

RESEARCH

Open Access



Preconditioned generalized orthogonal matching pursuit

Zhishen Tong^{1,2†}, Feng Wang^{3†}, Chenyu Hu^{1,2}, Jian Wang^{4,5*} and Shensheng Han^{1,2,6}

*Correspondence:

jian_wang@fudan.edu.cn

[†]Zhishen Tong and Feng Wang contributed equally to this work.

⁴School of Data Science, Fudan University, 200433 Shanghai, China

⁵ZJLab, Fudan-Xinzailing Joint Research Centre for Big Data, Shanghai Key Lab of Intelligent Information Processing, 200433 Shanghai, China

Full list of author information is available at the end of the article

Abstract

Recently, compressed sensing (CS) has aroused much attention for that sparse signals can be retrieved from a small set of linear samples. Algorithms for CS reconstruction can be roughly classified into two categories: (1) optimization-based algorithms and (2) greedy search ones. In this paper, we propose an algorithm called the preconditioned generalized orthogonal matching pursuit (Pre-gOMP) to promote the recovery performance. We provide a sufficient condition for exact recovery via the Pre-gOMP algorithm, which says that if the mutual coherence of the preconditioned sampling matrix Φ satisfies

$$\mu(\Phi) < \frac{1}{SK - S + 1},$$

then the Pre-gOMP algorithm exactly recovers any K -sparse signals from the compressed samples, where $S (> 1)$ is the number of indices selected in each iteration of Pre-gOMP. We also apply the Pre-gOMP algorithm to the application of ghost imaging. Our experimental results demonstrate that the Pre-gOMP can largely improve the imaging quality of ghost imaging, while boosting the imaging speed.

Keywords: Compressed sensing, Preconditioning, Generalized orthogonal matching pursuit, Ghost imaging, Mutual coherence

1 Introduction

Recently, compressed sensing (CS) has gained a lot of interests and promoted the applications of many fields, such as the imaging signal processing, applied mathematics, and statistics [1–5]. The main goal of CS is to estimate a high dimensional K -sparse signal vector $\mathbf{x} \in \mathcal{R}^n$ ($(\|\mathbf{x}\|_0 = K \ll n)$) from a small number of linear samples:

$$\mathbf{y}_0 = \Psi \mathbf{x}, \quad (1)$$

where $\Psi \in \mathcal{R}^{m \times n}$ is often called the sampling matrix. Although the Eq. 1 is underdetermined, owing to the sparsity prior, \mathbf{x} can be accurately recovered from its samples \mathbf{y}_0 by solving the ℓ_0 -minimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y}_0 = \Psi \mathbf{x}. \quad (2)$$

There has been much effort in solving (2), which can be roughly classified into two categories: (i) those relying on optimization and (ii) those using greedy search. The

optimization-based approaches relaxes the ℓ_0 -norm to the ℓ_1 -norm and solves the convex optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y}_0 = \Psi \mathbf{x}. \quad (3)$$

A well-known algorithm solving (3) is called basis pursuit (BP) [2] which can reliably recover sparse signals under appropriate constraints on the sampling matrix [2]. On the other hand, greedy search algorithms have received considerable attention due to their computational simplicity. Examples includes the matching pursuit (MP) [6], orthogonal matching pursuit (OMP) [7], and orthogonal least squares (OLS) [8]. To improve the computational efficiency and recovery performance, there have also been many studies on the modification of OMP. As a representative variant, generalized OMP (gOMP) [9] chooses S columns of Ψ that are maximally correlated with the residual vector at each iteration, which exhibits computational advantages over the conventional OMP algorithm.

As mentioned, the property of sampling matrix has a great influence on the recovery performance. To evaluate the property of Ψ , the mutual coherence property has been widely used [10], which is defined as

$$\mu(\Psi) = \max_{1 \leq i \neq j \leq n} \frac{|\langle \Psi_i, \Psi_j \rangle|}{\|\Psi_i\|_2 \|\Psi_j\|_2}. \quad (4)$$

Generally speaking, a smaller μ contributes to better performance on signal recovery. One way to reduce the mutual coherence is to multiply a matrix \mathbf{P} on both sides of (2), i.e.,

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{P}\mathbf{y}_0 = \mathbf{P}\Psi \mathbf{x}. \quad (5)$$

In doing so, we wish $\mu(\mathbf{P}\Psi)$ to be smaller than that of the original one. This operation is commonly referred to as preconditioning in numerical linear algebra, where the matrix \mathbf{P} is called preconditioner [11]. There has been much evidence that preconditioning is useful to promote the recovery quality of sparse signals.

In this paper, we propose a preconditioned gOMP (Pre-gOMP) algorithm for the recovery of sparse signals. As shown in Algorithm 1, the Pre-gOMP algorithm consists of (i) a preconditioning step and (ii) a conventional signal reconstruction step. The primary contributions of this paper are summarized as follows:

1. Based on the mutual coherence framework, we develop a sufficient condition for the Pre-gOMP algorithm. Specifically, we show that

$$\mu(\mathbf{P}\Psi) < \frac{1}{SK - S + 1}$$

is sufficient for Pre-gOMP to exactly recover any K -sparse vector in K iterations.

2. To evaluate the recovery performance of the Pre-gOMP algorithm. We apply it to imaging objects in the application of ghost imaging (GI). Our experimental results reveal that Pre-gOMP algorithm can largely improve the imaging quality compared to the existing methods.

The rest of this paper is organized as follows. In Section 2, we introduce the Pre-gOMP algorithm and analyze it under the mutual coherence framework. Section 3 provides simulation and the setup of GI. Section 4 presents simulated results and experimental results for the propose algorithm. We conclude our work in Section 5.

Algorithm 1 The Pre-gOMP Algorithm

Input: sampling matrix $\Psi \in \mathcal{R}^{m \times n}$, sampling vector $\mathbf{y}_0 \in \mathcal{R}^m$, sparsity level K , residual tolerant ϵ , and number of indices for each selection $S \leq \min\{K, \frac{m}{K}\}$.

Output: the support of estimated signal $\hat{T} = \arg \min_{S: |S|=K} \|\mathbf{x}^k - \mathbf{x}_S^k\|_2$, the estimated signal $\hat{\mathbf{x}}$

satisfying $\hat{\mathbf{x}}_{\Omega \setminus \hat{T}} = \mathbf{0}$ and $\hat{\mathbf{x}}_{\hat{T}} = \Phi_{\hat{T}}^\dagger \mathbf{y}$.

```

1: function PRE-GOMP( $\Psi, \mathbf{y}_0, K, \epsilon, S$ )
2:   preconditioning matrix  $\mathbf{P} \leftarrow \Psi^T (\Psi \Psi^T)^{-1}$ ,
3:   preconditioned samples  $\mathbf{y} \leftarrow \mathbf{P} \mathbf{y}_0$ .
4:   preconditioned sampling matrix  $\Psi \leftarrow \mathbf{P} \Psi$ .
5:   iteration count  $k \leftarrow 0$ ,
6:   estimated support  $\Lambda_0 \leftarrow \emptyset$ ,
7:   residual vector  $\mathbf{r}_0 \leftarrow \mathbf{y}$ .
8:   while  $\|\mathbf{r}_k\|_2 > \epsilon$  and  $k < \min\{K, \frac{m}{K}\}$  do
9:      $k = k + 1$ .
10:    (Identification) Select indices  $\{\phi(i)\}_{i=1,2,\dots,S}$ , corresponding to  $S$  largest
    entries (in magnitude) in  $\Psi^T \mathbf{r}_{k-1}$ .
11:    (Augmentation)  $\Lambda_k = \Lambda_{k-1} \cup S$ .
12:    (Estimation)  $\mathbf{x}^k = \arg \min_{\mathbf{u}: \text{supp}(\mathbf{u})=\Lambda_k} \|\mathbf{y} - \Psi \mathbf{u}\|_2$ .
13:    (Residual Update)  $\mathbf{r}_k = \mathbf{y} - \Psi \mathbf{x}^k$ .
14:  end while
15:  return the support of estimated signal  $\hat{T} = \arg \min_{S: |S|=K} \|\mathbf{x}^k - \mathbf{x}_S^k\|_2$ , the estimated
    signal  $\hat{\mathbf{x}}$  satisfying  $\hat{\mathbf{x}}_{\Omega \setminus \hat{T}} = \mathbf{0}$  and  $\hat{\mathbf{x}}_{\hat{T}} = \Psi_{\hat{T}}^\dagger \mathbf{y}$ .
16: end function

```

2 Method

2.1 Notations

Let $\Omega = \{1, \dots, n\}$. $T = \text{supp}(\mathbf{x}) = \{i | i \in \Omega, x_i \neq 0\}$ is the support set of \mathbf{x} . $S \subseteq \Omega$ is the set of selected indices in each iteration and $|S|$ is the cardinality of S . $T \setminus S = \{i | i \in T, i \notin S\}$. $\Lambda_k = \Lambda_{k-1} \cup S$ is the estimated support set at the k th iteration of Pre-gOMP. $\mathbf{x}_S \in \mathcal{R}^{|S|}$ is the subset of \mathbf{x} indexed by S . Similarly, $\Psi_S \in \mathcal{R}^{m \times |S|}$ is a submatrix of Ψ that contains columns of Ψ indexed by S . If Ψ_S has full column rank, then $\Psi_S^\dagger = (\Psi_S^T \Psi_S)^{-1} \Psi_S^T$ is the pseudoinverse of Ψ_S . $\text{span}(\Psi_S)$ is the span of columns in Ψ_S . $\mathcal{P}_S = \Psi_S \Psi_S^\dagger$ is the projection matrix onto $\text{span}(\Psi_S)$. $\mathcal{P}_S^\perp = \mathbf{I} - \mathcal{P}_S$ is the projection matrix onto the orthogonal complement of $\text{span}(\Psi_S)$ where \mathbf{I} is the identity matrix.

2.2 The pre-gOMP algorithm

As mentioned, the Pre-gOMP algorithm consists of two parts: (i) the preconditioning operation and (ii) the signal reconstruction step. The preconditioning operation aims to reduce the mutual coherence of the sampling matrix. In this paper, we adopt the operation in [12], in which the preconditioner \mathbf{P} is given in closed-form as

$$\mathbf{P} = \Psi^T (\Psi \Psi^T)^{-1}, \quad (6)$$

which has been shown to very effective in improving the mutual coherence. Interested readers are referred to [12] for a detailed description and theoretical analysis of the preconditioner. A similar treatment has also been proposed in [13].

In the signal reconstruction step, the gOMP algorithm is used, where the preconditioned samples $\mathbf{y} = \mathbf{P}\mathbf{y}_0$ and the preconditioned sampling matrix $\Psi = \mathbf{P}\Psi$ are the inputs. We would like to mention two advantages of the Pre-gOMP algorithm. Firstly, the preconditioning operation leads to a reduction of the mutual coherence, which is useful for promoting the recovery accuracy. Secondly, the signal reconstruction step can be very efficient because the gOMP algorithm essentially carries out a parallel processing to identify support indices of \mathbf{x} , as pointed out in [9]. The computational benefit is no doubt helpful in the application of GI.

The analysis of Pre-gOMP algorithm consists of two parts: (i) the reduction on the mutual coherence after preconditioning and (ii) the sufficient condition analysis for Pre-gOMP in terms of mutual coherence.

2.3 The reduction on μ after preconditioning

Lemma 1 (Preconditioning [12]) *Given a sampling matrix $\Psi \in \mathcal{R}^{m \times n}$ with $m \leq n$. The preconditioned matrix $\mathbf{P}\Psi$ with $\mathbf{P} = \Psi^T(\Psi\Psi^T)^{-1}$ is a Parseval tight frame.*

One can interpret from Lemma 1 that the preconditioned sampling matrix $\mathbf{P}\Psi$ has identical non-zero singular values. As stated in [14], the larger the smallest non-zero singular value of a matrix is, the smaller mutual coherence of the matrix is. Therefore, the preconditioner \mathbf{P} can be useful in improving the mutual coherence.

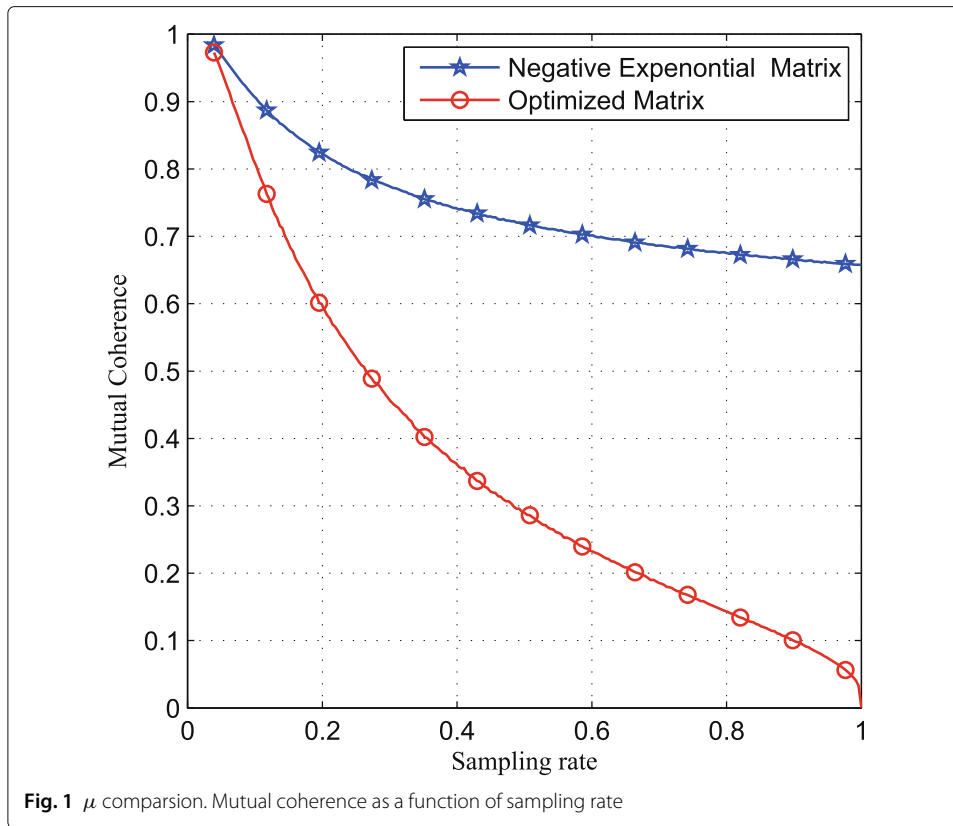
To test the effectiveness of the preconditioning method, we perform simulation and experiment. In our simulation, random negative exponential sampling matrices, which are commonly used in GI [15], is considered. The entries of random negative exponential sampling matrix Ψ are drawn independently from the negative exponential distribution $p(x) \sim \frac{1}{x} \exp(-\frac{x}{x})$. The size of the sampling matrix is $m \times n$ with fixed $n = 256$ and m ranges from 10 to 256. For each sampling number, 500 independent trials are performed, and the mean mutual coherence of the matrix is calculated. In Fig. 1, we plot the mutual coherence as a function of the sampling rate r , which is defined as $r = m/n$, where the blue pentagram line describes $\mu(\Psi)$ as a function of the sampling rate r , while the red circle line represents $\mu(\Psi)$ (denoted as optimized matrix in Fig. 1) as a function of the sampling rate r . It is observed that the μ decreases as the sampling rate r increases. In particular, the $\mu(\Psi)$ is uniformly smaller than $\mu(\Psi)$ for all region of sampling rate, which clearly validates the effectiveness of our preconditioning method.

2.4 Sufficient condition for pre-gOMP based on μ

Theorem 1 *Let $\Psi \in \mathcal{R}^{m \times n}$ be the preconditioned sampling matrix. Then, Pre-gOMP exactly recovers any K -sparse signal $\mathbf{x} \in \mathcal{R}^n$ from its preconditioned samples $\mathbf{y} = \Phi\mathbf{x}$ under*

$$\mu < \begin{cases} \frac{1}{2K-1}, & \text{when } S = 1, \\ \frac{1}{(K-1)S+1}, & \text{when } S \geq 2, \end{cases} \quad (7)$$

where $S (\geq 1)$ is the number of indices selected in each selection in Pre-gOMP algorithm.



Remark 1 When $S = 1$, the sufficient condition for Pre-gOMP algorithm is the same as that for OMP algorithm [16]. When $S \geq 2$, the bound in (7) decreases monotonically in S . Namely, the larger S is, the more restrictive the requirement on the preconditioned sampling matrix Ψ would be. Nevertheless, the large S can largely reduce the number of iterations, which is useful for improving the computational complexity of the algorithm.

2.5 Proof of Theorem 1

Before proving Theorem 1, we give some lemmas that are useful in the proof.

Lemma 2 (norm inequality [17]) For matrices $\mathbf{A}, \mathbf{B} \in \mathcal{R}^{m \times n}$, and $\mathbf{u} \in \mathcal{R}^m$, the following inequalities hold:

$$|\|\mathbf{A}\|_2 - \|\mathbf{B}\|_2| \leq \|\mathbf{A} + \mathbf{B}\|_2 \leq \|\mathbf{A}\|_2 + \|\mathbf{B}\|_2, \quad (8a)$$

$$\frac{\|\mathbf{u}\|_1}{\sqrt{m}} \leq \|\mathbf{u}\|_2 \leq \sqrt{m} \|\mathbf{u}\|_\infty. \quad (8b)$$

Lemma 3 (Lemma 6 in [12]) For two disjoint sets $I_1, I_2 \subset \{1, 2, \dots, n\}$ and Ψ_{I_1}, Ψ_{I_2} are the corresponding subsets of Ψ , then

$$\|\Psi_{I_1}^T \Psi_{I_2}\|_2 \leq \mu \sqrt{|I_1| |I_2|}, \quad (9)$$

where μ is the mutual coherence of Ψ .

Lemma 4 (Consequences of RIP [18, 19]) : Let $\mathcal{S} \subseteq \Omega$, if $\delta_{|\mathcal{S}|} \in [0, 1)$, then for any vector $\mathbf{u} \in \mathcal{R}^{|\mathcal{S}|}$,

$$(1 - \delta_{|\mathcal{S}|}) \|\mathbf{u}\|_2 \leq \|\Psi_{\mathcal{S}}^T \Psi_{\mathcal{S}} \mathbf{u}\|_2 \leq (1 + \delta_{|\mathcal{S}|}) \|\mathbf{u}\|_2 \quad (10a)$$

$$\frac{\|\mathbf{u}\|_2}{1 + \delta_{|\mathcal{S}|}} \leq \|(\Psi_{\mathcal{S}}^T \Psi_{\mathcal{S}})^{-1} \mathbf{u}\|_2 \leq \frac{\|\mathbf{u}\|_2}{1 - \delta_{|\mathcal{S}|}} \quad (10b)$$

$$\delta_{|\mathcal{S}|} \leq (|\mathcal{S}| - 1)\mu \quad (10c)$$

Now, we proceed to prove Theorem 1 via mathematical induction, which is a similar strategy as in [9, 20, 21] but with extension to the mutual coherence framework. Suppose that the Pre-gOMP algorithm has performed k iterations successfully, i.e., Λ_k contains at least k correct indices. Then, in the k th iteration, the residual is

$$\mathbf{r}_k = \mathcal{P}_{\Lambda_k}^\perp \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k}.$$

Let β_1 be the maximal absolute inner product of the residual signal \mathbf{r}_k and correct atoms $\Psi_i, i \in T$. Let $\alpha_i, i = 1, 2, \dots, S$ be the S largest absolute inner product of the residual signal \mathbf{r}_k and incorrect atoms $\Psi_i, i \in T^c$. We arrange $\alpha_i, i = 1, 2, \dots, S$, according to their magnitude in the descending order ($\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_S$). Following the strategy in [12], we build the sufficient condition by showing that

$$\beta_1 > \alpha_S, \quad (11)$$

which guarantees at least one correct atom selected at the $(k + 1)$ -th iteration of Pre-gOMP.

In the $(k + 1)$ -th ($0 \leq k \leq K - 1$) iteration, one has

$$\begin{aligned} \beta_1 &= \|\Psi_{T \setminus \Lambda_k}^T \mathbf{r}_k\|_\infty \\ &\stackrel{(8b)}{\geq} \frac{\|\Psi_{T \setminus \Lambda_k}^T \mathbf{r}_k\|_2}{\sqrt{|T \setminus \Lambda_k|}} \\ &= \frac{\|\Psi_{T \setminus \Lambda_k}^T \mathcal{P}_{\Lambda_k}^\perp \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k}\|_2}{\sqrt{|T \setminus \Lambda_k|}} \\ &\stackrel{(8a)}{\geq} \frac{\|\Psi_{T \setminus \Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k}\|_2 - \|\Psi_{T \setminus \Lambda_k}^T \mathcal{P}_{\Lambda_k} \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k}\|_2}{\sqrt{|T \setminus \Lambda_k|}}, \end{aligned} \quad (12)$$

where

$$\begin{aligned} &\left\| \Psi_{T \setminus \Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\ &\stackrel{(10a), (10c)}{\geq} (1 - (|T \setminus \Lambda_k| - 1)\mu) \|\mathbf{x}_{T \setminus \Lambda_k}\|_2, \end{aligned} \quad (13)$$

and

$$\begin{aligned}
& \left\| \Psi_{T \setminus \Lambda_k}^T \mathcal{P}_{\Lambda_k} \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&= \left\| \Psi_{T \setminus \Lambda_k}^T \Psi_{\Lambda_k} \left(\Psi_{\Lambda_k}^T \Psi_{\Lambda_k} \right)^{-1} \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(9)}{\leq} \mu \sqrt{|T \setminus \Lambda_k| |\Lambda_k|} \left\| \left(\Psi_{\Lambda_k}^T \Psi_{\Lambda_k} \right)^{-1} \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(10b), (10c)}{\leq} \frac{\mu \sqrt{|T \setminus \Lambda_k| |\Lambda_k|}}{1 - (|\Lambda_k| - 1) \mu} \left\| \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(9)}{\leq} \frac{\mu \sqrt{|T \setminus \Lambda_k| |\Lambda_k|}}{1 - (|\Lambda_k| - 1) \mu} \sqrt{|T \setminus \Lambda_k| |\Lambda_k|} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&= \frac{|T \setminus \Lambda_k| |\Lambda_k| \mu}{1 - (|\Lambda_k| - 1) \mu} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2. \tag{14}
\end{aligned}$$

Combining (13) with (14), we can get

$$\begin{aligned}
\beta_1 &\geq \frac{1}{\sqrt{|T \setminus \Lambda_k|}} \left(1 - (|T \setminus \Lambda_k| - 1) \mu - \frac{|T \setminus \Lambda_k| |\Lambda_k| \mu}{1 - (|\Lambda_k| - 1) \mu} \right) \\
&\quad \cdot \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2. \tag{15}
\end{aligned}$$

Our next job is to calculate α_S . Before doing so, we observe that

$$\begin{aligned}
& \left\| \Psi_{U_k}^T \mathbf{r}_k \right\|_2 \\
&= \left\| \Psi_{U_k}^T (\mathbf{I} - \mathcal{P}_{\Lambda_k}) \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\leq \left\| \Psi_{U_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 + \left\| \Psi_{U_k}^T \mathcal{P}_{\Lambda_k} \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2, \tag{16}
\end{aligned}$$

where $U_k := \arg \max_{s \subset \Omega \setminus (T \cup \Lambda_k), |s|=S} \left\| \phi_s^T \mathbf{r}_k \right\|_1$. The first and second term on the right-hand side of (16) can be rewritten as

$$\begin{aligned}
\left\| \Psi_{U_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 &\stackrel{(9)}{\leq} \mu \sqrt{|U_k| |T \setminus \Lambda_k|} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&= \mu \sqrt{S |T \setminus \Lambda_k|} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2, \tag{17}
\end{aligned}$$

and

$$\begin{aligned}
& \left\| \Psi_{U_k}^T \mathcal{P}_{\Lambda_k} \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&= \left\| \Psi_{U_k}^T \Psi_{\Lambda_k} \left(\Psi_{\Lambda_k}^T \Psi_{\Lambda_k} \right)^{-1} \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(9)}{\leq} \mu \sqrt{|U_k| |\Lambda_k|} \left\| \left(\Psi_{\Lambda_k}^T \Psi_{\Lambda_k} \right)^{-1} \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(10b), (10c)}{\leq} \frac{\mu \sqrt{|U_k| |\Lambda_k|}}{1 - (|\Lambda_k| - 1) \mu} \left\| \Psi_{\Lambda_k}^T \Psi_{T \setminus \Lambda_k} \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&\stackrel{(9)}{\leq} \frac{\mu \sqrt{|U_k| |\Lambda_k|}}{1 - (|\Lambda_k| - 1) \mu} \mu \sqrt{|T \setminus \Lambda_k| |\Lambda_k|} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2 \\
&= \frac{|\Lambda_k| \sqrt{|U_k| |T \setminus \Lambda_k|} \mu^2}{1 - (|\Lambda_k| - 1) \mu} \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2, \tag{18}
\end{aligned}$$

respectively. Combining (17) with (18), one gets

$$\begin{aligned} \left\| \Psi_{U_k}^T \mathbf{r}_k \right\|_2 &\leq \left(\mu \sqrt{S |T \setminus \Lambda_k|} + \frac{|\Lambda_k| \sqrt{S |T \setminus \Lambda_k|} \mu^2}{1 - (|\Lambda_k| - 1) \mu} \right) \\ &\quad \cdot \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2. \end{aligned} \quad (19)$$

On the other hand,

$$\begin{aligned} \left\| \Psi_{U_k}^T \mathbf{r}_k \right\|_2 &\geq \frac{1}{\sqrt{S}} \left\| \Psi_{U_k}^T \mathbf{r}_k \right\|_1 \\ &\geq \frac{1}{\sqrt{S}} S \alpha_S \\ &\geq \sqrt{S} \alpha_S. \end{aligned} \quad (20)$$

From (19) and (20), one further has

$$\begin{aligned} \alpha_S &\leq \left(\mu \sqrt{|T \setminus \Lambda_k|} + \frac{|\Lambda_k| \sqrt{|T \setminus \Lambda_k|} \mu^2}{1 - (|\Lambda_k| - 1) \mu} \right) \\ &\quad \cdot \left\| \mathbf{x}_{T \setminus \Lambda_k} \right\|_2. \end{aligned} \quad (21)$$

By combining (15) with (21), we obtain the sufficient condition of (11) as

$$\begin{aligned} \frac{1}{\sqrt{|T \setminus \Lambda_k|}} \left(1 - (|T \setminus \Lambda_k| - 1) \mu - \frac{|T \setminus \Lambda_k| |\Lambda_k| \mu^2}{1 - (|\Lambda_k| - 1) \mu} \right) \\ > \mu \sqrt{|T \setminus \Lambda_k|} + \frac{|\Lambda_k| \sqrt{|T \setminus \Lambda_k|} \mu^2}{1 - (|\Lambda_k| - 1) \mu}, \end{aligned} \quad (22)$$

which can be simplified as

$$|T \setminus \Lambda_k| < \frac{1}{2} \left(1 + \frac{1}{\mu} - |\Lambda_k| \right). \quad (23)$$

Since $|T \setminus \Lambda_k| \leq K - k$ and $|\Lambda_k| = Sk$, to guarantee (23) holds, it requires

$$K < \frac{1}{2} \left(1 + \frac{1}{\mu} \right) - \left(\frac{S}{2} - 1 \right) k. \quad (24)$$

When $S = 1$, sufficient condition of (24) becomes

$$K < \frac{1}{2} \left(1 + \frac{1}{\mu} \right) + \frac{1}{2} k. \quad (25)$$

Since $0 \leq k \leq K - 1$, (25) is guaranteed by

$$K < \frac{1}{2} \left(1 + \frac{1}{\mu} \right), \quad (26)$$

or equivalently,

$$\mu < \frac{1}{2K - 1}. \quad (27)$$

When $S \geq 2$, the sufficient condition of (24) can be given by

$$K < \frac{1}{2} \left(1 + \frac{1}{\mu} \right) - \left(\frac{S}{2} - 1 \right) (K - 1), \quad (28)$$

that is,

$$\mu < \frac{1}{S(K - 1) + 1}. \quad (29)$$

Finally, by combining (27) with (29), the sufficient condition for Pre-gOMP can be given by

$$\mu < \begin{cases} \frac{1}{2K-1}, & S = 1, \\ \frac{1}{(K-1)S+1}, & S \geq 2. \end{cases} \quad (30)$$

3 Experiments

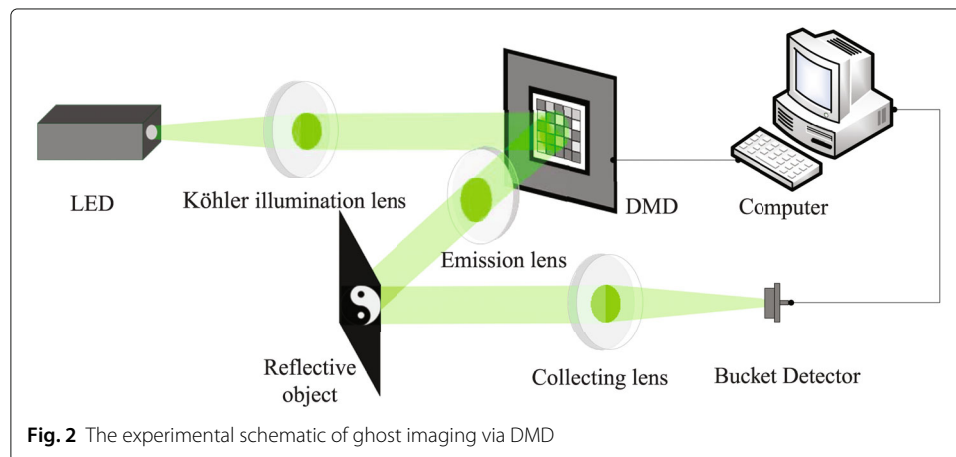
In this section, we carry out simulation and experiment to test the performance of Pre-gOMP. We also apply the Pre-gOMP algorithm to recover image signals in the application of GI.

3.1 Simulation experiments

In the simulation, we use the testing strategy in [22, 23] which measures the effectiveness of recovery algorithms by checking the empirical frequency of exact reconstruction in the noiseless case. For comparison, we adopt the OMP, iterative hard thresholding (IHT) [24], iterative soft thresholding (IST) [25], BP, gOMP, and Pre-gOMP algorithm to recover signals. In each trial, we construct $m \times n$ ($m = 128$ and $n = 256$) random negative exponential sampling matrix Ψ with entries drawn independently from the negative exponential distribution $\exp(1)$. Moreover, we generate K -sparse vector \mathbf{x} whose support is chosen at random. Three types of sparse signals are taken into account: (i) sparse Gaussian signals, (ii) sparse pulse amplitude modulation (PAM) signals, and (iii) sparse two-valued signals, whose non-zeros elements are selected from $\mathcal{N}(0, 1)$, $\{\pm 1, \pm 3\}$, and $\{0, 255\}$, respectively.

3.2 Setup of GI

Figure 2 presents a typical schematic of computational ghost imaging [26]. A light-emitting diode (LED) with wavelength $\lambda = 532nm$ is used as the light source. The light beam is uniformly projected on the digital micromirror device (DMD) by the means of Köhler illumination through the Köhler illumination lens. Here, a series of desired random coded patterns is prebuilt by the DMD, which controls the direction of light by the micro mirrors and the gray value of the light pattern is achieved by controlling the integration time of the DMD. Then, the coded patterns are projected on the object through



the emission lens and photons transmitted through the object are gathered by a bucket detector through a conventional imaging lens.

In the GI system, the DMD consists of 1024×786 pixels and the size of each pixel is $13\mu m$, and only 252×252 pixels on the center of the DMD are used to generate the coded pattern. The size of object is $9mm \times 9mm$. The object's image $O_{GI}(x_r, y_r)$ can be retrieved via conventional the GI and CS algorithm, respectively. In the GI algorithm, $O_{GI}(x_r, y_r)$ is retrieved from the correlation between test arm light intensity B^s and the reference arm light intensity $I_r^s(x_r, y_r)$ [27]:

$$O_{GI}(x_r, y_r) = \frac{1}{m} \sum_{s=1}^m I_r^s(x_r, y_r) - \langle I_r^s(x_r, y_r) \rangle B^s, \quad (31)$$

where s denotes the s th sampling, m is the total sampling number, and $\langle I_r^s(x_r, y_r) \rangle = \sum_{s=1}^m I_r^s(x_r, y_r) / m$ represents the ensemble average of $I_r^s(x_r, y_r)$.

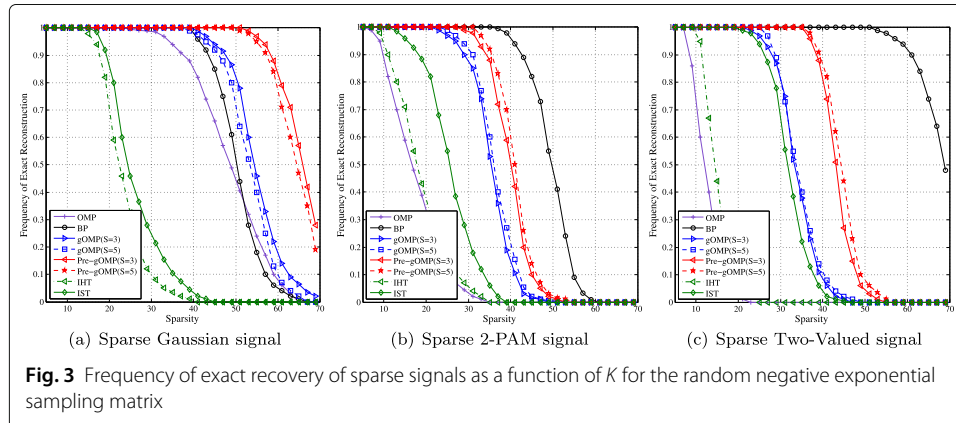
In the CS algorithm, $O_{GI}(x_r, y_r)$ is recovered by solving the problem,

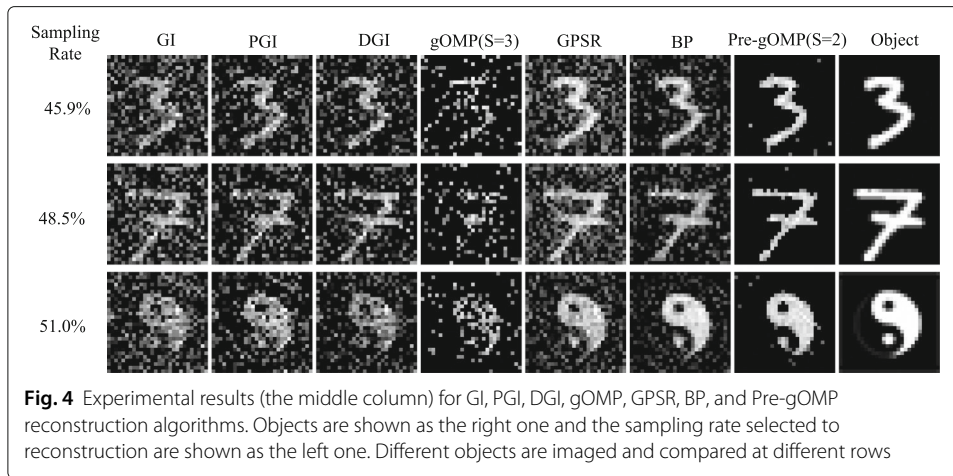
$$\min_{\mathbf{O}} \|\mathbf{O}\|_0 \quad \text{subject to} \quad \mathbf{y} = \Psi \mathbf{O}, \quad (32)$$

where $\mathbf{y} = [B^1 \dots B^s \dots B^m]^T \in \mathcal{R}^m$ denotes the sampling vector, the object \mathbf{O} is reshaped into a column vector, and the Ψ is related to the coded patterns $I_r^s(x_r, y_r)$, $s = 1, \dots, m$.

4 Results and discussion

In Fig. 3, we perform 500 independent trials for each sparsity and plot the empirical frequency of exact reconstruction as a function of the sparsity level. In the gOMP algorithm, we choose $S = 3, 5$ in our simulation. As observed, Fig. 3a, b, and c tell the recovery performance for recovering sparse Gaussian signals, sparse PAM signals, and sparse two-valued signals, respectively. The results reveal that the critical sparsity of the Pre-gOMP algorithm is larger than that of the gOMP algorithm, which implies the preconditioning method indeed promotes the recovery performance. It can also be observed that the Pre-gOMP algorithm outperforms the OMP and the thresholding algorithms. Even when compared with BP, the Pre-gOMP still shows quite competitive recovery performance. Overall, we observe that the Pre-gOMP is effective for recovering all three types of sparse signals.





In the GI imaging experiment, we adopt GI, differential GI (DGI) [27], pseudo-inverse GI (PGI) [28], gOMP, gradient projection for sparse reconstruction (GPSR) [29], BP, and Pre-gOMP algorithms to recover the image of object and compare their performances. The reconstruction results of different objects via different reconstruction algorithms are shown in Fig. 4. In Fig. 4, some results at some specific sampling rate are selected and the sampling rate is labeled in the left column. The reconstruction results are shown in the middle column, the original objects are in the right column, and the reconstruction results via different algorithms are compared in adjacent columns. It is observed that the reconstruction results of GI is significantly improved by the Pre-gOMP algorithm.

Figure 5 shows the recovery performance with respect to different sampling rates. To quantitatively measure the recovery quality, the peak signal noise rate (PSNR) is adopted, which is defined as $PSNR = 10 \log \left(\frac{MAX_I^2}{MSE} \right)$, where the MAX_I is the maximum value in the reconstructed image and the MSE is the mean square error between the reconstructed image and the original object. Larger PSNR generally implies better recovery quality. As observed in Fig. 5, the PSNR increases with the sampling rate. The Pre-gOMP algorithm indeed improves the recovery quality of GI compared with other algorithms. In particular, the recovery quality via Pre-gOMP is slightly better than that via BP within a range of sampling rate. The recovery quality via Pre-gOMP is improved over 2 dB to that via gOMP.

Figure 6 represents the running time via different recovery algorithms. The running time is measured by using the MATLAB program under quad-core 64-bit processor and Windows 10 environment. In Fig. 6, the result of GI algorithm is not included because the computational complexity of GI is similar to that of DGI. Overall, it is observed that the running time of the DGI and PGI is smaller than that of CS algorithms, and the Pre-gOMP(S=2) algorithm is faster than the BP, gOMP(S=3), and GPSR algorithm. Both simulation and experimental results demonstrate that the Pre-gOMP algorithm exhibits competitive performance in the signal reconstruction, while with fast running time.

5 Conclusion

In this paper, we have proposed an algorithm called the Pre-gOMP algorithm for the recovery of sparse signals. Using the mutual coherence framework, we have developed a

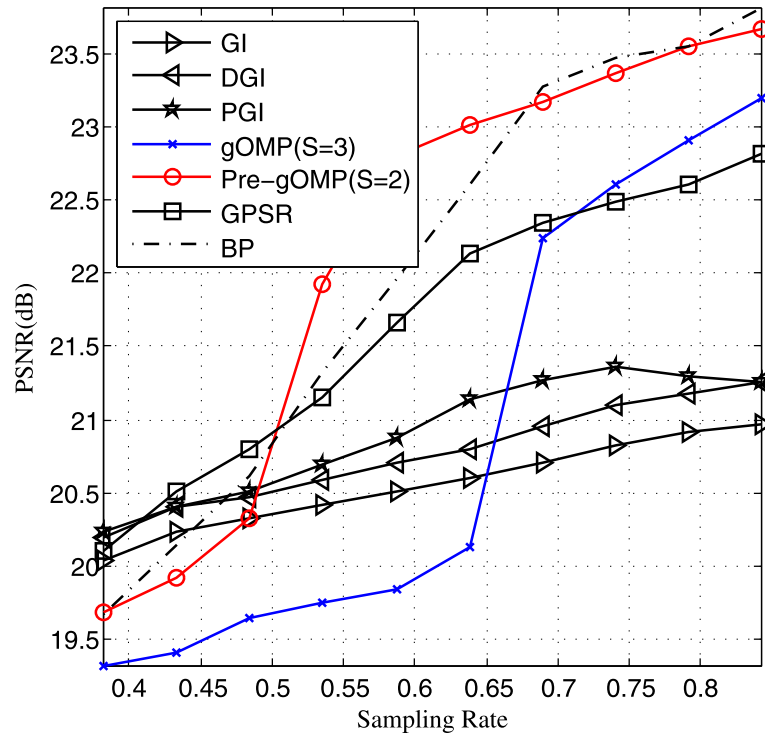


Fig. 5 The reconstruction performance as a function of the sampling rate is shown. Different reconstruction algorithms are compared

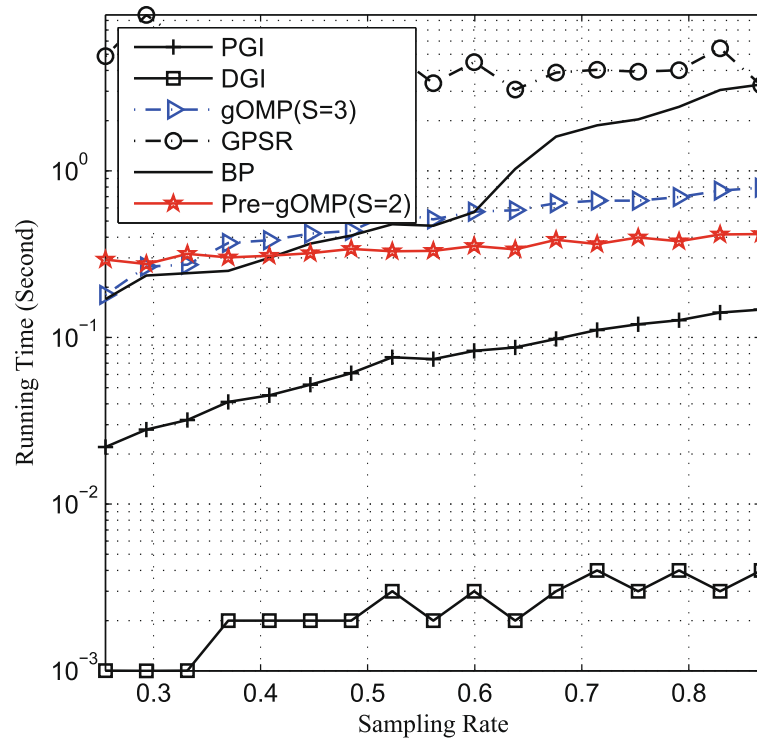


Fig. 6 Running time as a function of sampling rate. Different reconstruction algorithms are compared

sufficient condition for Pre-gOMP to exact reconstruct any K -sparse signal. It is shown that if the μ of the preconditioned sampling matrix satisfies $\mu < 1/(KS - S + 1)$, ($S > 1$), then the Pre-gOMP algorithm perfectly recovers any K -sparse signals from its preconditioned samples. Furthermore, we apply the Pre-gOMP algorithm to recover the image signal in the application of GI. Our experimental results demonstrate that the Pre-gOMP can largely improve the imaging quality of GI, while boosting the recovery speed.

Acknowledgements

The authors would like to thank the National Natural Science Foundation of China and the Youth Innovation Promotion Association of the Chinese Academy of Sciences for their support and to thank anyone who supports this paper to be published.

Authors' contributions

Tong, Hu, Han, and Wang conceived of the algorithm and designed the experiments. Wang and Wang revised the manuscript. All authors read and approved the final manuscript. Zhishen Tong and Feng Wang contribute equally to the first author.

Authors' information

Not applicable.

Funding

This work is supported in part by the National Natural Science Foundation of China (61971146, U1509217) and the Shanghai Municipal Science and Technology Major Project (2018SHZDZX01), and in part by the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2017-2013162).

Availability of data and materials

Please contact the author for data requests.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Key Laboratory for Quantum Optics and Center for Cold Atom Physics of CAS, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, No. 390 Qinghe Road, Jiading District, Shanghai, 201800 Shanghai, China. ²Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, 100049 Beijing, China. ³Department of Management, Shanghai Business School, 200235 Shanghai, China. ⁴School of Data Science, Fudan University, 200433 Shanghai, China. ⁵ZJLab, Fudan-Xinzailing Joint Research Centre for Big Data, Shanghai Key Lab of Intelligent Information Processing, 200433 Shanghai, China. ⁶Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, 310024 Hangzhou, China.

Received: 16 July 2019 Accepted: 6 April 2020

Published online: 07 May 2020

References

1. B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, M. J. Padgett, 3D computational imaging with single-pixel detectors. *Science*. **340**(6134), 844–847 (2013)
2. S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit. *SIAM Rev.* **43**(1), 129–159 (2001)
3. D. L. Donoho, Compressed sensing. *IEEE Trans. Inf. Theory*. **52**(4), 1289–1306 (2006)
4. E. J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*. **52**(2), 489–509 (2006)
5. E. Candes, T. Tao, The Dantzig selector: Statistical estimation when p is much larger than n . *Ann. Stat.* **35**(6), 2313–2351 (2007)
6. S. G. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries. *IEEE Trans. Sig. Process.* **41**(12), 3397–3415 (1993)
7. Y. C. Pati, R. Rezaifar, P. S. Krishnaprasad, in *Proceedings of 27th Asilomar conference on signals, systems and computers*, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition (IEEE, 1993). <https://doi.org/10.1109/acssc.1993.3424651993>
8. S. Chen, S. A. Billings, W. Luo, Orthogonal least squares methods and their application to non-linear system identification. *Int. J. Control.* **50**(5), 1873–1896 (1989)
9. J. Wang, S. Kwon, B. Shim, Generalized orthogonal matching pursuit. *IEEE Trans. Sig. Process.* **60**(12), 6202–6216 (2012)
10. E. J. Candès, The restricted isometry property and its implications for compressed sensing. *C. R. Math.* **346**(9–10), 589–592 (2008)
11. M. Benzi, Preconditioning techniques for large linear systems: A Survey. *J. Comput. Phys.* **182**(2), 418–477 (2002)

12. Z. Tong, J. Wang, S. Han, Preconditioned multiple orthogonal least squares and applications in ghost imaging via sparsity constraint. arXiv:1910.04926 (2019)
13. X. Liao, H. Li, L. Carin, Generalized alternating projection for weighted- $\ell_{2,1}$ minimization with applications to model-based compressive sensing. *SIAM J. Imaging Sci.* **7**(2), 797–823 (2014)
14. S. Ubaru, A. K. Seghouane, Y. Saad, Improving the incoherence of a learned dictionary via rank shrinkage. *Neural Comput.* **29**(1), 263–285 (2017)
15. F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla, L. A. Lugiato, High-resolution ghost image and ghost diffraction experiments with thermal light. *Phys. Rev. Lett.* **94**(184), 1836024 (20054)
16. J. A. Tropp, A. C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Trans. Inf. Theory.* **53**(12), 4655–4666 (2007)
17. G. H. Golub, C. F. Van Loan, *Matrix computations*. (Johns Hopkins University Press, Baltimore, 1983)
18. S. Kwon, J. Wang, B. Shim, Multipath matching pursuit. *IEEE Trans. Inform. Theory.* **60**(5), 2986–3001 (2014)
19. H. Li, J. Wen, A new analysis for support recovery with block orthogonal matching pursuit. *IEEE Sig. Process. Lett.* **26**(2), 247–251 (2018)
20. H. Li, L. Wang, X. Zhan, D. K. Jian, On the fundamental limit of orthogonal matching pursuit for multiple measurement vector. *IEEE Access.* **7**, 48860–48866 (2019)
21. H. Li, J. Wen, Generalized covariance-assisted matching pursuit. *Sig. Process.* **163**, 232–237 (2019)
22. W. Dai, O. Milenkovic, Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inform. Theory.* **55**(5), 2230–2249 (2009)
23. E. Candes, M. Rudelson, T. Tao, R. Vershynin, in *IEEE 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, Error correction via linear programming (IEEE, 2005). <https://doi.org/10.1109/sfcs.2005.5464411>
24. B. Thomas, D. E. Mike, Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* **27**(3), 265–274 (2009)
25. K. Bredies, A. D. Lorenz, Linear convergence of iterative soft-thresholding. *J. Fourier Anal. Appl.* **14**(5), 813–837 (2008)
26. H. J. Shapiro, Computational ghost imaging. *Phys. Rev. A.* **78**(6), 061802–061809 (2008)
27. F. Ferri, D. Magatti, L. A. Lugiato, A. Gatti, Differential ghost imaging. *Phys. Rev. Lett.* **104**(25), 253603 (2010)
28. C. Zhang, S. Guo, J. Cao, J. Guan, F. Gao, Object reconstitution using pseudo-inverse for ghost imaging. *Opt. Express.* **22**(24), 30063–30073 (2014)
29. M. A. Figueiredo, R. D. Nowak, S. J. Wright, Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J. Sel. Topics Sig. Process.* **1**(4), 586–597 (2007)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

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

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)