

RESEARCH

Open Access



Cosparsity-based Stagewise Matching Pursuit algorithm for reconstruction of the cosparsity signals

Di Wu¹, Yuxin Zhao^{1*}, Wenwu Wang² and Yanling Hao¹**Abstract**

The cosparsity analysis model has been introduced as an interesting alternative to the standard sparse synthesis model. Given a set of corrupted measurements, finding a signal belonging to this model is known as analysis pursuit, which is an important problem in analysis model based sparse representation. Several pursuit methods have already been proposed, such as the methods based on l_1 -relaxation and greedy approaches based on the cosparsity of the signal. This paper presents a novel greedy-like algorithm, called Cosparsity-based Stagewise Matching Pursuit (CSMP), where the cosparsity of the target signal is estimated adaptively with a stagewise approach composed of forward and backward processes. In the forward process, the cosparsity is estimated and the signal is approximated, followed by the refinement of the cosparsity and the signal in the backward process. As a result, the target signal can be reconstructed without the prior information of the cosparsity level. Experiments show that the performance of the proposed algorithm is comparable to those of the l_1 -relaxation and Analysis Subspace Pursuit (ASP)/Analysis Compressive Sampling Matching Pursuit (ACoSAMP) in noiseless case and better than that of Greedy Analysis Pursuit (GAP) in noisy case.

Keywords: Cosparsity analysis model, Greedy pursuit, Adaptive cosparsity, Stagewise strategy

1 Introduction

Many high-dimensional signals, such as natural images, can be represented in a low-dimensional space. A variety of methods have been proposed in the past for low-dimensional representations of high-dimensional signals. Sparse representation is one of such techniques that have been studied extensively. This issue can be described as recovering an unknown signal \mathbf{x} of a high-dimension from a low-dimensional signal \mathbf{y} having a limited set of measurements:

$$\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e} \quad (1)$$

where $\mathbf{M} \in \mathcal{R}^{m \times d}$ is a known linear operator and $\mathbf{e} \in \mathcal{R}^m$ is additive noise bounded by $\|\mathbf{e}\|_2 \leq \epsilon^2$. In a noiseless case, \mathbf{e} is set to be a zero vector. One such example is the problem of compressed sensing where \mathbf{M} is a measurement matrix. For $m < d$, this is an ill-posed underdetermined problem and thus has an infinite number of solutions, and extra

priors or constraints need to be imposed on the model in order to limit the range of possible solutions to \mathbf{x} .

Sparsity is usually considered as an effective prior in both the synthesis and analysis models [1]. The synthesis model assumes that \mathbf{x} can be represented by $\mathbf{x} = \mathbf{D}\mathbf{z}$, where $\mathbf{D} \in \mathcal{R}^{d \times n}$ is an over-complete dictionary with $n > d$ and \mathbf{z} is the sparse representation coefficient. The original signal \mathbf{x} can be recovered by solving the following optimization problem:

$$\hat{\mathbf{z}} = \underset{\mathbf{z}}{\operatorname{argmin}} \|\mathbf{z}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{M}\mathbf{D}\mathbf{z}\|_2 \leq \epsilon \quad (2)$$

where $\|\cdot\|_0$ is a l_0 pseudo-norm counting the number of non-zero elements in its argument vector.

Since solving (2) is a non-deterministic polynomial (NP)-hard problem [2], many approximation techniques have been proposed to recover \mathbf{x} . Basis pursuit (BP) [3], which is based on the l_1 -minimization using linear programming (LP), is a well-known reconstruction algorithm. Another option for approximating (2) is to use a family of greedy-like algorithms, such as Orthogonal Matching Pursuit [4] or the thresholding technique [5–10].

* Correspondence: zhaoyuxin@hrbeu.edu.cn

¹College of Automation, Harbin Engineering University, Harbin 150001, China
Full list of author information is available at the end of the article

In the analysis model, recovering \mathbf{x} from the incomplete measurements is achieved by solving the following minimization problem [11]:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\Omega\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2 \leq \varepsilon \quad (3)$$

where $\Omega \in \mathbb{R}^{p \times d}$ is a fixed analysis operator which is also referred to as the analysis dictionary. Typically, the dimensions are $m \leq d \leq p, n$. In the analysis model, the cosparsity l (used to distinguish from the term ‘‘sparsity’’ in the synthesis model) is defined as

$$l = p - \|\Omega\mathbf{x}\|_0. \quad (4)$$

The role of cosparsity in the analysis model is similar to the role of sparsity in the synthesis model. The level of sparsity in the synthesis model indicates the number of non-zeros in the representation vector \mathbf{z} in (2), while in the analysis model, the cosparsity l is used to indicate the number of zeros in the vector $\Omega\mathbf{x}$, as defined in (4). In other words, the quantity l denotes the number of rows of Ω that are orthogonal to the signal.

Solving problem (3) is NP-complete, just as in the synthesis case; thus, approximation methods are required for reconstructing \mathbf{x} . As before, l_1 -relaxation [13, 14] can be used to replace the l_0 pseudo-norm-based optimization and the problem can then be solved by linear programming. With l_1 -relaxation, only a small number of measurements are required to achieve a high reconstruction rate. However, the computational complexity of this method may limit its practical use in large-scale applications. The restricted isometry property (RIP) is commonly used by both the synthesis- and analysis-based algorithms to govern the recovery condition of the sparse or cosparse signals. The measurement matrix \mathbf{M} has the Ω -RIP property with a constant δ_b , if δ_l is the smallest constant that satisfies

$$(1 - \delta_l) \|\mathbf{v}\|_2^2 \leq \|\mathbf{M}\mathbf{v}\| \leq (1 + \delta_l) \|\mathbf{v}\|_2^2 \quad (5)$$

whenever $\Omega\mathbf{v}$ has at least l zeros [12].

Another popular class of cosparse reconstruction algorithms is based on the idea of iterative greedy pursuit, such as Greedy Analysis Pursuit (GAP) [11, 15, 16]. As compared with l_1 -relaxation, GAP has better reconstruction performance and, to some degree, a lower computational complexity. Analysis Iterative Hard Thresholding (AIHT) and Analysis Hard Thresholding Pursuit (AHTP) [12, 17, 18] have been proposed by incorporating the idea of backtracking, which enables the wrong cosupports obtained in the previous iteration to be pruned in the current iteration and offers strong theoretical guarantees. Experiments show that both of them recover the signal faster than the GAP algorithm. Nevertheless, they require a relatively large number of measurements for exact reconstruction.

Recently, more sophisticated greedy algorithms have been developed, such as Analysis Subspace Pursuit (ASP) and Analysis Compressive Sampling Matching Pursuit (ACoSAMP) [12, 19]. They employ the backtracking strategy and offer strong theoretical guarantees. ASP and ACoSaMP with a candidate set size of $2l - p$ have good performance on reconstructing the signal when l is close to d , but they require more measurements for an exact reconstruction with an increasing level of cosparsity. ASP and ACoSaMP with a candidate set size of l provide a comparable reconstruction quality to that of the l_1 -relaxation methods with a lower reconstruction complexity. Other recent methods include a Bayesian method [20] where the model parameters are estimated by a Bayesian algorithm for the reconstruction of the signal under consideration.

Although all these greedy pursuit methods achieve signal reconstruction with a high accuracy, they require the cosparsity l to be known a priori for signal recovery. However, l may not be available in many practical applications. For example, most natural image signals are only cosparse when represented by an analysis operator such as a two-dimensional Fourier transform. It is difficult to define a cosparsity that exactly matches the signal under consideration. The inaccurate cosparsity may degrade the performance of the signal reconstruction algorithm, as demonstrated in the next section.

In this paper, a new greedy algorithm named Cosparsity-based Stagewise Matching Pursuit (CSMP) is proposed for the case where l is unknown. By analyzing the projection from the signal under consideration to the analysis operator, CSMP estimates the cosparsity with a pre-set step size stage by stage without the cosparsity knowledge. Cosupport and measurement residual are estimated alternately in the forward stage and fine-tuned in the backward stage, and the signal approximation is obtained at the end of the procedure. Our experiments show that the proposed algorithm has a reconstruction performance comparable to ASP and ACoSaMP, but without the knowledge of the cosparsity.

This paper is organized as follows. In Section 2, we present the motivation of this work. The CSMP algorithm is detailed in Section 3 and a theoretical analysis to the algorithm is also provided. The simulation results are given in Section 4, followed by concluding remarks and future work in Section 5.

2 Motivation

In analysis cosparse representation, it is important to establish the cosupport accurately for an exact signal reconstruction. A common approach is to adopt the ‘‘correlation’’ term defined as follows:

$$\alpha_i = |\Omega_i \mathbf{M}^T \mathbf{y}|, \quad i \in \{1, 2, \dots, p\} \quad (6)$$

where $M^T y$ resembles the original signal, Ω_i denotes the i -th row of Ω , the superscript T denotes matrix transpose, and $|\cdot|$ takes the absolute value of its argument. Obviously, a growing α_i implies the increase of the correlation between the signal and Ω_i . When the signal is orthogonal to Ω_i and $M = I$, we get $\alpha_i = 0$, where I is an identity matrix. The existing recovery methods based on the analysis model can be categorized into bottom-up and top-down methods, respectively. The bottom-up methods, such as GAP, prune one or more rows of the analysis operator which correspond to the entries of the largest correlation in each iteration. It could lead to unreliable reconstruction if the cosupports have been removed incorrectly. In contrast, the top-down methods, such as ACoSaMP and ASP [12], employ the backtracking technique to establish the cosupports. Although such cosupport refinement technique improves the performance of ACoSaMP and ASP significantly, it requires the cosparsity l as a priori for an exact recovery of the target signal. In practical applications, such information is often unavailable and must be pre-set in advance. If l is set to an inappropriate value, the performance of the algorithm for signal recovery in terms of both accuracy and robustness could be degraded significantly.

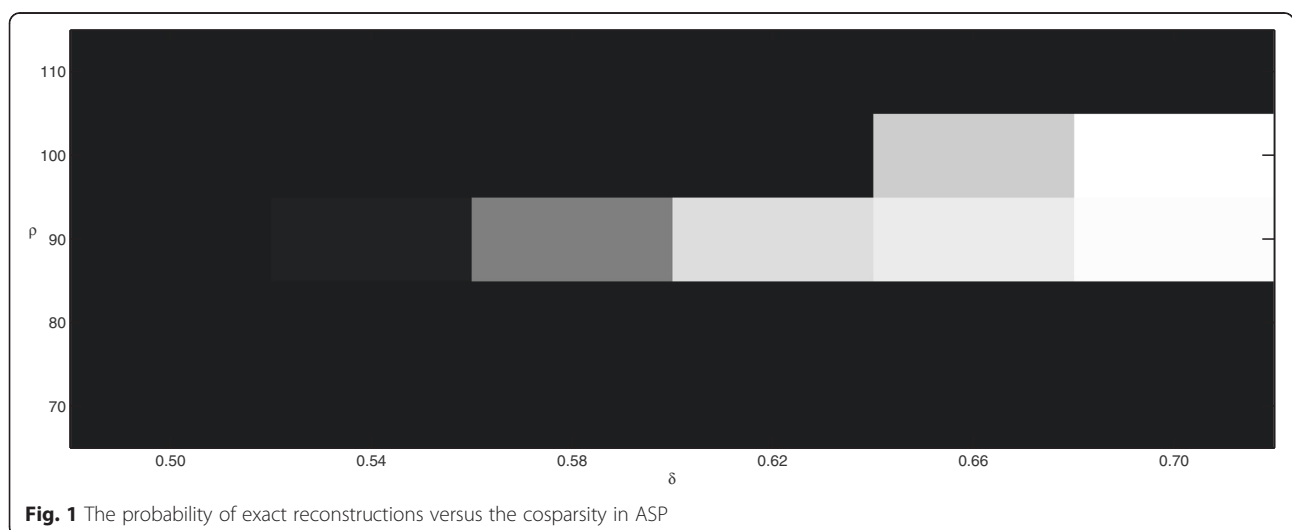
To see this, we first perform an experiment for the ASP algorithm, when l is not given accurately. Here, the analysis operator is a two-dimensional finite difference operator $\Omega \in \mathbb{R}^{p \times d}$, where $p = 144$ and $d = 120$; x is a Gaussian random signal of length $d = 120$ and a cosparsity of $l = 90$.¹ In this experiment, the sampling rate δ , which is defined as $\delta = m/d$ [12], is chosen from the set $\{0.50, 0.54, 0.58, 0.62, 0.66, 0.70\}$. The cosparsity l_{est} which denotes the estimated cosparsity of the signal, is chosen from the set $\{110, 100, 90, 80, 70\}$. We draw a phase transition diagram [12] for this algorithm.

The vertical and horizontal axes of the diagram are l_{est} and δ , respectively. For each pair of l_{est} and δ , we repeat the experiment 50 times. In each experiment, we check whether we have an exact reconstruction in the sense that the energy of the difference between the reconstruction and the original signal is smaller than 10^{-6} [12]. Note that the choice of this threshold value is followed from the existing algorithms in the literature, such as GAP, ASP, ACoSaMP, AIHT, and AHTP, for fair comparisons to be demonstrated in our experiments later. White cells in the diagram denote a completely exact reconstruction, and black cells denote a total failure in the exact reconstruction. Figure 1 shows the probability of exact reconstruction changes with respect to the sampling rate δ and the cosparsity l_{est} . We can see that the performance of ASP drops significantly if an inaccurate cosparsity l is used.

3 Cosparsity-based Stagewise Matching Pursuit

3.1 Algorithm description

To address the above issue, we propose a novel greedy algorithm for *blind* sparse reconstruction, where the cosupports are refined iteratively and the information on cosparsity is extracted automatically. The proposed CSMP algorithm, as shown in Algorithm 1, is composed of two processes, namely, the forward and backward processes. The forward process estimates the cosparsity, constructs the cosupport starting from a cosupport with all rows in Ω , and updates the measurement residual simultaneously. The procedure ends with a backward process that tries to add the rows in Ω with a smaller correlation until the terminating condition is reached. The terminating condition for CSMP is controlled by a threshold, which ensures that the estimated cosparsity is fairly close to the actual one and the target signal has been well reconstructed. The main steps of CSMP are summarized in Algorithm 1.



Algorithm 1. Cosparsity-based Stagewise Matching Pursuit

| | |
|---|--|
| Input: M, Ω, y , where $y = Mx + e$ and $\ e\ _2 = \varepsilon$, step size s , step size q ; | |
| Output: \hat{x} : cosparse approximation of x . | |

| | |
|--|-------------------------|
| Initialization: $\hat{x} = 0$, the measurement residual $y_r = y$, the cosupport $\Lambda^0 = \{1, 2, \dots, p\}$, the initial estimated cosparsity $l_{est} = d - s$, the index number $num_b = 1$ for the backward stage, the iteration number $k = 1$. | |
| Repeat until the termination condition becomes true | {Forward process} |
| 1. $\alpha_i = \Omega_i M^T y_r , i \in \{1, 2, \dots, p\}$ | “Correlation” |
| 2. $\Gamma = \text{Min}(\alpha, 2l_{est} - p)$ | Construct candidate set |
| 3. $\tilde{\Lambda}^k = \Lambda^{k-1} \cap \Gamma$ | Temporary cosupport |
| 4. $x_{temp} = \arg \min_x \ y - Mx\ _2^2$ s.t. $\Omega_{\tilde{\Lambda}^k} x = 0$ | Temporary approximation |
| 5. $\Lambda^k = \text{cosupp}(x_{temp}, l_{est})$ | Update cosupport |
| 6. $\hat{x}_{temp} = \arg \min_x \ y - Mx\ _2^2$ s.t. $\Omega_{\Lambda^k} x = 0$ | Update approximation |
| 7. $y_{temp} = y - M\hat{x}_{temp}$ | Update residual |
| 8. $\Lambda^\Delta = \text{index}(\Omega \hat{x}_{temp}, q)$ | |
| 9. if $\ y_{temp}\ _2 \geq \ y_r^{k-1}\ _2$, then | Stage switching |
| $l_{est} = l_{est} - s$; | |
| $k = k + 1$; | |
| else | |
| $y_r = y_{temp}, \hat{x} = \hat{x}_{temp}$; | |
| $k = k + 1$; | |
| end | |
| End repeat | {End forward process} |
| Repeat until $num_b \geq \lceil s/q \rceil$ | {Backward process} |
| 10. $\tilde{\Lambda}^{k+1} = \Lambda^k \cup \Lambda^\Delta$ | Temporary cosupport |
| 11. $l_{est} = l_{est} + q$ | |
| 12. $x_{temp} = \arg \min_x \ y - Mx\ _2^2$ s.t. $\Omega_{\tilde{\Lambda}^{k+1}} x = 0$ | Temporary approximation |
| 13. $\Lambda^{k+1} = \text{cosupp}(x_{temp}, l_{est})$ | Update cosupport |
| 14. $\hat{x}_{temp} = \arg \min_x \ y - Mx\ _2^2$ s.t. $\Omega_{\Lambda^{k+1}} x = 0$ | Update approximation |
| 15. $\Lambda^\Delta = \text{index}(\Omega \hat{x}_{temp}, q)$ | |
| 16. stop if the termination condition is true; | |
| 17. $num_b = num_b + 1$; | |
| $k = k + 1$, | |
| $y_r = y_{temp}, \hat{x} = \hat{x}_{temp}$; | |
| End repeat | {End backward process} |
| Return \hat{x} | |

In Algorithm 1, $\text{cosupp}(x) = \{i: \Omega_i x = 0\}$, $\text{cosupp}(x, l_{est})$ returns the index set of l_{est} smallest (in absolute value) elements in Ωx , $2l_{est} - p$ is the size of $|\Gamma|$ and for a vector α , function $\text{Min}(\alpha, 2l_{est} - p)$ returns the $2l_{est} - p$ indices corresponding to the $2l_{est} - p$ smallest values of α , $\text{index}(\Omega x, q)$ returns q elements from the $(l_{est} + 1)$ -th to $(l_{est} + q)$ -th smallest (in absolute value) rows in Ωx , and $\Gamma, \Lambda^\Delta, \tilde{\Lambda}^k$ and Λ^k are subsets of $\{1, 2, \dots, p\}$, and $\lceil s/q \rceil$ returns the smallest integer which is larger than s/q .

The CSMP algorithm adopts a stagewise approach [6] to estimate the real cosparsity in each stage in the forward process, which only requires the step size s to be set in initialization. Here, l is defined as the real cosparsity of the original signal, and s should not be larger than $d - l$ normally [11]. The initial cosupport Λ^0 has the maximum number of rows in Ω that enables CSMP to recover the signal. To avoid under-estimation of the cosparsity, a safe choice is $s = 1$ if l is unknown.

Nevertheless, there is a tradeoff between s and the recovery speed as a smaller s requires more stages. We can see that GAP and ASP/ACoSaMP are the special cases of the proposed algorithm when CSMP has a step size of $s = 1$ and $s = d - l$, respectively.

Suppose CSMP has a cosparsity of l_{est} in the forward process. With this cosparsity, the candidate set is constructed by selecting some rows in Ω with smaller correlations. Here, we should explain the relation between greedy synthesis algorithms and their analysis counterparts. Given two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{R}^n$ such that $\Lambda_1 = \text{cosupp}(\Omega \mathbf{v}_1)$ and $\Lambda_2 = \text{cosupp}(\Omega \mathbf{v}_2)$. Assuming that $\|\Lambda_1\|_0 \leq (l_{\text{est}})_1$ and $\|\Lambda_2\|_0 \leq (l_{\text{est}})_2$, it holds that $\|\Lambda_1 \cap \Lambda_2\|_0 \geq (l_{\text{est}})_1 + (l_{\text{est}})_2 - p$. For the case $\|\Lambda_1\|_0 = \|\Lambda_2\|_0 = l_{\text{est}}$, we have $2l_{\text{est}} - p \leq \|\Lambda_1 \cap \Lambda_2\|_0 \leq l_{\text{est}}$. So $2l_{\text{est}} - p$ is a reasonable size of the candidate set for CSMP, which corresponds to the candidate set size of $2k$ of CoSaMP in the synthesis model. Denoting $T_1 = \text{supp}(\Omega \mathbf{v}_1)$ and $T_2 = \text{supp}(\Omega \mathbf{v}_2)$, it is clear that $\text{supp}(\Omega(\mathbf{v}_1 + \mathbf{v}_2)) \subseteq T_1 \cup T_2$. Noticing that $\text{supp}(\cdot) = \text{cosupp}(\cdot)^C$, we get $\text{cosupp}(\Omega(\mathbf{v}_1 + \mathbf{v}_2)) \supseteq (T_1 \cup T_2)^C = T_1^C \cap T_2^C = \Lambda_1 \cap \Lambda_2$, where the superscript C denotes the complementary set. This implies that the union of the supports in the synthesis case is parallel to the intersection of the cosupports in the analysis case.

Now, with the candidate set, we begin to construct the cosupport and update the measurement residual. The rows in the analysis operator, which correspond to the smallest l_{est} components in the temporarily estimated signal (calculated in Step 4 in Algorithm 1), are used to form the cosupport. Then, we approximate the signal and update the measurement residual of the current iteration with this cosupport. The backtracking strategy [12, 19], which can be described as not only selecting rows that match the current residual signal better from the candidate set in each iteration but also excluding the other rows from the cosupport, provides the basis for constructing a more accurate cosupport and obtaining a smaller measurement residual. Here, $\text{index}(\Omega \hat{\mathbf{x}}_{\text{temp}}, q)$ is reserved for sharing in the backward process. An efficient mechanism is required for stage switching when each stage finishes till $l_{\text{est}} < l$. This can be performed if the current measurement residual energy no longer decreases when compared with that in the last iteration. From Step 9 in Algorithm 1, we can see that the algorithm will run for some iterations with the same cosparsity l_{est} until it reaches the stage switching condition. The proposed algorithm will perform with a cosparsity of $l_{\text{est}} - s$ in the next stage. The forward process will not stop until the measurement residual reaches a pre-set threshold such as $\|\mathbf{y}_r\|_2 / \|\mathbf{y}\|_2 \leq 10^{-6}$ for the noiseless case or $\|\mathbf{y}_r\|_2 \leq \varepsilon$ for the noisy case.

In the backward process, the algorithm tries to further increase the cosparsity by adding the less used rows into the cosupport. These q rows have been chosen in the last iteration of the forward process. The choice of the value of q can be made in terms of the value of s . As a rule of thumb, a small q is chosen when s is relatively small, and likewise, a greater q should be chosen if s is large. Typically, in our experiments, we choose $q = 1$ when $s = 10$, but we select $q = 50$ when s is in the order of thousands. With this strategy, we can get a more accurate cosupport for signal approximation. The iterations stop as soon as the measurement residual reaches the terminating threshold used in the forward process. The backward process needs to repeat less than $\lceil s/q \rceil$ times since it is enough for the cosparsity to change from l_{est} to $l_{\text{est}} + s$ in this process.

3.2 Relaxed versions for high-dimensional problems

In CSMP, the constrained minimization problem

$$\min_x \|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 \quad \text{s.t.} \quad \Omega_\Lambda \mathbf{x} = \mathbf{0} \quad (7)$$

is hard to solve for high-dimensional signals, and we propose to replace it with the minimization of the following cost function in Steps 4, 6, 12, and 14 in Algorithm 1:

$$\|\mathbf{y} - \mathbf{M}\mathbf{x}\|_2^2 + \lambda \|\Omega_\Lambda \mathbf{x}\|_2^2 \quad (8)$$

where λ is a relaxation constraint and we choose $\lambda = 0.001$ as in [12] in our experiments. This results in a relaxed version of the algorithm. We refer hereafter to this version as relaxed CSMP (RCSMP).

3.3 Theoretical performance analysis

This section describes our theoretical analysis of the behavior of CSMP for the cosparse model in both noiseless and noisy cases. Because the proposed algorithm has a similar strategy of backtracking which is used in ASP, the proofs are mainly based on the proof framework of ASP/ACoSaMP. The following theorems are formed in parallel with those in [12], except for the unknown cosparsity and the initial cosupport and measurement vectors.

To show the ability of exact and stable recovery of cosparse signals by the CSMP algorithm, we define all variables involved in the process of signal reconstruction as follows:

Definition 3.1 [12] Let $\mathbf{Q}_\Lambda = \mathbf{I} - \Omega_\Lambda^+ \Omega_\Lambda$ be the orthogonal projection onto Ω_Λ , where Ω_Λ is a sub-matrix of the analysis operator $\Omega \in \mathcal{R}^{p \times d}$ and Λ is a subset of

$\{1, 2, \dots, p\}$. A cosupport \hat{S}_l which implies a near-optimal projection is defined as

$$\hat{S}_l(\mathbf{v}) = \underset{\Lambda \in L_l}{\operatorname{argmin}} \|\mathbf{v} - \mathbf{Q}_\Lambda \mathbf{v}\|_2^2 \quad (9)$$

where $\mathbf{v} \in \mathfrak{R}^d$ and $L_l = \{\Lambda \in [1, 2, \dots, p], |\Lambda| \geq l\}$ l -cosparsely cosupports.

Definition 3.2 (Problem p) [12] Consider a measurement vector $\mathbf{y} \in \mathfrak{R}^m$ such that $\mathbf{y} = \mathbf{M}\mathbf{x} + \mathbf{e}$, where $\mathbf{x} \in \mathfrak{R}^d$ is l -cosparsely, $\mathbf{M} \in \mathfrak{R}^{m \times d}$ is a measurement matrix, and $\mathbf{e} \in \mathfrak{R}^m$ is a bounded additive noise. The largest singular value of \mathbf{M} is σ_M and its Ω -RIP constant is δ_l . The analysis operator $\Omega \in \mathfrak{R}^{p \times d}$ is given and fixed. Define C_l to be the fraction of the largest and the smallest eigenvalues (which are not zero) of the sub-matrix composed of l rows from Ω . Assume that $\hat{S}_l = \operatorname{cosupp}(\Omega \mathbf{x}, l)$. According to Definition 3.1, the cosupport of \mathbf{x} is a near-optimal projection. Our task is to recover \mathbf{x} from \mathbf{y} . The recovery result is denoted by $\hat{\mathbf{x}}$.

Now, we give the guarantees on exact recovery and stability of CSMP for recovering the cosparsely signals.

Theorem 3.3 (Convergence for cosparsely recovery) Consider the problem p , and suppose that there exists $\gamma > 0$ such that

$$(1 + C_{2l_q-p}) \left(1 - \left(\frac{C_{2l_q-p}}{(1+\gamma)^2} - (C_{2l_q-p} - 1) \sigma_M^2 \right) \right) < 1, \quad (10)$$

then there exists $\delta(C_{2l_q-p}, \sigma_M^2, \gamma) > 0$, whenever $\delta_{4l_q-3p} \leq \delta(C_{2l_q-p}, \sigma_M^2, \gamma)$, such that the k -th iteration of the algorithm satisfies

$$\begin{aligned} \|\mathbf{x} - \hat{\mathbf{x}}^k\|_2 &\leq \frac{1 + \delta_{2l_q-p}}{1 - \delta_{2l_q-p}} \rho_1 \rho_2 \|\mathbf{x} - \hat{\mathbf{x}}^{k-1}\|_2 \\ &+ \left(\frac{1 + \delta_{2l_q-p}}{1 - \delta_{2l_q-p}} (\eta_1 + \rho_1 \eta_2) + \frac{2}{1 - \delta_{2l_q-p}} \right) \|\mathbf{e}\|_2, \end{aligned} \quad (11)$$

where

$$\begin{aligned} \eta_1 &\triangleq \frac{\sqrt{\frac{2 + C_{l_q}}{1 + C_{l_q}} + 2\sqrt{C_{l_q}} + C_{l_q}} \sqrt{1 + \delta_{3l_q-2p}}}{1 - \delta_{4l_q-3p}}, \\ \eta_2 &\triangleq \left(\frac{1 + \delta_{3l_q-2p}}{\gamma(1 + \alpha)} + \frac{(1 + \delta_{2l_q-p}) C_{2l_q-p}}{\gamma(1 + \alpha)(1 + \gamma)} + \frac{(C_{2l_q-p} - 1)(1 + \gamma) \sigma_M^2}{(1 + \alpha)(1 + \gamma)\gamma} \right), \\ \rho_1 &\triangleq \frac{1 + 2\delta_{4l_q-3p} \sqrt{C_{l_q}} + C_{l_q}}{1 - \delta_{4l_q-3p}^2}, \\ \rho_2 &\triangleq 1 - \left(\sqrt{\delta_{4l_q-3p}} - \sqrt{\frac{C_{2l_q-p}}{(1 + \gamma)^2} \left(1 - \sqrt{\delta_{2l_q-p}} \right)^2 - (C_{2l_q-p} - 1)(1 + \delta_{2l_q-p}) \sigma_M^2} \right)^2 \end{aligned}$$

and

$$\alpha = \frac{\sqrt{\delta_{4l_q-3p}}}{\sqrt{\frac{C_{2l_q-p}}{(1 + \gamma)^2} \left(1 - \sqrt{\delta_{2l_q-p}} \right)^2 - (C_{2l_q-p} - 1)(1 + \delta_{2l_q-p}) \sigma_M^2} - \sqrt{\delta_{4l_q-3p}}} \quad (12)$$

Moreover, when $\rho_1^2 \rho_2^2 < 1$, the iteration converges. The constant γ gives a tradeoff between satisfying the theorem conditions and the noise level, and the conditions for the noiseless case are achieved when γ tends to zero.

Theorem 3.4 (Exact recovery for cosparsely signals) Consider the problem p when $\|\mathbf{e}\|_2 = 0$. Let $l_s = d - s \lceil (d - l)/s \rceil$ and $l_q = l_s + q \lfloor (l - l_s)/q \rfloor$. If the measurement matrix \mathbf{M} satisfies the Ω -RIP with parameter $\delta_{4l_q-3p} \leq \delta(C_{2l_q-p}, \sigma_M^2, \gamma)$, where C_{2l_q-p} and γ are as in Theorem 3.3 and $\delta(C_{2l_q-p}, \sigma_M^2, \gamma)$ is a constant guaranteed to be greater than zero whenever (10) is satisfied, the CSMP algorithm guarantees an exact recovery of \mathbf{x} from \mathbf{y} via a finite number of iterations.

The proof is mainly based on the following lemma:

Lemma 3.5 If the measurement matrix \mathbf{M} satisfies the Ω -RIP with the same conditions as in Theorem 3.4, then

- The $\lceil (d - l)/s \rceil + \lfloor (l - l_s)/q \rfloor$ -th stage of the algorithm is equivalent to the ASP algorithm with estimated cosparsity l_q except that they have different initial cosupports and initial measurement vectors.
- CSMP recovers the target signal exactly after completing the $\lceil (d - l)/s \rceil + \lfloor (l - l_s)/q \rfloor$ -th stage.

In the Appendix, Lemma 3.5 is proved in detail.

Lemma 3.5 describes that CSMP has a process of signal reconstruction that is equivalent to ASP, and could complete the exact recovery of the cosparsely signals in a finite number of stages. To complete the proof, it is sufficient to show that the CSMP algorithm never gets stuck at any iteration of either stage, i.e., it takes a finite number of iterations up to $\lceil (d - l)/s \rceil + \lfloor (l - l_s)/q \rfloor$ stages.

At each stage, the cosupport (whose size is assumed to be l_{est}) adds and discards some rows of Ω , and the number of rows is fixed and finite. Hence, there are a finite number of combinations, at most, $\binom{d}{l_{\text{est}}}$, where d is the length of the signal. Thus, if CSMP takes an infinite number of iterations in this stage, the construction of cosupport would be repeated after at most $\binom{d}{l_{\text{est}}}$ iterations. Hence, Theorem 3.6 follows.

Lemma 3.6 (Stability for cosparsity recovery) *Consider the problem p . If (10) holds and $\delta_{4l_q-3p} \leq \delta(C_{2l_q-p}, \sigma_M^2, \gamma)$, where γ is as in Theorem 3.3 and $\delta(C_{2l_q-p}, \sigma_M^2, \gamma)$ is a constant guaranteed to be greater than zero whenever $\frac{(C_{2l_q-p}^2-1)\sigma_M^2}{C_{2l_q-p}} < 1$ is satisfied, then for any*

$$k \geq \left| \frac{\log(\|\mathbf{x}\|_2 / \|\mathbf{e}\|_2)}{\log\left(1 / \frac{1+\delta_{2l_q-p}}{1-\delta_{2l_q-p}} \rho_1 \rho_2\right)} \right|,$$

we have

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq \left(1 + \frac{1 - \left(\frac{1+\delta_{2l_q-p}}{1-\delta_{2l_q-p}} \rho_1 \rho_2\right)^k}{1 - \frac{1+\delta_{2l_q-p}}{1-\delta_{2l_q-p}} \rho_1 \rho_2} \cdot \left(\frac{1+\delta_{2l_q-p}}{1-\delta_{2l_q-p}} (\eta_1 + \rho_1 \eta_2) + \frac{2}{1-\delta_{2l_q-p}} \right) \right) \|\mathbf{e}\|_2 \quad (13)$$

implying that CSMP leads to a stable recovery. The constants η_1 , η_2 , ρ_1 , and ρ_2 are the same as in Theorem 3.3.

Similarly, the proof of Lemma 3.6 is based on Lemma 3.5 and the corresponding theorems of ASP algorithm in [12], and we omit the detailed proof here.

The above theorems are sufficient conditions of CSMP for exact recovery and stability. They are slightly more restrictive than the corresponding results of ASP algorithms because the true cosparsity level l is always larger than or equal to the estimated one l_q . This may be regarded as an additional cost for not having precise information of cosparsity. On the other hand, the proofs also show that these sufficient conditions may not be optimal or tight enough because they only consider the final stage and ignore the influence of previous stages on the performance of the algorithm.

4 Experiments

In this section, we evaluate the performance of the proposed algorithm, as compared with several baseline algorithms. To this end, we repeat some of the experiments

performed in [12] for the noiseless case ($e = 0$) and noisy case.

4.1 Phase transition diagrams for synthetic signals in the noiseless case

We show the performance of the proposed algorithm as compared with six baseline methods, namely, AIHT, AHTP, ASP, ACoSaMP, l_1 -relaxation, and GAP, using the same experiments as performed in [12] for the noiseless case. We begin with synthetic signals and test the performance of CSMP with $s = 1$, $s = 5$, and $s = 10$, respectively. The results of the proposed algorithm are compared with those of AIHT and AHTP with an adaptively changing step size, ASP and ACoSaMP with $a = 1$ and $a = \frac{2-p}{l}$, l_1 -relaxation, and GAP. We use a random matrix \mathbf{M} , where each entry is drawn independently from a Gaussian distribution, and a random tight frame Ω of size $d = 120$ and $p = 144$.

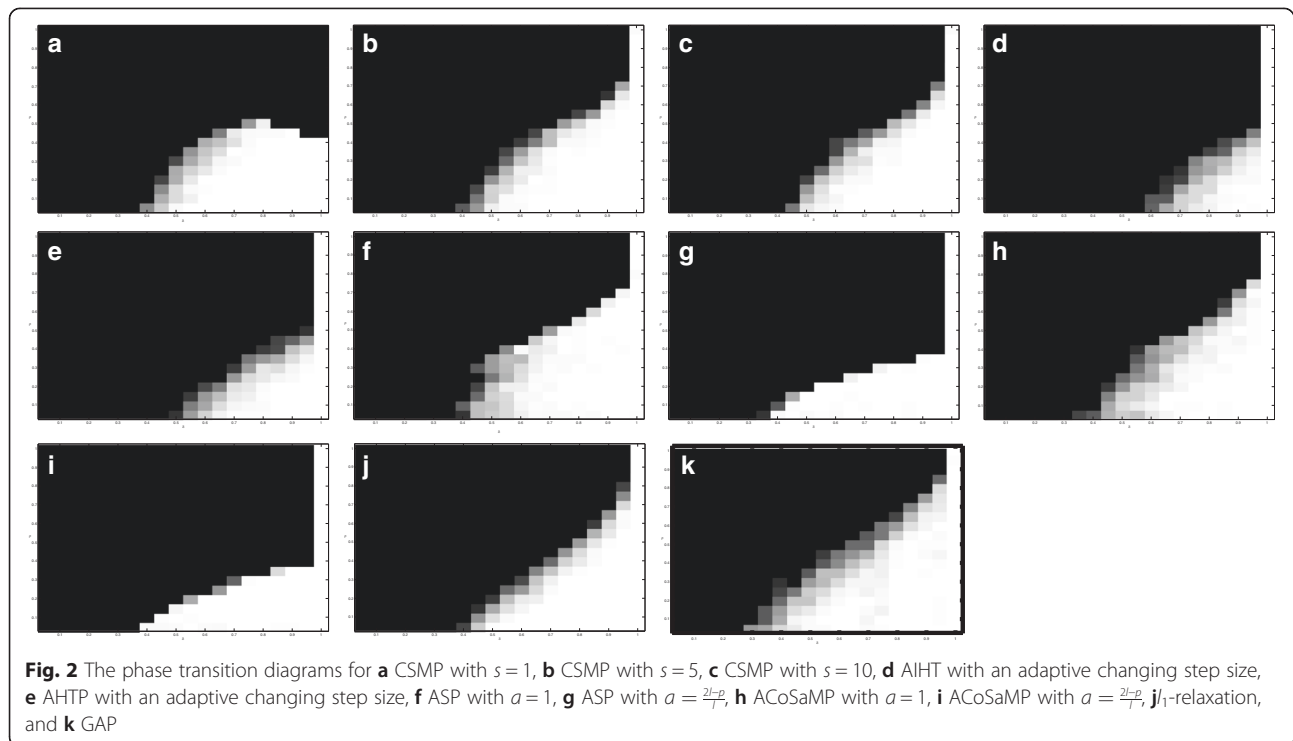
We draw a phase transition diagram [12] for each of the algorithms. In each phase transition diagram, 20 different possible values of m and 20 different values of l are tested. For each pair of m and l_{est} , we repeat the experiment for 50 times, with the value of l_{est} selected according to the formula in [12]:

$$l = d - \rho m \quad (14)$$

where ρ is the ratio between the cosparsity of the signal and the number of measurements, shown as the vertical axis of the phase diagram. The sampling rate δ is defined as $\delta = m / d$ and shown as the horizontal axis. There are 400 cells in each phase transition diagram, and the gray level of the cell shows the exact reconstruction rate of its recovery algorithm. White cells in the diagram denote a completely exact reconstruction, and black cells denote a total failure in the exact reconstruction.

The reconstruction results of the proposed algorithm and the baseline algorithms are shown in Fig. 2.

In Fig. 2, experiments with $s = 1$, $s = 5$, and $s = 10$ are performed for CSMP, respectively. From Fig. 2, it can be seen that CSMP has better results than those of AIHT and AHTP with an adaptively changing step size, and ASP and ACoSaMP with $a = \frac{2-p}{l}$ when l_{est} is far from d . In addition, the proposed algorithm with $s = 5$ and $s = 10$ provides comparable performance to ASP and ACoSaMP with $a = 1$ and the l_1 -relaxation when the knowledge of cosparsity is unknown. Although the accurate recovery rates of GAP for experiments of all pairs of δ and ρ are higher than those of CSMP, the number of white cells is 59 in Fig. 2k, which is less than 67 in Fig. 2b and 63 in Fig. 2c.



4.2 Reconstruction of high-dimensional images in the noiseless and noisy cases

We now test the methods for high-dimensional signals. We use RCSMP (relaxed versions of CSMP defined in Section 3.2) for the reconstruction of the *Shepp-Logan phantom* from a few number of measurements. RCSMP is computationally more efficient than CSMP for high-dimensional signals and is thus chosen in this experiment. The sampling operator M is a two-dimensional Fourier transform that measures only a certain number of radial lines from the Fourier transform. The cosparsity operator is a two-dimensional finite difference operator Ω_{2D-DIP} of which the number of rows is $p = 130,560$ and the real cosparsity of the signal under this operator is $l = 128,014$. The original phantom image is presented in Fig. 3a. Using the RCSMP with $s = 4000$, $q = 50$, we get perfect reconstructions using only 12 radial lines just as RASP (relaxed Analysis Subspace Pursuit), i.e., only $m = 3032$ measurements out of $d = 65,536$ which is less than 4.63 % of the data in the original image. The algorithm requires less than 15 iterations which are less than those required by RASP for achieving this recovery percentage. The reconstruction results of CSMP are shown in Fig. 3c.

We now turn to test the method for the noisy case; we perform reconstruction using RCSMP with $s = 4000$, $q = 50$ of a noisy measurement of the phantom with 22 radial lines (in Fig. 4b) and signal-to-noise ratio (SNR) at 20 dB. Figure 4c presents the noisy image,

which is the result of applying inverse Fourier transform on the measurements. We get the reconstruction results by the proposed method with a peak SNR (PSNR) of 37.11 dB in Fig. 4e. The recovery image using GAP is shown in Fig. 4f, with a little worse PSNR of 34.34 dB. Note that for the minimization process, we solve conjugate gradients in Steps 4, 6, 12, and 14 in each iteration in Algorithm 1, take only the real part of the result, and crop the values of the resulting image to be in the range of $[0, 1]$ [12].

5 Conclusions

We have presented a novel greedy pursuit algorithm CSMP for the cosparsity analysis model. With the proposed algorithm, the information of the cosparsity of the target signal is not required as a priori. It addresses a common limitation associated with the existing greedy pursuit algorithms. The underlying intuition of CSMP is to obtain the cosparsity estimation and the signal approximation in the forward process and refine them in the backward process. Borrowing the idea from ASP, a theoretical study of the proposed algorithm has been performed to give guarantees for stable recovery under the assumption of the Ω -RIP and the existence of an optimal or a near-optimal projection. Experiments have confirmed that the proposed algorithm gives competitive results for signal recovery as compared with those of l_1 -relaxation and ACoSaMP/ASP in the noiseless case and better results than GAP in the noisy case.

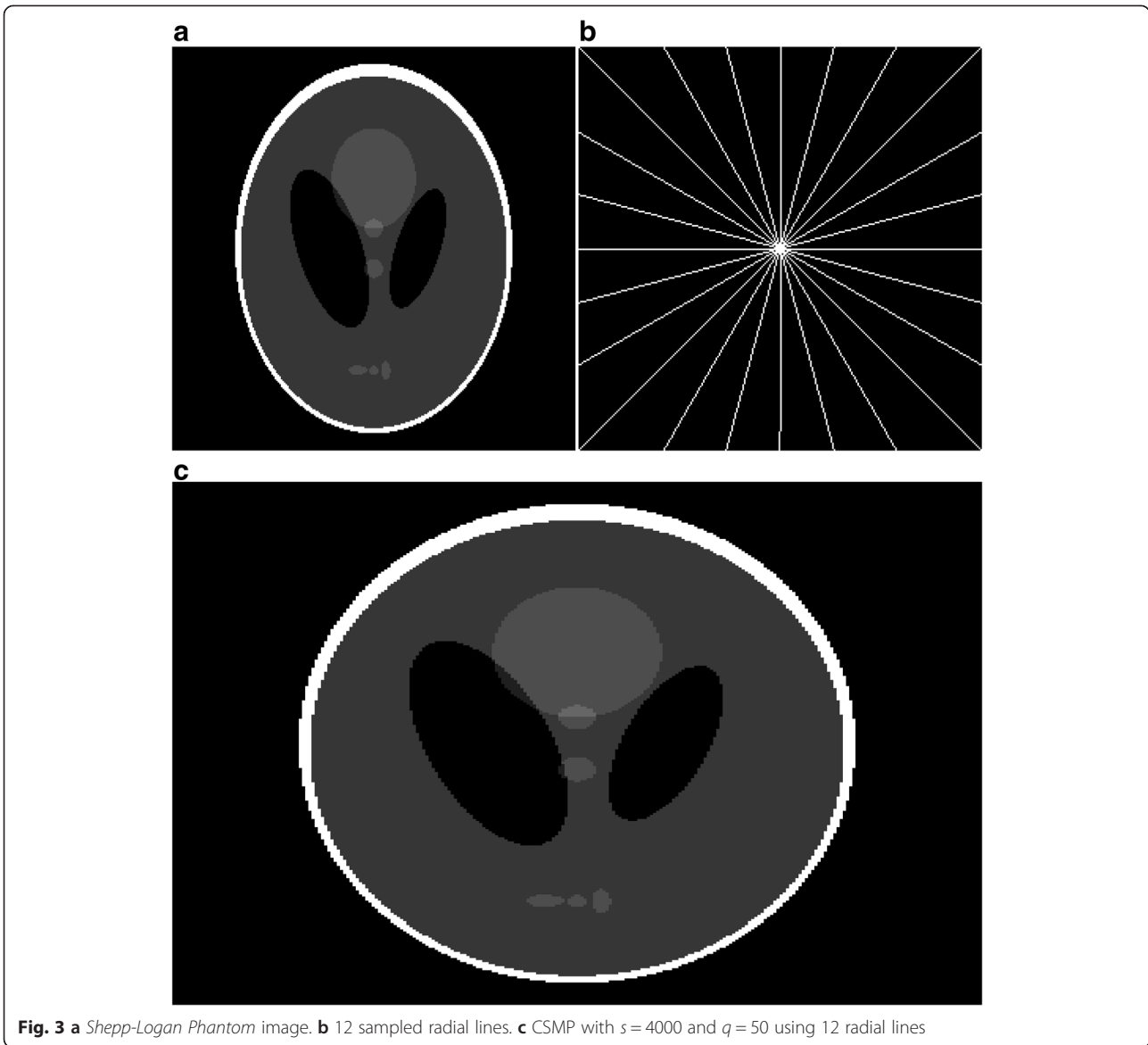


Fig. 3 a Shepp-Logan Phantom image. b 12 sampled radial lines. c CSMP with $s = 4000$ and $q = 50$ using 12 radial lines

6 Endnotes

¹To form a Gaussian random signal \mathbf{x} of length of $d = 120$ and a cosparsity of $l = 90$, we choose l rows from Ω randomly to form $\Omega_\Lambda \in \mathcal{R}^{l \times d}$, where Λ is composed of the indices of the chosen rows. We apply singular value decomposition (SVD) to $\Omega_\Lambda = U \cdot D \cdot V^{-1}$,

where $D \in \mathcal{R}^{l \times d}$ and $D(:, l + 1 : d) = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} = \mathbf{0}$.

Define Nullspace = $V(:, l + 1 : d)$, and let $\mathbf{r} \in \mathcal{R}^{d-l-1}$ be an i.i.d. Gaussian random variable of identical variance. Further define $\mathbf{x} = \text{Nullspace} \cdot \mathbf{r}$. We have

$$\Omega_\Lambda \cdot \mathbf{x} = U \cdot D \cdot V^{-1} \cdot V(:, l + 1 : d) \cdot \mathbf{r} = U \cdot D \cdot \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \cdot \mathbf{r} \quad ,$$

where $\mathbf{I} \in \mathcal{R}^{(d-l-1) \times (d-l-1)}$ is a unit matrix, so we have

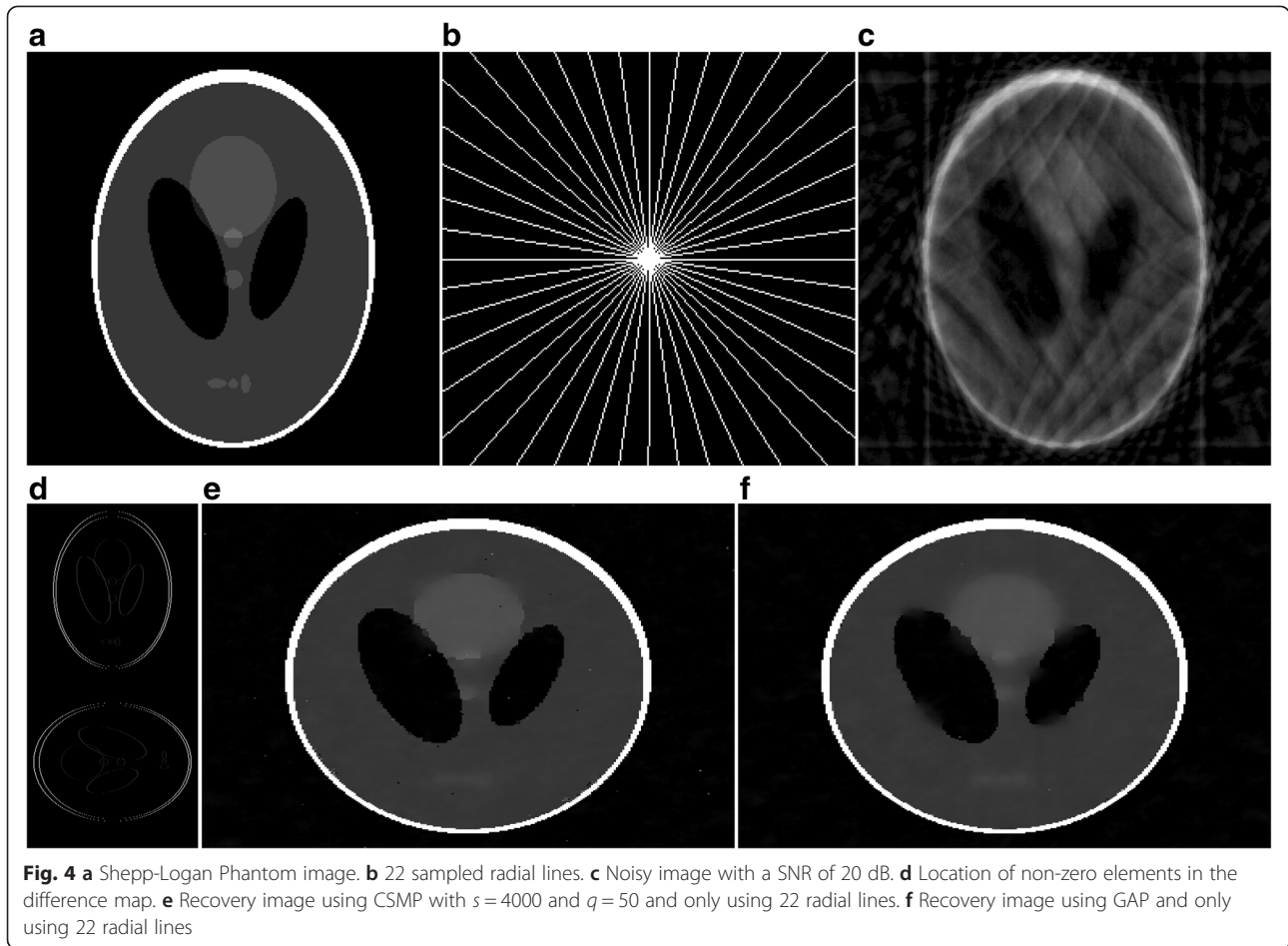
$\Omega_\Lambda \cdot \mathbf{x} = U \cdot D(:, l + 1 : d) \cdot \mathbf{r} = U \cdot \mathbf{0} \cdot \mathbf{r} = \mathbf{0}$ and $\|\Omega \cdot \mathbf{x}\|_0 \leq p - l$. Then, we have generated a signal \mathbf{x} of length of $d = 120$ and a cosparsity of $l = 90$.

7 Appendix

7.1 Proof of Lemma 3.5

Proof of Lemma 3.5 Assuming that $l_s = d - s \lceil (d-l)/s \rceil$ and $l_q = l_s + q \lfloor (l-l_s)/q \rfloor$, we first discuss the relation between l , s , and q , which includes the following two cases.

Case 1 When $(d-l)/s$ is an integer, we have $\lfloor (l-l_s)/q \rfloor = 0$ and $l_q = l_s = l$. It means that after $(d-l)/s$ stages, an estimated cosparsity l_s can be obtained by the proposed algorithm which is equal to the real cosparsity l in the forward process; the backward process will not be performed in this situation.



Case 2 If $\lceil (d-l)/s \rceil > (d-l)/s$, we have $l_s < l$. After $\lceil (d-l)/s \rceil$ stages, an estimated cosparsity l_s can be obtained by the proposed algorithm which is smaller than the real cosparsity l in the forward process. This is followed by the backward process. If $(l-l_s)/q$ is an integer, we have $l_q = l$. At the $(l-l_s)/q$ stage of the backward process, the cosparsity l_q increases to l and the CSMP algorithm will be performed as ASP with cosparsity with l_q . If $\lfloor (l-l_s)/q \rfloor < (l-l_s)/q$, we have $l_s < l_q$.

For both cases, we summarize that CSMP will recover a signal exactly with an estimated cosparsity of l_q ($l_s \leq l_q$) after $\lceil (d-l)/s \rceil + \lfloor (l-l_s)/q \rfloor$ stages.

The analysis operator has a property that for $\tilde{l} \leq l$, it holds that $\delta_{\tilde{l}} \leq \delta_{\tilde{l}}$ [12]. Via (4) and this property, the inequality $(1 - \delta_{4l_s - 3p}) \|\mathbf{v}\|_2^2 \leq (1 - \delta_{4l_q - 3p}) \|\mathbf{v}\|_2^2 \leq \|\mathbf{M}\mathbf{v}\|_2^2 \leq (1 + \delta_{4l_q - 3p}) \|\mathbf{v}\|_2^2 \leq (1 + \delta_{4l_s - 3p}) \|\mathbf{v}\|_2^2$ holds; hence, we choose the parameter $\delta_{4l_q - 3p}$ to satisfy the Ω -RIP instead of $\delta_{4l_s - 3p}$.

At the last stage, we know that the construction of the candidate set and cosupport of CSMP is similar to that of ASP with the corresponding value of

cosparsity. The only difference is that while the ASP algorithm has full initial cosupport and measurement data, CSMP has the cosupport and measurement residual of the last iteration as its initialization. This is the first part of Lemma 3.5.

The second part of Lemma 3.5 is derived from the fact that the convergence condition of the ASP algorithm in [12] does not depend on those initial values but the construction of the candidate set and cosupport. In particular, it is based on the following observations:

- The energy of the part of the signal \mathbf{x} not captured by the current cosupport is a constant of times smaller than that of the signal \mathbf{x} not captured by the cosupport in the previous iteration.
- The energy of the measurement residual of the current iteration is a constant of times smaller than that of the previous iteration.

When the condition of Ω -RIP is satisfied, both the above constants are smaller than the one that results in the exact recovery after a finite number of iterations. This is the main result of Theorem 6.13 in [12]. Because

the final stage in the proposed algorithm is equivalent to ASP with estimated sparsity level l_q , it is obvious that the target signal will be exactly recovered after this stage if the condition on Ω -RIP of parameter l_q is satisfied.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The work presented in this paper was carried out in collaboration between all authors. YZ and YH conceived the research theme. YZ and DW designed the algorithm, analyzed the algorithm, performed the experiments, and discussed the results. The theorem proofs were given by DW. WW contributed to the revision of the overall work and manuscript and provided insightful comments and suggestions.

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant 51379049 and the Engineering and Physical Sciences Research Council (EPSRC) of the UK under Grant EP/K014307/1. The authors wish to thank the Associate Editor and the anonymous reviewers for their contributions to improving this paper.

Author details

¹College of Automation, Harbin Engineering University, Harbin 150001, China.

²Department of Electronic Engineering, University of Surrey, Guildford GU2 7XH, UK.

Received: 9 June 2015 Accepted: 6 November 2015

Published online: 01 December 2015

References

1. M Elad, P Milanfar, R Rubinstein, Analysis versus synthesis in signal priors. *Inverse Probl.* **23**(3), 947–968 (2007)
2. M Bruckstein, DL Donoho, M Elad, From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.* **51**(1), 34–81 (2009)
3. SS Chen, DL Donoho, MA Saunders, Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.* **20**(1), 33–61 (1998)
4. Y Pati, R Rezaifar, P Krishnaprasad, Orthonormal matching pursuit: recursive function approximation with applications to wavelet decomposition, in *Proc. of the 27th Annual Asilomar Conf. on Signals, Systems and Computers*, 1993, pp. 40–44
5. D Needell, J Tropp, CoSaMP: iterative signal recovery from incomplete and inaccurate samples. *Appl. Comput. Harmon. Anal.* **26**(3), 301–320 (2009)
6. TT Do, L Gan, N Nguyen, TD Tran, Sparsity adaptive matching pursuit algorithm for practical compressed sensing, in *Proc. of the 42nd Asilomar Conf. on Signals, Systems and Computers, IEEE*, 2008, pp. 581–587
7. T Blumensath, M Davies, Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* **27**(3), 265–274 (2009)
8. S Foucart, Hard thresholding pursuit: an algorithm for compressive sensing. *SIAM J. Numer. Anal.* **49**(6), 2543–2563 (2011)
9. H Wu, S Wang, Adaptive sparsity matching pursuit algorithm for sparse reconstruction. *IEEE Signal Proc. Lett.* **19**(8), 471–474 (2012)
10. D Wu, K Wang, Y Zhao, W Wang, L Chen, Stagewise regularized orthogonal matching pursuit algorithm. *Opt. Precis. Eng.* **22**(5), 1395–1402 (2014)
11. S Nam, M Davies, M Elad, R Gribonval, The cosparsity analysis model and algorithms. *Appl. Comput. Harmon. Anal.* **34**(1), 30–56 (2013)
12. R Giryes, S Nam, M Elad, R Gribonval, ME Davies, Greedy-like algorithms for the cosparsity analysis model. *Linear Algebra. Appl.* **441**, 22–60 (2014)
13. EJ Candes, YC Eldar, D Needell, P Randall, Compressed sensing with coherent and redundant dictionaries. *Appl. Comput. Harmon. Anal.* **31**(1), 59–73 (2011)
14. S Vaiteer, G Peyré, C Dossal, J Fadili, Robust sparse analysis regularization. *IEEE Trans. Inf. Theory* **59**(4), 2001–2016 (2013)
15. S Nam, M Davies, M Elad, and R Gribonval, Cosparsity analysis modeling, Workshop on Signal Processing with Adaptive Sparse Structured Representations. 2011.
16. S Nam, M Davies, M Elad, R Gribonval, Recovery of cosparsity signals with greedy analysis pursuit in the presence of noise, in *Proc. 4th IEEE Int. Workshop on CAMSAP*, 2011, pp. 361–364.
17. R Giryes, S Nam, R Gribonval, and ME Davies, Iterative cosparsity projection algorithms for the recovery of cosparsity vectors, Signal Processing Conference, 2011 19th European. IEEE, 2011: 1460–1464.
18. R Giryes, Greedy algorithm for the analysis transform domain. arXiv Preprint, arXiv **1309**, 7298 (2013)
19. R Giryes and M Elad, CoSaMP and SP for the cosparsity analysis model, Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European. IEEE, 2012: 964–968.
20. J Turek, I Yavneh, M Elad, On MAP and MMSE estimators for the co-sparsity analysis model. *Digit. Signal. Process.* **28**, 57–74 (2014)

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com