COSAMP AND SP FOR THE COSPARSE ANALYSIS MODEL

Raja Giryes and Michael Elad

Computer Science Department
Technion - Israel Institute of Technology

ABSTRACT

CoSaMP and Subspace-Pursuit (SP) are two recovery algorithms that find the sparsest representation for a given signal under a given dictionary in the presence of noise. These two methods were conceived in the context of the synthesis sparse representation modeling. The cosparse analysis model is a recent construction that stands as an interesting alternative to the synthesis approach. This new model characterizes signals by the space they are orthogonal to. Despite the similarity between the two, the cosparse analysis model is markedly different from the synthesis one. In this paper we propose analysis versions of the CoSaMP and the SP algorithms, and demonstrate their performance for the compressed sensing problem.

***Index Terms*—** Sparse representations, Compressed Sensing, Synthesis, Analysis, CoSaMP, Subspace-Pursuit.

1. INTRODUCTION

In many signal and image processing applications we encounter the following problem: recovering an original signal $\mathbf{x} \in \mathbb{R}^d$ from a set of noisy measurements

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}, \quad (1.1)$$

where $\mathbf{M} \in \mathbb{R}^{m\times d}$ is a known linear operator and $\mathbf{e} \in \mathbb{R}^d$ is an additive bounded noise, i.e., $\|\mathbf{e}\|_2^2 \le \epsilon^2$. In many cases $m < d$ and thus (1.1) has infinite number of solutions. One such example is the problem of compressed sensing where $\mathbf{M}$ is the measurement matrix. Other examples include signal interpolation and inpainting. Thus, we should rely on priors that we may have on $\mathbf{x}$, in order to be able to recover the signal.

A popular and very effective prior is one that is based on sparsity. This assumption leads to two possible models, the synthesis and the analysis models [1]. The synthesis model, which received great attention in the past decade, assumes that $\mathbf{x}$ has a k-sparse representation $\boldsymbol{\alpha}$ under a given dictionary $\mathbf{D} \in \mathbb{R}^{d\times n}$ [2]. In other words, there exists a vector $\boldsymbol{\alpha} \in \mathbb{R}^n$ such that $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$ and $\|\boldsymbol{\alpha}\|_0 = k$, i.e., $\boldsymbol{\alpha}$ has k non-zero elements. Having the synthesis constraint we can recover $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$ by solving

$$\hat{\boldsymbol{\alpha}}_{\ell_0} = \operatorname{argmin} \|\boldsymbol{\alpha}\|_0 \quad s.t. \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|_2 \le \epsilon. \quad (1.2)$$

Since solving (1.2) is an NP-hard problem [2], approximation techniques are required for recovering $\mathbf{x}$. One strategy is by using relaxation, replacing the ℓ_0 with ℓ_1 norm, resulting with the ℓ_1-synthesis problem

$$\hat{\boldsymbol{\alpha}}_{\ell_1} = \operatorname{argmin} \|\boldsymbol{\alpha}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\boldsymbol{\alpha}\|_2 \le \epsilon. \quad (1.3)$$

Another option for approximating (1.2) is using a greedy strategy. Two recently introduced methods that use this strategy are CoSaMP [3] and subspace pursuit (SP) [4], both described in algorithm 1. In the algorithm $\operatorname{supp}(\cdot, K)$ returns the set of K largest elements, and the subscript T denotes taking the elements (columns) in places corresponding to T in a vector (a matrix).

For a unitary $\mathbf{D}$, a vector $\mathbf{x}$ with a k-sparse representation $\boldsymbol{\alpha}$ and $\hat{\mathbf{x}}_{\ell_1} = \mathbf{D}\hat{\boldsymbol{\alpha}}_{\ell_1}$ it was shown that if $\delta_{2k} < \delta_{\ell_1}$ then

$$\|\hat{\mathbf{x}}_{\ell_1} - \mathbf{x}\|_2 \le c_{\ell_1}\epsilon, \quad (1.4)$$

implying a stable recovery, where δ_{2k} is the constant of the restricted isometry property (RIP) of $\mathbf{M}$ for $2k$ sparse signals, and c_{ℓ_1} $(\ge \sqrt{2})$ and δ_{ℓ_1} $(\simeq 0.4652)$ are constants [5, 6]. CoSaMP and SP were the first greedy methods shown to have recovery guarantees in the form of (1.4) assuming $\delta_{4k} < \delta_{\text{CoSaMP}}$ and $\delta_{3k} < \delta_{\text{SP}}$ where δ_{3k} and δ_{4k} are the constants of the RIP of $\mathbf{M}$ for $3k$ and $4k$ sparse signals respectively, and δ_{CoSaMP} and δ_{SP} are reference constants [3, 4]. Following these results, the iterative hard thresholding (IHT) [7] and hard thresholding pursuit (HTP) [8] were shown to have similar reconstruction guarantees under similar conditions. The above were extended also for incoherent redundant dictionaries [9].

In the above results no dependencies were allowed in the dictionary $\mathbf{D}$. Candès *et. al.*[10] considered

$$\hat{\mathbf{x}}_{A-\ell_1} = \operatorname*{argmin}_{\tilde{\mathbf{x}}} \|\mathbf{D}^*\tilde{\mathbf{x}}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\|_2 \le \epsilon, \quad (1.5)$$

with the assumptions that $\mathbf{x}$ has a k-sparse representation under a tight frame $\mathbf{D}$ and $\mathbf{M}$ has a $\mathbf{D}$-RIP, an extension of the RIP for the case that $\mathbf{D}$ is non-unitary, with $\delta_{7k} \le 0.6$. We

This work was supported in part by the EU SMALL FET-Open program project (grant #225913). R. Giryes is also grateful to the Azrieli Foundation for the award of an Azrieli Fellowship.

Algorithm 1 Subspace Pursuit (SP) and CoSaMP

Require: $k, \mathbf{M}, \mathbf{D}, \mathbf{y}, a$ where $\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}$ and $\mathbf{x} = \mathbf{D}\boldsymbol{\alpha}$, k is the cardinality of $\boldsymbol{\alpha}$ and $\mathbf{e}$ is the additive noise. $a = 1$ (SP), $a = 2$ (CoSaMP).

Result: $\hat{\mathbf{x}}_{\text{CoSaMP}}$ or $\hat{\mathbf{x}}_{\text{SP}}$: k-sparse approximation of $\mathbf{x}$.

Initialize the support $T^0 = \emptyset$, the residual $\mathbf{y}_r^0 = \mathbf{y}$ and set $t = 0$.

while halting criterion is not satisfied **do**

$t = t + 1$.

Find new support elements:

$$T_\Delta = \text{supp}(\mathbf{D}^*\mathbf{M}^*\mathbf{y}_r^{t-1}, ak).$$

Update the support: $\tilde{T}^t = T^{t-1} \cup T_\Delta$.

Compute a temporal estimate: $\boldsymbol{\alpha}_p = (\mathbf{M}\mathbf{D}_{\tilde{T}^t})^\dagger \mathbf{y}$.

Prune small entries: $T^t = \text{supp}(\boldsymbol{\alpha}_p, k)$.

Calculate a new estimate: $\hat{\mathbf{x}}_{\text{CoSaMP}}^t = \mathbf{D}_{T^t}(\boldsymbol{\alpha}_p)_{T^t}$ for CoSaMP, and $\hat{\mathbf{x}}_{\text{SP}}^t = \mathbf{D}(\mathbf{M}\mathbf{D}_{T^t})^\dagger\mathbf{y}$ for SP.

Update the residual: $\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}_{\text{CoSaMP}}^t$ for CoSaMP, and $\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}_{\text{SP}}^t$ for SP.

end while

Form final solution $\hat{\mathbf{x}}_{\text{CoSaMP}} = \hat{\mathbf{x}}_{\text{CoSaMP}}^t$ for CoSaMP and $\hat{\mathbf{x}}_{\text{SP}} = \hat{\mathbf{x}}_{\text{SP}}^t$ for SP.

---

say that a matrix $\mathbf{M}$ has a $\mathbf{D}$-RIP with a constant δ_k if for any signal $\tilde{\mathbf{x}}$ that has a k-sparse representation under $\mathbf{D}$

$$(1 - \delta_k)\left\|\tilde{\mathbf{x}}\right\|_2^2 \leq \left\|\mathbf{M}\tilde{\mathbf{x}}\right\|_2^2 \leq (1 + \delta_k)\left\|\tilde{\mathbf{x}}\right\|_2^2. \quad (1.6)$$

The authors in [10] showed that under the above assumptions

$$\left\|\hat{\mathbf{x}}_{A-\ell_1} - \mathbf{x}\right\|_2 \leq \tilde{c}_{\ell_1}\epsilon + \left\|\mathbf{D}^*\mathbf{x} - [\mathbf{D}^*\mathbf{x}]_k\right\|_1/\sqrt{k}, \quad (1.7)$$

where $\tilde{c}_{\ell_1}$ is a constant and $[\cdot]_k$ is a hard thresholding operator that keeps the largest k elements and zeros the rest. In other words, this result gives a reconstruction guarantee for redundant dictionaries with linear dependencies at the cost of limiting the family of signals to those for which $\left\|\mathbf{D}^*\mathbf{x} - [\mathbf{D}^*\mathbf{x}]_k\right\|_1$ is small.

The minimization problem (1.7) is a special case of the ℓ_1-analysis minimization problem in which $\mathbf{D}^*$ is replaced by a general operator $\boldsymbol{\Omega} \in \mathbb{R}^{p \times d}$ [1]. Similar to what we have in the synthesis framework, the ℓ_1-analysis is a relaxation strategy for solving the analysis problem:

$$\underset{\tilde{\mathbf{x}}}{\text{argmin}} \left\|\boldsymbol{\Omega}\tilde{\mathbf{x}}\right\|_0 \quad s.t. \quad \left\|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\right\|_2 \leq \epsilon. \quad (1.8)$$

Though similar to (1.2), the analysis problem implies another sparsity model altogether [11]. In this model, instead of a dictionary $\mathbf{D}$ with columns that synthesize the signal, we use an operator $\boldsymbol{\Omega}$ that analyzes the signal by checking which of its rows are orthogonal to the signal. Instead of looking at the non-zero elements of the representation of $\mathbf{x}$ under $\mathbf{D}$, we look at the zero coefficients of $\boldsymbol{\Omega}\mathbf{x}$. Each zero element in $\boldsymbol{\Omega}\mathbf{x}$ implies a subspace orthogonal to $\mathbf{x}$. An example for $\boldsymbol{\Omega}$ is the two-dimensional finite difference operator, known also as the two dimensional total variation (2D-TV). More details about the subspaces implied by this operator can be found in [1, 11].

Algorithm 2 Analysis Subspace Pursuit (ASP) and Analysis CoSaMP (ACoSaMP)

Require: $l, \mathbf{M}, \boldsymbol{\Omega}, \mathbf{y}, a$ where $\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}$, l is the cosparsity of $\mathbf{x}$ under $\boldsymbol{\Omega}$ and $\mathbf{e}$ is the additive noise. $a = 1$ (ASP), $a = \frac{2l-p}{l}$ (ACoSaMP).

Result: $\hat{\mathbf{x}}_{\text{ACoSaMP}}$ or $\hat{\mathbf{x}}_{\text{ASP}}$: l-cosparse approximation of $\mathbf{x}$.

Initialize the cosupport $\Lambda^0 = \{i, 1 \leq i \leq p\}$, the residual $\mathbf{y}_r^0 = \mathbf{y}$ and set $t = 0$.

while halting criterion is not satisfied **do**

$t = t + 1$.

Find new cosupport elements:

$$\Lambda_\Delta = \text{cosupp}(\boldsymbol{\Omega}\mathbf{M}^*\mathbf{y}_r^{t-1}, al).$$

Update the cosupport: $\tilde{\Lambda}^t = \Lambda^{t-1} \cap \Lambda_\Delta$.

Compute a temporal estimate:

$$\mathbf{x}_p = \text{argmin}_{\tilde{\mathbf{x}}} \left\|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\right\|_2^2 \text{ s.t. } \boldsymbol{\Omega}_{\tilde{\Lambda}^t}\tilde{\mathbf{x}} = 0.$$

Enlarge cosupport: $\Lambda^t = \text{cosupp}(\boldsymbol{\Omega}\mathbf{x}_p, l)$.

Calculate a new estimate: $\hat{\mathbf{x}}_{\text{ACoSaMP}}^t = \mathbf{Q}_{\Lambda^t}\mathbf{x}_p$ for ACoSaMP, and $\hat{\mathbf{x}}_{\text{ASP}}^t = \text{argmin}_{\tilde{\mathbf{x}}} \left\|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\right\|_2^2$ s.t. $\boldsymbol{\Omega}_{\Lambda^t}\tilde{\mathbf{x}} = 0$ for ASP.

Update the residual: $\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}_{\text{ACoSaMP}}^t$ for ACoSaMP, and $\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}_{\text{ASP}}^t$ for ASP.

end while

Form final solution $\hat{\mathbf{x}}_{\text{ACoSaMP}} = \hat{\mathbf{x}}_{\text{ACoSaMP}}^t$ for ACoSaMP and $\hat{\mathbf{x}}_{\text{ASP}} = \hat{\mathbf{x}}_{\text{ASP}}^t$ for ASP.

A vector $\tilde{\mathbf{x}}$ with l zero coefficients in $\boldsymbol{\Omega}\tilde{\mathbf{x}}$ is denoted as an l-cosparse vector and the set of indices of the corresponding l rows in $\boldsymbol{\Omega}$ as the cosupport set Λ. If $\boldsymbol{\Omega}_\Lambda$, a sub-matrix of $\boldsymbol{\Omega}$ with rows that belong to Λ, has a full-rank then all vectors with cosupport Λ belong to a subspace of dimension $d - l$.

As in the synthesis case, solving (1.8) is an NP-hard problem [11] and thus approximations are needed. One option is using an ℓ_1 relaxation. Another approach is using a greedy strategy. The first algorithm to take this approach was greedy analysis pursuit (GAP) [11, 12]. This algorithm starts with a full cosupport and in each iteration removes one element from it. Other two methods are analysis IHT (AIHT) and analysis HTP (AHTP), which have been shown to have reconstruction guarantees similar to the RIP-based guarantees of IHT and HTP in the synthesis context [13, 14].

In this paper we present analysis versions for CoSaMP and SP. We demonstrate their performance for inverse problems and more specifically for compressed sensing. Note that in this paper we state a theoretical result that guarantees the performance of these newly formed algorithms. The proofs and more appear in [14].

Analysis CoSaMP (ACoSaMP) and analysis SP (ASP) are presented in Section 2. In section 3 we repeat the experiments in [11] presenting phase transition diagrams for these techniques and a reconstruction of a high dimensional image from few measurements. Section 4 concludes the paper.

Stage	Synthesis	Analysis	Explanation	$\mathbf{D} = \mathbf{\Omega} = \mathbf{I}$
Choice of a	$a = 1$ for SP and 2 for CoSaMP	$a = 1$ for ASP and $\frac{2l-p}{l}$ for ACoSaMP	Since $p > l$, we have that $\frac{2l-p}{l} < 1$. This implies that Λ_Δ in ACoSAMP is smaller (compared to ASP), leading to a larger subspace implied by the intersection set.	Both the synthesis and analysis algorithms lead to the same support/cosupport.
Finding new elements	Largest elements in $\mathbf{D}^*\mathbf{M}^*\mathbf{y}_r$	Smallest elements in $\mathbf{\Omega}\mathbf{M}^*\mathbf{y}_r$	Synthesis: we focus on the non-zeros, and thus take the largest elements. Analysis: we focus on the zeros and thus take the smallest elements.	Selecting the largest elements for T_Δ is the same as selecting the smallest elements for $\Lambda_\Delta = T_\Delta^C$.
Updating the (co)support	$T^{t-1} \cup T_\Delta$	$\Lambda^{t-1} \cap \Lambda_\Delta$	Union of support sets is equivalent to intersection of cosupport sets.	$\Lambda^{t-1} \cap \Lambda_\Delta = (T^{t-1})^C \cap T_\Delta^C = T^{t-1} \cup T_\Delta$.
Compute a temporal estimate	$(\mathbf{M}\mathbf{D}_{\tilde{T}^t})^\dagger \mathbf{y}$	$\min_{\tilde{\mathbf{x}}} \|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\|_2$ s.t. $\mathbf{\Omega}_{\tilde{\Lambda}^t}\tilde{\mathbf{x}} = 0$	$(\mathbf{M}\mathbf{D}_{\tilde{T}^t})^\dagger \mathbf{y}$ is the solution to $\min_{\tilde{\mathbf{x}}} \|\mathbf{y} - \mathbf{M}\mathbf{D}\tilde{\boldsymbol{\alpha}}\|_2$ s.t. $\tilde{\boldsymbol{\alpha}}_{(\tilde{T}^t)^C} = 0$ and $\tilde{\boldsymbol{\alpha}}_{(\tilde{T}^t)^C} = 0$ is parallel to $\mathbf{\Omega}_{\tilde{\Lambda}^t}\tilde{\mathbf{x}} = 0$.	$\mathbf{I}_{\tilde{\Lambda}^t}\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_{(\tilde{T}^t)^C}$.
Pruning/ Enlarging	$\mathrm{supp}(\boldsymbol{\alpha}_p, k)$	$\mathrm{cosupp}(\mathbf{\Omega}\mathbf{x}_p, l)$	Same argument as for finding new elements.	$\mathrm{supp}(\boldsymbol{\alpha}_p, k) = \mathrm{supp}(\mathbf{x}_p, k) = \mathrm{cosupp}(\mathbf{x}_p, l)^C$.
New estimate for SP/ASP	$(\mathbf{M}\mathbf{D}_{T^t})^\dagger \mathbf{y}$	$\min_{\tilde{\mathbf{x}}} \|\mathbf{y} - \mathbf{M}\tilde{\mathbf{x}}\|_2$ s.t. $\mathbf{\Omega}_{\Lambda^t}\tilde{\mathbf{x}} = 0$	Same argument as for computing the temporal estimate.	$\mathbf{I}_{\Lambda^t}\tilde{\mathbf{x}} = \tilde{\mathbf{x}}_{T^{tC}}$.
New estimate for CoSaMP/ ACoSaMP	$(\boldsymbol{\alpha}_p)_{T^t}$	$\mathbf{Q}_{\Lambda^t}\mathbf{x}_p$	$(\boldsymbol{\alpha}_p)_{T^t}$ is a projection to the subspace of k-sparse vectors supported on T^t. Its equivalence in the analysis model is $\mathbf{Q}_{\Lambda^t}\mathbf{x}$ which projects to the subspace of l-cosparse vectors cosupported on Λ^t.	$\mathbf{Q}_{\Lambda^t}\mathbf{x}_p = \mathbf{x}_p - (\mathbf{I}_{(T^t)^C})^\dagger \mathbf{I}_{(T^t)^C}\mathbf{x}_p = (\mathbf{x}_p)_{T^t} = (\boldsymbol{\alpha}_p)_{T^t}$.
Residual	$\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}^t$	$\mathbf{y}_r^t = \mathbf{y} - \mathbf{M}\hat{\mathbf{x}}^t$	Same operation in both versions.	

Table 1. Comparison between CoSaMP and SP and their analysis versions

2. NEW ANALYSIS ALGORITHMS

The proposed ACoSaMP and ASP are presented in algorithm 2. $\mathbf{Q}_\Lambda = \mathbf{I} - \mathbf{\Omega}_\Lambda^\dagger \mathbf{\Omega}_\Lambda$ denotes the projection onto the orthogonal complement of $\mathrm{range}(\mathbf{\Omega}_\Lambda^T)$, $\mathbf{I}$ is the identity matrix and $\mathrm{cosupp}(\cdot, l)$ returns the set of l smallest elements. As a stopping criteria one can look at the convergence rate or the residual size.

To see the analogy between CoSaMP and SP and their analysis versions, we consider the following: Given two cosupport sets Λ and $\tilde{\Lambda}$, where $|\Lambda| = l$ (the size of Λ is l) and $|\tilde{\Lambda}| = \tilde{l}$, it holds that $|\Lambda \cap \tilde{\Lambda}| \geq l + \tilde{l} - p$. This implies that for the case $|\Lambda| = |\tilde{\Lambda}| = l$ we have $l \geq |\Lambda \cap \tilde{\Lambda}| \geq 2l - p$. In the synthesis case, the result of adding two k-sparse signals is a signal with sparsity of at most $2k$. However, in the analysis case the result of adding two l-cosparse signals is a signal with cosparsity of at least $2l - p$. This implies that union of support sets in the synthesis case is parallel to intersection of cosupport sets in the analysis. Based on the above observations, Table 1 presents the analogy between the methods.

The last column of the table refers to the case $\mathbf{D} = \mathbf{\Omega} = \mathbf{I}$, where ACoSaMP and ASP become the same as CoSaMP and SP. In this case we have $p = d$, $k = d - l$, $T^C = \Lambda$, $\tilde{T}^C = \tilde{\Lambda}$, $\mathbf{Q}_\Lambda \mathbf{x} = \mathbf{x}_T$ and $T \cup \tilde{T} = \Lambda \cap \tilde{\Lambda}$. In this case the recovery guarantees of CoSaMP and SP apply also for ACoSaMP and ASP in a trivial way. Thus, it is tempting to assume that ACoSaMP and ASP should have similar guarantees given that the $\mathbf{\Omega}$-RIP constant, the equivalent property of the $\mathbf{D}$-RIP [10] in the analysis case, is small. We say that a matrix $\mathbf{M}$ has an $\mathbf{\Omega}$-RIP with a constant δ_l if for any l-cosparse signal $\tilde{\mathbf{x}}$

$$(1 - \delta_l) \|\tilde{\mathbf{x}}\|_2^2 \leq \|\mathbf{M}\tilde{\mathbf{x}}\|_2^2 \leq (1 + \delta_l) \|\tilde{\mathbf{x}}\|_2^2 . \quad (2.1)$$

This property alone is not enough for having recovery guarantees. Though we have many similarities between the two versions of the algorithms, there is a vital difference in the case where $\mathbf{\Omega} \neq \mathbf{I}$. Given a vector $\mathbf{z} \in \mathbb{R}^d$, finding the support of its closest (in the ℓ_2-norm sense) k-sparse vector is done by simply taking the support of the largest k elements. However, in the analysis case finding an l-cosparse vector that is closest to the original vector is a combinatorial problem. Choosing the l smallest elements in $\mathbf{\Omega}\mathbf{z}$ is not necessarily the optimal solution. For this reason we introduce the definition of near-optimal projection [13].

Definition 2.1 *Given a projection $\mathbf{P}_{\mathbf{\Omega},l}$ that projects to an l-cosparse subspace, we say that it is near-optimal with a constant C_l (≥ 1) if for any $\mathbf{z} \in \mathbb{R}^d$*

$$\|\mathbf{z} - \mathbf{P}_{\mathbf{\Omega},l}\mathbf{z}\|_2^2 \leq C_l \min_{\tilde{\mathbf{x}} \text{ } l\text{-cosparse}} \|\tilde{\mathbf{x}} - \mathbf{z}\|_2^2 . \quad (2.2)$$

The ACoSaMP and ASP uses the following projection:

$$\tilde{\mathbf{P}}_{\mathbf{\Omega},l}\mathbf{z} = \mathbf{Q}_{\Lambda}\mathbf{z} \text{ where } \Lambda = \text{cosupp}(\mathbf{\Omega}, l). \quad (2.3)$$

In the general case, the near-optimality constant C_l for this projection is bounded by the RIP constant of $\mathbf{\Omega}^*$. More details about bounding C_l can be found in [13]. We should note here that this is not the only possible choice for a projection and other cosupport selection methods can be used within the ACoSaMP and ASP techniques.

Armed with the above definition we present the following theorem that provides guarantees for ACoSaMP and ASP with $a = \frac{2l-p}{l}$. More details can be found in [14].

Theorem 2.2 (Theorem 4.1 in [14]) *Consider* $\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}$ *where* $\mathbf{x}$ *is an l-cosparse vector. Apply either ACoSaMP or ASP with* $a = \frac{2l-p}{l}$*, obtaining* $\hat{\mathbf{x}}^t$ *after t iterations. If*

$$(1 + \tilde{C})\left(1 - \left(\tilde{C} - (\tilde{C} - 1)\sigma^2_{\mathbf{M}}\right)\right) < 1, \quad (2.4)$$

and $\delta_{4l-3p} < \delta_2(\tilde{C}, \sigma^2_{\mathbf{M}})$*, where* $\tilde{C} = \max(C_l, C_{2l-p})$ *and* $\delta_2(\tilde{C}, \sigma^2_{\mathbf{M}})$ *is a constant guaranteed to be greater than zero whenever* (2.4) *is satisfied, then after a finite number of iterations* t^*

$$\left\|\mathbf{x} - \hat{\mathbf{x}}^{t^*}\right\|_2 \leq c \left\|\mathbf{e}\right\|_2. \quad (2.5)$$

implying that these algorithms lead to a stable recovery. The constant c is a function of δ_{4l-3p}*,* C_l*,* C_{2l-p} *and* $\sigma^2_{\mathbf{M}}$*, where* $\sigma^2_{\mathbf{M}}$ *is the largest singular value of* $\mathbf{M}$ *and* C_l *and* C_{2l-p} *are the constants of the near optimal projection* $\mathbf{P}_{\mathbf{\Omega},l}$*.*

Before we move to the next section we present a variation of the ASP and ACoSaMP. In the algorithms' steps we need to solve the constrained optimization problem $\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2$ s.t. $\|\mathbf{\Omega}_{\Lambda}\mathbf{x}\|_2^2 = 0$. In high dimensional problems this task is hard to solve and we propose a relaxed version of the algorithms that minimizes $\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \|\mathbf{\Omega}_{\Lambda}\mathbf{x}\|_2^2$ instead. We denote these relaxed versions as relaxed ASP (RASP) and relaxed CoSaMP (RACoSaMP).

3. EXPERIMENTS

In this section we repeat some of the experiments performed in [11] for the noiseless case ($\mathbf{e} = 0$) and some of the experiments reported in [12] for the noisy case. We begin with synthetic signals in the noiseless case. We test the performance of ASP with $a = 1$ and ACoSaMP with $a = \frac{2l-p}{l}$ and $a = 1$. The performance of RACoSaMP and RASP are similar to those of ACoSaMP and ASP and thus omitted. We compare the results to those of ℓ_1-minimization [1], GAP [11] and AHTP [13]. We use a random matrix $\mathbf{M}$, where each entry in the matrix is drawn independently from the Gaussian distribution, and a random tight frame $\mathbf{\Omega}$ of size $d = 200$ and $p = 240$.

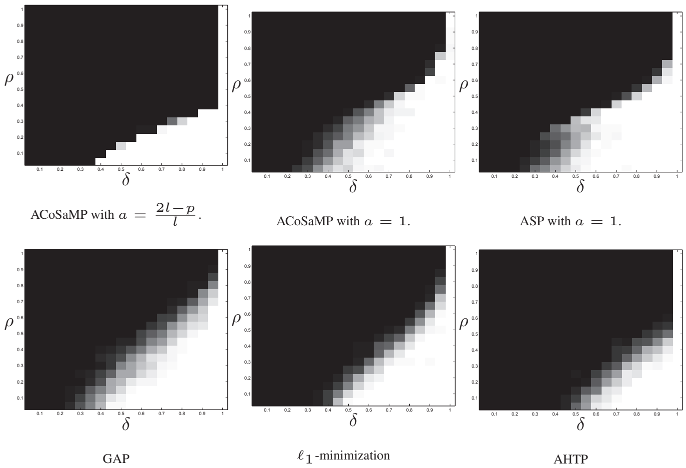

Fig. 1. Phase transitions for ACoSaMP with $a = \frac{2l-p}{l}$, ACoSaMP with $a = 1$, ASP with $a = 1$, GAP, ℓ_1-minimization and AHTP.

We draw a phase transition diagram for each of the algorithms. We test 20 different possible values of m and 20 different values of l and for each pair repeat the experiment 50 times. In each experiment we check whether we have a perfect reconstruction. White cells in the diagram denote a perfect reconstruction in all the experiments of the pair and black cells denote total failure in the reconstruction. The values of m and l are selected according to the formula

$$m = \delta d \qquad l = d - \rho m, \quad (3.1)$$

where δ, the sampling rate, is the x-axis of the phase diagram and ρ, the ratio between the cosparsity of the signal and the number of measurements, is the y-axis.

Figure 1 presents the reconstruction results for ASP with $a = 1$ and ACoSaMP with $a = 1$ and $\frac{2l-p}{l}$. As a reference we present also the results of the other algorithms. It should be observed that ACoSaMP behaves better with $a = 1$ than with $a = \frac{2l-p}{l}$. Compared to the other algorithms, the reconstruction results of the proposed algorithms with $\alpha = 1$ are better than those of ℓ_1-minimization and AHTP, and competitive to those of GAP. Note that the proposed analysis algorithms are effective especially when l is close to d, leading to a non-empty intersection in the cosupport update stage.

We turn now to test the methods for a high dimensional signal. We use RASP and RACoSaMP for the reconstruction of the *Shepp-Logan phantom* from few number of measurements. The sampling operator is a two dimensional Fourier transform that measures only a certain number of radial lines from the Fourier transform and the cosparse operator is the 2D-TV. The phantom image, its differences image and an example of 22 sampled radial lines are presented in Fig. 2 (a) and (b). Using RACoSaMP/RASP we get a perfect reconstruction using only 15/12 radial lines, i.e., only $m = 3782/3032$ measurements out of $d = 65536$ which is less

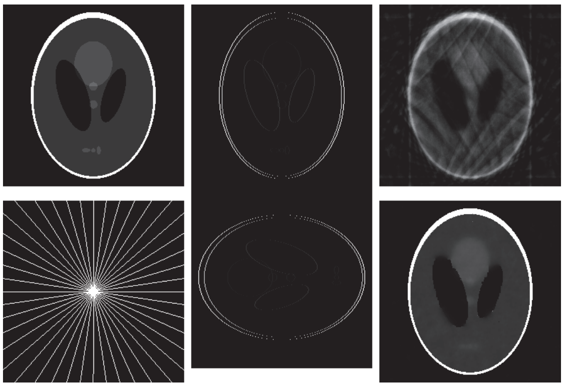

Fig. 2. Shepp-Logan phantom image (top left), 22 sampled radial lines (bottom left), Location of non zero elements in the difference map (middle), Noisy phantom with SNR $=20dB$ (top right) and recovered image (bottom right).

than $5.77\%/4.63\%$ of the data in the original image. The algorithms require less than 20 iterations for getting this perfect recovery.

Exploring the noisy case, we perform the reconstruction using RASP of a noisy measurement of the phantom with 22 radial lines and signal to noise ratio (SNR) of 20. Figure 2 (c) presents the result of applying inverse Fourier transform on the measurements, and its reconstruction result. Note that for the minimization process we use the conjugate gradient algorithm, and in each iteration we take only the real part of the result and crop the values of the resulted image to be in the range of $[0, 1]$. We get a peak SNR (PSNR) of $36.5dB$. We should note that in a similar setting GAPN gives a comparable result with $35.5dB$ [12].

4. CONCLUSION

In this work we have presented two novel algorithms for the analysis model, ACoSaMP and ASP. We have demonstrated their empirical performance showing competitive recovery results compared to other methods. ACoSaMP and ASP are the parallel to CoSaMP and SP and the work in [14] shows that they have similar theoretical recovery guarantees.

5. REFERENCES

[1] M. Elad, P. Milanfar, and R. Rubinstein, “Analysis versus synthesis in signal priors,” *Inverse Problems*, vol. 23, no. 3, pp. 947–968, June 2007.

[2] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” *SIAM Review*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[3] D. Needell and J.A. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301 – 321, May 2009.

[4] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Trans. Inf. Theory.*, vol. 55, no. 5, pp. 2230 –2249, May 2009.

[5] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inf. Theory.*, vol. 52, no. 12, pp. 5406 –5425, Dec. 2006.

[6] S. Foucart, “Sparse recovery algorithms: sufficient conditions in terms of restricted isometry constants,” in *Approximation Theory XIII*. 2010, pp. 65–77, Springer Proceedings in Mathematics.

[7] T. Blumensath and M.E. Davies, “Iterative hard thresholding for compressed sensing,” *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265 – 274, 2009.

[8] S. Foucart, “Hard thresholding pursuit: an algorithm for compressive sensing,” *SIAM J. Numer. Anal.*, vol. 49, no. 6, pp. 2543–2563, 2011.

[9] H. Rauhut, K. Schnass, and P. Vandergheynst, “Compressed sensing and redundant dictionaries,” *IEEE Trans. Inf. Theory.*, vol. 54, no. 5, pp. 2210 –2219, May 2008.

[10] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, “Compressed sensing with coherent and redundant dictionaries,” *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, pp. 59 – 73, 2011.

[11] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, “The cosparse analysis model and algorithms,” *Applied and Computational Harmonic Analysis*, 2012.

[12] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, “Cosparse analysis modeling,” in *9th International Conference on SamplingTheory and Applications (sampta-2011)*, Singapore, 2011.

[13] R. Giryes, S. Nam, R. Gribonval, and M. E. Davies, “Iterative cosparse projection algorithms for the recovery of cosparse vectors,” in *The 19th European Signal Processing Conference (EUSIPCO-2011)*, Barcelona, Spain, 2011.

[14] R. Giryes, S. Nam, M. Elad, R. Gribonval, and M. E. Davies, “Greedy-like algorithms for the cosparse analysis model,” *Submitted to the Special Issue in Linear Algebra and its Applications on Sparse Approximate Solution of Linear Systems*, 2012.