*, 2010.
- [13] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Info. Th.*, vol. 51(12), pp. 4203 - 4215, Dec. 2005.
- [14] 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.