

Research Article

Improved Stochastic Gradient Matching Pursuit Algorithm Based on the Soft-Thresholds Selection

Zhao Liqun  and Hu Yunfeng

College of Information Engineering, Northeast Electric Power University, Jilin 132012, China

Correspondence should be addressed to Zhao Liqun; zhao_liquan@163.com

Received 19 February 2018; Accepted 11 July 2018; Published 24 September 2018

Academic Editor: Panajotis Agathoklis

Copyright © 2018 Zhao Liqun and Hu Yunfeng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The preliminary atom set exits redundant atoms in the stochastic gradient matching pursuit algorithm, which affects the accuracy of the signal reconstruction and increases the computational complexity. To overcome the problem, an improved method is proposed. Firstly, a limited soft-threshold selection strategy is used to select the new atoms from the preliminary atom set, to reduce the redundancy of the preliminary atom set. Secondly, before finding the least squares solution of the residual, it is determined whether the number of columns of the measurement matrix is smaller than the number of rows. If the condition is satisfied, the least squares solution is calculated; otherwise, the loop is exited. Finally, if the length of the candidate atomic index set is less than the sparsity level, the current candidate atom index set is the support atom set. If the condition is not satisfied, the support atom index set is determined by the least squares solution. Simulation results indicate that the proposed method is better than other methods in terms of the reconstruction probability and shorter running time than the stochastic gradient matching pursuit algorithm.

1. Introduction

Compressed sensing (CS) [1, 2] theory has three core problems: sparse representation of signals, and design of the measurement matrix, and design of reconstruction algorithms. The reconstruction algorithm is directly related to the accuracy of the recovery signal and the convergence rate of the algorithm, which determines whether the theory is feasible. The recovery algorithm is described as the recovery of the high-dimensional original signal from the low-dimensional measurement vectors. At present, many reconstruction algorithms have been proposed to recover the signal, which include convex optimization methods, combinational methods, and greedy pursuit methods. Convex optimization methods approximate the signal by transforming nonconvex problems into convex ones, which include basis pursuit (BP) [3], the gradient projection for sparse reconstruction (GPSR) [4] algorithm, iterative threshold (IT) [5], interior-point method [6], Bergman iteration (BT) [7], and total variation (TV) [8]. Although

convex optimization algorithms have fewer observations, their higher complexity is not suitable for practical applications.

Combinational methods include Fourier sampling [9,10], chaining pursuit (CP) [11], heavy hitters on steroids pursuit (HHSP) [12]. Combinational methods are low in complexity; however, the accuracy of the recovery signal is not as good as convex optimization algorithms. Greedy algorithms have many advantages, such as simple structure, fast convergence, and low complexity, becoming the first choice for the recovery algorithm.

To date, many greedy pursuit algorithms have been proposed, including the first, the matching pursuit (MP) [13] algorithm. Based on this algorithm, the orthogonal matching pursuit (OMP) [14] algorithm was proposed to optimize MP via orthogonalization of the atoms of the support set. However, the OMP algorithm only selects one of the atoms of the support set at each round of iteration, and the efficiency is lower. Successors include the regularized OMP (ROMP) [15], subspace pursuit (SP) [16] algorithm,

compressive sampling matching pursuit (CoSaMP) [17, 18] algorithm, and stagewise OMP (StOMP) [19] algorithm. The ROMP algorithm realizes the effective selection of the atom via the regularized rule, to improve the speed of the OMP. The SP and CoSaMP algorithms are similar. Both are proposed with the backtracking strategy. The difference between the SP and CoSaMP is the number of atoms that are selected from measurement matrix to compose preliminary atomic set in each round of iteration. The SP selects s atoms, while the CoSaMP selects $2s$ atoms. The StOMP algorithm selects multiple atoms or columns of the measurement matrix in each round of iteration via a threshold parameter. The greedy algorithms mentioned above all require the sparsity information of the signal, and the choice of the type of atoms is inflexible, which affects the convergence speed, robustness, and reconstruction performance of the algorithms.

Because the traditional greedy pursuit algorithm needs to compute the inverse matrix of the sensing matrix, this process requires a significant amount of computation time and storage space, resulting in lower reconstruction probability. In recent years, some research workers have proposed the fast version of the greedy algorithm, which avoids computation of the inverse matrix [20]. The GP algorithm uses the gradient idea of the unconstrained optimization method to replace the computation of the inverse matrix and to reduce the computational complexity and storage space of the traditional greedy algorithms [21]. However, the GP algorithm has a slow convergence and lower efficiency. In addition, the GP algorithm cannot solve the problem of large-scale data recovery. To improve those problems, the conjugate gradient pursuit (CGP) [22] algorithm and approximate conjugate gradient pursuit (ACGP) [23] algorithm are proposed, respectively. Although CGP and ACGP algorithms effectively reduce the computational complexity and storage space of traditional greedy algorithms, the reconstruction performance still needs to be improved. Based on the GP algorithm, the stagewise weak selection (SWGSP) [24] algorithm was proposed to improve the reconstruction performance and convergence speed of the GP algorithm, via introduction of the conjugate direction and weak selection strategy. Motivated by the stochastic gradient descent methods, the stochastic gradient matching pursuit (StoGradMP) algorithm [25] was recently proposed for the optimization problem with sparsity constraints. The atomic selection method of the fixed number in the StoGradMP algorithm (that is, selecting $2s$ atoms at the preliminary stage of each iteration) will lead to the redundant atoms of the preliminary atomic set. When joined with the candidate atomic set, this will reduce the accuracy of the least squares solution and the inaccuracy of the support atomic set estimation, which affects the precision of the signal reconstruction and increases the computational complexity of the algorithm. In this study, we use the limited soft-threshold selection strategy to realize the second selection of the preliminary atom set after the preliminary stage, which will improve the reconstruction accuracy. The combination with reliability verification conditions ensures the correctness and effectiveness of the proposed method.

2. Compressed Sensing Theory

The recent work in the area of the compressed sensing [26] demonstrated that it was possible to algorithmically recover sparse (and, more generally compressible) signals from incomplete observations. The simplest model is an n -dimensional signal x with a small number of nonzero entries under no noise conditions:

$$x \in R^n, \quad \|x\|_0 \leq s \leq n, \quad (1)$$

where $\|x\|_0$ and s are the number of nonzero entries and sparsity level of the signal, respectively. Such signals are called s -sparse. Any signal in R^n can be represented in terms of $n \times 1$ vectors $\{\psi_i\}_{i=1}^n$. For simplicity, we assume that the basis is orthonormal. Forming the $n \times n$ basis matrix $\Psi = [\psi_1, \psi_2, \dots, \psi_n]$ by stacking the vectors $\{\psi_i\}$ as columns, we can express any signal x as

$$x = \sum_{i=1}^n \theta_i \psi_i = \Psi \theta, \quad (2)$$

where $\theta = \Psi^T x$ is the $n \times 1$ column vector of projection coefficients, $\theta_i = \langle x, \psi_i \rangle = \psi_i^T x$ is the projection coefficient. Obviously, x and θ are the equivalent representations of the same signals in the different domains. That is, x and θ are the signals in the time domain and Ψ domain, respectively.

When x is s -sparse, let $\Psi = I$, I is the unit matrix, and $x = \theta$. We use a measurement matrix $\Phi \in R^{m \times n}$ ($m \leq n$) to make a linear measurement of the projection coefficient vector θ , obtaining an observation vector $y \in R^{m \times 1}$, expressed as

$$y = \Phi x, \quad (3)$$

where Equation (3) is a linear projection of the original signals x on Φ . Note that the measurement process is nonadaptive; that is, Φ does not depend in any way on the signal x . Clearly, the dimension of y is much lower than the dimension of x . That is, this problem is an under-determined problem; Equation (3) has infinitely many solutions. It is difficult to recover the projection coefficient vector θ directly from the observation vector y . However, the original signal x is s -sparse and Φ satisfies certain conditions in Equation (3); x can be recovered by solving the l_0 -minimization problem:

$$\begin{aligned} \min_x \quad & \|x\|_0, \\ \text{s.t.} \quad & \Phi x = y, \end{aligned} \quad (4)$$

where $\|x\|_0$ is the l_0 -norm of the vector, representing the number of nonzero entries. Candes and Tao demonstrated that if the s -sparse signal x is to be accurately recovered, the number of measurements m (or the dimension of observation vector y) must satisfy $m = O(s \ln(n))$, and the measurement matrix must satisfy the restricted isometric property (RIP) [27, 28].

When x is not s -sparse in the time domain, the signal recovery process cannot be directly used in the Equation (4). The signal x can be sparse representations on the sparse basis matrix Ψ . Combining Equations (2) and (3), we obtain

$$y = \Phi\Psi^T x = \tilde{\Phi}\theta, \quad (5)$$

where $\tilde{\Phi} = \Phi\Psi^T \in R^{m \times n}$ is the sensing matrix [29]. According to [30], the equivalent condition of the RIP is that the measurement matrix Φ is not correlated with the sparse basis matrix Ψ . Note that if the sensing matrix $\tilde{\Phi}$ also satisfies the RIP, the recovery signal (or projection coefficient) θ can be obtained by solving an optimal l_0 -norm problem similar to Equation (4):

$$\begin{aligned} \min_{\theta} \quad & \|\theta\|_0, \\ \text{s.t.} \quad & \tilde{\Phi}\theta = y. \end{aligned} \quad (6)$$

From Equation (5), we see that Ψ is fixed, so that Φ also satisfies the RIP. The measurement matrix Φ must meet certain conditions. It is shown in [31] that when the measurement matrix is a Gaussian random matrix, the sensing matrix $\tilde{\Phi}$ can satisfy the RIP condition with a larger probability.

However in most scenarios, the signal contains noise. In this case, the measurement process can be expressed as

$$y = \tilde{\Phi}\theta + e, \quad (7)$$

where the $m \times n$ matrix $\tilde{\Phi}$ is the sensing matrix and e is the m -dimensional noise vector. θ is the s -sparse signal in the Ψ domain. In this study, for simplicity, we assume that the signal x is s -sparse, that is, $x = \theta$ and $\Phi = \tilde{\Phi}$. Therefore, Equation (7) can be written as $y = \Phi x + e$, and we minimize the following formula to recover x :

$$\begin{aligned} \min_{x \in R^n} \quad & \frac{1}{2m} \|y - \Phi x\|_2^2, \\ \text{s.t.} \quad & \|x\|_0 \leq s, \end{aligned} \quad (8)$$

where s controls the sparsity of the solution to Equation (9).

To analyze (8), we combine Equation (2), where θ_i are the projection coefficients of signal x . x is considered sparse with respect to Ψ if s is relatively small compared to the ambient dimension n . Therefore, we can express the optimization (9) in the form as follows:

$$\begin{aligned} \min_x \quad & \frac{1}{M} \sum_{i=1}^M f_i(x) \\ & \underbrace{\hspace{10em}}_{F(x)}, \\ \text{s.t.} \quad & \|x\|_{0,\Psi} \leq s, \end{aligned} \quad (9)$$

where $f_i(x)$, $x \in R^n$, is a smooth function that is nonconvex; $\|x\|_{0,\Psi}$ is defined as the norm that captures the sparsity level of x . In particular, $\|x\|_{0,\Psi}$ is the smallest number of atoms in Ψ such that x can be represented as

$$\|x\|_{0,\Psi} = \min_x \left\{ s : x = \sum_{i \in T} \theta_i \psi_i \text{ with } |T| = s \right\}. \quad (10)$$

For sparse signal recovery, the set Ψ consists of n basic vectors, each of size n in the Euclidean space. This problem can be seen as a special case of Equation (10), with $f_i(x) = (y_i - \langle \varphi_i, x \rangle)^2$ and $M = m$. We decompose the vector $y^{m \times 1}$

into nonoverlapping vectors y_{b_i} with a size of b and denote Φ_{b_i} as $\Phi_{b_i \times n}$. $\Phi_{b_i \times n}$ is the submatrix of measurement matrix Φ . According to Equation (9), the objective function is $F(x) = (1/2m) \|y - \Phi x\|_2^2$. We then denote $F(x)$ as

$$F(x) = \frac{1}{M} \sum_{i=1}^M \frac{1}{2b} \|y_{b_i} - \Phi_{b_i} x\|_2^2 = \frac{1}{M} \sum_{i=1}^M f_i(x), \quad (11)$$

where $M = m/b$, which is a positive integer. We treat each function $f_i(x)$ as $f_i(x) = (1/2b) \|y_{b_i} - \Phi_{b_i} x\|_2^2$, and each function $f_i(x)$ represents a collection (or block) of observations of size b . Here, we split the function $F(x)$ into multiple subfunctions $f_i(x)$ or block the measurement matrix Φ into the submatrix Φ_{b_i} , which will be beneficial for the calculation of the gradient.

3. StoGradMP Algorithm

CoSaMP [17] is a popular algorithm for recovering a sparse signal from its linear measurements. The idea of the CoSaMP algorithm is generalized to prove the GradMP algorithm that solves a broader class of sparsity-constrained problems. In this section, we describe the stochastic version of the GradMP, the StoGradMP [25] algorithm, where at each iteration, only the evaluation of the gradient of a function f_i is required.

The StoGradMP algorithm is described in Algorithm 1, which consists of the following steps in each iteration:

Randomize. The measurement matrix Φ is subject to random block processing, to determine the region of the submatrix $\Phi_{b_i \times n}$. Then, according to Equation (11), calculates the subfunction $f_{i_k}(x_k)$.

Proxy. Compute the gradient of $f_{i_k}(x_k)$. Here, the gradient is an $n \times 1$ column vector.

Identify. Select the column indexes of the submatrix $\Phi_{b_i \times n}$ corresponding to the maximum $2s$ components in the gradient vector, forming a preliminary index set S_k .

Merge. Merge the preliminary index set S_k and the support index set F_{k-1} of the previous iteration to form a candidate index set C_k .

Estimation. The estimation of signal b_k by a sub-optimization method is determined, which is the least squares method. Generally, b_k is the transition signal.

Prune. Select the column indexes of the measurement matrix Φ corresponding to the maximum s components in the signal estimation vector b_k that forms a support index set F .

Update. Update the signal estimation $\hat{x} = x_k$.

Check. When the residual is less than the tolerance error of the proposed method iteration, stop the iteration. If the loop index k is greater than the maximum number of iterations s , the proposed method ends and the approximation of signal $\hat{x} = x_k$ is output. If the iteration ends, the condition is not satisfied. Otherwise, continue the iteration until the halting condition is met.

Input:

Sparsity level s
 Sensing matrix $\Phi^{m \times n}$
 Observation vector $y^{m \times 1}$
 Block size b
 Tolerance used to exit loop tol
 Maximum number of iterations maxIter

Initialization parameter:

$\hat{x} = 0$ {initialize signal approximation}
 $k = 0$ {loop index}
 $\text{done} = 0$ {while loop flag}
 $S_0 = \emptyset$ {empty preliminary index set}
 $C_0 = \emptyset$ {empty candidate index set}
 $F_0 = \emptyset$ {empty support index set}
 $M = \text{floor}(m/b)$ {number of block}

While ($\sim \text{done}$)

$k = k + 1$

(1) Randomize

$ii = \text{ceil}(\text{rand} * M)$

$\text{block} = b * (ii - 1) + 1 : b * ii$

$f_{i_k}(x_k) = (1/2b) \|y_{b_{i_k}} - \Phi_{b_{i_k}} x\|_2^2$

(2) Computation of gradient

$r_k = \nabla f_{i_k}(x_k) = -2 * \Phi_{b_{i_k}}^T (y_{b_{i_k}} - \Phi_{b_{i_k}} x_{k-1})$

(3) Identify the large $2s$ components

$S_k = \max(|r_k|, 2s)$

(4) Merge to update candidate index set

$C_k = F_{k-1} \cup S_k$

(5) Signal estimation by the least square method

$b_k = \Phi_{C_k}^+ y$

(6) Prune to obtain current support index set

$F = \max(|b_k|, s)$

(7) Update

$x_k = b_{kF}$

$F_k = F$

(8) Check the iteration condition

If $(\|y - \Phi x_k\|_2 \leq \text{tol} \text{ or } k \geq \text{maxIter})$

$\text{done} = 1$ quit iteration

end

end

Output: $\hat{x} = x_k$ (s -sparse approximation of signal x)

ALGORITHM 1: StoGradMP algorithm.

4. Improved StoGradMP Algorithm

The StoGradMP algorithm selects $2s$ atoms in each iteration, where s is a fixed number. The choice of atoms is inflexible and will increase the redundancy of the preliminary atom set. This affects the reconstruction and computational speed. To solve this problem, we use the limited soft selection strategy to select atoms from the preliminary atom set.

Firstly, according to Equation (11), we randomly block the measurement matrix to obtain a stochastic block matrix (or submatrix) $\Phi_{\text{block} \times n}$, which can be expressed as

$$\begin{aligned} ii &= \text{ceil}(\text{rand} * M), \\ \text{block} &= b * (ii - 1) + 1 : b * ii, \end{aligned} \quad (12)$$

where M is the number of the block measurement matrix, $M = \text{floor}(m/b)$, and $b = \min(s, m)$ is the number of rows of the block matrix. $\text{floor}(\cdot)$ and $\text{rand}(\cdot)$ generate a maximum

integer smaller than m/b and uniformly distributed pseudorandom numbers, respectively. $\text{ceil}(\cdot)$ obtains a minimum integer greater than $\text{rand} * M$. block is the row index of the measurement matrix Φ . From Equation (12), we know that the area of the block matrix is randomly determined. For simplicity, we express $\Phi_{\text{block} \times n}$ and y_{block} as Φ_{b_i} and y_{b_i} , respectively. Next, we compute the subfunction $f_i(x)$ as follows:

$$f_{i_k}(x_k) = \frac{1}{2b} \|y_{b_{i_k}} - \Phi_{b_{i_k}} x_{k-1}\|_2^2, \quad (13)$$

where k is the number of iterations, $i \in [M]$. Φ_{b_i} is the i -th submatrix of the measurement matrix Φ , which is stochastically determined and of size $b_i \times n$. The subfunctions $f_i(x)$ are also stochastically determined, and $f_i(x) \in F(x)$. Here, $f_i(x)$ is the i -th subfunction of $F(x)$.

After determining the area of submatrix Φ_{b_i} and the subfunction $f_i(x)$, we calculate the gradient of the subfunction, which is expressed as

$$r_k = \nabla f_{i_k}(x_k) = -2 * \Phi_{b_{i_k}}^T (y_{b_{i_k}} - \Phi_{b_{i_k}} x_{k-1}), \quad (14)$$

where r_k is the gradient of the subfunction $f_i(x)$ at the k -th iteration and the gradient is an $n \times 1$ column vector. $\nabla(\cdot)$ and $\Phi_{b_i}^T$ represent the derivative of the subfunction $f_i(x)$ and the transpose matrix of submeasurement matrix Φ_{b_i} , respectively. x_{k-1} is the approximation of sparse signal x at the $k-1$ -th iteration.

In the atomic preliminary selection stage of the stochastic algorithm, we obtain the gradient vector of the function through the derivation of randomly determined subfunctions, and then select the $2s$ largest gradient values from the gradient vector r_k , thereby determining the column index of the measurement matrix corresponding to the values of the gradient vector, and form a preliminary atom index set S_k as

$$S_k = \max(|r_k|, 2s), \quad (15)$$

where $|r_k|$ is the absolute value of the gradient vector r_k , k is the number of iterations, and $\max(|r_k|, 2s)$ is the atomic (or column) index of the measurement matrix corresponding to the maximal value from r_k that forms a preliminary atom index set S_k .

After the preliminary atomic stage, the preliminary atomic set exits the redundant atoms, which will reduce the accuracy of the least square solutions of signal and the inaccuracy of the support atomic set estimation. This will eventually affect the precision of signal reconstruction, and increase the computational complexity of the algorithm. To improve the reconstruction accuracy, we use the limited soft-threshold selection strategy to realize the second selection of the preliminary atom set.

In the atomic index set S_k , the atoms corresponding to the index are expressed as Γ , $\Gamma = [\Gamma_1, \Gamma_2, \dots, \Gamma_{2s}]$. We select the atoms that are larger than the threshold in the atomic set Γ , to form a new atomic set. The index of the new atomic set forms the new atomic index set S_k^* . This process is described as

$$\begin{aligned} \text{sigma} &= \max(|r_k|), \\ S_k^* &= \text{Find}(\|r_k\|_{2s} \geq t * \text{sigma}), \end{aligned} \quad (16)$$

where sigma is the maximum gradient value at the k -th iteration. $\|r_k\|_{2s}$ represents the largest $2s$ gradient values from the absolute value of the gradient vector, which corresponds to the atomic index set S_k . t is the threshold, whose value range is generally $[0.1 \sim 1.0]$. It should be noted that if the size of the threshold t is greater than 1, the soft-thresholds selection strategy fails. That is to say that the selection of the preliminary atomic index set S_k^* cannot be completed. Meanwhile, if the value of the threshold t is too small, it will lead to the redundant atoms cannot be effectively eliminated in the preliminary atomic set. $t * \text{sigma}$ is the soft-thresholds selection strategy. $\text{Find}(\cdot)$ searches the corresponding index that satisfies the soft-threshold condition and forms a new atomic index set S_k^* .

After the soft-threshold selection strategy is completed, we merge the new atomic set $\Phi_{S_k^*}$ and the support set $\Phi_{F_{k-1}}$ to update the candidate atom set Φ_{C_k} , which can be expressed as

$$\begin{aligned} C_k &= F_{k-1} \cup S_k^*, \\ \Phi_{C_k} &= \Phi_{F_{k-1}} \cup \Phi_{S_k^*}, \end{aligned} \quad (17)$$

where C_k , F_{k-1} , and S_k^* represent the candidate atomic index set, the support atomic index set, and the preliminary atomic index set, respectively. Φ_{C_k} is the candidate atomic set at the k -th iteration. $\Phi_{F_{k-1}}$ is the support atomic set at the previous iteration. $\Phi_{S_k^*}$ is the new atomic set corresponding to atomic index set S_k^* at the k -th iteration.

Before solving the suboptimization method, we must ensure that the number of rows is greater than the number of columns in the candidate atomic matrix Φ_{C_k} ; that is, Φ_{C_k} is a full column-rank matrix. Therefore, we provide the first reliability verification condition; that is, if

$$\text{length}(C_k) \leq m, \quad (18)$$

then

$$A_k = \Phi_{C_k}, \quad (19)$$

where A_k is the candidate atomic set (or matrix) at the k -th iteration. If the condition is not satisfied, namely, the number of rows is smaller than the number of columns, the matrix is not inverted. If this occurs, we exit the loop. Let $b_k = 0$. Next, solving the least squares solution of the sparse approximation signal \hat{x} , it can be expressed as

$$b_k = A_{C_k}^+ y, \quad (20)$$

where b_k is the estimation vector of the sparse signal x at the k -th iteration. This is known as the transition signal. $A_{C_k}^+$ is the inverse matrix of the candidate atomic set (or matrix), and y is the observing vector.

Since the soft-threshold selection strategy is used to complete the second selection of the preliminary atomic set, the size of the candidate atomic index set C_k may be less than s . Therefore, to ensure the correctness and effectiveness of

the proposed method, we provide the second reliability verification condition, that is, if

$$\text{length}(C_k) < s, \quad (21)$$

then

$$F = C_k, \quad (22)$$

where s is the size of sparsity of the signal. $\text{length}(C_k)$ is the size of the candidate atomic index set C_k . If

$$\text{length}(C_k) \geq s, \quad (23)$$

then

$$F = \max(|b_k|, s), \quad (24)$$

where $|b_k|$ is the absolute value of the transition signal at the k -th iteration. $\max(|b_k|, s)$ determines the atomic (or column) index of the measurement matrix corresponding to the maximum value from b_k and the former support atomic index set F .

Next, we update the approximation of sparse signal x , which is expressed as

$$x_k = b_{kF}, \quad (25)$$

where x_k is the approximation of the signal at the k -th iteration, and b_{kF} is the recovery signal corresponding to the support atomic index set F .

Finally, we check the iteration stopping condition. If

$$\|y - \Phi x_k\|_2 < \text{tol} \text{ or } k \geq \text{maxIter}, \quad (26)$$

then

$$\text{done} = 1, \quad (27)$$

where $y - \Phi x_k$ represents the residual at the k -th iteration and tol is the tolerance error of the algorithm iteration. maxIter is the maximum number of iteration, with special value $500 * M$ in this study. $\text{done} = 1$ represents the algorithm stops and outputs the signal approximation $\hat{x} = x_k$. If the iteration stopping criteria is not satisfied, then the iteration is continued until the condition is satisfied. The entire procedure is shown in Algorithm 2.

5. Results and Discussion

In this section, we used the signal with s -sparsity as the original signal. The measurement matrix is randomly generated with a Gaussian distribution. All performance is an average after running the simulation 100 times using a computer with a quadcore, 64-bit processor, and 4G memory.

In Figure 1, we compare the reconstruction probability of different sparsity of the proposed algorithm with different measurements in different threshold conditions. We set the sparsity level set as $s \in [8, 12, 16, 20]$. The threshold parameter set is $t \in [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]$. From Figures 1(a)–1(e), we can see that when the size of the threshold parameter is 0.1, 0.2, 0.3, 0.4, and 0.5, the reconstruction probability of the proposed method is very

Input:
 Sparsity level s
 Sensing matrix $\Phi^{m \times n}$
 Observation vector $y^{m \times 1}$
 Block size b
 Tolerance used to exit loop tol
 Maximum number of iterations maxIter

Initialize parameter:
 $\hat{x} = 0$ {initialize signal approximation}
 $k = 0$ {loop index}
 $\text{done} = 0$ {while loop flag}
 $S_0 = \emptyset$ {empty preliminary index set}
 $C_0 = \emptyset$ {empty candidate index set}
 $F_0 = \emptyset$ {empty support index set}
 $M = \text{floor}(m/b)$ {number of block}

While ($\sim \text{done}$)
 $k = k + 1$

(1) **Randomize**
 $ii = \text{ceil}(\text{rand} * M)$
 $\text{block} = b * (ii - 1) + 1 : b * ii$
 $f_{i_k}(x_k) = (1/2b) \|y_{b_{i_k}} - \Phi_{b_{i_k}} x_{k-1}\|_2^2$

(2) **Computation of gradient**
 $r_k = \nabla f_{i_k}(x_k) = -2 * \Phi_{b_{i_k}}^T (y_{b_{i_k}} - \Phi_{b_{i_k}} x_{k-1})$

(3) **Identify the large $2s$ components**
 $S_k = \text{max}(|r_k|, 2s)$

(4) **Soft-threshold selection strategy**
 $\text{sigma} = \text{max}(|r_k|)$
 $S_k^* = \text{Find}(\|r_k\|_{2s} \geq t * \text{sigma})$

(5) **Merge to update the candidate index set**
 $C_k = F_{k-1} \cup S_k^*$
Reliability verification conditions 1
If $\text{length}(C_k) \leq m$
 $A_k = \Phi_{C_k}$
else
 if $k == 1$
 $b_k = 0$
end
 break;
end

(6) **Estimation of signal by least square method**
 $b_k = A_{C_k}^+ y$

(7) **Prune to obtain current support index set**
Reliability verification conditions 2
If $(\text{length}(C_k) < s)$
 $F = C_k$
else
 $F = \text{max}(|b_k|, s)$
end

(8) **Update**
 $x_k = b_{kF}$
 $F_k = F$

(9) **Check the iteration stopping condition**
If $(\|y - \Phi x_k\|_2 < \text{tol} \text{ or } k \geq \text{maxIter})$
 $\text{done} = \text{quit iteration}$
end

end
Output: $\hat{x} = x_k$ (s -sparse approximation of signal x)

ALGORITHM 2: Proposed algorithm.

close, with almost no difference. In Figure 1(f), when the size of the threshold parameter is 0.6, the proposed method can complete signal reconstruction with less measurements, compared with other thresholds, when the size of sparsity is 16 and 20. In Figures 1(g)–1(j), we see that when the size of the threshold is 0.7, 0.8, 0.9, and 1.0, the proposed method requires more measurements to complete the signal reconstruction under the same sparsity level.

Figure 2 compares the reconstruction probability of different measurements of the proposed method with different thresholds under the same sparsity level. We set the sparsity set as $s = [8, 12, 16, 20]$. The threshold parameter set is consistent with the threshold parameter set in Figure 1. From Figure 2, we see that different threshold parameters have certain influences on the reconstruction probability of the signal for different sparsity levels. In Figures 2(a) and 2(b), we can see that when the sparsity levels are $s = 8$ and $s = 12$, the reconstruction probability of all thresholds conditions is very close, except for the thresholds 0.8, 0.9, and 1.0. Meanwhile, from Figures 2(c) and 2(d), we can see that when the sparsity levels are $s = 16$ or $s = 20$ and the sizes of the threshold are 0.5, 0.6, and 0.7, the reconstruction performance of the proposed method is better than the reconstruction probability of the other thresholds for different measurements. Therefore, from Figure (2), it demonstrates that the soft-threshold strategy is more advantageous for larger sparsity. In particular, when the sparsity level $s = 20$ and threshold $t = 0.6$, the reconstruction probability of the proposed method is better than the reconstruction probability of other threshold conditions. Based on the analysis of Figures 1 and 2, we can see that that when the threshold is $t = 0.6$, the proposed method has better performance. Therefore, in the following simulations, we set the threshold as 0.6.

In Figure 3, we compared the average runtime of different thresholds of the proposed algorithm with different measurements. From Figures 1 and 2, we see that the reconstruction probability of the proposed method is 100% when the number of the measurements is greater than or equal to 145. In particular, when the sparsity level is $s = 20$ and the threshold is 0.6, the reconstruction probability is better than the reconstruction probability of other thresholds. Therefore, we set the range of the number of the measurements as [145 200] in the simulation of Figure 3. From this, we see that the proposed algorithm with $t = 0.8$ has the shortest runtime, and the next shortest are the proposed algorithm with $t = 0.7$ and the proposed algorithm with $t = 0.6$. The proposed method with $t = 0.9$ has the longest runtime. This means that the selection of the threshold parameter t is important to the runtime of the proposed method.

From the aspects of reconstruction probability and runtime of the proposed method, we conclude that when the size of threshold parameter is $t = 0.6$ and the sparsity level is $s = 20$, reconstruction performance of the proposed method

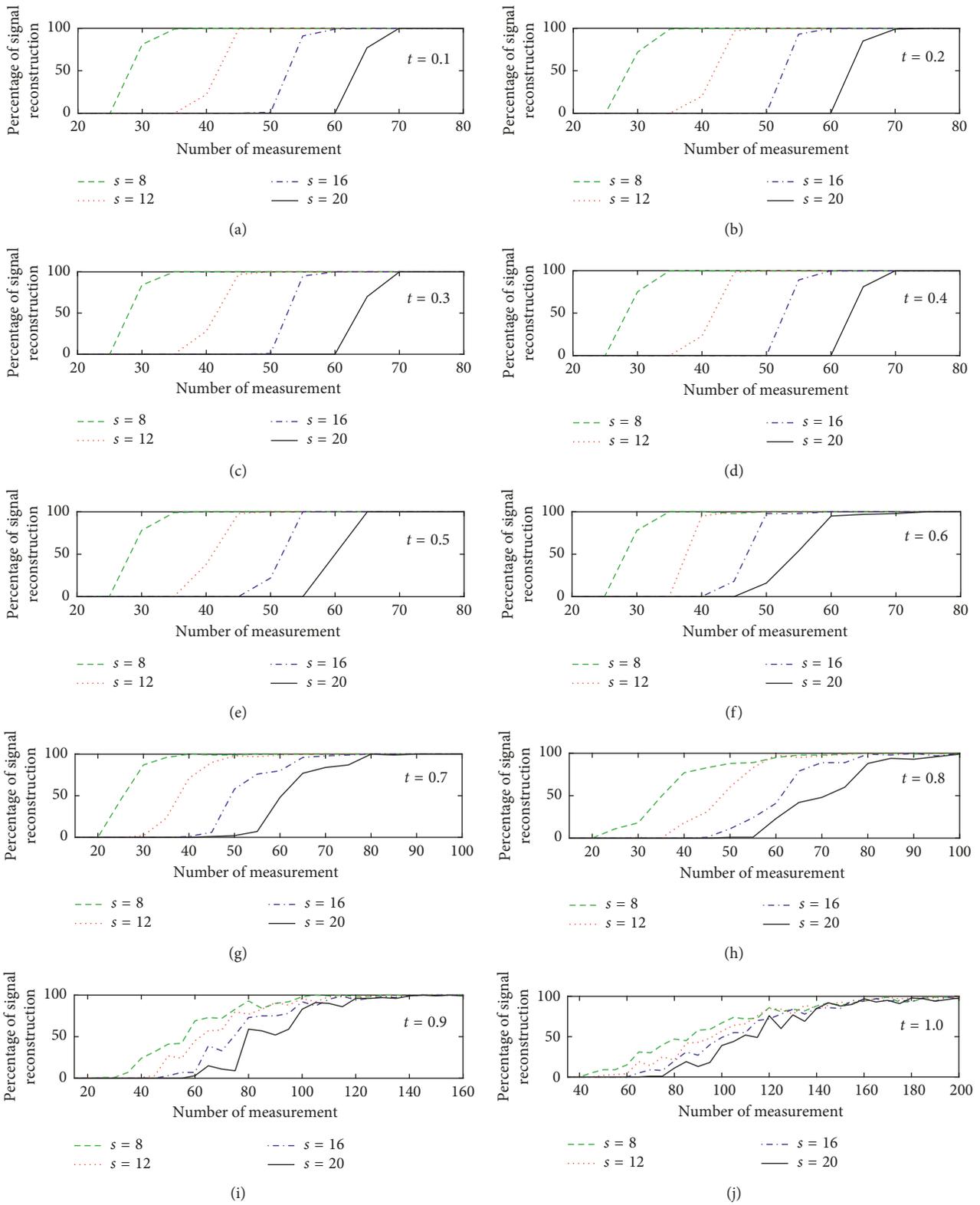


FIGURE 1: Reconstruction probability of different sparsity with different measurements in different threshold conditions ($n = 256$, $s \in [8, 12, 16, 20]$, $t \in [0.1 \ 1.0]$, Gaussian signal).

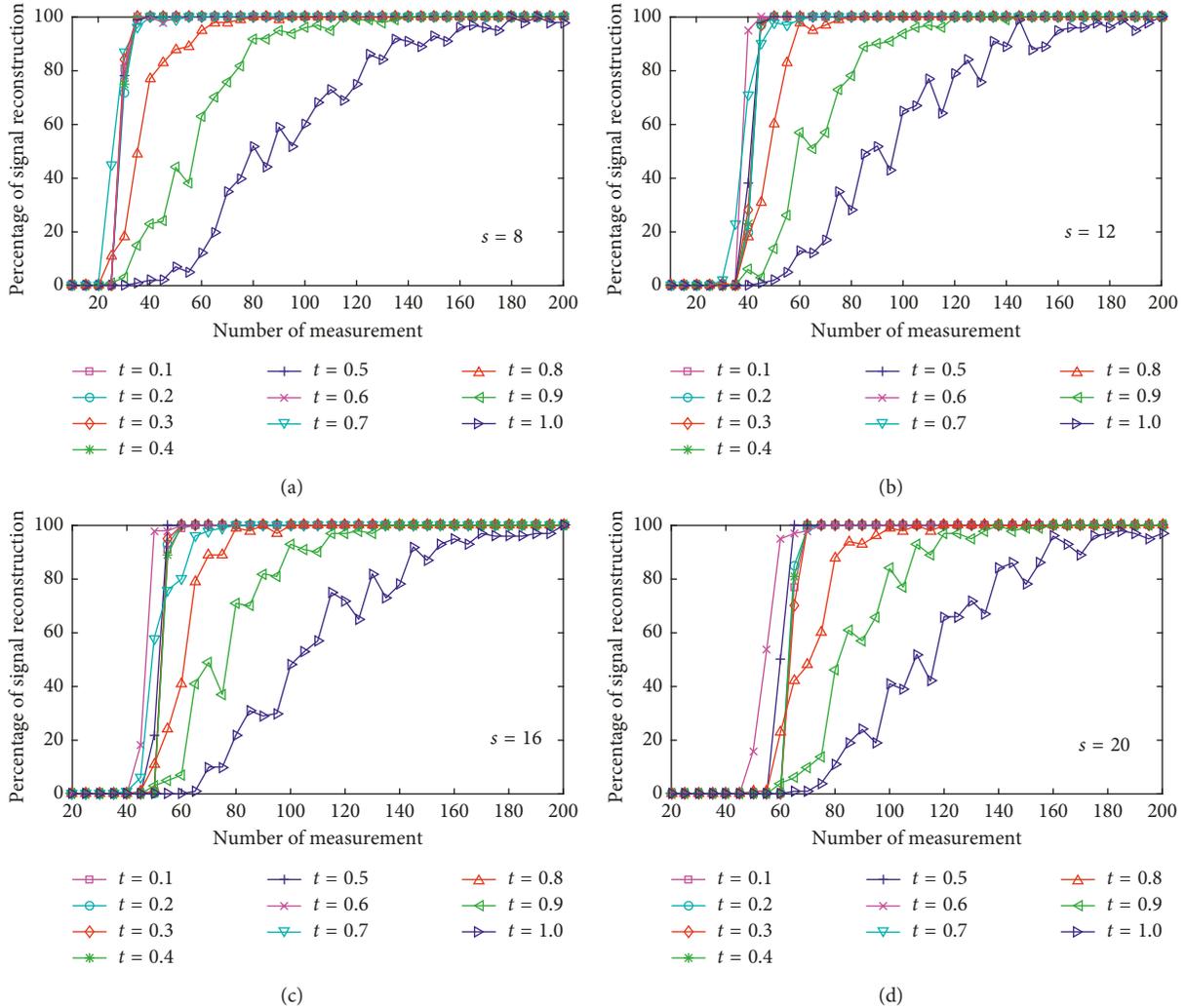


FIGURE 2: Reconstruction probability of different thresholds with different measurements in different sparsity conditions ($n = 256$, $t \in [0.1 \ 1.0]$, $s \in [8, 12, 16, 20]$, Gaussian signal).

is better than the reconstruction performance of the other threshold conditions for different measurements. Therefore, in the following without special instructions, the default threshold is 0.6 and the sparsity level is $s = 20$.

In Figure 4, we compared the reconstruction performance of the proposed algorithm in the single reconstruction. From Figure 4, we see that the recovery error is less than 5×10^{-15} , which is much less than the tolerance error of the proposed method iteration. This shows that the reconstruction of the proposed method is ideal.

In Figure 5, we compare the reconstruction probability of different sparsities of the proposed method and the StoGradMP in different measurements. We set the sparsity set as $s = [8, 12, 16, 20]$. The size of the threshold is 0.6. From Figure 5, we see that when the sparsity level is $s = 8$, the reconstruction probability of the proposed algorithm and the StoGradMP algorithm is nearly identical for all measurements. When the sparsity level is $s = 8$, the reconstruction probability of the StoGradMP algorithm and

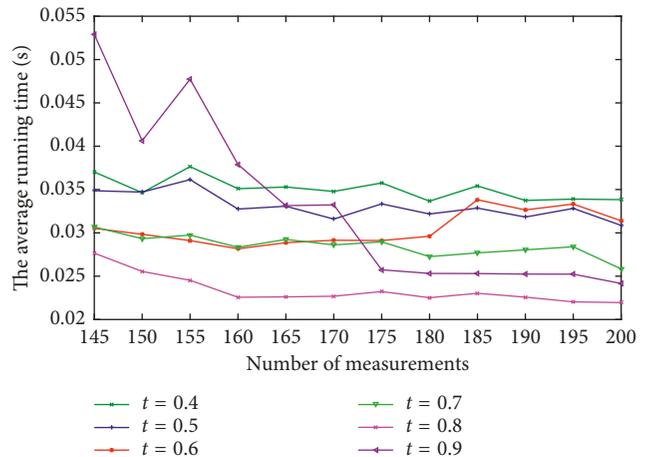


FIGURE 3: The average running time of different thresholds of the proposed algorithm with different measurements ($s = 20$, $n = 256$, $t \in [0.4 \ 0.9]$, Gaussian signal).

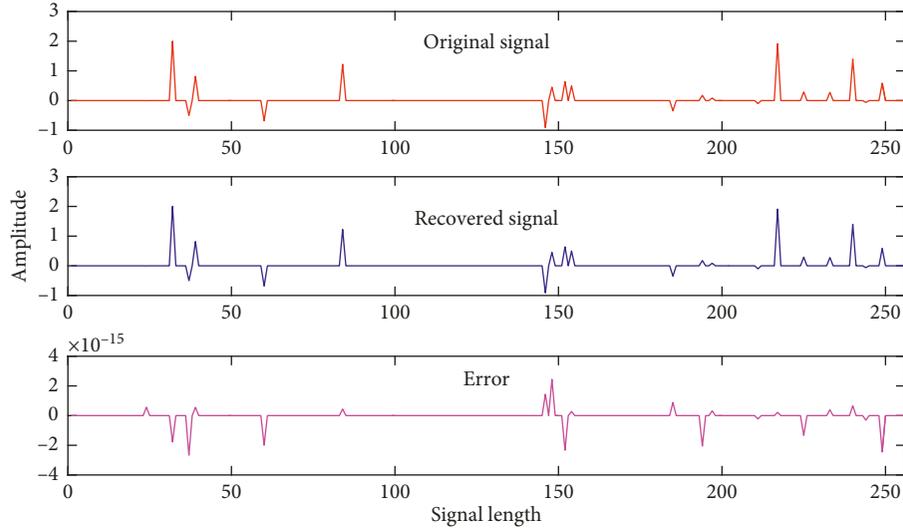


FIGURE 4: Comparison of original signal and recovery signal in the single reconstruction ($n = 256, s = 20, m = 100, t = 0.6$).

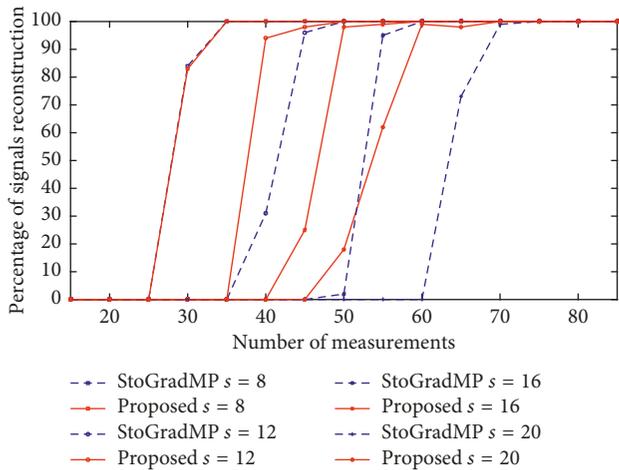


FIGURE 5: Reconstruction probability of different sparsities of the StoGradMP and proposed algorithms with different measurements ($n = 256, t = 0.6, s \in [8, 12, 16, 20]$, Gaussian signal).

the proposed algorithm is nearly identical. When $s = 12, 16$, and 20 , the reconstruction probability of the proposed method is higher than the StoGradMP algorithm under the same sparsity level. Among them, when the sparsity level $s = 20$, the difference between the reconstruction probability of the two algorithms is the largest.

In Figure 6, we compared the reconstruction probability of the proposed algorithm with StoGradMP, GradMP, and StoIHT algorithms for different measurements. From Figure 5, we see that the reconstruction probability is 0% when the sparsity is equal to 20 and the number of the measurements is smaller than or equal to 40. Therefore, we set the range of the measurements as $[40 \ 130]$ in the simulation of Figure 6. From Figure 6, we see that when $40 \leq m \leq 46$, the reconstruction probability of almost all algorithms is equal to 0%. For $46 \leq m \leq 64$, the reconstruction probability of the proposed algorithm ranges from 0 to 97% and has a higher reconstruction probability than other methods. When

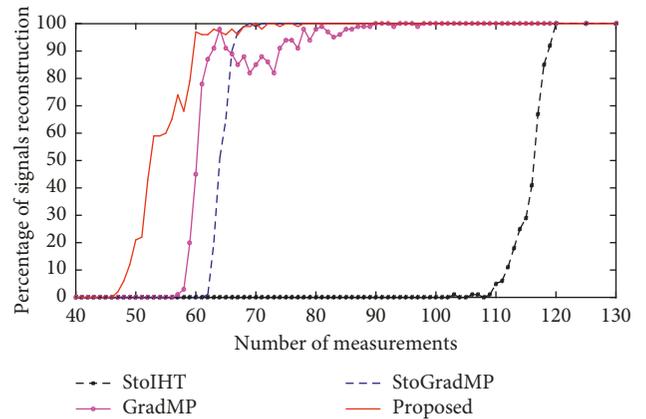


FIGURE 6: Reconstruction probability of different algorithms with different measurements ($n = 256, s = 20, t = 0.6$, Gaussian signal).

$56 \leq m \leq 64$, the reconstruction probability of t GradMP begins to increase, from approximately 0% to 97%, while the reconstruction probability ranges from 0 to 51%, and the reconstruction probability is 0%. When $64 \leq m \leq 73$, the reconstruction probability of the proposed method and StoGradMP ranges from 97–100% and 51–100%, respectively. However, the reconstruction probability of the GradMP algorithm declines from 98% to 82%. When $73 \leq m \leq 108$, the reconstruction probability of the GradMP algorithm increases from 82% to 100% and the reconstruction probability of the proposed method and StoGradMP algorithm is 100%, with almost no change. However, the reconstruction probability of the StoIHT algorithm is still 0%. When $108 \leq m \leq 120$, the reconstruction probability of the StoIHT is increased from 0% to 100%. When $120 \leq m$, all algorithms can complete the reconstruction.

In Figure 7, we compare the average runtime of different algorithms for different measurements. From Figure 6, we see that the reconstruction probability of all recovery

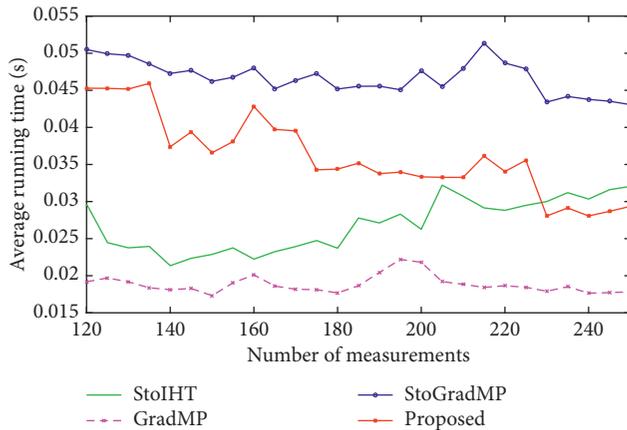


FIGURE 7: The average running time of different algorithms with different measurements ($n = 256$, $s = 20$, $t = 0.6$, Gaussian signal).

algorithms is 100% when the sparsity is equal to 20 and the number of measurements is greater than or equal to 120. Therefore, we set the range of the number of the measurements as [120 250] in the simulation of Figure 7, and the sparsity level is equal to 20. We see that the GradMP algorithm has the lowest runtime, and the next lowest are the StoIHT and proposed algorithms. The StoGradMP algorithm has the highest running time. From Figures 6 and 7, we see that although the runtime of the proposed method is much more than the GradMP and StoIHT algorithms, its reconstruction probability is much better than the other two algorithms. This means that the proposed algorithm has lower complexity than other algorithms, except for the GradMP and StoIHT algorithms in different measurements.

Based on the above analysis, the proposed method with threshold $t = 0.6$ has better reconstruction performance for different measurements than others and lower computational complexity than StoGradMP.

6. Conclusion

In this study, an improved reconstruction method was proposed. The proposed method utilizes the limited soft-threshold selection strategy to select the most relevant atoms from the preliminary atom set, which could reduce the redundancy of the preliminary atoms set and improve the accuracy of the support atom set estimation, thereby improving the reconstruction precision of the signal and reducing the computational complexity of the algorithm. In addition, the combination with reliability verification conditions ensured the correctness and effectiveness of the proposed method. The simulation results proved that the proposed method has better performance than other recovery methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61271115) and the Foundation of Jilin Educational Committee (JJKH20180336KJ). The authors would also like to thank Yanfei Jia for proof reading of this paper.

References

- [1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2004.
- [2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *Siam Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [4] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 586–597, 2007.
- [5] I. Daubechies, M. Defrise, and C. D. Mol, "An iterative threshold algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2003.
- [6] K. Koh, S. J. Kim, and S. Boyd, "An interior-point method for large-scale l_1 -regularized logistic regression," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2008.
- [7] S. Osher, Y. Mao, B. Dong, and Y. Wotao, "Fast linearized bregman iteration for compressive sensing and sparse denoising," *Mathematics of Computation*, vol. 8, no. 1, pp. 93–111, 2008.
- [8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," in *Proceedings of Eleventh International Conference of the Centre for Nonlinear Studies on Experimental Mathematics*, pp. 259–268, Los Alamos, NM, USA, 1992.
- [9] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-optimal sparse Fourier representations via sampling," in *Proceedings of ACM Symposium on Theory of Computing*, pp. 152–161, Montréal, Québec, Canada, May 2002.
- [10] A. C. Gilbert and M. Strauss, "Improved time bounds for near-optimal sparse Fourier representations," in *Proceedings of SPIE Wavelets XI*, pp. 1–15, Bellingham, WA, USA, September 2005.
- [11] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "Algorithmic linear dimension reduction in the l_1 norm for sparse vectors," in *Proceedings of 44th Annual Allerton Conference on Communication, Control, and Computing*, pp. 1–27, Allerton, IL, USA, September 2006.
- [12] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "One sketch for all: fast algorithms for compressed sensing," in *Proceedings of 39th Annual ACM Symposium on Theory of Computing*, pp. 237–246, San Diego, CA, USA, June 2007.

- [13] S. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [14] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [15] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, 2010.
- [16] D. Wei and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.
- [17] D. Needell and J. A. Tropp, "CoSaMP: iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.
- [18] M. A. Davenport, D. Needell, and M. B. Wakin, "Signal space CoSaMP for sparse recovery with redundant dictionaries," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 6820–6829, 2012.
- [19] D. L. Donoho, Y. Tsaig, I. Drori, and J. L. Starck, "Sparse solution of underdetermined systems of linear equations by stage-wise orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1094–1121, 2012.
- [20] A. Sancetta, "Greedy algorithms for prediction," *Bernoulli*, vol. 22, no. 2, pp. 1227–1277, 2016.
- [21] T. Blumensath and M. E. Davies, "Gradient pursuits," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2370–2382, 2008.
- [22] T. Blumensath and M. E. Davies, "Gradient pursuit for non-linear sparse signal modeling," in *Proceedings of 16th European Signal Processing Conference*, pp. 1–5, Lausanne, Switzerland, 2008.
- [23] T. Blumensath and M. E. Davies, "Fast greedy algorithms for large sparse inverse problem," pp. 1–5, 2012, see.ed.ac.uk.
- [24] T. Blumensath and M. E. Davies, "Stage-wise weak gradient pursuits," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4333–4346, 2009.
- [25] N. Nguyen, D. Needell, and T. Woolf, "Linear convergence of stochastic iterative greedy algorithms with sparse constraints," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 6869–6895, 2017.
- [26] E. J. Candès, "Compressive sampling," *Marta Sanz-Solé*, vol. 17, no. 4, pp. 1433–1452, 2006.
- [27] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [28] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [29] R. G. Baraniuk, "A lecture on compressive sensing," *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 1–9, 2007.
- [30] R. Baraniuk, "Compressive sensing," in *Proceedings of 42nd Annual Conference on Information Sciences and Systems*, pp. 1–9, Princeton, NJ, USA, March 2008.
- [31] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2005.



Hindawi

Submit your manuscripts at
www.hindawi.com

