Exact Reconstruction Conditions and Error Bounds for Regularized Modified Basis Pursuit (Reg-Modified-BP)

Wei Lu and Namrata Vaswani

Department of Electrical and Computer Engineering, Iowa State University, Ames, IA

Email: {luwei,namrata}@iastate.edu

Abstract—We study the problem of reconstructing a sparse signal from a limited number of linear measurements, *when a part of its support and the signal estimate on it are known*. The support and signal estimate can be obtained from prior knowledge, e.g., in a real-time dynamic MRI application, they could be the support and signal estimate from the previous time instant. We propose regularized Modified Basis Pursuit (Reg-mod-BP). We also provide the exact reconstruction conditions and we argue that they can be weaker than modified-CS. We then bound its reconstruction error when exact reconstruction can not happen and we show that the bound is much smaller than modified-CS when the available measurements are few. We also use Monte Carlo to verify that reg-mod-BP has better exact reconstruction conditions than other methods with very few measurements. We also compare the average errors when exact reconstruction can not be achieved and show that the errors are smaller than other methods.

Index Terms—Compressive sensing, Sparse reconstruction

I. INTRODUCTION

In this work, we study the problem of sparse reconstruction with partial, and partly erroneous, knowledge of support and of the signal values on the "known" support. In practical applications, this may be available from prior knowledge, or in recursive reconstruction applications, e.g. recursive dynamic MRI, and one can use the support and signal estimate from the previous time instant for this purpose[2].

The problem of sparse reconstruction with partial knowledge of support was first simultaneously introduced in our work (modified-CS) [1] and in [9]. An earlier work [8] proposed an approach similar to modified-CS but did not analyze it and also did not show careful numerical or real experiments either. Denote the support "knowledge" by T. Modified-CS finds a signal that is sparsest outside of T and satisfies the data constraint. The work of [9] puts a probabilistic prior on the support and obtains a solution. Both [1] and [9] obtain exact reconstruction conditions for the respective approaches.

When both partial support knowledge, T, and signal value estimates on this support, μ_T are available, modCS can be augmented by also imposing a constraint on the ℓ_∞ distance of the solution from μ_T. We refer to this as "reg-mod-BP". In this work, we derive its exact reconstruction condition and bound the reconstruction error. We provide both exact reconstruction comparisons with existing work as well as reconstruction error comparisons. We also show comparisons with the work of [9]

and with modCS. Other somewhat related work includes [7], [10], [11], [12].

This paper is organized as follows. We introduce reg-mod-BP in Sec. II. We use different strategies to obtain the exact reconstruction conditions and reconstruction error bounds. The exact reconstruction condition is obtained using a similar way as [1] in Sec. III. In Sec. IV, we bound the reconstruction error when exact reconstruction can not occur in a fashion as in [5]. We do not use the method of [5], [6] to get exact reconstruction conditions in Sec. III because it can not give better conditions. Finally, simulations are shown to demonstrate reg-mod-BP gets better exact reconstruction conditions and smaller reconstruction errors in Sec. V.

II. REGULARIZED MODIFIED BASIS PURSUIT

In [1], we proposed modCS which only used partially known support. ModCS solves

$$\min_\beta ||\beta_{T^c}||_1, \ s.t. \ y = A\beta \qquad (1)$$

However, modCS puts no constraint on x_T. Thus, when few measurements are available, x_T can be arbitrarily large resulting in very bad reconstruction. Hence, we propose reg-mod-BP to also constrain x_T by bounding $||x_T - \mu_T||_\infty$, i.e.

$$\min_\beta ||\beta_{T^c}||_1, \ s.t. \ y = A\beta, \ ||\beta_T - \mu_T||_\infty \le \rho \qquad (2)$$

where T is the support estimate and μ_T is the signal estimate on T. Let N denote the support of x. Define $\Delta \triangleq N \setminus T$ and $\Delta_e \triangleq T \setminus N$. The cost function is to find the sparsest solution outside T and the first constraint gives the data constraint while the second one imposes the closeness to μ along T.

III. EXACT RECONSTRUCTION

In this section, we try to obtain exact reconstruction conditions for reg-mod-BP that are weaker than those for mod-CS in terms of the number of measurements required. The key idea is to use the following fact which is observed from simulation experiments. If one or more of the prior constraints, $||\beta_T - \mu_T||_\infty \le \rho$, are active, then reg-mod-BP needs fewer measurements for exact reconstruction. Hence, let us define the following active sets T_{a_1}, T_{a_2} and the inactive set T_{in}.

$$T_{a_1} \triangleq \{i \in T : x_i - \mu_i = \rho\}$$
$$T_{a_2} \triangleq \{i \in T : x_i - \mu_i = -\rho\}$$
$$T_{in} \triangleq \{i \in T : |x_i - \mu_i| < \rho\}$$

and their sizes $k \triangleq |T|$, $u \triangleq |\Delta|$, $k_{in} \triangleq |T_{in}|$. Then we can obtain the following Theorem.

Theorem 1 (Exact Reconstruction Conditions): Given a sparse vector, x, whose support,

$$N = T \cup \Delta \setminus \Delta_e \tag{3}$$

where Δ and T are disjoint and $\Delta_e \subseteq T$. Also, x satisfies

$$y = Ax \tag{4}$$

Let

$$M \triangleq I - A_{T_{in}}(A'_{T_{in}} A_{T_{in}})^{-1} A'_{T_{in}} \tag{5}$$

x is the unique minimizer of (2) if

1) μ is such that

$$T_{a_1} \subseteq T_+ = \{i \in T : A'_i M A_\Delta (A'_\Delta M A_\Delta)^{-1} sgn(x_\Delta) > 0\}$$
$$T_{a_2} \subseteq T_- = \{i \in T : A'_i M A_\Delta (A'_\Delta M A_\Delta)^{-1} sgn(x_\Delta) < 0\}$$
$$\|\mu_{T_{in}} - x_{T_{in}}\|_\infty < \rho \tag{6}$$

2) n is large enough such that

$$\delta_{k+u} < 1 \text{ and } \delta_{2u} + \delta_k + \theta_{k,2u}^2 < 1$$
$$\text{and } a_k(2u,u) + a_{k_{in}}(u,u) < 1 \tag{7}$$

$$\text{where } a_k(S, \check{S}) \triangleq \frac{\theta_{\check{S},S} + \frac{\theta_{\check{S},k} \theta_{S,k}}{1-\delta_k}}{1 - \delta_S - \frac{\theta_{S,k}^2}{1-\delta_k}} \tag{8}$$

Recall that $k \triangleq |T|$, $u \triangleq |\Delta|$, $k_{in} \triangleq |T_{in}|$.

The proof is given in Appendix. Notice that modCS is reg-mod-BP with $\rho = \infty$. Thus, for modCS, $k_{in} = k$. With $\rho = \infty$ and $k_{in} = k$, the above result is the same as that of Theorem 1 of [1]. When $\rho < \infty$, and if one or more of the prior constraints are active, then $k_{in} < k$. In this case, reg-mod-BP needs weaker conditions on the measurement matrix than modCS. But, it also needs that (a) the constraint is active on a part of T and (b) the active sets T_{a_1} and T_{a_2} satisfy condition (6). In other words, it needs a very specific type of prior information is available. This fact is also verified by our simulations.

Requiring that μ satisfy these conditions is somewhat restrictive. But as we show in the next section and also in our simulations, the error bound for reg-mod-BP is smaller than that for modCS (and for other methods) even without requiring any constraint to be active.

IV. RECONSTRUCTION ERROR BOUND

When exact reconstruction cannot be achieved, we want to bound the error of $h = \hat{x} - x$. We adapt the approach of [5], [6] to bound the ℓ_2 norm of the error $\|h\|_2$. First consider modCS, i.e. (1). When exact reconstruction condition does not hold, the following lemma provides one way to bound the error.

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

$$\|x - \hat{x}\|_2 \leq$$
$$\frac{4\sqrt{1+\delta_{|\tilde{T}|+2|\tilde{\Delta}|}}}{1-(\sqrt{2}+1)\delta_{|\tilde{T}|+2|\tilde{\Delta}|}}\epsilon + \frac{1-\delta_{|\tilde{T}|+2|\tilde{\Delta}|}}{1-(\sqrt{2}+1)\delta_{|\tilde{T}|+2|\tilde{\Delta}|}} \cdot \frac{2\|x_{(\tilde{T}\cup\tilde{\Delta})^c}\|_1}{\sqrt{|\tilde{\Delta}|}} \tag{9}$$

As long as the true x is always part of the feasible set of (2), i.e. as long as $\|x_T - \mu_T\|_\infty \leq \rho$, the above lemma also holds for reg-mod-BP. In the next lemma we also use this prior constraint to obtain another error bound for reg-mod-BP, which is tighter than that of Lemma 1 when ρ is small enough, i.e. prior information is strong.

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

$$\|x - \hat{x}\|_2 \leq (\frac{2\sqrt{k}\delta_{k+2u}}{1-(\sqrt{2}+1)\delta_{2u}} + 2)\rho \tag{10}$$

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

Theorem 2 (Reconstruction Error Bound): Let $\hat{x}$ solve (2). If $\|x_T - \mu_T\|_\infty \leq \rho$ and if $\delta_{2u} \leq \sqrt{2} - 1$ and $\delta_{k+2u} < 1$, then

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

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

V. EXPERIMENTAL RESULTS

In this section, we use simulations to verify the conclusions from previous sections. We compare reg-mod-BP with modCS and weighted ℓ_1 reconstruction as well as CS. Weighted ℓ_1 reconstruction is to solve

$$\min_\beta \quad \|\beta_{T^c}\|_1 + \gamma\|\beta_T\|_1 \quad \text{s.t. } y = A\beta \tag{11}$$

We first demonstrate that reg-mod-BP can obtain exact reconstruction with much fewer measurements under which other methods can not. Then, we will show the reconstruction errors are smaller for reg-mod-BP than other methods.

A. Exact Reconstruction Comparisons

We compare the smallest number of measurements for exact reconstruction of reg-mod-BP with that of modCS and weighted ℓ_1 reconstruction. We use the following procedure to compare all four algorithms when exact reconstruction always occurs, for a given A (i.e. we average over the joint distribution of x and y given A). In the following simulation steps, the

notation $z \sim iid(\pm a)$ means that we generate z as an i.i.d. random vector and each element is either $+a$ or $-a$ with probability $1/2$, while $z \sim unif(a, b)$ generates a uniform random variable distributed in $[a, b]$.

1) Fix signal length, $m = 256$, its support size, $|N| = 0.1m = 26$ and $u = 0.04|N|$, $e = 0.1|N|$. Select n, .
2) Generate the $n \times m$ random-Gaussian matrix, A (generate an $n \times m$ matrix with independent identically distributed (i.i.d.) zero mean Gaussian entries and normalize each column to unit ℓ_2 norm)
3) Repeat the following $tot = 100$ times
   a) Generate the support, N, of size s, uniformly at random from $[1, n]$.
   b) Generate Δ of size u uniformly at random from the elements of N.
   c) Generate Δ_e of size e, uniformly at random from the elements of $[1, n] \setminus N$.
   d) Let $T = N \cup \Delta_e \setminus \Delta$. Generate $x_{N \cap T} \sim iid(\pm 1)$ and $x_\Delta \sim iid(\pm 0.25)$. Set $x_{N^c} = 0$. Generate $y = Ax$.
   e) Generate T_{in} uniformly in T with $|T_{in}| = 0.4|N|$. Compute T_+ and T_- as in (6) and set $T_{a_1} \leftarrow T_+ \setminus T_{in}$ and $T_{a_2} \leftarrow T_- \setminus T_{in}$.
   f) Fix $\rho = 0.1$. Generate $\hat{\mu}_{T_{a_1}} = x_{T_{a_1}} - \rho$ and $\hat{\mu}_{T_{a_2}} = x_{T_{a_2}} + \rho$. For any $j \in T_{in}$, generate $\hat{\mu}_{T_j} \sim x_{T_j} + 0.9\rho \times unif(-1, 1)$. Set $\hat{\mu}_{T^c} = 0$.
   g) Run CS, modCS and reg-mod-BP to obtain $\hat{x}_{CS}$, $\hat{x}_{modCS}$ and $\hat{x}_{regmodBP}$.
   h) Run weighted ℓ_1 reconstruction with different choices of γ and store each one to $\hat{x}_{wl1}$.
4) Estimate the probability of exact reconstruction using CS by counting the number of times $\hat{x}_{CS}$ was equal to x ("equal" was defined as $\|\hat{x}_{CS} - x\|_2/\|x\|_2 < 10^{-5}$) and dividing by $tot = 100$.
5) Do the same for modCS, reg-mod-BP and weighted ℓ_1 reconstruction for each γ.
6) Repeat for various values of n.

We tabulate our results in Table I. We can see that reg-mod-BP only needs 11% measurements for exact reconstruction, but both modCS and weighted ℓ_1 need 17% measurements and of course CS needs much more measurements. Hence, reg-mod-BP can give better exact reconstructions if μ meets the specific requirements as in (6).But if we compute the probability of exact reconstruction using a given small number of measurements, then weighted ℓ_1 has a higher probability than that of modCS but lower than that of reg-mod-BP. For example, when $n = 0.14m$, the probability of exact reconstruction for modCS is 0.8 and weighted ℓ_1 has 0.88 while reg-mod-BP achieves 100% probability of exact reconstruction.

	CS	modCS	wℓ_1	RM
n	$0.42m$	$0.17m$	$0.17m$	$0.11m$

TABLE I
MINIMUM NUMBER OF MEASUREMENTS REQUIRED TO ACHIEVE EXACT RECONSTRUCTION ALWAYS FOR CS, MODCS, WEIGHTED ℓ_1 AND REG-MOD-BP.

(a) $n = 0.11m, |\Delta| = |\Delta_e| = 10\%|N|, \rho_a = 0.1$					
modCS	**wℓ_1 (0.1)**	**wℓ_1 (0.01)**	**RM$(0.5\rho_a)$**	**RM(ρ_a)**	**RM$(1.5\rho_a)$**
0.4559	0.3984	0.38	0.0734	0.0958	0.1249

(b) $n = 0.15m, |\Delta| = |\Delta_e| = 10\%|N|, \rho_a = 0.1$					
modCS	**wℓ_1 (0.1)**	**wℓ_1 (0.01)**	**RM$(0.5\rho_a)$**	**RM(ρ_a)**	**RM$(1.5\rho_a)$**
0.1152	0.1365	0.1245	0.0489	0.034	0.0661

TABLE II
RECONSTRUCTION ERROR (N-RMSE) COMPARISON OF MODCS, WEIGHTED ℓ_1 AND REG-MOD-BP. IN BOTH CASES, REG-MOD-BP HAS MUCH SMALLER ERROR THAN THOSE OF MODCS AND WEIGHTED ℓ_1 FOR DIFFERENT CHOICES OF ρ.

B. Reconstruction Error Comparisons

We use the same simulation approach as above but with steps 3e and 3f replaced by $\mu = x + unif(-\rho_a, \rho_a)$ where $\rho_a = 0.1$, i.e. the restrictive assumption on μ is removed. We also fix $|\Delta| = |\Delta_e| = 10\%|N|$ and use different choices of ρ and γ for reg-mod-BP and weighted ℓ_1 each to test their robustness to parameters. Then we compute N-RMSE ($\sqrt{\frac{E[\|\hat{x} - x\|_2^2]}{E[\|x\|_2^2]}}$). The results are summarized in Table II. We can see that reg-mod-BP has the smallest reconstruction error for different choices of ρ when the measurements $n = 0.11m$ or $n = 0.15m$.

VI. CONCLUSIONS

We proposed a modification of the modified-CS idea, called regularized modified Basis Pursuit, for sparse reconstruction when a part of the support and the signal estimate on it are known. We obtained the exact reconstruction conditions for reg-mod-BP and bounded its reconstruction error when exact reconstruction cannot be achieved. We demonstrated that reg-mod-BP can have better exact reconstruction than modCS and weighted ℓ_1 reconstruction as well as CS when the signal estimate satisfies the required structure. We then also showed that reg-mod-BP has the smallest reconstruction bounds and reconstruction errors and this does not require the specific type of signal estimate. In summary, reg-mod-BP significantly improves upon modCS, weighted ℓ_1 and CS in terms of reconstruction error bound with fewer measurements. If a very specific prior information is available so that the prior constraint is active on part of T and satisfies (6), then it also achieves exact reconstruction with fewest measurements.

VII. APPENDIX

A. Key Lemmas

To obtain Theorem 1, we need the following three lemmas to complete the proof of Theorem 1.

Lemma 3: The sparse signal, x, is the unique minimizer of (2) if $\delta_{|T|+|\Delta|} < 1$ and if we can find a vector w satisfying the following condition

1) $w'A_j \geq 0$ if $j \in T_{a_1}$ and $w'A_j \leq 0$ if $j \in T_{a_2}$
2) $w'A_j = 0$ if $j \in T_{in}$
3) $w'A_j = sgn(x_j)$ if $j \in \Delta$
4) $|w'A_j| < 1$ if $j \notin T \cup \Delta$

Lemma 4: Given the known part of the support, T, of size k. Let S, $\check{S}$ be such that $k + S + \check{S} \leq n$ and $\delta_S + \delta_k + \theta_{k,S}^2 < 1$. Let c be a vector supported on a set T_d, that is disjoint with

T, of size $|T_d| \leq S$. Then there exists a vector $\tilde{w}$ and an exceptional set, $\check{E}$, disjoint with $T \cup T_d$, s.t.

$$
\begin{aligned}
A_j'\tilde{w} &= 0, \; \forall \, j \in T \\
A_j'\tilde{w} &= c_j, \; \forall \, j \in T_d \\
|E| &< \check{S} \\
\|A_E'\tilde{w}\|_2 &\leq a_k(S, \check{S})\|c\|_2 \\
|A_j'\tilde{w}| &\leq \frac{a_k(S, \check{S})}{\sqrt{\check{S}}}\|c\|_2 \; \forall j \notin T \cup T_d \cup E \;\; \text{and} \\
\|\tilde{w}\|_2 &\leq K_k(S)\|c\|_2
\end{aligned}
\tag{13}
$$

where $a_k(S, \check{S})$ is defined in (8) and

$$
K_k(S) := \frac{\sqrt{1+\delta_S}}{1 - \delta_S - \frac{\theta_{S,k}^2}{1-\delta_k}} \tag{14}
$$

Lemma 5: Given the known part of the support, T, of size k and given $T_{in} \subseteq T$, $T_{a_1} \subseteq T$ and $T_{a_2} \subseteq T$ which are as defined in Theorem 1. Also let k, u, k_{in} and the matrix M be as defined in Theorem 1. Let $S, \check{S}$ be such that $k+S+\check{S} \leq n$. Let c be a vector supported on a set T_d, that is disjoint with T, of size $|T_d| \leq S$. If $\delta_S + \delta_k + \theta_{k,S}^2 < 1$ and if

$$
\begin{aligned}
T_{a_1} &\subseteq T_+ \triangleq \{i \in T : A_i' M A_{T_d}(A_{T_d}' M A_{T_d})^{-1}c > 0\}, \\
T_{a_2} &\subseteq T_- \triangleq \{i \in T : A_i' M A_{T_d}(A_{T_d}' M A_{T_d})^{-1}c < 0\}
\end{aligned}
\tag{15}
$$

then, there exists a vector $\tilde{w}$ and an exceptional set, E, disjoint with $T \cup T_d$, s.t.

$$
\begin{aligned}
A_j'\tilde{w} &= 0, \; \forall \, j \in T_{in} \\
A_j'\tilde{w} &> 0, \; \forall \, j \in T_{a_1} \\
A_j'\tilde{w} &< 0, \; \forall \, j \in T_{a_2} \\
A_j'\tilde{w} &= c_j, \; \forall \, j \in T_d \\
|E| &< \check{S} \\
\|A_E'\tilde{w}\|_2 &\leq a_{k_{in}}(S, \check{S})\|c\|_2 \\
|A_j'\tilde{w}| &\leq \frac{a_{k_{in}}(S, \check{S})}{\sqrt{\check{S}}}\|c\|_2 \; \forall j \notin T \cup T_d \cup E \;\; \text{and} \\
\|\tilde{w}\|_2 &\leq K_{k_{in}}(S)\|c\|_2
\end{aligned}
\tag{16,17}
$$

where $a_k(S, \check{S})$ is defined in (8) and

$$
K_k(S) := \frac{\sqrt{1+\delta_S}}{1 - \delta_S - \frac{\theta_{S,k}^2}{1-\delta_k}} \tag{18}
$$

The proof of Lemma 3 is given in the next subsection. The proof of Lemma 4 is in [1] and the proof of Lemma 5 is given in the long version[13].

B. Proof of Lemma 3

The Lagrange multiplier is

$$
J(\beta) \triangleq \|\beta_{T^c}\|_1 + \Lambda_1'(\beta_T - \mu_T - \rho\mathbf{1}) + \Lambda_2'(-\beta_T + \mu_T - \rho\mathbf{1}) + w'(y - A\beta)
$$

where $\Lambda_1 \succeq 0$ and $\Lambda_2 \succeq 0$. Also, we have for any $j \in T$

$$
\lambda_{1,j}(\beta_j - \mu_j - \rho) = 0 \quad \lambda_{2,j}(-\beta_j + \mu_j - \rho) = 0 \tag{19}
$$

Therefore, for this convex problem, the minimizer β satisfies the following equality

$$
\nabla J(\beta) = \begin{bmatrix} 0_T \\ sgn(\beta_\Delta) \\ g_{(T \cup \Delta)^c} \end{bmatrix} + \begin{bmatrix} (\Lambda_1 - \Lambda_2) \\ 0_{T^c} \end{bmatrix} - A'w = 0 \tag{20}
$$

If x is a minimizer of this problem, then we can get

1) For $j \in T$, $w'A_j = \lambda_{1,j} - \lambda_{2,j}$ where $\lambda_{1,j} \geq 0$ and $\lambda_{2,j} \geq 0$ and they also satisfy

$$
\lambda_{1,j}(x_j - \mu_j - \rho) = 0 \quad \lambda_{2,j}(-x_j + \mu_j - \rho) = 0 \tag{21}
$$

2) For $j \in \Delta$, $w'A_j = sgn(x_\Delta)$
3) For $j \notin T \cup \Delta$, $|w'A_j| \leq 1$

Now, we need to prove that x is the minimizer by showing that given the conditions of Lemma 3, a unique minimizer exists which is equal to x. Assume there is another minimizer β, thus,

$$
\|\beta_{T^c}\|_1 \leq \|x_{T^c}\|_1 = \sum_{j \in \Delta} |x_j| \tag{22}
$$

Then, we try to prove that $\|x_{T^c}\|_1 \leq \|\beta_{T^c}\|_1$.

$$
\begin{aligned}
\|\beta_{T^c}\|_1 &= \sum_{j \in \Delta} |x_j + (\beta_j - x_j)| + \sum_{j \notin T \cup \Delta} |\beta_j| \\
&\geq \sum_{j \in \Delta} |x_j + (\beta_j - x_j)| + \sum_{j \notin T \cup \Delta} w'A_j\beta_j \\
&\geq \sum_{j \in \Delta} sgn(x_j)(x_j + (\beta_j - x_j)) + \sum_{j \notin T \cup \Delta} w'A_j\beta_j \\
&= \sum_{j \in \Delta} |x_j| + \sum_{j \in \Delta} w'A_j(\beta_j - x_j) + \sum_{j \notin T \cup \Delta} w'A_j(\beta_j - x_j) \\
&\quad + \sum_{j \in T} w'A_j(\beta_j - x_j) - \sum_{j \in T} w'A_j(\beta_j - x_j) \\
&= \|x_{T^c}\|_1 + w'(A\beta - Ax) - \sum_{j \in T} w'A_j(\beta_j - x_j) \\
&= \|x_{T^c}\|_1 - \sum_{j \in T}(\lambda_{1,j} - \lambda_{2,j})(\beta_j - \mu_j + \mu_j - x_j) \tag{23}
\end{aligned}
$$

By (21), we know $\sum_{j \in T}(\lambda_{1,j} - \lambda_{2,j})(\mu_j - x_j) = -\rho(\lambda_{1,j} + \lambda_{2,j})$. Hence, (23) becomes

$$
\|x_{T^c}\|_1 - \sum_{j \in T}(\lambda_{1,j} - \lambda_{2,j})(\beta_j - \mu_j) + \sum_{j \in T}\rho(\lambda_{1,j} + \lambda_{2,j}) \tag{24}
$$

Since $|\lambda_{1,j} - \lambda_{2,j}| \leq \lambda_{1,j} + \lambda_{2,j}$ and $|\beta_j - \mu_j| \leq \rho$, we get

$$
-\sum_{j \in T}(\lambda_{1,j} - \lambda_{2,j})(\beta_j - \mu_j) + \sum_{j \in T}\rho(\lambda_{1,j} + \lambda_{2,j}) \geq 0 \tag{25}
$$

Therefore, we get

$$
\|\beta_{T^c}\|_1 \geq \|x_{T^c}\|_1 \tag{26}
$$

This can only happen when $\|\beta_{T^c}\|_1 = \|x_{T^c}\|_1$. Consider the first inequality, since $|w'A_j| < 1$ for $j \notin T \cup \Delta$, then $\beta_j = 0$ for all $j \notin T \cup \Delta$. Since $A\beta - Ax = 0$ and they are both supported on $T \cup \Delta$, we know $A_{T \cup \Delta}(\beta_{T \cup \Delta} - x_{T \cup \Delta}) = 0$. Since $\delta_{|T|+|\Delta|} < 1$, $A_{T \cup \Delta}$ has full rank. Therefore, $\beta_{T \cup \Delta} = x_{T \cup \Delta}$. Finally, we can conclude that $\beta = x$, thus, x is the unique minimizer.

Now, by complementary slackness, we know $\lambda_{1,j} = 0$ and $\lambda_{2,j} = 0$ if $|x_j - \mu_j| \neq \rho$. Therefore, $w'A_j = 0$ for $j \notin (T_{a_1} \cup T_{a_2})$. When $x_j - \mu_j = \rho$, $\lambda_{2,j} = 0$; when $x_j - \mu_j = -\rho$, $\lambda_{1,j} = 0$. Define

$$
T_{a_1} = \{j : x_j - \mu_j = \rho\}, \quad T_{a_2} = \{j : x_j - \mu_j = -\rho\} \tag{27}
$$

Therefore, we know

$$
\begin{aligned}
w'A_j &= \lambda_{1,j} \text{ if } j \in T_{a_1} \tag{28} \\
w'A_j &= -\lambda_{2,j} \text{ if } j \in T_{a_2} \tag{29} \\
w'A_j &= 0 \text{ if } j \in T_{in} \tag{30}
\end{aligned}
$$

Finally, we can get Lemma 3.

C. Proof Outline of Theorem 1

To prove Theorem 1, apply Lemma 4 iteratively, in a fashion similar to that of the proof of [[1],Theorem 1]. The main idea is as follows. At iteration zero, apply Lemma 5 with $T_d \equiv \Delta$ (so that $S \equiv u$), $c_j \equiv sgn(x_j) \ \forall \ j \in \Delta$ (so that $\|c\|_2 = \sqrt{u}$), and with $\check{S} \equiv u$, to get a w_1 and an exceptional set $T_{d,1}$, disjoint with $T \cup \Delta$, of size less than $\check{S} = u$. Lemma 5 can be applied because $\delta_u + \delta_k + \theta_{k,u}^2 < 1$ (follows from condition 7) and condition 6 holds. At iteration $r > 0$, apply Lemma 4 with $T_d \equiv \Delta \cup T_{d,r}$ (so that $S \equiv 2u$), $c_j \equiv 0 \ \forall \ j \in \Delta$, $c_j \equiv A_j'w_r \ \forall \ j \in T_{d,r}$ and $\check{S} \equiv u$ to get w_{r+1} and an exceptional set $T_{d,r+1}$. Lemma 4 is applicable in the above fashion because condition 7 of Theorem 1 holds. Define $w := \sum_{r=1}^{\infty}(-1)^{r-1}w_r$ .We then argue that if condition 7 of Theorem 1 holds, w satisfies the conditions of Lemma 3. Applying Lemma 3, the result follows. We give the entire proof in the long version[13].

D. Proof of Lemma 2

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

$$\|h_T\|_2 \le \|x_T - \mu_T\|_2 + \|\hat{x}_T - \mu_T\|_2 \le 2\rho\sqrt{k} \qquad (31)$$

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

$$\|h_{(T \cup \Delta \cup \Delta_1)^c}\|_2 \le \sum_{j \ge 2}\|h_{\Delta_j}\|_2 \le \frac{1}{\sqrt{u}}\|h_{(T \cup \Delta)^c}\|_1 \qquad (32)$$

Since $\hat{x} = x + h$ is the minimizer of (2) and since both x and $\hat{x}$ are feasible,

$$\begin{aligned}\|x_{T^c}\|_1 &\ge \|(x+h)_{T^c}\|_1 \\ &\ge \|x_\Delta\|_1 - \|h_\Delta\|_1 + \|h_{(T \cup \Delta)^c}\|_1 - \|x_{(T \cup \Delta)^c}\|_1 (33)\end{aligned}$$

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

$$\|h_{(T \cup \Delta)^c}\|_1 \le \|h_\Delta\|_1 \qquad (34)$$

Combining this with (32), and using $\frac{\|h_\Delta\|_1}{\sqrt{u}} \le \|h_\Delta\|_2$, we get

$$\|h_{(T \cup \Delta \cup \Delta_1)^c}\|_2 \le \sum_{j \ge 2}\|h_{\Delta_j}\|_2 \le \|h_\Delta\|_2$$

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

$$Ah = A(\hat{x} - x) = 0 \qquad (35)$$

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

$$(1 - \delta_{2u})\|h_{\Delta \cup \Delta_1}\|_2^2 \le \|Ah_{\Delta \cup \Delta_1}\|_2^2 \qquad (36)$$

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

$$\|Ah_{\Delta \cup \Delta_1}\|_2^2 = <Ah_{\Delta \cup \Delta_1}, Ah> -$$
$$\sum_{j \ge 2}<Ah_{\Delta \cup \Delta_1}, Ah_{\Delta_j}> - <Ah_{\Delta \cup \Delta_1}, Ah_T> \qquad (37)$$

Using (35),

$$|<Ah_{\Delta \cup \Delta_1}, Ah>| = 0 \qquad (38)$$

Using RIP and (35),

$$\begin{aligned}|\sum_{j \ge 2}<Ah_{\Delta \cup \Delta_1}, Ah_{\Delta_j}>| &\le |\sum_{j \ge 2}<Ah_\Delta, Ah_{\Delta_j}>| \\ +|\sum_{j \ge 2}<Ah_{\Delta_1}, Ah_{\Delta_j}>| &\le \sqrt{2}\delta_{2u}\|h_{\Delta \cup \Delta_1}\|_2\|h_\Delta\|_2 \qquad (39)\end{aligned}$$

Finally, using RIP and (31),

$$|<Ah_{\Delta \cup \Delta_1}, Ah_T>| \le \delta_{k+2u}\|h_{\Delta \cup \Delta_1}\|_2\|h_T\|_2 \qquad (40)$$
$$\le \delta_{k+2u}\|h_{\Delta \cup \Delta_1}\|_2 2\rho\sqrt{k} \qquad (41)$$

Combining the above 5 equations, we get

$$(1 - \delta_{2u})\|h_{\Delta \cup \Delta_1}\|_2 \le 2\delta_{k+2u}\rho\sqrt{k} + \sqrt{2}\delta_{2u}\|h_\Delta\|_2 \qquad (42)$$

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

$$\|h_{\Delta \cup \Delta_1}\|_2 \le \frac{2\sqrt{k}\delta_{k+2u}}{1 - (\sqrt{2}+1)\delta_{2u}}\rho, \text{ where}$$

Combining with (35) and (31), we get

$$\begin{aligned}\|h\|_2 &\le \|h_{\Delta \cup \Delta_1}\|_2 + \|h_{(T \cup \Delta \cup \Delta_1)^c}\|_2 + \|h_T\|_2 \qquad (43)\\ &\le 2\|h_{\Delta \cup \Delta_1}\|_2 + 2\rho \le B_1 \qquad (44)\end{aligned}$$

REFERENCES

[1] N. Vaswani and W. Lu, "Modified-CS: Modifying Compressive Sensing for Problems with Partially Known Support", IEEE Trans. Signal Processing, Sept. 2010.

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

[3] W. Lu and N. Vaswani, "Modified Basis Pursuit Denoising (Modified-BPDN) For Noisy Compressive Sensing With Partially Known Support", ICASSP, 2010.

[4] E. Candes and T. Tao, "Decoding by Linear Programming", IEEE Trans. Info. Th., 51(12):4203 - 4215, Dec. 2005.

[5] E. Candes, "The restricted isometry property and its implications for compressed sensing", Compte Rendus de lAcademie des Sciences, Paris, Serie I, pp. 589C592, 2008

[6] L. Jacques, "A short Note on Compressed Sensing with Partially Known Signal Support", ArXiv preprint, arXiv:0908.0660v1, 2009.

[7] A. Carmi and P. Gurfil and D. Kanevsky, "A simple method for sparse signal recovery from noisy observations using Kalman filtering", IBM Technical report, Dec. 2008

[8] R. von Borries, C. J. Miosso, and C. Potes, "Compressed sensing using prior information", CAMSAP, 2007.

[9] A. Khajehnejad, W. Xu, A. Avestimehr, B. Hassibi, "Weighted l1 Minimization for Sparse Recovery with Prior Information", IEEE Intl. Symp. Info. Theory(ISIT),2009

[10] M. Asif and J. Romberg, "Dynamic updating for sparse time varying signals", CISS, 2009.

[11] D. Angelosante, E. Grossi, G. B. Giannakis, "Compressed Sensing of Time-varying Signals", DSP 2009

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

[13] W. Lu, N. Vaswani, "Exact Reconstruction Conditions and Error Bounds for Regularized Modified Basis Pursuit (Reg-Modified-BP)", long version, http://www.ece.iastate.edu/~luwei/modcs/regmodBPlong.pdf