Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **6**, November **2018**

A Multi-stage and Variable Step Size Sparsity Adaptive Generalized Orthogonal Matching Pursuit Method

Liquan Zhao[1], Yunfeng Hu[1], Yulong Liu[2], Changyin Liu[3]

[1]College of Information and Communication Engineering
Northeast Electric power University
No.169 Changchun Road, Chuanying District, Jilin City, China
[2]Nanning power supply Bureau
Guangxi power grid corporation
No.2 Zhongjian Road,Qingxiu District, Nanning City, China
[3] Qitaihe power supply station
Heilongjiang power grid corporation
No.35 Shanhu Road, Qitaihe City, China
zhao_liquan@163.com

Received March, 2017; Revised June, 2018

Abstract. *The generalized orthogonal matching pursuit method supposes that the sparsity is known, however, the sparsity is unknown in many practices. To overcome the problem, an improved method that can reconstruct signal without the prior sparsity information is proposed. The proposed method divided the signal reconstruction into two stages. The first stage will be stopped until the residual equal or less than the chosen threshold, and the size of the current estimated support set is used as an initial valued of adaptive parameter. In the second stage, a backtracking step is executed. We separately use the size of current adaptive parameter and updated adaptive parameter that is the sum of current adaptive parameter and the number of chosen indices, as the size of estimated support set, and compare their corresponding residual. If the residual of former is larger than the latter, the updated adaptive parameter is used as the new adaptive parameter; else we do not update the adaptive parameter. The simulation results demonstrated that the proposed method could reconstruct signal without requiring prior information of the sparsity. This method had a better recovery performance than the original algorithm..*
Keywords: Generalized Orthogonal Matching Pursuit; Reconstruction performance; Sparsity information; Threshold.

---

1. **Introduction.** In recent years, we have seen significant interests and research progress in the area of compressive sensing (CS) [1], which can surpass the limits of the Nyquist sampling rate to exactly reconstruct signal from a small number of random projections of a sparse signal which contain enough information for exact signal reconstruction. CS differs from the traditional Nyquist sampling theory and includes three procedures: sparse representation, non-related linear measurement, and signal reconstruction. The reconstruction algorithm aims to recover signals accurately from the measurements, and this step is one of most important parts of CS.

The challenge of compressed sensing is in solving the nonlinear optimization problem which is NP hard. Many reconstruction algorithms have been proposed to solve the problem and obtain the original sparse signal from measurements. Two major classes of reconstruction algorithms are l_1-minimization and greedy pursuit algorithms.

Common l_1-minimization approaches include basis pursuit (BP) [2], Gradient projection for sparse reconstruction (GPSR) [3], iterative threshold (IT) [4], and other algorithms. Those algorithms proposes good performance in solving a convex minimization problem, but they have a higher computational complexity.

Greedy algorithms have excellent performance and small cost in recovering sparse signals from compressed measurements. A greedy algorithm proposed early was the matching pursuit algorithm (MP) [5] which uses greedy heuristics to select the basis that spans the space for non-zero elements. Building on the MP algorithm, the orthogonal matching pursuit algorithm (OMP) [6] was proposed to optimize the MP via orthogonalization of the estimate support set. The OMP has become a well-known greedy algorithm with wide application. The regularized orthogonal matching pursuit algorithm (ROMP) [7] was developed to refine the selected columns of the measurement matrix with a regularized rule to improve the speed of OMP. The stage wise orthogonal matching pursuit (StOMP) [8] selects multiple columns in each iteration via a presupposed threshold. The subspace pursuit (SP) [9] and compressive sampling matching pursuit (CoSaMP) [10] proposed similar improvement methods. Both of these algorithms were proposed with the idea of backtracking, and the difference is that SP selects k columns from the sensing matrix in each iteration, while CoSaMP selects 2k. The generalized orthogonal matching pursuit (GOMP) was proposed by Wang [11, 12]. The algorithm selects S($S \leq K$) columns in each iteration. The generalized OMP (GOMP) has received increasing attention in recent years, because the method can enhance the recovery performance of OMP. Several papers have been published on the analysis of the theoretical performance of GOMP [11, 12, 13, 14]. However, all the above greedy algorithms require the sparsity K as prior information which may not be available in practical applications. The sparsity adaptive greedy algorithm represented by the sparsity adaptive matching pursuit (SAMP) [15] to overcome this drawback. Those sparsity adaptive greedy algorithms can reconstruct the signals without knowing the sparsity.

2. **GOMP algorithm.** The GOMP is a variation of OMP algorithm. Compared to OMP, which selects only one column in each iteration, GOMP changes the number of columns that are selected in each iteration to improve the computational efficiency and recovery performance.

In kth iteration, GOMP firstly computes the correlation between the columns of the sensing matrix Ψ and the residual vector r^{k-1} by $\Phi' r^{k-1}$, and r^k denote residual vector in kth iteration. Then indices of the columns corresponding to S maximal correlation are chosen as the new elements of the estimated support set Λ^k in each iteration, where Λ^k is the estimated support set in kth iteration. Next to obtain $\hat{x}^k$ using the least square method (LS), where $\hat{x}^k$ is the new approximation of x in kth iteration. The residual $r^k \in R^M$ is revised by $\Phi_{\Lambda^k}\hat{x}_{\Lambda^k}$ from y:

$$r^k = y - \Phi_{\Lambda^k}\hat{x}_{\Lambda^k} \tag{1}$$

where $y = \Phi x$. These operations are repeated until either the iteration number reaches the maximum $k_{\max} = \min(K, M/S)$ where K is the sparsity of$x$, or the l_2 -norm of the residual falls below a threshold ε ($||r^k||_2 \leq \varepsilon$). Ease of understanding, we describe the GOMP algorithm in Table 1 according to [11, 12].

3. **Sparsity adaptive GOMP method.** The GOMP only modifies on the identification step of OMP. OMP selects only one column as candidate in each iteration but GOMP

Table 1. GOMP ALGORITHM

Input:	measurements $y \in R^M$, sensing $\Phi \in R^{M\times N}$, sparsity K, number of indices of columns for each selection $S(S \le K)$.
Initialize:	iteration count $k = 0$, residual vector $r^0 = y$, estimated support set $\Lambda^0 = \emptyset$.
While	$\|\|\mathrm{r}^{\mathrm{k}}\|\| > \varepsilon$ and $k < \min\{K, M/S\}$ **do** $k = k + 1$. **(Identification)** Select S largest entries (in magnitude) from $\Phi' r^{k-1}$. Then record the $\{\varphi(i)\}_{i=1,2,3......S}$ corresponding to the entries. **(Augmentation)** $\Lambda^k = \Lambda^{k-1} \cup \{\varphi(1), \varphi(2)......\varphi(S)\}$. **(Estimation of x_{Λ^k})** $\hat{x}_{\Lambda^k} = \arg\min_{\sup p(u)=\Lambda^k} \|y - \Phi u\|_2$. **(Residual Update)** $r^k = y - \Phi_{\Lambda^k}\hat{x}_{\Lambda^k}$.
End	
Output	The estimated support $\hat{\Lambda} = \underset{\mathrm{T}:\|\mathrm{T}\|=K}{\arg\min} \|\hat{x}_{\Lambda^k} - \hat{x}_{\mathrm{T}}\|_2$, and signal $\hat{x}_{\hat{\Lambda}} = \Phi^{\dagger}{}_{\hat{\Lambda}}y$.

selects S columns as candidates in each iteration. When S=1, GOMP is exactly the same as OMP. That is to say OMP is a special case (S=1) of GOMP. The GOMP made a simply modification to OMP, however, the promotion to OMP in computational efficiency and recovery performance is great. This is because GOMP increases the probability of selecting correct candidate by selecting more candidates. At the same time, GOMP may select more than one correct candidate in each iteration. That makes it own faster speed to find the correct support set and exactly reconstruct signal.

Although GOMP own excellent computational efficiency and recovery performance, it has its drawback. Firstly, the GOMP algorithm needs the sparsity as prior information to reconstruct signal. Table1 shows that the GOMP algorithm requires sparsity K of the signal to serve as the iteration stop condition. The S, chosen in the GOMP, also requires sparsity K to serve as the limiting condition. However, sparsity K is not always available in many practical applications. Besides, at the same time to select more correct candidates, GOMP also adds more error candidates into the estimation support set. Because the GOMP can not remove those error candidate out the estimation support set, those accumulated error candidates will remain in the estimation support set throughout the remainder of the reconstruction process. Those accumulated error candidates will make the size of estimation support set bigger than the real support set. This will further increase the cost of the algorithm. To overcome these two problems, we propose a novel sparsity adaptive GOMP method for signal recovery when the K is unknown, in this section.

The sparsity K is equal to the size of true support set. For the first problem, this new method uses the estimated size of support set as the substitute of sparsity K. So, the new method can exactly reconstruct signal using a suitable estimation support set and the problem how to obtain the sparsity K is converted to how to obtain a suitable estimation support set. For the second problem, this new method adopts backtracking to enhance the reliability of indices in the estimation support set. However, when the number of candidates in the estimation support set is far less than true support set, to execute the backtracking step and judge whether the size of the estimation support set is suitable is unnecessary. In order to avoid unnecessary computation cost, this new method adopts the divide and conquer principle and divide the reconstruction process into two steps using a specific threshold t.

In the first step, when the residual norm $\|r^k\|_2$ was larger than the specific threshold t, the method computes the correlation between the columns of the sensing matrix Φ and the residual vector r^{k-1} by $\Phi' r^{k-1}$. Then indices of the columns corresponding to S maximal correlation are chosen as the new reliable candidates and are added into the estimated support setΛ^k. Because the larger the correct candidates in the estimation support set, the smaller residual norm $\|r\|_2$ norm .When the residual $\left\|r^k\right\|_2$ norm is smaller than t , this suggested there are enough correct candidates in the estimation support set for current stage. At this point, the new method calculates the size of current estimation support set as an initial parameter P and entered the next stage.

In the second step, the new method then updates the candidates of the estimation support set using backtracking to reduce the number of error candidates. The backtracking would select L ($L = P$) largest elements from $|x_{\Lambda^k}|$. The indices are reported as they correspond to the elements. The Λ_k is updated by using those indices. The new method utilized the new estimation support set to generate a new residual r_{new}. If $\|r_{new}\|_2 < \left\|r^{k-1}\right\|_2$, we consider the L to be suitable. If $\|r_{new}\|_2 > \left\|r^{k-1}\right\|_2$, we consider the L to be too small for the current iteration and update L to $L = P + S \times step$. It is a dilatation step size with initial value $step$= 0. When $\|r_{new}\|_2 \geq \left\|r^{k-1}\right\|_2$, $step$=$step$+1. When L was considered too small, the method adopted a stage-wise approach that expanded the estimated support set stage by stage. When the new method completes the backtracking, it updates r^k=r_{new} and proceeds to the next iteration. This continued until the iteration number reached M ($k_{\max} = M$) or the l_2-norm of the residual fell below a thresholdε. The process of the proposed algorithm was expressed in table II.

The new method could reconstruct the signal without the sparsity K from the measurements. This showed that the new method achieved the sparsity adaptive function. The proposed method also had a better reconstruction performance than the GOMP because the backtracking was added. Experimental evidence demonstrated that our changes improved the performance.

Table 2. SPARSITY ADAPTIVE GOMP METHOD

Input:	measurements $y \in R^M$, sensing $\Phi \in R^{M \times N}$, sparsity K, number of indices of columns for each selection $S(S \leq K)$. threshold t.
Initialize:	iteration count $k = 0$, residual vector $r^0 = y$, estimated support set $\Lambda^0 = \emptyset$, dilatation step size $step = 1$.
While	$\|\|r^k\|\|_2 > \varepsilon$ and $k < M$ do $k = k + 1$;
First Stage:	1.Select S largest entries (in magnitude) from $\Phi' r^{k-1}$. Then the indices were recorded $\{\varphi(i)\}_{i=1,2,3......S}$ corresponding to the entries. 2. $\Lambda^k = \Lambda^{k-1} \cup \{\varphi(1), \varphi(2)......\varphi(S)\}$ 3. $x_{\Lambda^k} = \arg\min_{\sup p(u)=\Lambda^k} \|y - \Phi u\|_2$ 4. $r_{temp} = y - \Phi_{\Lambda^k} \hat{x}_{\Lambda^k}$ 5. If $\|r_{temp}\|_2 < t$, the algorithm goes into the second stage, else $r^k = r_{temp}$.
Second Stage:	1. When the new method enters the second stage in the first time, calculate size of current estimation support set Λ^f. $P =$ size (Λ^f), $L = P + S \times step$. Select L largest elements of$\|x_{\Lambda^k}\|$ .Then record the indices that correspond to the elements and renew the Λ_k with those indices. 2. $r_{new} = y - \Phi_{\Lambda^k} \hat{x}_{\Lambda^k}$ 3. If $\|r_{new}\|_2 > \left\|r^{k-1}\right\|_2$, $step = step + 1$, $r^k = r_{new}$.
End	
Output	The estimated support $\hat{\Lambda} = \arg\min \|\hat{x}_{\Lambda^k} - \hat{x}_{\mathrm{T}}\|_2$ and signal $\hat{x}_{\hat{\Lambda}} = \Phi^{\dagger}{}_{\hat{\Lambda}} y$.

4. **Simulation and Discussion.** The reconstruction performance of the new method with sparse signals was evaluated. In each trial, we generated a K-sparse vector $x \in R^N$ whose support was chosen at random. Additionally, we constructed a $M \times N$ sensing matrix Φ with entries drawn independently from a Gaussian distribution $N(0, 1/M)$. $t = 10$ was empirically chosen. To ensure compatibility with gOMP, the trials used the same s, Φ ($M = 128$, $N = 256$). The gOMP threshold $\varepsilon = 10^{-6}$ was adopted. We used MATLAB 7.0 with a quad-core 64-bit processor (Windows 7) for each algorithm. This was repeated 600 times. The probability of the exact reconstructions and the average running time for each K was recorded. The probability of the reconstruction and the running were selected as the two major criterions.

In Figure 1, we compared the probability of successful reconstruction of the proposed method(S=3) with the ROMP [7], OMP [6], SP [9], CoSaMP [10], StOMP [8], and gOMP [12](S=3) algorithms. The proposed method had the best reconstruction performance for probability of reconstruction. When K=45, our proposed method remained near 100% for reconstruction probability. When the reconstruction probability of our proposed method decreased with a rise in K, it remained the highest of the algorithms with the same K. We specially chooses S=3 for both the gOMP and the proposed method to reduce the differences. The trial result demonstrated that the proposed method was more effective than these mainstream algorithms.

The reason that the proposed algorithm is superior to other algorithms is because its particular backtracking method.Compared to these algorithms without backtracking that are OMP,ROMP,StOMP and gOMP,the backtracking of the proposed algorithm can reevaluate the reliability of candidates of the estimation support set according to their contribution to the estimated signal. Then the backtracking will remove those unreliable candidate out the estimation support set.This can reduce the number of error indices in the estimated support set then improve the signal reconstruction effect.It's worth nothing that CoSaMP and SP also adopt backtracking.Compared to CoSaMP and SP, the proposed algorithm adopts particular backtracking method.In each iteration, the SP algorithm adds K new candidates into the estimation support set , the CoSaMP algorithm adds 2K new candidates, but the proposed algorithm only adds $S(S \leq K)$ new candidates.It means the proposed algorithm has a higher accuracy in the identification of candidates than CoSaMP and SP. In order to ensure a higher computational efficiency, the proposed algorithm innovatively divide the reconstruction process into two steps. In first step,the proposed algorithm only selects S new candidates into the estimation support set.When the residual $\left\|r^k\right\|_2$ norm is smaller than t , this suggested there are enough correct candidates in the estimation support set for current stage.Then the proposed algorithm enter to the second step and backtracking to the candidates of the estimation support set.The particular backtracking method is compatible with computational efficiency and accuracy.

We compared the performances of the proposed method and the gOMP with smaller values for the parameter S(S=3,5,7) in Figures 2 and 3. In Figure 2, we compared the probability of reconstruction of the proposed method and the gOMP. We compared the two algorithms from K=35 to K=70. This was because the probability of reconstruction of both algorithms was 100% for K>10 and close to zero for K>70. The best performance was the proposed method, with S=5 and S=7. The two curves were near identical. The proposed method at S=3 has a slightly worse performance than the proposed method with S=5 and S=7. The performance of the gOMP at S=3 was the best. The performance of the gOMP at S=5 was worse, however, the performance of the gOMP at S=7 was the worse overall. The proposed method demonstrated obvious advantages in the probability of reconstruction.

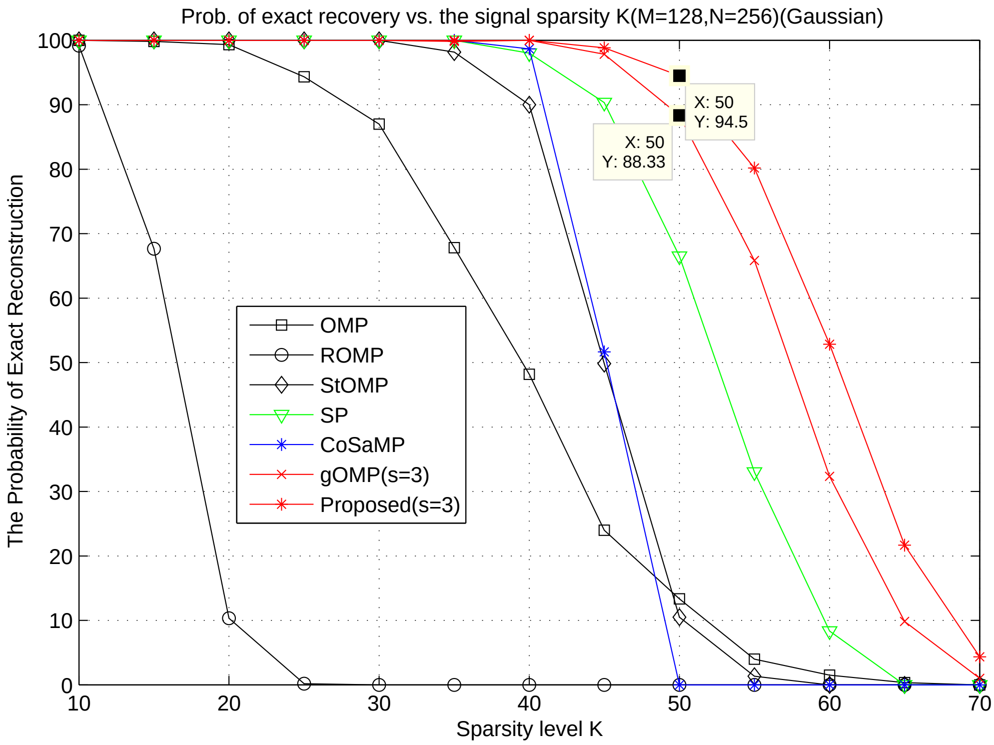

FIGURE 1. Reconstruction performance for the K-sparse Gaussian signal vector as a function of the sparsity K.

In this experiment, we compared our proposed method and the GOMP. We set two signals with different sizes ($N = 256, 512$) and three different $S(S = 3, 10, 20)$. In order to compare two algorithms with different S, we set the mean of two algorithms to be observed in the experiment.

In Figure 3, we compared the average exact running time of the proposed method and the gOMP. The effective running time should be computed in order to achieve a successful reconstruction. Figure 2 illustrates that the two algorithms ensured 100% exact reconstruct signal when K<40. We compared the average exact running time of the proposed method and the gOMP from K=10 to K=40. In Figure 3, the gOMP at S=5 had the shortest running time. The running time of the gOMP at S=7 was slightly higher than gOMP at S=5. The running time of the gOMP at S=3 was higher than the gOMP at S=7. The trend of the proposed method showed that when the S was larger, the running time was shorter. Figure 3 shows that the average exact running time of the proposed method was slightly higher than the gOMP. The proposed effective sparsity adaptive strategy found it difficult to avoid increasing the running time. Given the effective sparsity adaptive capacity and the increased probability of reconstruction, the increased the running time of the proposed method was not unacceptable. The proposed method remains competitive.

In Figure 4 and Figure 5, we compared the performances of the proposed method and the gOMP with the larger parameter values S(S=10,15,20).In Figure 4, we compared the probability of reconstruction of the proposed method and the gOMP. We compared the two algorithms from K=25 to K=70. Figure 4 shows that the larger S, the worse the recovering probability performances were for both the algorithms. This was because as the S increased, the probability of choosing an error increased. When a larger S was

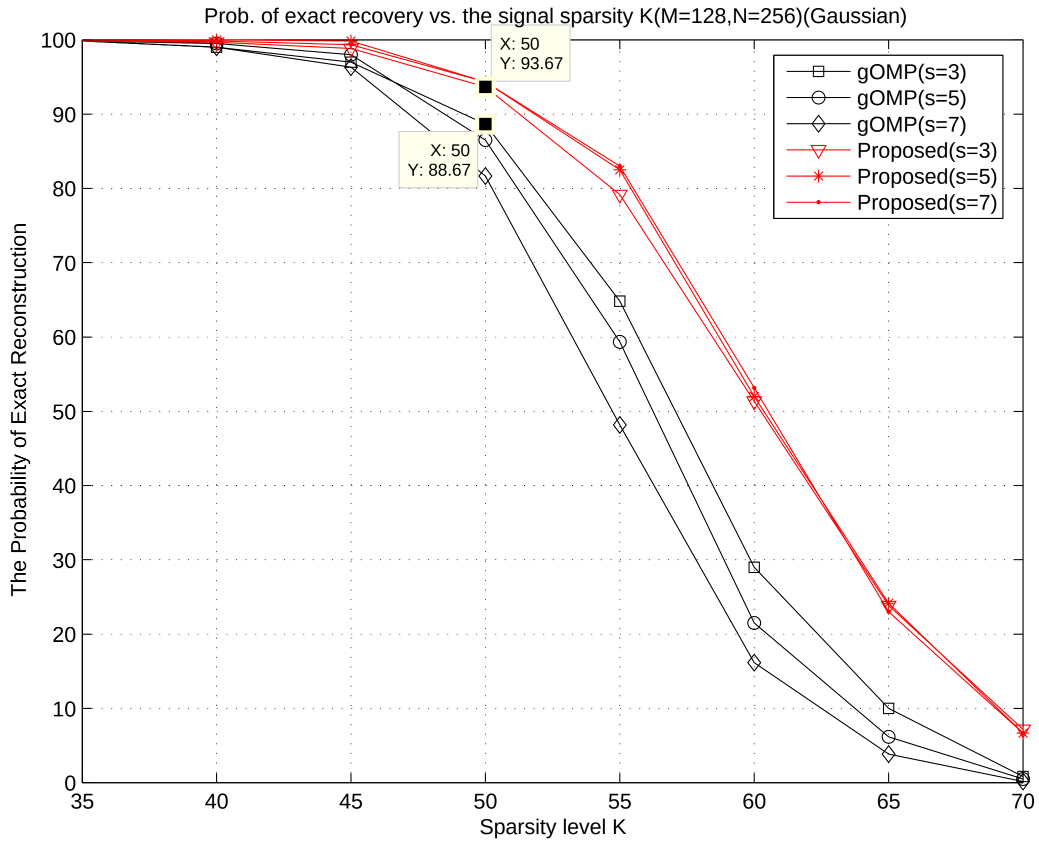

FIGURE 2. Reconstruction probability of the GOMP and the proposed method with a small S value.

chosen, those selected error candidates could reduce the probability of reconstruction. The backtracking in the sparsity adaptive strategy of the proposed method caused the proposed method to get rid of some unreliable candidates from the estimation support set. The performances of the proposed method was better than the gOMP when a larger S was chosen. Figure 4 shows that the proposed method (S=10) had the best performances for the recovering probability. The proposed method (S=10) had the best performance, followed by the proposed method(S=15), the proposed method(S=20), the gOMP(S=10), the gOMP(S=15), and the gOMP(S=20).

In Figure 5, we compared the average exact running times of the proposed method and the gOMP. We compared the two algorithms from K=10 to K=30. From Figure 5, we found that the gOMP had no absolute advantage in running time when a larger S was chosen. The gOMP(S=10) had the shortest running time. The proposed method(S=10) had the second shortest running time and the gOMP(S=20) had the longest running time. The running time of the gOMP(S=15), the proposed method(S=15), and the proposed method (S=20) fell in between the proposed method (S=10) and the gOMP(S=20). These three curves intertwine with each other and had similar performances.

These simulation results showed that the proposed method had a higher probability of reconstruction without the sparsity information. This makes the proposed method more suitable for practice application.

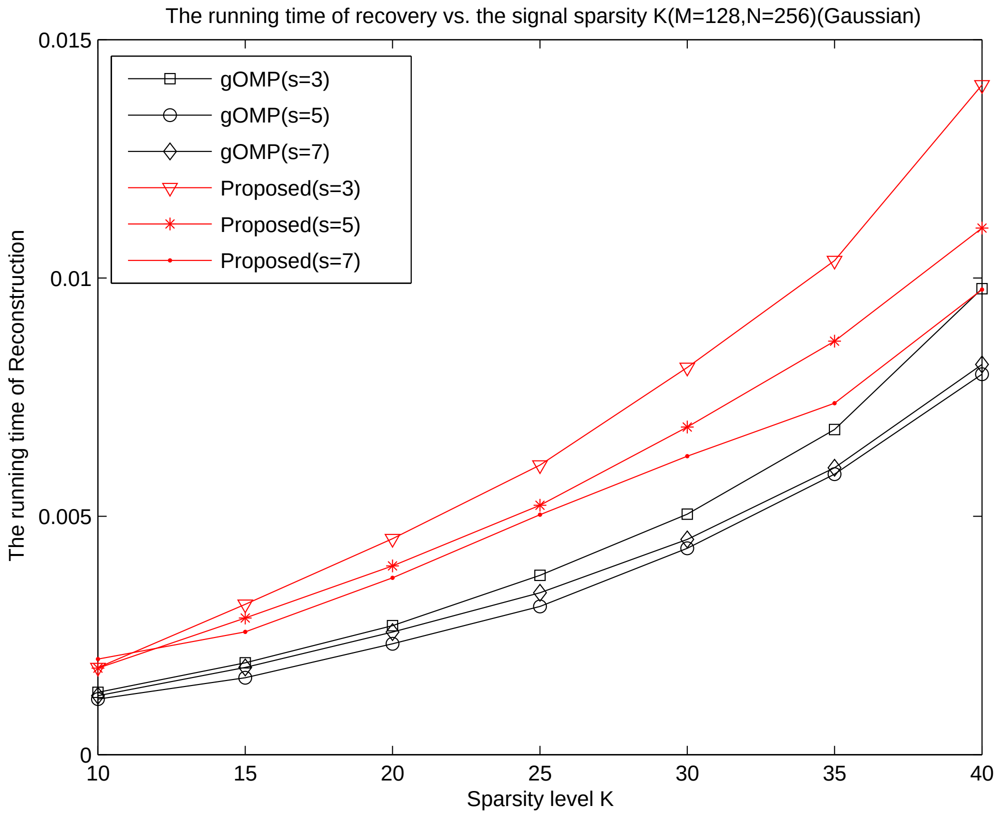

FIGURE 3. The average running time of the gOMP and the proposed method with a small S value

5. **Conclusion.** In this paper, we proposed a novel sparsity adaptive generalized orthogonal matching pursuit method to for reconstruction of the sparse signal when the sparsity was unknown. The proposed method initially chose S candidates in order to estimate the support set of each iteration and set a specifically chosen threshold to estimate the number of correct candidates. When the residual dropped below the threshold, the proposed method considered there was to be enough candidates selected for the estimation support set. The proposed method then executed a backtracking step at various parameters in order to generate a new residual. The proposed method used the varying parameter to control the size of estimation support set. If the norm of the new residual was smaller than the residual of the previous iteration, the proposed method considered the varying parameter to be suitable. The proposed method then executed the backtracking step with a varying parameter. If the norm of the new residual was larger than the residual of the previous iteration, the proposed method considered the varying parameter to be too small and the parameter was enlarged. The new method estimated the suitable size of estimation support set step by step in order to consider the sparsity adaptive capability. The simulation results showed that the proposed method considered the sparsity adaptive function and had a better performance than the gOMP in terms of the probability of exact reconstruction.

REFERENCES

[1] D. L. Donoho, Compressed Sensing, *IEEE Transactions on Information Theory*, vol.52, no.4, pp.1289–1306, 2006.

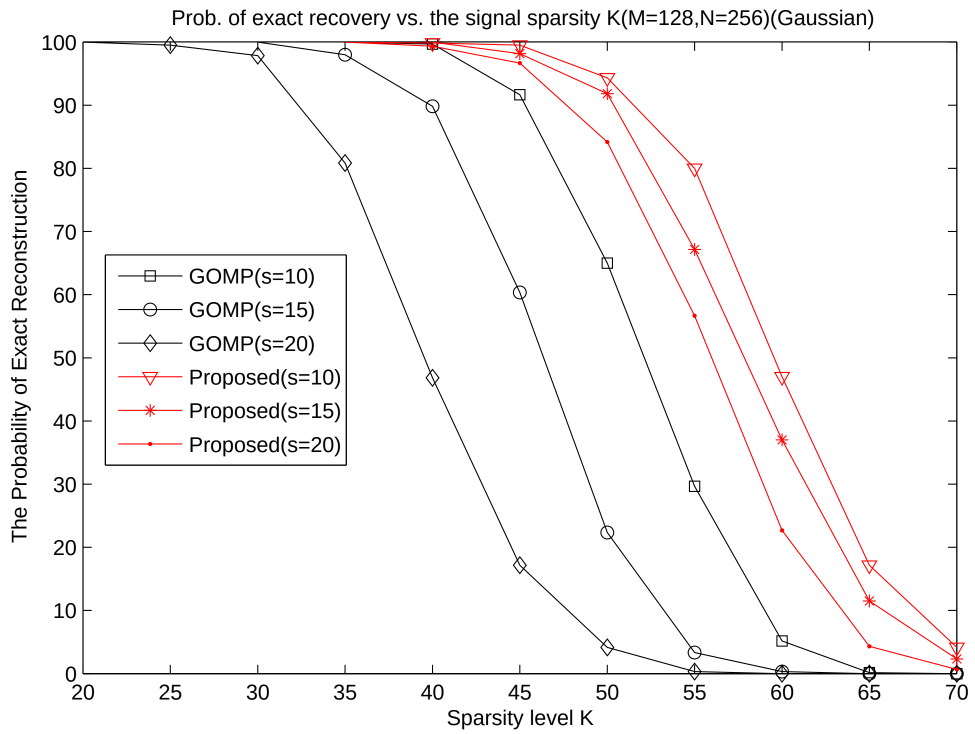

FIGURE 4. Reconstruction probability of the GOMP and the proposed method with a larger S value.

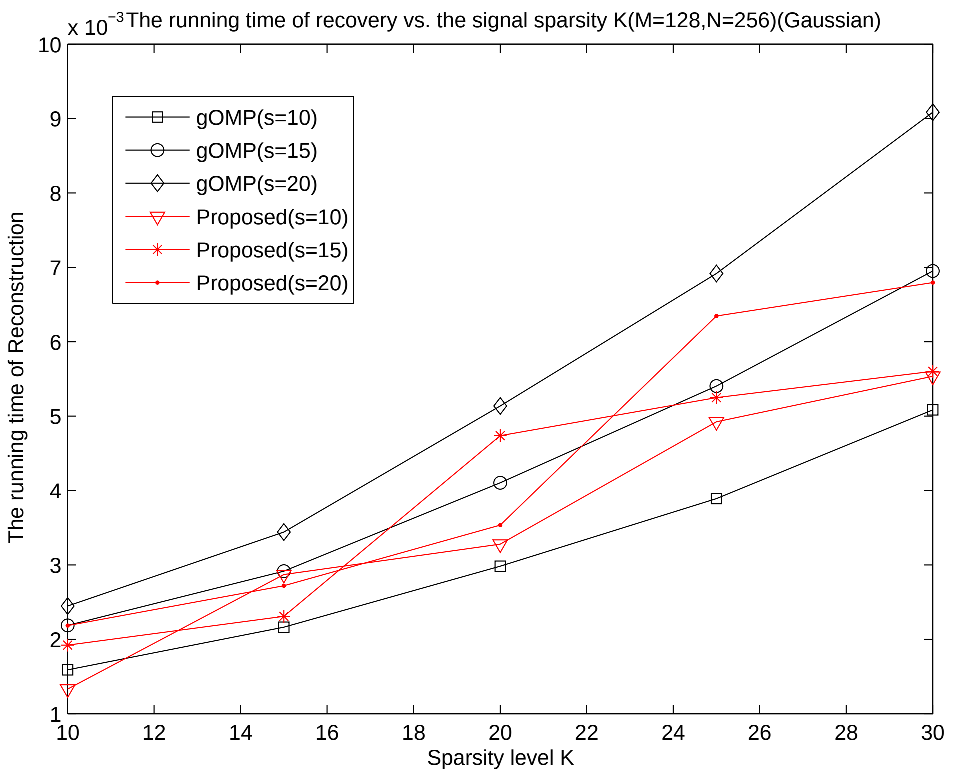

FIGURE 5. The average running time of the gOMP and the proposed method with a larger S value

[2] 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.

[3] 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.

[4] I. Daubechies, M. Defrise, and C. D. Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Communications on Pure and Applied Mathematics*, vol.57, no.11, pp.1413?1457, 2004.

[5] S. Mallat and, Z. Zhang, Matching pursuit with time-frequency dictionaries, *IEEE Transactions on Signal Processing*, vol.41, no.12, pp.3397–3415, 1993.

[6] 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.

[7] 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, 2007.

[8] D. L. Donoho, Y. Tsaig, and I. Drori, Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit, *IEEE Transactions on Information Theory*, vol.58, no.2, pp.1094–1121, 2012.

[9] W. Dai, O. R. Milenkovic, and A. Spanias, Subspace pursuit for compressive sensing signal reconstruction, *IEEE Transactions on Information Theory*, vol.55, no.5, pp.2230–2249, 2009.

[10] 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, 2008.

[11] J. Wang, S. Kwon, and B. Shim, Generalized orthogonal matching pursuit, *IEEE Transaction on Signal Processing* , vol.60, no.12, pp.6202–6216, 2012.

[12] J. Wang, S. Kwon, and P. Li, Recovery of Sparse Signals via Generalized Orthogonal Matching Pursuit: A New Analysis, *IEEE Transactions on Signal Processing* , vol.64, no.4, pp.1076–1089, 2016.

[13] S. Satpathi ,R. L. Das, and M. Chakraborty, Improving the Bound on the RIP Constant in Generalized Orthogonal Matching Pursuit, *IEEE Signal Processing Letters*, vol.20, no.20, pp.1074–1077, 2013.

[14] Y. Shen, B. Li, and Pan. W, et al, Analysis of generalised orthogonal matching pursuit using restricted isometry constant, *Electronics Letters*, vol.54, no.14, pp.1020–1022, 2014.

[15] Y. Shen, B. Li, and Z. Wu, Sufficient conditions for generalized Orthogonal Matching Pursuit in noisy case, *Signal Processing*, vol.108, no.108, pp.111–123, 2015.