



Design and Application of a Greedy Pursuit Algorithm Adapted to Overcomplete Dictionary for Sparse Signal Recovery

Shengjie Zhao^{1*}, Jianchen Zhu², Di Wu³

¹ Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200082, China

² School of Software Engineering and Key Laboratory of Embedded System and Service Computing, Ministry of Education, Tongji University, Shanghai 200082, China

³ College of Electronics and Communication Engineering, Tongji University, Shanghai 200082, China

Corresponding Author Email: 1610486@tongji.edu.cn

<https://doi.org/10.18280/ts.370504>

ABSTRACT

Received: 17 May 2020

Accepted: 26 August 2020

Keywords:

overcomplete dictionary, hard thresholding pursuit, projections

Compressive sensing (CS) is a novel paradigm to recover a sparse signal in compressed domain. In some overcomplete dictionaries, most practical signals are sparse rather than orthonormal. Signal space greedy method can derive the optimal or near-optimal projections, making it possible to identify a few most relevant dictionary atoms of an arbitrary signal. More practically, such projections can be processed by standard CS recovery algorithms. This paper proposes a signal space subspace pursuit (SSSP) method to compute sparse signal representations with overcomplete dictionaries, whenever the sensing matrix satisfies the restricted isometry property adapted to dictionary (D-RIP). Specifically, theoretical guarantees were provided to recover the signals from their measurements with overwhelming probability, as long as the sensing matrix satisfies the D-RIP. In addition, a thorough analysis was performed to minimize the number of measurements required for such guarantees. Simulation results demonstrate the validity of our hypothetical theory, as well as the superiority of the proposed approach.

1. INTRODUCTION

Compressive sensing (CS) [1, 2] is a novel paradigm that recovers signals that are sparse in a certain domain, from a small set of compressed measurements. This paradigm has been widely applied in various signal processing applications [3-8], ranging from image [9, 10], audio [11-13], to video [14-19].

The CS aims to sample a sparse signal $x \in \mathbb{R}^n$ from a limited number m of compressed measurements:

$$y = Ax + e \quad (1)$$

where, $y \in \mathbb{R}^m$ is the measurement vector; $A \in \mathbb{R}^{m \times n}$ is an $m \times n$ sensing matrix whose entries are random Gaussian variables; e is the vector of the measurement error induced by the signal structure.

It is assumed that x is sparse in some bases, such that $x = \psi a$ with a has at most k nonzero entries, i.e., $\|x\|_0 = |\text{supp}(x)| \leq k \ll n$. In the CS framework, such an unknown signal can be correctly recovered from y and A , only if $m = O(k \log n)$. The recovery has a low cost and limited computing load.

It is a highly nonlinear process to find a sparse signal from linear measurements, that is, a random linear projection acquisition. A naive approach is to mathematically capture the sparse solution via the l_1 -minimization optimization:

$$\hat{x} = \arg \min \|x\|_1 \text{ s.t. } \|y - Ax\|_2 \leq \varepsilon \quad (2)$$

where, $\|x\|_1 = \sum |x_i|$ is the l_1 -norm; $\|\cdot\|_2$ is the standard Euclidean norm; $\varepsilon > 0$ is decided by the upper bound of $\|e\|_2$. In Problem (2), the accurate recovery of signal x , which is not entirely but nearly sparse, from observation y requires the sufficient condition of restricted isometry property (RIP).

Definition 1 [20]: The sensing matrix $A \in \mathbb{R}^{m \times n}$ follows the RIP of order k if the smallest restricted isometry constant (RIC) $\delta_k \in (0, 1)$, that is,

$$(1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k) \|x\|_2^2 \quad (3)$$

holds for all k sparse signals with $\|x\|_0 \leq k$. Note that, when m is reasonably large, many kinds of matrices, whose entries are random Gaussian variables, are highly likely to have very small RICs.

Let us denote the RIP by A . Then, Problem (2) can be solved stably and robustly based on noisy measurements $y = Ax + e$. Then the solution \hat{x} to Problem (2) can be expressed as:

$$\|\hat{x} - x\|_2 \leq c_1 \frac{\|x - x_k\|_1}{\sqrt{k}} + c_2 \varepsilon \quad (4)$$

where, $c_1 = \frac{2 + (2\sqrt{2} - 2)\delta_{2k}}{1 - (\sqrt{2} + 1)\delta_{2k}}$, and $c_2 = \frac{4\sqrt{1 + \delta_{2k}}}{1 - (\sqrt{2} + 1)\delta_{2k}}$ are positive constants; x_k is an approximation to x with k largest nonzero entries, i.e., $x - x_k$; δ_{2k} is the $2k$ -order RIC of A .

Formula (4) illustrates that the recovery error is proportional to the noise level and the signal tail, that is, the coefficients of the compressible signals follow a power-law decay. It is safe

to say that \hat{x} is approximately x .

Following this train of thought, several conventional/heuristic methods have been proposed to address the non-deterministic polynomial-time (NP)-hard problem depending on A and y , which is computationally infeasible with the growing dimension of signals. These methods include linear programming [2], convex optimization [21, 22], gradient descent [23], and greedy algorithms [24-29]. Among them, the most popular approaches are based on basis pursuit (BP) or matching pursuit (MP).

The abovementioned methods applicable to signals with sparse representation under the standard coordinate basis or several suitable orthonormal bases. However, the actual signals are not necessarily sparse in an orthonormal basis [30-36]. In particular, the results of the above methods do not hold when ψ is an orthonormal basis, but hold for an overcomplete dictionary. This means that signal $f \in R^n$ can be expressed as $x=Da$, where $D \in R^{n \times d}$ ($d > n$) is a given overcomplete dictionary serving as the sparsifying basis, with $a \in R^d$ being the vector of sparse coefficient.

Overcomplete dictionaries are widely used in many signal processing applications, namely, Gabor functions [37], algebraic coded systems [38], wavelets and sinusoids [39], as well as multiscale windowed ridgelet [40]. Although the classical greedy pursuit algorithm provides theoretical guarantees, the practical signals usually fail due to the over-completeness of D . Hence, it is natural to extend the standard CS framework to the signals that are sparse over some overcomplete dictionaries. That is why the authors intended to find the signal directly instead of its coefficient vector provided by the D-RIP.

This paper proposes a variant of MP called signal space subspace pursuit (SSSP), and presents its theoretical guarantees. Similar to MP, the SSSP is a hard thresholding pursuit algorithm that selects atoms from an arbitrary overcomplete dictionary instead of an orthonormal basis for signal representation, thereby minimizing the l_1 -norm of the sparse coefficient vector. By the SSSP, the sample vectors derived from the sparse dictionary signals can be recovered from a small set of linear measurements, which depend on the information level and noise level. Assuming that the sensing matrix satisfies the D-RIP, two relatively simple families of methods, namely, BP-and MP-based methods [41, 42], were adopted to calculate the required near-optimal projections within the preset number of iterations. In addition, the authors analyzed how the minimal number of measurements required for stable recovery of sparse dictionary signals is associated with the concentration inequalities derived from the D-RIP. Finally, the proposed method was proved to achieve much better recovery performance than standard CS recovery algorithms.

2. PROBLEM FORMULATION

Our problem is to recover sparse dictionary signals from a set of noisy measurements like the framework proposed by Zhang and Li [43]. Let $x \in R^n$ be a k sparse input vector, where k is the known level of sparsity. It is assumed that x can be sparsely expressed as the linear combination of atoms in an overcomplete dictionary:

$$x = Da, s. t. \|a\|_0 \leq k \quad (5)$$

where, $D \in R^{n \times d}$ is the sparsifying basis, namely, an overcomplete dictionary; l_0 -norm is all the nonzero entries in the sparse coefficient vector a . D is assumed to be a tight frame such that $DD^T=I_n$, with T being the transpose operator. Similar to the traditional CS, the sampling of each component of the input signal x is realized using a linear-operator A . The process and domains of the CS is illustrated in Figure 1.

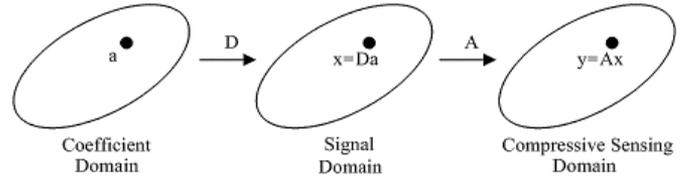


Figure 1. The CS process and its domains

As shown in Figure 1, the random linear projection can be expressed in the form of a matrix as

$$y = Ax + e \quad (6)$$

where, $y \in R^m$ is the observations; $A \in R^{m \times n}$ is a known linear function; $e \in R^m$ is a measurement error, whose entries are independent and identically distributed (I.I.D) Gaussian or sub-Gaussian random variables with zero mean and variance $\delta^2 I_m$, i.e., $e \sim \mathbf{N}(0, \delta^2 I_m)$.

According to advanced hypotheses in the field of CS and recovery of sparse dictionary signals, the sensing matrix A must be extremely incoherent with overcomplete dictionary D , i.e., the dictionary should not have any uncorrelated columns, while preserving as much salient information as possible. Formally, mutual coherence can be defined as:

$$\mu(A, D) = \frac{| \langle A_i, D_j \rangle |}{\|A_i\|_2 \|D_j\|_2} \quad (7)$$

where, A_i and D_j are the columns of A and D , respectively. The coherence provides a fundamental way to measure the maximum magnitude of elements of the inner product AD . Their mutual coherence is proportional to the entries contained in A and D , which depends on μ . If μ is sufficiently large, a dictionary is coherent; otherwise, the dictionary is incoherent. If any pair of columns in D are highly coherent, mutual coherence might not be necessary for recovering the proper signal Da from the coefficient vector a .

To the best of our knowledge, D-RIP, as the natural extension of the standard RIP, is a popular and useful tool to estimate the quality of a sensing matrix for the full recovery of a nearly sparse signal x , which is corrupted by the additive noise from its observations y .

Definition 2 [41]: Let $D \in R^{n \times d}$ be an overcomplete dictionary and a tight frame. Suppose a sensing matrix A follows the D-RIP with the smallest constant δ_k , such that

$$(1 - \delta_k) \|Da\|_2^2 \leq \|Ada\|_2^2 \leq (1 + \delta_k) \|Da\|_2^2 \quad (8)$$

holds for all k sparse signals with $\|a\|_0 \leq k$.

Formula (8) indicates that nearly all the random matrices, whose entries are drawn from Gaussian, sub-Gaussian, or Bernoulli distribution with measurement number on the order of $\|a\|_0 \leq k$, are very likely to satisfy the D-RIP, which exploits the nearly mutually orthogonal of the columns of the sensing matrix.

Suppose A and D exhibit the D-RIP. Under a stronger assumption that x is k -analysis-sparse, i.e., $D^T x \in R^d$ is k -sparse, the highly correlated property of D can be used to recover x from its noisy measurements $y=ADa+e$ by solving the relaxed l_1 -minimization problem:

$$\hat{x} = \underset{x \in R^n}{\operatorname{argmin}} \|D^T \tilde{x}\|_1, \text{ s. t. } \|y - A\tilde{x}\|_2 \leq \varepsilon \quad (9)$$

where, $\varepsilon > 0$ is an upper bound proportional to the noise level $\|e\|_2$.

Then, the solution of Problem (9) can be expressed as:

$$\|\hat{x} - x\|_2 \leq c_1 \frac{\|D^T x - (D^T x)_k\|_1}{\sqrt{k}} + c_2 \varepsilon \quad (10)$$

where, $c_1 > 0$ and $c_2 > 0$ are numerical constants solely depending on δ_{2k} ; $(D^T x)_k$ is the best approximation of $D^T x$ within all the k -largest nonzero entries in l_1 -norm, analogous to the solution (4). Thus, the recovery error $\|\hat{x} - x\|_2$ depends on the noise level $\|e\|_2$ and the tail of the analysis vector $\frac{\|D^T x - (D^T x)_k\|_1}{\sqrt{k}}$, and decays rapidly with the growing number of iterations. Since these results hold even if the dictionary has highly correlated columns, it is possible to derive the results mathematically, in a way similar to MP-based methods, under the previous constraints. Without loss of generality, the convergence for the algorithm is similar to that for the l_1 -analysis methods.

3. SSSP ALGORITHM

This section derives the SSSP algorithm from MP-based methods under the constraint that D is an overcomplete dictionary rather than an orthonormal basis. Assuming that the sensing matrix obeys the D-RIP, a guarantee was designed to minimize the measurement number required to recover a sparse dictionary signal. Finally, a rigorous recovery error bound was derived and analyzed theoretically to verify the algorithm performance.

3.1 Algorithm design

The key of the underlying CS argument was explained on the simple case of a sparse signal, using a famous sparsifying basis, i.e., the orthonormal basis. The current CS theory holds that sparse signal can be represented through convex programming, such as match pursuit (MP) and its variants. For signal representation, MP-based methods iteratively select the columns of the basis closest to the original signal for projection onto the subspace, until meeting the preset termination condition. More specifically, x is recovered by solving the following constrained optimization problem based on least-squares:

$$\hat{x} = \underset{z}{\operatorname{argmin}} \|y - Az\|_2, z \in R(\Psi_A) \quad (11)$$

where, Ψ_A is the matrix Ψ constraint to the columns indexed by A ; $R(\Psi_A)$ is the span of Ψ_A . Drawing on the concept of pseudo inverse, Problem (11) can be solved by:

$$\hat{x} = \Psi_A \tilde{A}_\Lambda^\dagger y = \Psi_A ((\tilde{A}_\Lambda^T \tilde{A}_\Lambda)^{-1} \tilde{A}_\Lambda^T y) \quad (12)$$

where, \dagger is the pseudo-inverse operator. However, actual signals are often compressible in some overcomplete dictionary, including oversampled discrete Fourier transform (DFT), Gabor frame, and curvelet coefficient sequence of an image. In the absence of dictionary orthogonal, the standard l_1 -minimization approach usually fails to pinpoint the support estimation Λ . Interestingly, a recent investigation reports that MP-based methods adapted to the setting of an arbitrary dictionary could recover the analysis-sparse vector $D^T x$ associated with signal space methods, which easily find near sparse solutions and guarantee the recovery performance.

As an MP variant, our algorithm recovers the x using D from the noisy measurements (Table 1).

Table 1. The SSSP algorithm

Algorithm 1: SSSP algorithm
Input: A, D, y, k , termination condition ε
Initialization: $l=0, I=\emptyset, r^0=y$, and $x^0=0$
while termination condition is not satisfied do
1: identify the index set $\Omega = \operatorname{supp}(D, 2k) (ATr)$
2: find the support estimation $T=\Omega \cup I$
3: compute the approximation: $\tilde{x} = \operatorname{argmin}_z \ y - Az\ _2$
s. t. $z \in R(D_T)$
4: shrink the index $I = \operatorname{supp}(x^*, k) = \text{index corresponding to the largest magnitude entries of estimation } \tilde{x}$.
5: compute the new signal estimation: $x^{l+1} = P_I \tilde{x}$
6: compute the latest residual: $r^{l+1} = y - Ax^{l+1}$
end while: $l = \operatorname{MaxIter}$ or $\ x^{l+1} - x^l\ _2 / \ x^l\ _2 \leq \varepsilon$ is satisfied
Output: $\hat{x} = x^{l+1} = \operatorname{SSSP}(A, D, y, k)$

The SSSP algorithm consists of six steps. The first and second steps are identification and merging. The third to last steps are iterative updates. In the first step, the support estimation Ω is identified by iteratively introducing multiple indices with $2k$ largest related values in magnitude provided by some inputs A, D and y , plus the initialization $x^0=0$. In the third step, the cardinality of the new support set I corresponding to the sparsity level k is updated, using the shrinking obtained by merging in the second step. This leads to a relatively sparse solution based on the trimmed set and a least-square solution.

The termination condition is of great importance to the algorithm. As outlined in the next section, the normalized relative error $\|x^{l+1} - x^l\|_2 / \|x^l\|_2 \leq \varepsilon$ was adopted as the termination condition in the simulations.

Like classical MP-based methods, the SSSP is a hard thresholding pursuit algorithm that makes support estimation by greedy strategy, performs such a projection, and represents the sparse signals within a finite number of iterations.

The main idea of the SSSP is to correctly choose the support estimation subject to an overcomplete dictionary constraint during each iteration. As shown in Table 1, the main difference between SSSP and traditional MP-based methods is the projection calculation by replacing simple hard thresholding:

$$\Lambda_{opt}(x, k) = \operatorname{argmin}_{\Lambda: \|\Lambda\|_0 \leq k} \|x - P_\Lambda x\|_2 \quad (13)$$

where, P_A is a hard thresholding operator applied to the optimal projection of a general vector $x \in R^n$ onto the columns of D subject to the index set A constraint. In particular, such a projection is calculated with the proxy $A^T r$ if it has at most k nonzero entries. This guarantees the approximation to the desired sparse solution.

Problem (13) is generally NP-hard compared to the classical CS problem. It is assumed that the calculation of the optimal projection $P_A x$ is unworkable and nontrivial provided by any A , D , and r . An alternative way is to apply near-optimal projection to capture a near approximation of $P_A x$, such that

$$\begin{aligned} \| P_{supp(D,k)(A^T r)} x - x \|_2 &\leq c_1 \| P_A x - x \|_2 \\ \| P_{supp(D,k)(A^T r)} x - x \|_2 &\leq c_2 \| P_A x \|_2 \end{aligned} \quad (14)$$

for some constants $c_1 \geq 0$ and $c_2 \geq 0$. Notably, near-optimal projection is equivalent to optimal projection, even if the columns of D are highly correlated under $P_{supp(D,k)(A^T r)} x = P_A x$. This clearly leads to highly accurate signal recovery.

Then, the least-squares problem based on the calculation of $P_{supp(D,k)(A^T r)} x$ was considered instead of that of $P_A x$. Specifically, the algorithm establishes the signal estimation \tilde{x} via the dominant least squares problem of the merging step provided by T such that

$$\hat{x} = \underset{z}{\operatorname{argmin}} \| y - Az \|_2, \text{ s. t. } z \in R(D_T) \quad (15)$$

with $2k$ largest nonzero entries of $D^T(A^T r)$. Similarly, the solution to the least-squares problem can be obtained by:

$$\hat{x} = D_T \tilde{A}_T^\dagger y = D_T ((\tilde{A}_T^T \tilde{A}_T)^{-1} \tilde{A}_T^T y) \quad (16)$$

where, A_T is the submatrix of A indexed by T . Many empirical evidences verify that the above standard CS recovery algorithms are suitable for calculating such a near-optimal projection.

3.2 Bound for measurement number

Setting aside the rich structure of dictionary D , the authors first explored the requirements on minimal measurement number, which depend on the sensing matrix A . Previous works have shown that x can be recovered stably from the compressive measurements, provided that A satisfies the RIP with the smallest constant $\delta_k \in (0,1)$. But numerous real signals are compressible under an overcomplete dictionary. Due to the overcomplete feature of D , the recovery error $\| x - \hat{x} \|_2$ in signal space is obviously smaller or larger than that $\| a - \hat{a} \|_2$ in coefficient space, i.e., $\| x - \hat{x} \|_2 = \| Da - \hat{D}\hat{a} \|_2 \neq \| a - \hat{a} \|_2$.

Next comes the situation where x can be sparsely represented by D . Specifically, the matrix A satisfies the D-RIP of order k , if there exists a constant $\delta_k \in (0,1)$ such that

$$\sqrt{1 - \delta_k} \leq \frac{\| ADa \|_2}{\| Da \|_2} \leq \sqrt{1 + \delta_k} \quad (17)$$

holds for all a with $\| a \|_0 \leq k$. The D-RIP ensures that the

norm preservation of all signals has a sparse representation $x = Da$, in contrast to the standard RIP.

Many methods are available to design matrices that satisfy the D-RIP. To the best of our knowledge, the sensing matrices whose entries are drawn from sub-Gaussian distribution are highly likely to satisfy the D-RIP. Here, the matrices are constructed in the following manner:

A matrix $A \in R^{m \times n}$ was generated by selecting the entries $A[M, N]$ as I.I.D random variables. Then, two conditions were imposed on the random distribution. First, the distribution must be centered and normalized such that $A[M, N] \stackrel{i.i.d}{\sim} N(0, \frac{1}{m})$. Second, the expected value of random variable $\| Ax \|_2^2$ must be $\| x \|_2^2$, which is given by:

$$E(\| Ax \|_2^2) = \| x \|_2^2 \quad (18)$$

Without loss of generality, any distribution with bounded support is sub-Gaussian, including Gaussian and uniform distributions. This paper utilizes a key property of such random variables: For a fixed vector (signal) x , any matrix $A \in R^{m \times n}$ which satisfies

$$Pr(\| Ax \|_2^2 - \| x \|_2^2 \geq \epsilon \| x \|_2^2) \leq 4e^{-c_0(\epsilon)m} \quad (19)$$

implies that a matrix A is highly likely to satisfy the D-RIP, if and only if m is at least on the order of $k \log(d/k)$. For this reason, the probability used in the paper was taken over all draws of A and the constant $c_0(\epsilon)$ based on the particular sub-Gaussian distribution and the range of ϵ . Perhaps the most fundamental for our purpose is the following lemma.

Lemma 1: Let χ denote any k -dimensional of R^n with the fixed δ , $a \in (0,1)$. Suppose that A is an $m \times n$ random matrix whose entries are selected from an I.I.D sub-Gaussian distribution satisfying (19), the minimum number of measurements for correct signal recovery can be obtained by:

$$m = O\left(\frac{2k \log(42/\delta) + \log(4/\alpha)}{c_0(\delta/\sqrt{2})}\right) \quad (20)$$

with probability exceeding $1 - \alpha$,

$$\sqrt{1 - \delta} \| x \|_2 \leq \| Ax \|_2 \leq \sqrt{1 + \delta} \| x \|_2 \quad (21)$$

for all $x \in \chi$.

Instead of focusing on a single k -dimensional subspace, Lemma 1 was used to consider all possible subspaces spanned by k columns of D to construct the D-RIP for A . Thus, the following lemma can be derived:

Lemma 2: Let D be an overcomplete dictionary of the dimension $n \times d$ with fixed δ , $a \in (0,1)$. The minimal number of measurements for correct signal recovery can be obtained by

$$m = O\left(\frac{2k \log(42ed/\delta k) + \log(4/\alpha)}{c_0(\delta/\sqrt{2})}\right) \quad (22)$$

where, e is the base of the natural logarithm. Then, A will satisfy the D-RIP of order k with a small constant δ at the probability of $1 - \alpha$.

As noted above, almost all random matrices help to implement the sensing process. The following parts will focus on the random matrices in the development of our theory.

3.3 Recovery of approximately sparse-dictionary signals from incomplete measurements

The assumption that signals have a sparse representation in an overcomplete dictionary D provides a theoretical guarantee for the accuracy recovery of sparse signals. Analogously to the guarantees of MP-based methods, the proof relies on iteration invariant, which indicates that the recovery error is mostly determined by the number of iterations.

Before stating the main result on the recovered error of the algorithm, the following lemma was presented:

Lemma 3: Assuming that A follows the D-RIP with the smallest constant δ_t , any signal x can be expressed as $x=Da$ with $a \in R^d$ provided by any support estimation β such that

$$|\langle x, (A^T A - I)x \rangle| \leq \delta_t \|x\|_2^2 \quad (23)$$

where $t = \text{card}\{\text{supp}(x)\}$.

So far, it is pointed out that, to bound the actual recovery error, a sequence approximation $x^l \in \{x^0, \dots, x^{l+1}\}$ with $\|x^l\|_0 \leq k$ is consecutively close to the original signal x established by Lemma 3. This leads to the following theorem.

Theorem 1: A procedure $\text{supp}_{(D,k)}(A^T r)$ corresponds to a near-optimal projection $P_{\text{supp}_{(D,k)}(A^T r)}$ when $(1 + c_1)(1 - \frac{c_2}{(1+\beta)^2}) < 1$, which indicates that

$$\|x - x^{l+1}\|_2 \leq \eta_1 \|e\|_2 \quad (24)$$

where, β is an arbitrary constant; η_1 is a constant depending on c_1, c_2 , and β . Inspired by previous works on signal space setting, the conditions of Theorem 1 on the near-optimal projections holds in cases where D cannot satisfy the traditional RIP. Thus, a stronger convergence for the algorithm was provided in the following theorem.

Theorem 2: Let A be a sensing matrix satisfying the D-RIP of order $4k$ or a coefficient vector a . Then, the signal estimation x^{l+1} after $l+1$ iterations of the algorithm satisfies

$$\|x - x^{l+1}\|_2 \leq C_1 \|x - x^l\|_2 + C_2 \|e\|_2$$

$$C_1 = ((2 + \lambda_1)\delta_{4k} + \lambda_1)(2 + \lambda_2) \frac{\sqrt{1 + \delta_{4k}}}{\sqrt{1 - \delta_{4k}}} \quad (25)$$

$$C_2 = \left(\frac{(2 + \lambda_2)((2 + \lambda_1)(1 + \delta_{4k}) + 2)}{\sqrt{1 - \delta_{4k}}} \right).$$

Note that constants C_1 and C_2 depend on the isometry constant δ_{4k} , as well as the approximation parameters λ_1 and λ_2 . Further, an immediate lemma associated with Theorem 2 can be derived:

Lemma 4: Under the conditions of Theorem 3, after a constant number of iterations $l + 1 = \lceil \frac{\log(\|x\|_2/\|e\|_2)}{\log(1/C_1)} \rceil$, there is:

$$\|x - x^{l+1}\|_2 \leq \left(1 + \frac{1 - C_1^{l+1}}{1 - C_1}\right) C_2 \|e\|_2 \quad (26)$$

Lemma 4 implies the results in Theorem 1 with $\eta_1 = (1 + \frac{1 - C_1^{l+1}}{1 - C_1}) C_2$. More specifically, through various combinations of C_1, C_2 and δ_{4k} , Theorem 2 shows that $C_1 < 1$ and the accuracy of the algorithm improves per iteration. Thus, it can be concluded that $C_1 \leq 0.5$ and $C_2 \leq 7.5$, if $\lambda_1 = \frac{1}{10}, \lambda_2 = 1$, and $\delta_{4k} \leq 0.1$. The

recursive nature of Theorem 2 is induced by the following lemma.

Lemma 5: Under the conditions of Theorem 2, the signal estimation x^l after l -th iteration satisfies

$$\|x - x^{l+1}\|_2 \leq 0.5 \|x - x^l\|_2 + 7.5 \|e\|_2.$$

Particularly,

$$\|x - x^l\|_2 \leq 2^{-l} \|x\|_2 + 15 \|e\|_2 \quad (27)$$

Each iteration of the algorithm is crucial to reducing the recovery error by a constant factor, while adding an noise component. After enough iterations l , the most term $2^{-l} \|x\|_2$ can be minimized, and the recovery error will solely depend on the noise level. The upper bound of recovery error (27) also applies to the accurate support estimation T .

3.4 Recovery of approximately arbitrary signals from incomplete measurements

Theorem 2 implies that the recovery accuracy is positively correlated with the values of C_1 and C_2 . It is possible to achieve accurate recovery by choosing a sufficiently small $\|e\|_2$. However, this is not the case if the signals do not have a sparse representation in D , that is, if

$$y = A(Da_k) + A(x - Da_k) + e = A(Da_k) + \hat{e} \quad (28)$$

where, $\hat{e} = A(x - Da_k) + e$ is the equivalent measurement error corrupted by the noise constraint to a sparse-dictionary signal with $\|a_k\|_0 \leq k$. In fact, such a noise \hat{e} limits the maximum achievable accuracy. For illustration, the relative conditions were treated as valid, and the following lemma was extended from Theorem 2.

Lemma 6: For the general CS model $y = A(Da_k) + \hat{e}$ (28), if $\delta_{4k} \leq 0.1$, the upper bound of recovery error can be expressed by

$$\|x - x^{l+1}\|_2 \leq 0.5 \|x - x^l\|_2 + \|x - Da_k\|_2 + 7.5 \|A(x - Da_k)\|_2 + 7.5 \|e\|_2.$$

Particularly,

$$\|x - x^l\|_2 \leq 2^{-l} \|Da_k\|_2 + \|x - Da_k\|_2 + 15 \|A(x - Da_k)\|_2 + 15 \|e\|_2 \quad (29)$$

where, a_k is an approximation of x with $\|a_k\|_0 \leq k$.

Note that the coefficient vector a_k minimizes the upper bound (29), highlighting the importance of a_k to sparse signal recovery from a small set of measurements. Based on Lemma 6, the term $\|A(x - Da_k)\|_2$ was used to prove the convergence for the algorithm if D is not unitary.

Then, (29) was modified into the following theorem, which implies the existence of an upper bound of the term $\|A(x - Da_k)\|_2$ in the signal space.

Theorem 3: Assuming that A satisfies the upper bound of D-RIP with the constant $\delta_{4k} \leq 0.1$, then the following holds for a general vector $x = Da_k \in R^T$:

$$\|Ax\|_2 \leq \sqrt{1 + \delta_k} (\|x\|_2 + \frac{1}{\sqrt{k}} \|x\|_1) \quad (30)$$

Bounding the right side of (29) with Theorem 3:

$$\begin{aligned} \|x - x^{l+1}\|_2 &\leq 0.5 \|x - x^l\|_2 + 7.5 \|e\|_2 \\ &\quad + (7.5\sqrt{1 + \delta_k} + 1) \|x - Da_k\|_2 \\ &\quad + \frac{7.5\sqrt{1 + \delta_k}}{\sqrt{k}} \|x - Da_k\|_1 \end{aligned}$$

Particularly,

$$\begin{aligned} \|x - x^l\|_2 &\leq 2^{-l} \|Da_k\|_2 + 15 \|e\|_2 \\ &\quad + (15\sqrt{1 + \delta_k} + 1) \|x - Da_k\|_2 \\ &\quad + \frac{15\sqrt{1 + \delta_k}}{\sqrt{k}} \|x - Da_k\|_1 \end{aligned} \quad (31)$$

Denote

$$\begin{aligned} M(x) := \inf_{a_k: \|a_k\|_0 \leq k} (\|x - Da_k\|_2 \\ + \frac{1}{\sqrt{k}} \|x - Da_k\|_1) \end{aligned} \quad (32)$$

is the model mismatch quantity (for any $x \in \mathbb{R}^n$). Notice that (30) and (32) have a similar form. Combining this result in (31), we have

$$\begin{aligned} \|x - x^{l+1}\|_2 &\leq 0.5 \|x - x^l\|_2 + 7.5 \|e\|_2 \\ &\quad + 8.5\sqrt{1 + \delta_k} M(x). \end{aligned}$$

Particularly,

$$\begin{aligned} \|x - x^l\|_2 &\leq 2^{-l} \|Da_k\|_2 + 15 \|e\|_2 \\ &\quad + 16\sqrt{1 + \delta_k} M(x) \end{aligned} \quad (33)$$

Note that this quantity measures the bound of the above recovery error. If $M(x)$ is sufficiently large, the signal is neither a k sparse signal nor a compressible signal, such that $x \neq Da_k$.

Since the term $\|A(x - Da_k)\|_2$ bounds the recovery error, (29) was modified into the following lemma, which implies the existence of an upper bound of the term $\|A(x - Da_k)\|_2$ in the coefficient space.

Lemma 7: If AD satisfies the condition of D-RIP with the constant $\delta_{4k} \leq 0.1$, (30) can be extended into:

$$\begin{aligned} \|A(x - x_k)\|_2 &= \|AD(a - a_k)\|_2 \\ &\leq \sqrt{1 + \delta_k} (\|a - a_k\|_2 + \frac{\|a - a_k\|_1}{\sqrt{k}}) \end{aligned} \quad (34)$$

By Lemma 7, the term $\|AD(a - a_k)\|_2$ bounds the recovery error in the coefficient space:

$$\begin{aligned} \|x - x^{l+1}\|_2 &\leq 0.5 \|x - x^l\|_2 + 7.5 \|e\|_2 \\ &\quad + \|x - Da_k\|_2 \\ &\quad + 7.5\sqrt{1 + \delta_k} (\|a - a_k\|_2 + \frac{1}{\sqrt{k}} \|a - a_k\|_1). \end{aligned}$$

Particularly,

$$\begin{aligned} \|x - x^l\|_2 &\leq 2^{-l} \|Da_k\|_2 + 15 \|e\|_2 \\ &\quad + \|x - Da_k\|_2 \\ &\quad + 15\sqrt{1 + \delta_k} (\|a - a_k\|_2 + \frac{1}{\sqrt{k}} \|a - a_k\|_1) \end{aligned} \quad (35)$$

where, a_k is a near k sparse approximation of a . If a_k is arbitrarily compressible, then $a = a_k$, indicating that the upper bound of (35) is reasonably small.

3.5 Computation complexity

This subsection discusses the subsequent results related to the convergence speed of the algorithm.

It is assumed that $\hat{x} = x^{l+1}$ is an output of the algorithm after $l+1$ iterations. For parameter $\eta > 0$, the algorithm produces a signal estimation \hat{x} after at most $O(\log \|x\|_2 / \eta)$ iterations such that

$$\|x - \hat{x}\|_2 = O(\eta + \|e\|_2) = O(\max\{\eta, \|e\|_2\}). \quad (36)$$

The computing cost is the sum of costs of its six steps. The first step acquires the signal estimation \tilde{x} via the LS provided by the proxy $A^T r$. The second step efficiently calculates the support estimation $\Omega = \text{supp}_{(D, 2k)}(A^T r)$ and its corresponding pruned set I with the standard CS recovery algorithms. The single runtime of these algorithms is proportional to $O(knd)$ or $O(nd)$ during each iteration. Therefore, the overall runtime of these algorithms is $O(knd \log \|x\|_2 / \eta)$ or $O(nd \log \|x\|_2 / \eta)$ as k increases. As depicted in Figure 2, the runtime was in line with advanced bounds, resulting in a relatively linear convergence.

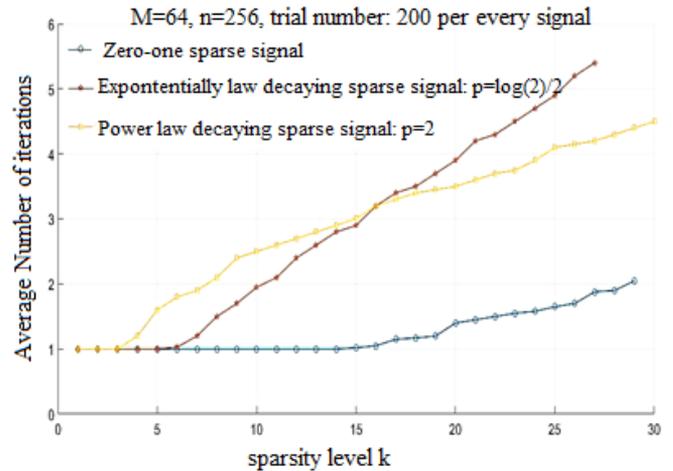


Figure 2. The convergence of the SSSP algorithm for three types of signals

The next section illustrates challenging task of computing the near-optimal projection for a near approximation to optimal projection is illustrated through simulations, and details the parameters related to sparse-dictionary signal recovery.

4. SIMULATION AND RESULTS ANALYSIS

This section compares the recovery performance of the proposed method with conventional baseline methods, and measures the signal recovery frequency with additive Gaussian noise. The proposed method was also applied to coded aperture spectral imaging (CASSI) on a hyperspectral data cube, in contrast to the said baseline methods.

4.1 Recovery performance under an overcomplete dictionary

In the first simulation, the original data were generated in three steps: (1) Generate a k sparse signal x with $n=256$ sparse in the dictionary D domain; (2) Generate a random sensing matrix; (3) Compute the compressive measurement. The simulation consists of 1,000 independent trials.

The recovery error R_{rec} were measured by two indices R_{rec} and R_{res} :

$$R_{rec}(x, \hat{x}) = \frac{\|x - \hat{x}\|_2}{\|x\|_2} \leq \varepsilon_1 \quad (37)$$

with $\|x - \hat{x}\|_2 \leq 10^{-4} \|x\|_2$.

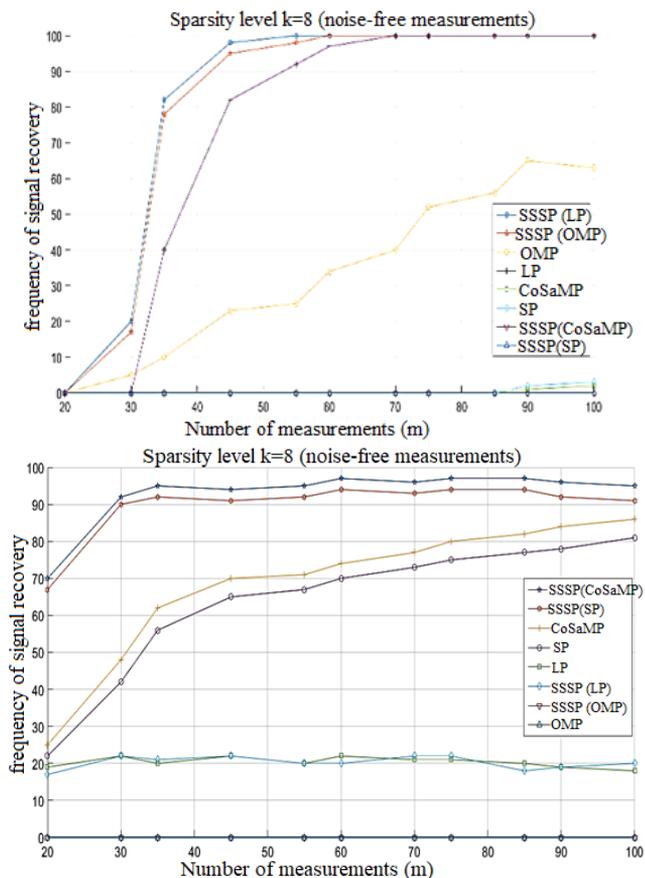


Figure 3. The frequency of signal recovery out of 1,000 trials for different SSSP variants, when the nonzero entries in a are well separated (a), and when the nonzero entries in a are clustered (b)

Note, $k=8$, $n=256$, $d=1,024$, the dictionary $D \in \mathbb{R}^{n \times d}$ is a $4 \times$ overcomplete DFT, and $A \in \mathbb{R}^{m \times n}$ is a Gaussian matrix.

Then, the effects of the algorithm with different support estimation techniques were checked for the case where the k nonzero entries of a are well separated, and the case where they are clustered, in comparison with four existing algorithms: Orthogonal matching pursuit (OMP), compressive sampling matching pursuit (CoSaMP), subspace pursuit (SP), and linear programming (LP). Note that D is a $4 \times$ overcomplete DFT dictionary. Thus, the neighboring columns in this dictionary are highly coherent. First, the frequency of signal recovery was investigated with sparsity level $k=8$ as a function of the measurement number m . The simulation results are shown in Figure 3.

As mentioned before, the main task of the algorithm is the calculation of $\Lambda_{opt}(x, k)$. Such projections are required for identification and update steps. To cope with the challenge, standard recovery algorithms like OMP, SP, CoSaMP, and LP were adopted to calculate the near-optimal projection $P_{supp(D, k)}(A^T r)$, which leads to the support estimation $supp_{(D, k)}(A^T r)$. For simplicity, the support estimations obtained by OMP, SP, CoSaMP, and LP are denoted as SSSP(OMP), SSSP(SP), SSSP(SP), and SSSP(CoSaMP), respectively.

Figure 3(a) compares the performance of eight different algorithms for the case where the nonzero entries of a are well separated. It can be seen that SSSP(LP) outperformed the other algorithms when using the support $supp_{(D, k)}(A^T r)$ derived by a classical algorithm like LP. This means LP is suitable for finding the exact $\Lambda_{opt}(x, k)$ when $x = P_{\Lambda_{opt}(x, k)} x$ and its corresponding nonzero entries of $\Lambda_{opt}(x, k)$ are well separated. It can also be seen that OMP, CoSaMP, and SP are not effective for signal recovery, due to the mutual coherence between A and D .

Figure 3(b) compares the performance of eight different algorithms for the case where the nonzero entries of a are clustered. It can be seen that SSSP(CoSaMP) outshined the other algorithms. Besides, SSSP(OMP) and OMP always failed with the increase of m , because OMP selects the largest index per iteration. This means the correct identification of support set hinges on the high coherence between close atoms in and around the cluster.

It can be thus concluded that the algorithm accurately recovers the signals, whereas LP and OMP performs poorly, when the supports of x are clustered together; the exact opposite phenomena occur if the supports are well separated. In general, the algorithm variants outperform the corresponding algorithm.

4.2 Recovery performance for overcomplete dictionaries in compressive spectral

Then, the signal recovery performance of our algorithm was verified on a small set of compressive CASSI measurements, which covers several practical issues. Specifically, a real hyperspectral data cube with a sparse representation in some dictionaries was adopted instead of an ideal sparse-dictionary signal whose coefficients are random sub-Gaussian variables.

In the second simulation, the final test dataset was sensed via a charge-coupled device (CCD) camera AVT Marlin F033B (range: 450-700nm; step size: 10nm). The hyperspectral images captured by the camera have a spatial resolution of 256×256 pixels, with $L=24$ bands. Figure 4 presents a red-green-blue (RGB) profile of the test data cube.

The $4 \times$ overcomplete dictionary of the dimension 64×256 can be sparsely represented by 1,024 nonoverlapping pixels for each image. The proposed algorithm was adopted to recover the underlying data cube from the CASSI measurements, corrupted by the sensing noise with zero-mean and variance $\delta^2=10^{-4}$. The same baseline methods were taken as the benchmarks for image recovery, due to their reasonable computation complexity. The peak signal-to-noise ratio (PSNR) was selected to measure the quality of each recovered image:

$$PSNR = 10 \log_{10} \frac{\|x\|_2}{\|x - \hat{x}\|_2} \quad (38)$$

where, x is the original image; \hat{x} is the recovered image.



Figure 4. A RGB decomposed representation of the original bands for the test dataset

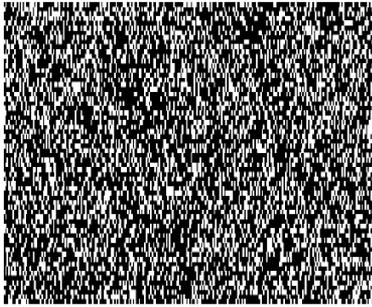


Figure 5. The random coded aperture

Table 2. The mean PSNR for the recovered bands of multispectral scene (unit: dB)

Method	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6
OMP	14.25	14.32	14.24	14.93	14.21	14.57
SP	16.32	16.58	16.43	16.59	16.42	16.58
CoSaMP	17.33	17.46	17.44	17.52	17.39	17.55
LP	27.33	27.31	27.75	27.81	27.77	27.41
SSSP	27.32	27.55	27.83	27.41	27.92	27.55



Figure 6. The reconstruction of a subset of bands for the hyperspectral images with different algorithms: (1) OMP, (2) SP, (3) CoSaMP, (4) LP, and (5) SSSP

As shown in Figure 6 and Table 2, the proposed algorithm achieved the best quality of recovered bands, provided by the structure of H using a random coded aperture (Figure 5), in terms of PSNR.

Note that all the entries lie in a random code obey sub-Gaussian distribution. The results visually show the bands recovered with the proposed algorithm are superior than those recovered with the benchmark algorithms. In contrast, the baseline methods performed poorly, because they neglect the near-optimal projection scheme.

It can also be seen from Figure 6 that, the PSNR achieved with our algorithm outperformed that of most baseline algorithms by up to 7dB. The only method that realized comparable effect is LP when using SSSP. The results verify that our algorithm can improve the recovery of hyperspectral images, while greatly reducing the computing complexity.

5. CONCLUSIONS

This paper presents a near-optimal projection scheme to solve the bottleneck in the setting of the overcomplete dictionary. The scheme develops such a projection to iteratively identify the atoms from the given dictionary. It was observed that the accuracy of the algorithm is solely controlled by the noise level due to its convergence. The behavior of the signal space method was also analyzed, even when the columns of dictionary too coherent to satisfy the general properties. It is verified that our method can obtain the optimal projection, and provide theoretical basis for clear explanation of the observed phenomena. Simulation results show that our method achieved outstanding recovery performance in noisy case.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61936014, and in part by the National Key Research and Development Project under Grants 2019YFB2102300 and 2019YFB2102301, and was supported by the Fundamental Research Funds for the Central Universities.

REFERENCES

- [1] Candes, E.J., Tao, T. (2005). Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(12): 4203-4215. <https://doi.org/10.1109/TIT.2005.858979>
- [2] Donoho, D.L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4): 1289-1306. <https://doi.org/10.1109/TIT.2006.871582>
- [3] Kokiopoulou, E., Frossard, P. (2008). Semantic coding by supervised dimensionality reduction. *IEEE Transactions on Multimedia*, 10(5): 806-818. <https://doi.org/10.1109/TMM.2008.922806>
- [4] Peng, J., Xie, Q., Zhao, Q., Wang, Y., Yee, L., Meng, D. (2020). Enhanced 3DTV regularization and its applications on HSI denoising and compressed sensing. *IEEE Transactions on Image Processing*, 29: 7889-7903. <https://doi.org/10.1109/TIP.2020.3007840>
- [5] He, Q., Chen, Z., Quek, T.Q., Choi, J., Li, S. (2018).

- Compressive channel estimation and user activity detection in distributed-input distributed-output systems. *IEEE Communications Letters*, 22(9): 1850-1853. <https://doi.org/10.1109/LCOMM.2018.2858241>
- [6] Ganguly, S., Ghosh, J., Srinivas, K., Kumar, P.K., Mukhopadhyay, M. (2019). Compressive sensing based two-dimensional DOA estimation using L-shaped array in a hostile environment. *Traitement du Signal*, 36(6): 529-538. <https://doi.org/10.18280/ts.360608>
- [7] Quan, T.M., Nguyen-Duc, T., Jeong, W.K. (2018). Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE Transactions on Medical Imaging*, 37(6): 1488-1497. <https://doi.org/10.1109/TMI.2018.2820120>
- [8] Tang, V.H., Bouzerdoum, A., Phung, S.L. (2020). Compressive radar imaging of stationary indoor targets with low-rank plus jointly sparse and total variation regularizations. *IEEE Transactions on Image Processing*, 29: 4598-4613. <https://doi.org/10.1109/TIP.2020.2973819>
- [9] Wen, B., Ravishankar, S., Pfister, L., Bresler, Y. (2020). Transform learning for magnetic resonance image reconstruction: From model-based learning to building neural networks. *IEEE Signal Processing Magazine*, 37(1): 41-53. <https://doi.org/10.1109/MSP.2019.2951469>
- [10] Jacob, M., Ye, J.C., Ying, L., Doneva, M. (2020). Computational MRI: Compressive sensing and beyond [from the guest editors]. *IEEE Signal Processing Magazine*, 37(1): 21-23. <https://doi.org/10.1109/MSP.2019.2953993>
- [11] Févotte, C., Kowalski, M. (2018). Estimation with low-rank time-frequency synthesis models. *IEEE Transactions on Signal Processing*, 66(15): 4121-4132. <https://doi.org/10.1109/TSP.2018.2844159>
- [12] Gerstoft, P., Mecklenbräuker, C.F., Seong, W., Bianco, M. (2018). Introduction to compressive sensing in acoustics. *Journal of the Acoustical Society of America*, 143(6): 3731-3736. <https://doi.org/10.1121/1.5043089>
- [13] Liu, X., Gönültaş, E., Studer, C. (2018). Analog-to-feature (A2F) conversion for audio-event classification. In 2018 26th European Signal Processing Conference (EUSIPCO), pp. 2275-2279. <https://doi.org/10.23919/EUSIPCO.2018.8553060>
- [14] León-López, K.M., Carreno, L.V.G., Fuentes, H.A. (2018). Temporal colored coded aperture design in compressive spectral video sensing. *IEEE Transactions on Image Processing*, 28(1): 253-264. <https://doi.org/10.1109/TIP.2018.2867171>
- [15] Wang, S., Yu, L., Xiang, S. (2019). A low complexity compressed sensing-based codec for consumer depth video sensors. *IEEE Transactions on Consumer Electronics*, 65(4): 434-443. <https://doi.org/10.1109/TCE.2019.2929586>
- [16] Guo, J., Song, B., Yu, F.R., Chi, Y., Yuen, C. (2019). Fast video frame correlation analysis for vehicular networks by using CVS-CNN. *IEEE Transactions on Vehicular Technology*, 68(7): 6286-6292. <https://doi.org/10.1109/TVT.2019.2916726>
- [17] Van Luong, H., Deligiannis, N., Seiler, J., Forchhammer, S., Kaup, A. (2018). Compressive Online Robust Principal Component Analysis via n - ℓ_1 Minimization. *IEEE Transactions on Image Processing*, 27(9): 4314-4329. <https://doi.org/10.1109/TIP.2018.2831915>
- [18] Zhou, Q., Ke, J., Lam, E.Y. (2019). Near-infrared temporal compressive imaging for video. *Optics Letters*, 44(7): 1702-1705. <https://doi.org/10.1364/OL.44.001702>
- [19] Iliadis, M., Spinoulas, L., Katsaggelos, A.K. (2018). Deep fully-connected networks for video compressive sensing. *Digital Signal Processing*, 72: 9-18. <https://doi.org/10.1016/j.dsp.2017.09.010>
- [20] Baraniuk, R., Davenport, M., DeVore, R., Wakin, M. (2008). A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3): 253-263. <https://doi.org/10.1007/s00365-007-9003-x>
- [21] Randall, P.A. (2009). Sparse recovery via convex optimization (Doctoral dissertation, California Institute of Technology). <https://doi.org/10.7907/3Z65-A925>
- [22] Candès, E.J., Recht, B. (2009). Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6): 717. <https://doi.org/10.1007/s10208-009-9045-5>
- [23] Blumensath, T., Davies, M.E. (2008). Gradient pursuits. *IEEE Transactions on Signal Processing*, 56(6): 2370-2382. <https://doi.org/10.1109/TSP.2007.916124>
- [24] Tropp, J.A., Gilbert, A.C., Strauss, M.J. (2006). Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Processing*, 86(3): 572-588. <https://doi.org/10.1016/j.sigpro.2005.05.030>
- [25] Mallat, S.G., Zhang, Z. (1993). Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12): 3397-3415. <https://doi.org/10.1109/78.258082>
- [26] Tropp, J.A., Gilbert, A.C. (2007). Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 53(12): 4655-4666. <https://doi.org/10.1109/TIT.2007.909108>
- [27] Needell, D., Vershynin, R. (2010). Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit. *IEEE Journal of Selected Topics in Signal Processing*, 4(2): 310-316. <https://doi.org/10.1109/JSTSP.2010.2042412>
- [28] Needell, D., Tropp, J.A. (2009). CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3): 301-321. <https://doi.org/10.1016/j.acha.2008.07.002>
- [29] Dai, W., Milenkovic, O. (2009). Subspace pursuit for compressive sensing signal reconstruction. *IEEE Transactions on Information Theory*, 55(5): 2230-2249. <https://doi.org/10.1109/TIT.2009.2016006>
- [30] Donoho, D.L. (2001). Sparse components of images and optimal atomic decompositions. *Constructive Approximation*, 17(3): 353-382. <https://doi.org/10.1007/s003650010032>
- [31] Gilbert, A.C., Tropp, J.A. (2005). Applications of sparse approximation in communications. In *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.* 1000-1004. <https://doi.org/10.1109/ISIT.2005.1523488>
- [32] Rao, B.D. (1998). Signal processing with the sparseness constraint. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, 3: 1861-1864. <https://doi.org/10.1109/ICASSP.1998.681826>
- [33] T, Berkin, Cilingiroglu, et al. Dictionary-based image reconstruction for superresolution in integrated circuit

- imaging. *Optics Express*, 2015.
- [34] Reyes, N.R., Candeas, P.V., Ferreras, F.L. (2010). Wavelet-based approach for transient modeling with application to parametric audio coding. *Digital Signal Processing*, 20(1): 123-132. <https://doi.org/10.1016/j.dsp.2009.04.011>
- [35] Lin, J.L., Hwang, W.L., Pei, S.C. (2006, May). Video compression based on orthonormal matching pursuits. In 2006 IEEE International Symposium on Circuits and Systems.
- [36] Malioutov D., Cetin M., Willsky A.S. (2005). A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Transactions on Signal Processing*, 53(8): 3010-3022.
- [37] Qian, S., Chen, D. (1994). Signal representation using adaptive normalized Gaussian functions. *Signal Processing*, 36(1): 1-11. [https://doi.org/10.1016/0165-1684\(94\)90174-0](https://doi.org/10.1016/0165-1684(94)90174-0)
- [38] Lee, C.T., Yang, Y.H., Chen, H.H. (2012). Multipitch estimation of piano music by exemplar-based sparse representation. *IEEE Transactions on Multimedia*, 14(3): 608-618.
- [39] Chen, S.S., Donoho, D.L., Saunders, M.A. (2001). Atomic decomposition by basis pursuit. *SIAM Review*, 43(1): 129-159. <https://doi.org/10.1137/S003614450037906X>
- [40] Starck, J.L., Candès, E.J., Donoho, D.L. (2002). The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11(6): 670-684. <https://doi.org/10.1109/TIP.2002.1014998>
- [41] Giryes, R., Needell, D. (2015). Greedy signal space methods for incoherence and beyond. *Applied and Computational Harmonic Analysis*, 39(1): 1-20. <https://doi.org/10.1016/j.acha.2014.07.004>
- [42] Kalluri, M., Jiang, M., Ling, N., Zheng, J., Zhang, P. (2018). Adaptive RD optimal sparse coding with quantization for image compression. *IEEE Transactions on Multimedia*, 21(1): 39-50. <https://doi.org/10.1109/TMM.2018.2847228>
- [43] Zhang, R., Li, S. (2015). Optimal D-RIP bounds in compressed sensing. *Acta Mathematica Sinica, English Series*, 31(5): 755-766. <https://doi.org/10.1007/s10114-015-4234-4>