IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 14, 2019, accepted February 2, 2019, date of publication February 12, 2019, date of current version March 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2898987

A Multiobjective Approach Based on Gaussian Mixture Clustering for Sparse Reconstruction

HUI LI[1], (Senior Member, IEEE), JIANYONG SUN[1], (Senior Member, IEEE), DEYU MENG[1], AND QINGFU ZHANG[2], (Fellow, IEEE)

[1]School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China
[2]Department of Computer Science, City University of Hong Kong, Hong Kong

Corresponding author: Hui Li (lihui10@mail.xjtu.edu.cn)

This work was supported by National Natural Science Foundation of China under Grant 61573279, Grant 61175063, Grant U1811461, Grant 11690011, and Grant 61721002. This work was also supported by a grant from ANR/RGC Joint Research Scheme sponsored by the Research Grants Council of the Hong Kong Special Administrative Region, China and France National Research Agency (Project No. A-CityU101/16).

ABSTRACT The application of multiobjective approaches for sparse reconstruction is a relatively new research topic in the area of compressive sensing. Unlike conventional iterative thresholding methods, multiobjective approaches attempt to find a set of solutions called Pareto front (PF) with different sparsity levels. The major focus of the existing sparse multiobjective approaches is to find the knee region of PF, where the K-sparse solution should reside. However, the strategies in these approaches for finding the knee region of PF are not very reliable due to the sensitivities on the setting of control parameters or noise levels. In this paper, we propose a new strategy based on Gaussian mixture models (GMMs) within a decomposition-based multiobjective framework for sparse reconstruction. The basic idea is to cluster the population found by a chain-based search procedure into two subsets via GMM. One of them with the small values of loss function should include the knee region. Our proposed algorithm was tested on a set of six artificial instance sets at four different noise levels. The experimental results showed that our proposed algorithm is superior to two existing sparse multiobjective approaches and one iterative thresholding algorithm.

INDEX TERMS Sparse optimization, iterative thresholding, multiobjective evolutionary approach, Gaussian mixture clustering.

I. INTRODUCTION

In the field of signal processing, signals are represented as a linear combination of some basis. It's been found that many coefficients in the signal representation are close or equal to zero. Compressive sensing [1], [2] aims to solve the following optimization problem for sparse signal reconstruction:

$$\min \|x\|_0 \quad \text{subject to } y = Ax + \epsilon \tag{1}$$

where

- $x \in \mathcal{R}^N$ is a N-dimensional vector, and $\|x\|_0$ (zero-norm) is defined as the number of nonzero components in x;
- $A \in \mathcal{R}^{M\times N}$ is a sensing matrix with $M \ll N$, and $y \in \mathcal{R}^M$ is an observation vector;
- $\epsilon \in R^M$ is a random vector following the normal distribution $\mathcal{N}(0, \sigma)$.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Chen.

In the noiseless case, i.e., $\epsilon = 0$, the optimal solution x^* of Problem (1) corresponds to the sparest member of the underdetermined system of linear equations $y = Ax$. If $\|x^*\|_0 = K$, then x^* is called a K-sparse solution. Optimization methods for sparse signal reconstruction can be classified into two categories - greedy heuristics and iterative thresholding methods. One well-known example for the former is the orthogonal match pursuit (OMP) [3], which builds a partial solution component by component in a greedy way. In contrast, iterative thresholding methods solve Problem (1) by minimizing the so-called regularization problem, which is stated as:

$$\min_{x\in\mathcal{R}^N} \|y - Ax\|^2 + \lambda\|x\|_q \tag{2}$$

where $q = 0$, 1 or $q \in (0, 1)$, and λ is a positive regularization parameter. The representative iterative thresholding methods include the iterative hard thresholding method ($q = 0$) [4], the iterative soft thresholding method ($q = 1$) [5] and the iterative half thresholding method ($q = 1/2$) [6].

The conversion of a constrained optimization into a bi-objective multiobjective optimization problem is an inclusive strategy for dealing with constraints in the area of evolutionary multiobjective optimization (EMO) [7]–[9]. Following this idea, the sparse optimization problem formulated in (1) can be naturally modeled as a bi-objective optimization problem, in which the sparsity function $\|x\|_0$ and the loss function $\|y - Ax\|^2$ are two objectives for minimization. In the conventional regularization-based methods, these two objectives are combined in a linear way via a regularization parameter. In fact, the setting of the regularization parameter is highly sensitive to the sparseness of the optimal solutions found by regularization-based methods. However, this parameter is not required in the bi-objective sparse optimization problem. It is easy to show that the K-sparse solution of Problem (1) is located within the knee region of the PF of the bi-objective sparse optimization problem [10].

The applications of multiobjective approaches for sparse reconstruction have attracted some research attention over the past few years [11]–[15]. Three representative multiobjective approaches for sparse signal reconstruction are as follows.

- In [11], StMOEA, a multiobjective approach based on NSGA-II [16] and soft thresholding iteration, was suggested to solve sparse optimization problem. In StMOEA, the whole PF is first approximated by evolving a population of solutions. Then, an angle-based spline method is used to find the knee solution in a posterior way. For each offspring solution generated by genetic operator in StMOEA, the iterative soft-thresholding is applied for local improvement.
- In [12], a multi-phase multiobjective approach based on MOEA/D [17] and iterative thresholding, called MOEA/D-L1/2, was developed to find the approximations of both the whole PF and the knee region. The weighted sum method is then applied to detect the position of the knee region.
- In [14], a decomposition-based multiobjective evolutionary approach called SPLS was suggested to find a set of sparse solutions in the knee region of PF in an online way. Compared with existing thresholding methods, SPLS is more powerful to reconstruct sparse signals for the sparse optimization problems with small number of observations or large scale signals. To detect the knee region, SPLS divides the population into two subsets via a horizontal thresholding line.

It should be pointed out that the strategies for finding the knee region mentioned in above three multiobjective approaches are not very reliable due to the existence of noises in sparse signals. To improve the reliability on sparsity detection, we propose an improved MOEA/D for sparse reconstruction with the following features:

- A two-phase iterative search procedure is suggested to find sparse solutions along the PF. In the first phase, the approximation of the whole PF is obtained via a chain-based search process. In the second phase, the knee region is examined with preference in a random way.
- Between the above two phases, the knee region is detected by clustering via Gaussian Mixture Models, which divides the approximation of the whole PF into two subsets.
- Extensive experiments are conducted in comparison of the proposed method with SPLS and MOEA/D-L1/2, as well as the HALF algorithm, on a set of six instance sets with different noise levels. The ability of MOEA/D-MGC for sparsity detection for noiseless sparse signals is also investigated.

The remainder of this paper is organized as follows. Section II provides a brief review on the HALF algorithm, sparse multiobjective model, and the strategies for finding knee in sparse multiobjective approaches. The new multiobjective approach based on mixture Gaussian clustering is presented in Section III. Experimental results are reported and discussed in Section V. The last section concludes this paper.

II. SPARSE MULTIOBJECTIVE OPTIMIZATION

In this section, the regularization frameworks for sparse optimization are first introduced. Then, the bi-objective formulation in sparse optimization is presented and analyzed. Afterwards, three existing strategies in sparse multiobjective approaches for sparsity detection are reviewed.

A. SPARSE REGULARIZATION

Regularization is a commonly-used strategy for solving ill-posed inverse problem or preventing overfitting in machine learning [18], [19]. One of the most important applications of regularization is the sparse signal reconstruction, where a regularization (sparsity) term is added into a loss term. In the regularization method based on ℓ_0, the following single objective optimization problem should be considered:

$$\min_{x \in \mathcal{R}^N} \|y - Ax\|^2 + \lambda \|x\|_0 \tag{3}$$

where $\lambda > 0$ is the regularization parameter. It balances the loss function $\|y - Ax\|^2$ and the sparsity of x. In fact, Problem (3) is a NP-hard problem due to the discreteness of $\|x\|_0$ [20]. To find the sparse optimal solution of Problem (3), alternative regularization problems are often obtained by replacing $\|x\|_0$ by $\|x\|_1$ (convex relaxation) and $\|x\|_{1/2}$ (non-convex relaxation). Note that the relaxed regularization problems can be solved via continuous optimizers, such as the SOFT algorithm [5] and the HALF algorithm [6].

In the typical iterative thresholding algorithms, three major steps, i.e., gradient descent, setting of regularization parameter and thresholding truncation, are often involved. Note that the regularization parameter λ is often set to an empirical value or determined by sorting the components of the solution obtained in gradient descent step, where the estimation of the sparsity level K should be provided in advance. In practice, cross validation is often adopted for the setting of the regularization parameters. Among all iterative thresholding

algorithms, the HALF algorithm has attracted an increasing research attention due to its powerful ability in finding high-precision sparse solutions [21]–[27]. The detailed implementation of the HALF algorithm is illustrated in **Algorithm** 1.

Algorithm 1 HalfThresholding$(x^0, \bar{K}, ls)$

Require: x^0: the initial solution; $\bar{K}$: the prefixed sparsity level; ls: the maximal number of iterations
Ensure: optimal x^*
1: Let $n = 0$.
2: **repeat**
3: **gradient descent and sort components**:

$$\tilde{x} = x^n + A^T(Ax^n - y)$$
$$|\tilde{x}_{i_1}| \geq \cdots \geq |\tilde{x}_{i_{\bar{K}}}| \geq |\tilde{x}_{i_{\bar{K}+1}}| \geq \cdots |\tilde{x}_{i_N}|$$

where $\{i_1, \ldots, i_N\}$is a permutation of $\{1, \ldots, N\}$.
4: **setting of regularization parameter**:

$$\lambda_n = \frac{\sqrt{96}}{9}|\tilde{x}_{i_{\bar{K}+1}}|^{\frac{3}{2}}$$

5: **thresholding truncation**:

$$x_i^{n+1} = \begin{cases} \psi(\tilde{x}_i) & \text{if } |\tilde{x}_i| > \frac{\sqrt[3]{54}}{4}\lambda_n^{\frac{2}{3}}, \\ 0 & \text{otherwise.} \end{cases}$$

where

$$\psi(\tilde{x}_i) = \frac{2}{3}\tilde{x}_i\left(1 + \cos\left(\frac{2\pi}{3} - \frac{2\phi(\tilde{x}_i)}{3}\right)\right)$$

with

$$\phi(\tilde{x}_i) = \arccos\left(\frac{\lambda_n}{8}\left(\frac{|\tilde{x}_i|}{3}\right)^{-\frac{3}{2}}\right)$$

$n \leftarrow n + 1$;
6: **until** $n \geq ls$
7: Return $x^* = x^n$.

B. MULTI-OBJECTIVE SPARSE OPTIMIZATION

In this subsection, the bi-objective sparse optimization model is first formulated and analyzed. Then, the strategies for finding the knee region of the PF are reviewed.

1) BI-OBJECTIVE FORMULATION

To remove the constraint in Problem (1), the following bi-objective sparse optimization problem is often considered:

$$\min F(x) = \{(f_1(x), f_2(x)\} = \{\|x\|_0, \|y - Ax\|^2\} \quad (4)$$

where y, A are the same as in (1). Note that the first objective function $\|x\|_0$ only takes discrete values, and the second objective function $\|y - Ax\|^2$ is a quadratic function. Similar to Problem (1), the above bi-objective optimization problem is also $\mathcal{NP}$-hard. Fig. 1 shows the distribution of all weakly Pareto solutions of Problem (4) with noise ($\epsilon \neq 0$) or without noise ($\epsilon = 0$). In the case of $\epsilon = 0$, it is easy to prove that

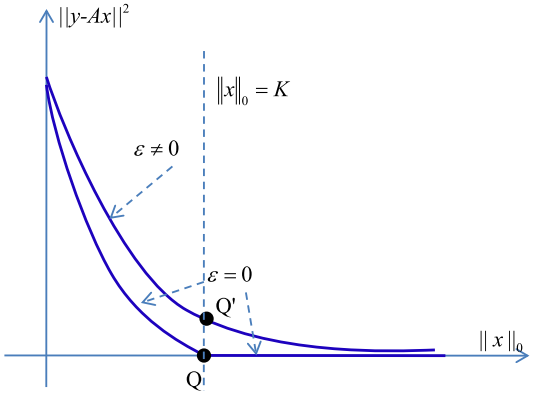


FIGURE 1. The distribution of Pareto-optimal solutions in the bi-objective sparse optimization problem.

the K-sparse solution is located at the intersection indicated by 'Q' between the strict PF part with $f_2 > 0$ and the linear PF part with $f_2 = 0$. The local area including the K-sparse solution x^* is so-called the knee region of the PF.

In real-world sparse optimization problems, the observation vector y is often corrupted with noise ($\epsilon \neq 0$). That is, $y = \bar{y} + \epsilon$, where $\bar{y} = Ax^*$. As shown in Fig.1, the PF part of Problem (4) on the right side of the K-sparse solution will not be parallel to the sparsity axes. It can also be observed that the distribution of the solutions along this part of PF is approximately linear. Moreover, the K-sparse Pareto solution is still located at the 'knee' region of the whole PF.

2) FINDING-KNEE STRATEGIES

Over the past several years, there have been some attempts on the use of multiobjective approaches for sparse optimization [10]–[12], [14]. All these methods are interested in finding the knee region of the PF, which should include some solutions that approximate the K-sparse solution of Problem (1) very well. In the following, three representative finding-knee strategies are reviewed.

- Angle-based spline method in StMOEA [11]: The angle-based method was first proposed to detect the knee solution in [28]. The idea of this method is depicted in Fig.2. In StMOEA, the spline curve fitting based

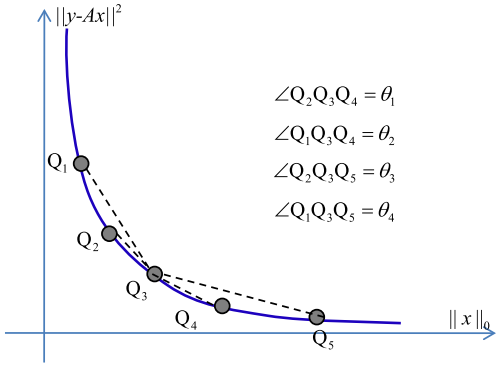


FIGURE 2. The angle-based method for finding the knee solution in StMOEA. The curvature of each point is measured by either θ_1 (standard version) or the maximum of $\theta_1 - \theta_4$(intensified version).

on angles is applied to detect the knee solution in the approximation of the whole PF, which is obtained by evolving a population via NSGA-II framework and soft thresholding. It should be pointed out that the angle-based spline method is not very reliable if the solutions in the knee region are not well approximated by the population of StMOEA.

- Weighted sum method in MOEA/D-L1/2 [12]: In MOEA/D-L1/2, the approximation of the whole PF is obtained by optimizing a number of subproblems defined by equally spaced sparsity levels. The knee solution is defined as the population member that minimizes a weighted sum function $\|y - Ax\|_2^2 + \alpha\|x\|_0$, which is shown in Fig.3. Therefore, the sparsity detection by the weighted sum method is very sensitive to the setting of α. Unlike SPLS, MOEA/D-L1/2 is applicable to both the noiseless signals and the noisy signals.
- Horizonal thresholding line method in SPLS [14]: According to our analysis on the distribution of the PF of Problem (4), the solutions in the linear part of PF often have very small values of f_2. Therefore, SPLS divides the population into two subsets via a horizonal thresholding line with $f_2 = f_2^{min} + \beta$, where f_2^{min} is the minimal f_2 value on all solutions in the population and β is a very small positive value. As shown in Fig.4, the solutions in the linear PF part locate inside $[f_2^{min} + \beta, f_2^{min}]$. Therefore, the knee region should include the solutions on the left side of the linear PF part. In the case of noiseless signals, SPLS is able to detect the sparsity level and reconstruct signals with high precision. However, the performance of SPLS is very sensitive to the setting of β when the noise is involved in signals.

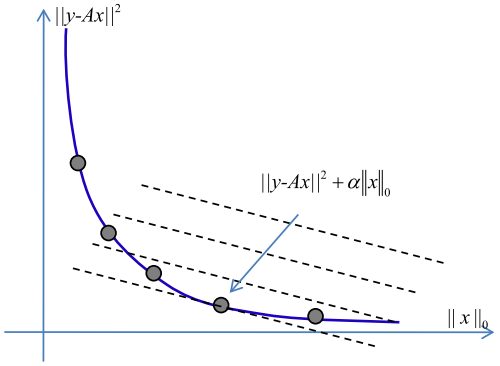


FIGURE 3. The weighted sum method for finding the knee solution in MOEA/D-L1/2.

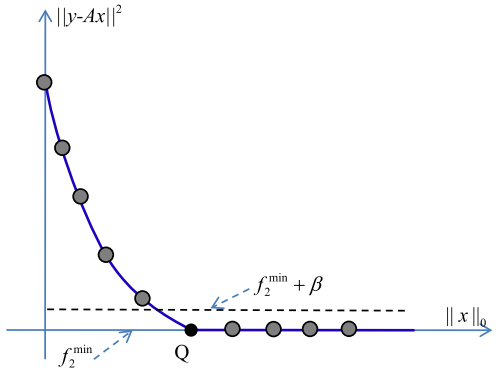


FIGURE 4. The horizontal thresholding line method for finding the knee solution in SPLS.

III. MOEA/D-MGC

In this section, the motivations on the use of mixture Gaussian clustering in MOEA/D for sparse optimization are first discussed. Then, the details of mixture Gaussian clustering are presented. At the end of this section, the algorithmic framework of our proposed method MOEA/D-MGC is illustrated.

A. MOTIVATIONS

As discussed in the previous section, the main tasks of sparse multiobjective approaches are the approximation of the whole PF, or the local PF, or both of them. One of the key issues in these methods lies in the detection of the knee region, which should include the optimal solution with the true sparsity level. In fact, each of the detection methods used in StMOEA, SPLS, and MOEA/D has its disadvantage.

- As analyzed in Section 2, the sparse solution of Problem (1) should locate in the position on the left side of the linear part of PF. Although this solution belongs to the knee region, it may not have the maximal curvature value cross the whole PF. Therefore, the angle-based method used in StMOEA has no ability to detect the exact position of the true sparsity level. On the other hand, the detection of knee region may not be reliable when the whole PF is not well approximated by StMOEA.
- In SPLS, the linear part of PF is detected by collecting the population members with small values of f_2. Such a detection requires the setting of the parameter β, which is used to divide the population into two subsets. When the linear part is parallel to the horizonal axis, SPLS can detect the knee region close to the true sparsity level. However, when the signal is involved with noise, the linear part of PF is not parallel to the horizonal axis any more. Therefore, the classification of population via a horizonal line defined by setting β is not reasonable.
- In MOEA/D-L1/2, the knee solution is set to the solution in the population minimizing the weighted sum function of the loss function and sparsity function. Like the ℓ_0 regularization, an empirical setting of the coefficient for the sparsity function should be provided. A large value of this coefficient will cause the detection of knee region with small value of sparsity. On the contrary, a small value of this coefficient will result in the detection of knee region with large value of sparsity. Like SPLS, the performance of MOEA/D-L1/2 is also sensitive to the setting of the coefficient for sparsity function in case of noisy signals.

According to our above analysis, none of three methods is reliable for the sparsity detection on both noiseless signal

and noisy signals. Therefore, the propose of a parameterless detection method is necessary and important.

B. GAUSSIAN MIXTURE CLUSTERING

The division of the population in SPLS into two subsets can be regarded as a simple clustering method. When signals involve no or very little noise, it can be observed that the solutions in the subset below the horizonal threshold line have a linear or approximately linear distribution. Gaussian Mixture Clustering (GMC) is a popular technique for clustering data in machine learning [29]. To capture the linear distribution of the PF in the bi-objective sparse optimization, Gaussian Mixture Models (GMM) seems more suitable than the horizonal threshold line used in SPLS. Assume that the approximation of the whole PF consists of P objective vectors $\{v_1, \ldots, v_P\}$. In GMM, each objective vector v follows the Gaussian mixture distribution stated as follows:

$$p(v) = \sum_{s=1}^{D} \pi_s \mathcal{N}(v|\mu_s, \Sigma_s) \tag{5}$$

where

- D is the number of clusters. It is set to 2 in this work.
- π_s is the probability of the data point v belonging to the s-th cluster;
- μ_s and Σ_s are the mean and the covariance matrix for the s-th cluster, respectively.

The parameters $\pi_s, \mu_s, \Sigma_s, s = 1, \ldots, D$, are obtained by maximizing the log of the likelihood function with the following form:

$$\ln p(v|\pi, \mu, \Sigma) = \sum_{p=1}^{P} \ln \left\{ \sum_{s=1}^{D} \pi_s \mathcal{N}(v_p|\mu_s, \Sigma_s) \right\} \tag{6}$$

To maximize the above likelihood function, the standard method, i.e., expectation-maximization (EM), is used in our work. The detailed framework of the EM algorithm is provided in **Algorithm** 2. In line 1, the parameters like means, covariances, and marginal probabilities, as well as the clusters are initialized. In line 2, the responsibility of the p-th data point for the s-th cluster is computed. In line 3, the parameters are updated. In line 4, the EM algorithm terminates if the value of the log maximum function has no any change or a very tiny change. In line 5, D clusters are then determined by computing the maximal responsibility of each data points. Fig.5 shows the plots of two clusters of data points divided by the **Algorithm** 2, which is able to capture a subset of data points with linear distribution.

C. ALGORITHMIC FRAMEWORK OF MOEA/D-MGC

Over the past ten years, MOEA/D has been widely studied in the area of EMO research [30]–[38]. Unlike other Pareto-based EMO algorithms, MOEA/D needs to optimize a number of single objective subproblems or multiple small multiobjective optimization problems. The weighted sum method and the weighted Tchebycheff method are two most

Algorithm 2 GaussianMixtureClustering

Require: The approximation of whole PF, $\{v_1, \ldots, v_P\}$; The number of clusters, D

Ensure: The set of D clusters: $C_1, \ldots, C_D$

1: Initialize μ_s, Σ_s, π_s, $C_s = \emptyset$, $s = 1, \ldots, D$, and evaluate the initial value of the log likelihood;

2: Evaluate the responsibilities by:

$$\gamma_{p,s} = \frac{\pi_s \mathcal{N}(v_p|\mu_s, \Sigma_s)}{\sum_{s=1}^{D} \pi_s \mathcal{N}(\mu_s, \Sigma_s)}$$

3: Re-estimate parameters:

$$P_s = \sum_{p=1}^{P} \gamma_{p,s}$$
$$\mu_s = \frac{1}{P_s} \sum_{p=1}^{P} \gamma_{p,s} v_p$$
$$\Sigma_s = \frac{1}{P_s} \sum_{p=1}^{P} \gamma_{p,s} (v_p - \mu_s)(v_p - \mu_s)^T$$
$$\pi_s = \frac{P_s}{P}$$

4: Re-calculate the value of the log likelihood function. If the stopping condition is met, then go to the next step; otherwise, go to line 2;

5: **for** $p = 1$ to P **do**

6: determine $\bar{j} = \text{argmax}_{j \in \{1,\ldots,D\}} \gamma_{p,j}$;

7: set $C_{\bar{j}} = C_{\bar{j}} \cup \{p\}$.

8: **end for**

9: Return $C_1, \ldots, C_D$.

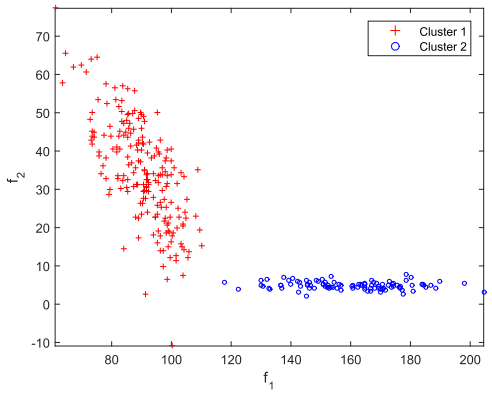


FIGURE 5. The plots of two clusters of data points obtained by Gaussian Mixture Clustering. The mean and covariance matrix of the left cluster are (90, 35) and 100[1 −1; −1 2] while those of the right cluster are (165, 5) and [300 0; 0 1].

commonly-used method for decomposition in MOEA/D. However, these two methods are not suitable for Problem (4). Instead, a number of subproblems with the same objective functions and different constraints are considered in our proposed algorithm MOEA/D-GMC. In detail, these subproblems are defined as follows:

$$\text{minimize} \|Ax - b\|^2, \quad \text{subject to } \|x\|_0 = S_i \tag{7}$$

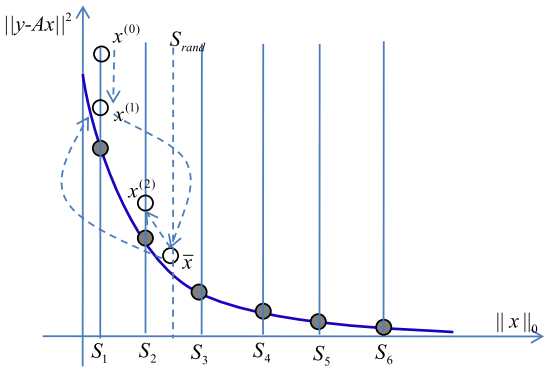


FIGURE 6. An example of updating population in the chain-based thresholding search. HalfThresholdingSearch is executed with a starting solution $x^{(1)}$ and a random sparsity level S_{rand} in $[S_2, S_3]$. The obtained solution $\bar{x}$ is used to update $x^{(1)}$ and $x^{(2)}$.

where the sparsity level $S_i, i = 1, \ldots, P$, takes the value in from a sparsity interval $[S_{min}, S_{max}]$. To approximate the whole PF, the values of $S_i, i = 1, \ldots, P$, are defined in the following way:

$$S_i = S_{min} + \left\lfloor (i-1) \times \frac{S_{max} - S_{min}}{P-1} \right\rfloor$$

Correspondingly, a population of P solutions, denoted by $X = \{x^1, \ldots, x^P\}$, is maintained. Each of them is associated with one subproblem. Note that a two-phase local search proposed for multi-objective combinatorial optimization problems approximates the PF in a chain way, where subproblems are defined via the weighted sum method [39]. In MOEA/D-GMC, the above P subproblems are also optimized in a chain way, where the sparsity level is increased from S_1 to S_P.

The algorithmic framework of MOEA/D-GMC is illustrated in **Algorithm** 3. It mainly consists of three steps: the approximation of the whole PF via a chain-based thresholding search step, the detection of knee solution via mixture gaussian clustering, and a random thresholding search step.

Algorithm 3 MOEA/D-GMC: A Multiobjective Approach Based on Decomposition and Mixture Gaussian Clustering

Require: The population size, P; The number of iterations in thresholding search, ls

Ensure: The approximation of the whole PF, X; The sparsity level of the knee solution, $\tilde{K}$

1: optimize P subproblems in a chain manner:

$$\textbf{ChainThresholdingSearch}(X, P, ls)$$

2: detect the knee solution by mixture gaussian clustering:

$$\tilde{K} \leftarrow \textbf{KneeSparsityDetection}(X)$$

3: optimize q local subproblems in a random:

$$\textbf{RandomThresholdingSearch}(X, \tilde{K}, ls)$$

4: Return $\tilde{K}$ and X.

Algorithm 4 ChainThresholdingSearch(X, P, ls)

1: Generate an initial solution $x^{(0)}$ with $\|x^{(0)}\|_0 = S_1$ randomly;

2: **for** $p = 1$ to P **do**

3: Choose a random sparsity level by:

$$S_{rand} = S_p + \left\lfloor rand * \frac{S_{max} - S_{min}}{P-1} \right\rfloor$$

4: Apply the HALF thresholding method by:

$$\bar{x} \leftarrow \textbf{HalfThresholding}(x^{(p-1)}, S_{rand}, ls)$$

5: Update the population by:

$$\textbf{UpdatePopulation}(\bar{x}, X, p).$$

6: **end for**

- Chain-based Thresholding Search (see **Algorithm** 4): In lines 3-5 of, P subproblems defined via P sparsity levels $S_1, \ldots, S_P$ are optimized in a chain way. Note that the current solution $x^{(p-1)}$ of the $(p-1)$-th subproblem is used as the starting solution for the optimization of its subsequent subproblem to obtain $x^{(p)}$. Instead of using the sparsity level S_p, a larger random value of sparsity level S_{rand} is set for the HALF thresholding iteration in line 3.

$$S_{rand} = S_p + \left\lfloor rand * \frac{S_{max} - S_{min}}{P-1} \right\rfloor$$

where *rand* is a random number in $[0, 1]$. In line 4, the random setting of sparsity levels for **HalfThresholdSearch** could provide a better diversity for the chain-based search. In line 5, the current solutions of subproblems are updated by truncating $\bar{x}$ at p sparsity levels, i.e., $S_1, \ldots, S_p$. The implementation of population updating is given in **Algorithm**5. A graphical example of population updating is shown in Fig. 6

- Gaussian Mixture Clustering (see **Algorithm** 6): In the second step, the mixture gaussian clustering is applied to cluster the approximation of PF regarding X into two subsets in line 1 of **Algorithm** 6. The subset $C_{\tilde{i}}$ including the linear PF part and the knee solution $\tilde{x}$ are then determined in line 2 and line 3 of **Algorithm** 6 respectively.
- Random Thresholding Search (see **Algorithm** 7): In this step, the subproblems with the sparsity levels on the right side of the knee region are optimized in a random order. Lines 4-6 are the same as lines 3-5 of **Algorithm** 4.

D. SOME REMARKS

1) THE CONNECTION BETWEEN MOEA/D-GMC VS. SPLS

Both MOEA/D-GMC and SPLS involve a chain-based search procedure, where the sparsity levels for the HALF thresholding search are changed with some probability. The random

Algorithm 5 UpdatePopulation$(\bar{x}, X, p)$

Sort the components of x^r:

$$|\bar{x}_{\pi(1)}| \geq |\bar{x}_{\pi(2)}| \geq \cdots \geq |\bar{x}_{\pi(N)}|$$

Set $p' = p$;
while $p' \geq 1$ **do**
  Set $\bar{x}_{\pi(j)} = 0, j = S_{p'} + 1, \ldots, N$;
  if $f_2(\bar{x}) < f_2(x^{(p')})$ **then**
    Replace $x^{(p')}$ with $\bar{x}$;
  else
    Break the while loop;
  end if
  Set $p' \leftarrow p' - 1$;
end while

Algorithm 6 KneeSparsityDetection(X)

Set $V = \{F(x)|x \in X\}$ and apply gaussian mixture clustering to divide V into two clusters:

$$\{C_1, C_2\} \leftarrow \textbf{GaussianMixtureClustering}(V, 2)$$

Determine the cluster including the linear PF part:

$$\tilde{i} = \text{argmax}_{i\in\{1,2\}} \left\{ \min_{s\in\{C_i\}} \|x^{(s)}\|_0 \right\}$$

Locate the knee solution:

$$\tilde{x} = \text{argmin}_{x\in\{x^{(i)}|j\in C_{\tilde{j}}\}} \|x\|_0$$

Return $\tilde{K} = \|\tilde{x}\|_0$.

change of sparsity levels in SPLS lies in the preference-based selection of population members, while in MOEA/D-GMC, it is implemented by selecting a random sparsity level between two neighboring sparsity levels. Note that the chain-based search step in MOEA/D-MGC aims at finding the approximation of the whole PF. In contrast, SPLS approximates the local PF close to the knee region in an online way. MOEA/D-MGC also has the ability to approximate the local PF near the knee region. Compared with the method in SPLS for sparsity detection, the GMC method doesn't contain any sensitive parameter.

2) THE CONNECTION BETWEEN MOEA/D-GMC VS. MOEA/D-L1/2

MOEA/D-MGC can be regarded as an improved version of MOEA/D-L1/2, which also aims at finding the approximation of the whole PF and the local PF in the knee region. The major difference between them lies in the detection of sparsity. Unlike the weighted sum method in MOEA/D-L1/2, there is no extra parameter in the GMC method when detecting the sparsity value. Moreover, MOEA/D-L1/2 needs to optimize all possible subproblems with the sparsity levels in the knee region while MOEA/D-GMC only considers the optimization

Algorithm 7 RandomThresholdingSearch$(X, \tilde{K}, ls)$

Set $L = \{S_i|S_i \geq \tilde{K}, i = 1, \ldots, P\}$;
repeat
  Pick one sparsity level S_q from L randomly;
  Choose a random sparsity level by:

$$S'_{rand} = S_q + \left\lfloor rand * \frac{S_{max} - S_{min}}{P - 1} \right\rfloor$$

  Perform the HALF method on $x^{(q)}$:

$$\bar{x} \leftarrow \textbf{HalfThresholding}(x^{(q)}, S'_{rand}, ls)$$

  Update the population by:

$$\textbf{UpdatePopulation}(\bar{x}, X_1, q).$$

until The stopping condition is satisfied.

of some subproblems with equally-spaced sparsity levels in the knee region.

3) THE TIME COMPLEXITY OF MOEA/D-GMC

Note that MOEA/D-GMC, MOEA/D-L1/2, and SPLS involve the same number of HALF thresholding iterations. Unlike the latter two algorithms, MOEA/D-GMC needs to classify the PF or weakly PF via GMM with the time complexity $O(PDM^3)$, where D and M are the number of clusters and the dimension of data points, respectively. Both MOEA/D-GMC and MOEA/D-L1/2 have the same time complexity for updating population, which involves $O(P)$ comparisons in terms of the loss function f_2.

IV. COMPUTATION EXPERIMENTS

In this section, the experimental settings are first presented. Then, the experimental results are reported and discussed.

A. EXPERIMENTAL SETTINGS

In our experiments, two existing multiobjective methods, i.e., SPLS and MOEA/D-L1/2, and the HALF method, are considered in the comparison of MOEA/D-GMC on the some artificial test instances for sparse signal reconstruction. To generate a test instance, the values of A, y, x^* are specified in the following three steps:

- a true sparse signal x^* with K nonzero components is randomly generated.
- a sensing matrix A is randomly generated by a normal distribution $\mathcal{N}(0, 2)$ and its columns are orthogonalized and normalized;
- an observation vector is computed by $y = Ax^*$.

Table 1 summarizes the configurations of six test instance sets used in our experiments. For each instance set, four noisy levels 0, 0.01, 0.05, 0.1, are considered. P1-P3 are three small test instance sets ($N = 512$) while P4-P6 with ten times of the settings of P1-P3 are three large instance sets ($N = 5120$). The total number of instance sets are 24. S_{min} and S_{max} are set to $0.1 \times K$ and $2 \times K$ respectively.

TABLE 1. The configurations of test instances in our experiments - *N* - length of signal, *M* - no. of observations, *K* - true sparsity.

Instance	N	M	K	σ
P1	512	300	130	(0, 0.01, 0.05, 0.1)
P2	512	270	130	(0, 0.01, 0.05, 0.1)
P3	512	240	130	(0, 0.01, 0.05, 0.1)
P4	5120	3000	1300	(0, 0.01, 0.05, 0.1)
P5	5120	2700	1300	(0, 0.01, 0.05, 0.1)
P6	5120	2400	1300	(0, 0.01, 0.05, 0.1)

To make a quantitative measurement on the quality of the solutions reconstructed by four algorithms, the mean square error (MSE) values between the knee solutions and the true sparse solution are calculated. Moreover, the relative sparsity error (RSE) values is defined as:

$$\mathrm{RSE}(K, \bar{K}) = \frac{|\bar{K} - K|}{K}$$

where $\bar{K}$ is the sparsity level of knee solution, and K is the true sparsity value. The smaller the value of RSE is, the better the considered algorithms perform in the detection of sparsity level.

The population sizes of MOEA/D-L1/2, SPLS, and MOEA/D-GMC is set to 30 for P1-P3 and 60 for P4-P6. The maximal number of iterations (i.e., the steps of gradient descent) is set to 3000 for P1-P3 and 6000 for P4-P6. The number of iterations *ls* used in **HalfThresholding** is set to 20 for P1-P6. In SPLS/HALF, the parameter for the horizon line β is set to 0.01 for the noiseless instance sets and 0.1 for the noisy instance sets. In MOEA/D-L1/2, the coefficient α in the weighted sum method is set to 0.01 for all instance sets. All algorithms are implemented in Matlab2018.

B. EXPERIMENTAL RESULTS

In Table 2, the median MSE values found by four considered sparse multiobjective algorithms (i.e., MOEA/D-GMC, MOEA/D-L1/2, SPLS/HALF, and HALF) on six instance sets with four different noise levels (i.e., $\sigma = 0.01, 0.02, 0.05, 0.1$) are provided. From this table, we have the following observations:

- It is clear that both MOEA/D-GMC and SPLS find smaller median MSE values than MOEA/D-L1/2 and HALF. MOEA/D-GMC performs better than the other three algorithms on 12 out of 24 instance sets while SPLS/HALF performs best on 11 out of 24 instance sets. However, MOEA/D-GMC performs worst on none of the instance sets while SPLS/HALF is outperformed by the other three algorithms on six instance sets. It can also be observed that MOEA/D-L1/2 performs best on none of the instance sets while HALF performs worst on 18 out of 24 instance sets.
- When the noise level is relatively small ($\sigma = 0.01, 0.02$), three multiobjective sparse algorithms, i.e., MOEA/D-GMC, MOEA/D-L1/2, and SPLS/HALF, performs very similarly in terms of the median MSE values on P1-P6. Note that all these three algorithms perform better than the HALF algorithm on 10 out of 12 instance sets. MOEA/D-GMC is slightly outperformed by SPLS/HALF. When the noise level is relatively large ($\sigma = 0.05, 0.1$), MOEA/D-GMC is able to find the smallest median MSE values on 8 out of 12 instance sets, while SPLS/HALF obtains the smallest median MSE values on 4 out of 12 instance sets.

From the above observations, we may conclude that the overall performance of MOEA/D-GMC is better than those of the

TABLE 2. The median MSE values found by MOEA/D-GMC, MOEA/D-L1/2, SPLS/HALF, and HALF on six instance sets with four different noise levels.

Noise	Instance	MOEA/D-GMC	MOEA/D-L1/2	SPLS/HALF	HALF
$\sigma = 0.01$	P1	0.000131	0.000172	0.000119	**0.000112**
	P2	0.020312	0.021077	**0.000206**	0.168768
	P3	0.000164	0.000190	**0.000141**	0.017733
	P4	**0.000074**	0.000178	0.000082	0.000259
	P5	**0.000101**	0.000266	0.000105	0.050018
	P6	0.001022	0.000811	**0.000153**	0.226330
$\sigma = 0.02$	P1	0.000392	0.000388	**0.000319**	0.000405
	P2	0.037799	0.037394	**0.000974**	0.155354
	P3	0.000487	0.000520	**0.000383**	0.023225
	P4	**0.000310**	0.000409	0.000707	0.000547
	P5	**0.000433**	0.000512	0.000884	0.040825
	P6	0.001091	0.001090	**0.001084**	0.210267
$\sigma = 0.05$	P1	**0.002580**	0.002929	0.004374	0.002842
	P2	0.063993	0.081346	**0.039585**	0.163912
	P3	**0.003889**	0.004161	0.005569	0.027931
	P4	**0.002437**	0.002969	0.005253	0.003605
	P5	**0.003704**	0.004064	0.006452	0.060650
	P6	0.034910	0.044925	**0.017143**	0.222924
$\sigma = 0.1$	P1	**0.013904**	0.020213	0.022010	0.016002
	P2	0.125428	0.127795	**0.117517**	0.187295
	P3	**0.022676**	0.027549	0.029316	0.057865
	P4	**0.012864**	0.019185	0.022336	0.020042
	P5	**0.024258**	0.028448	0.030456	0.086044
	P6	0.122636	0.128867	**0.120344**	0.250405

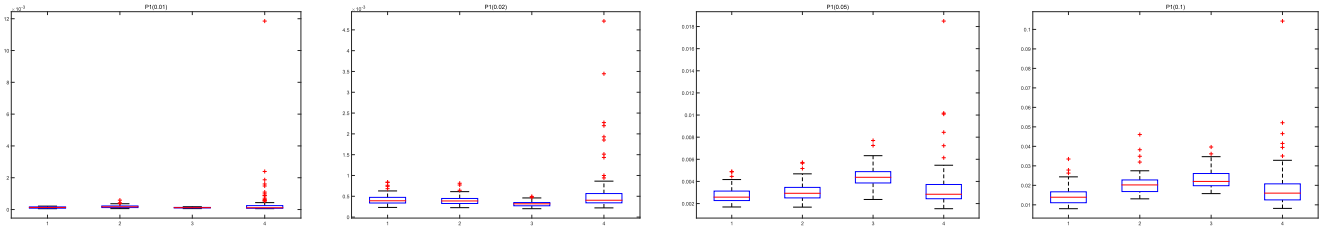


FIGURE 7. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P1 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

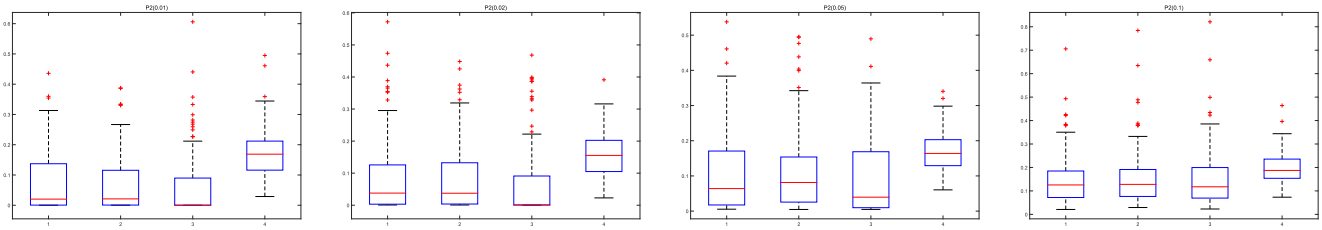


FIGURE 8. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P2 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

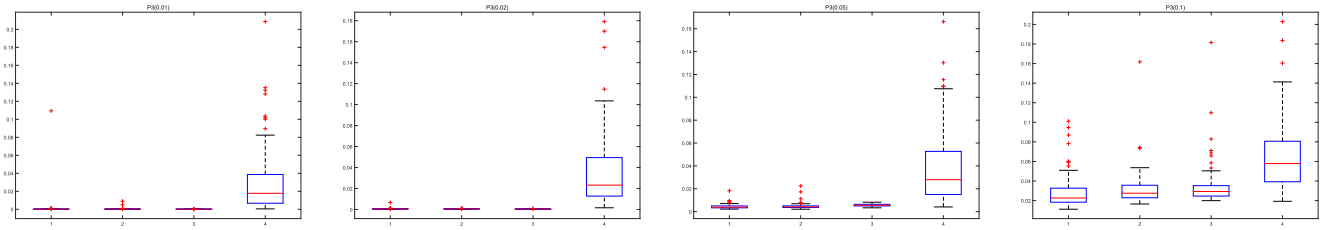


FIGURE 9. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P3 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

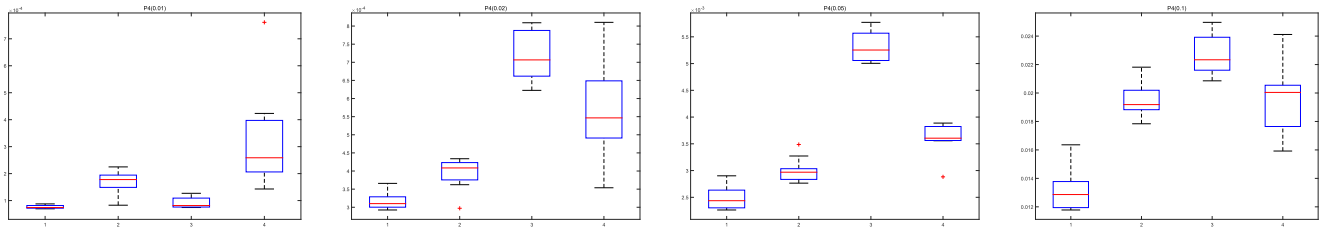


FIGURE 10. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P4 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

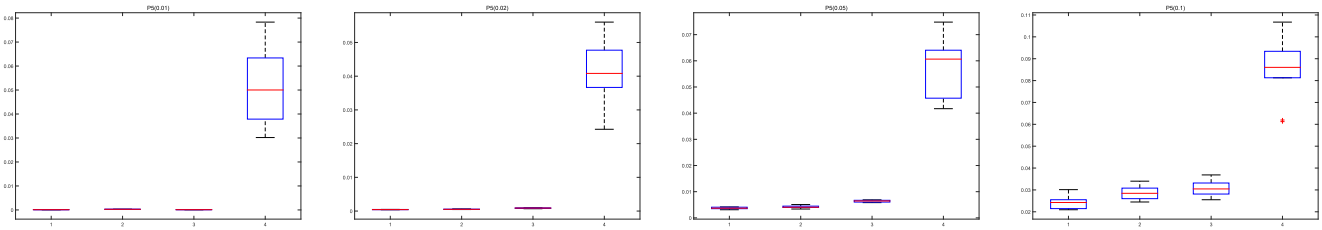


FIGURE 11. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P5 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

other three algorithms in terms of the median MSE values at different noise levels.

Figs. 7-12 show the box plots of the MSE values found by four algorithms on 24 instance sets. From the median, the upper and lower quartile values reported in these figures, we can make the following observations on the stability of the four algorithms.

- On two short-length instance sets P1 and P2 with a large observation ratio (M/N), all four algorithms have the similar performance on the stabilities in terms of the

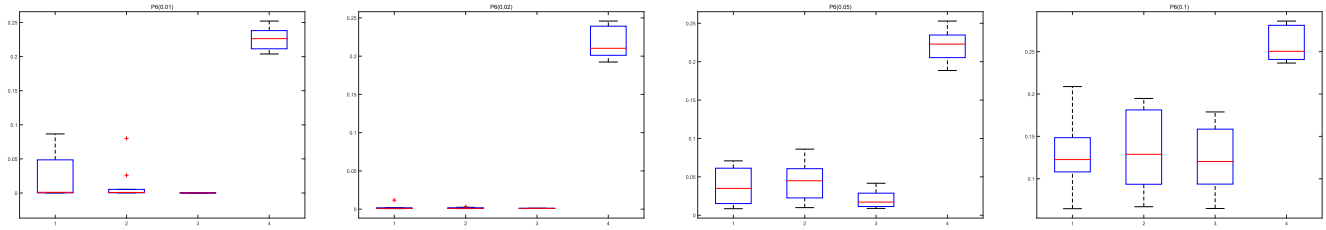


FIGURE 12. The boxplots of the MSE values found by MOEA/D-GMC('1'), MOEA/D-L1/2('2'), SPLS/HALF('3'), and HALF('4') on P6 with four different noise levels 0.01, 0.02, 0.05, 0.1 (from left to right).

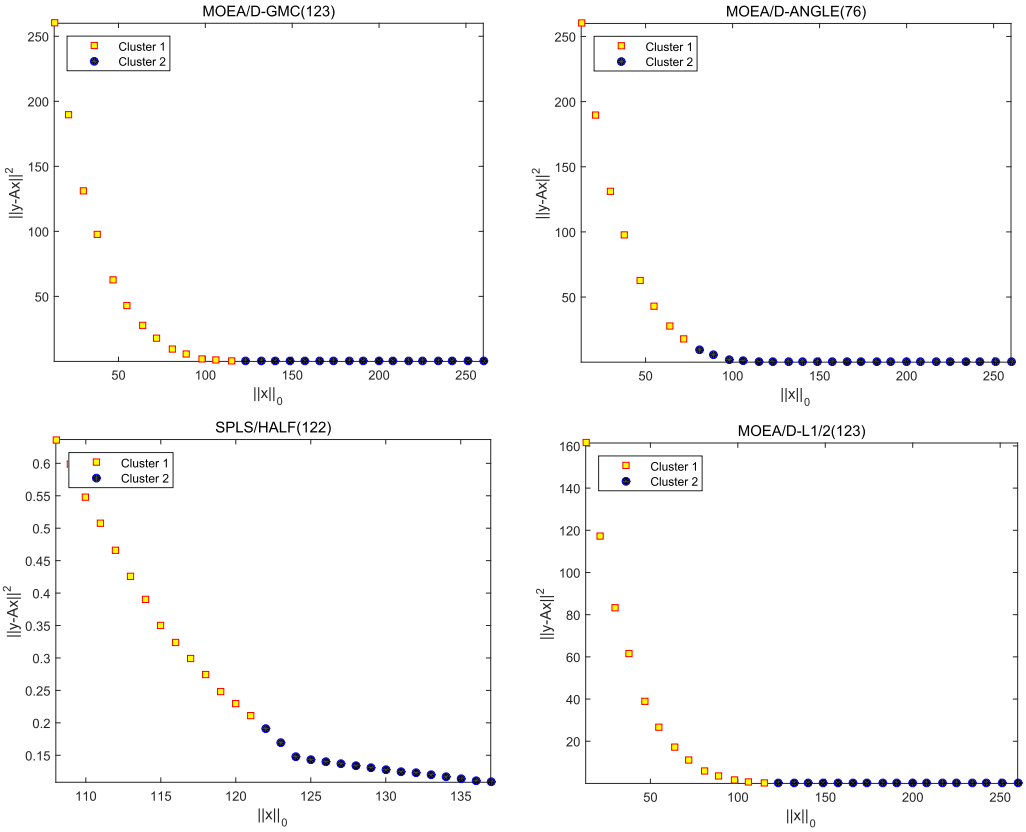


FIGURE 13. The plots of final solutions in two clusters obtained by four strategies in MOEA/D-GMC, ANGLE, SPLS, and MOEA/D-L1/2 on P1 with $\sigma = 0.01$.

upper and lower quartile values at different noise levels. This is reasonable since both P1 and P2 are relatively easy to solve. On the short-length instance set P3 with a small observation ratio, all three multiobjective algorithms performed much better than the HALF algorithm in stability.

- On the long-length instance set P4 with the largest observation ratio (3000/5120), both MOEA/D-GMC and MOEA/D-L1/2 performed better in terms of stability. On the long-length instance set P5 with the medium observation ratio (2700/5120), all three multiobjective algorithms have better stabilities than the HALF algorithm. On the long-length instance set P6 with the smallest observation ratio (2400/5120), none of the algorithms performs clearly better than others.

In summary, we may conclude that sparse multiobjective methods have advantages against the HALF algorithm in stability at different instance scales and noise level. The overall stability of MOEA/D-GMC is better than those of the other three algorithms.

Compared with the HALF algorithm, all three multiobjective methods have the ability to estimate the underlying true sparsity level by finding the knee position in the PF. Table 3 reported that the mean RSE values found by four different strategies, i.e., the mixture gaussian clustering MOEA/D-GMC, the weighted sum method in MOEA/D-1/2, the thresholding method in SPLS/HALF, and the spline-based method.[1] From this table, we can have the following observations:

- On all six instance sets P1-P6 with the smaller noise levels, i.e., $\sigma = 0.01, 0.02$, both MOEA/D-GMC and

[1]To make a fair comparison, the spline-based method proposed in StMOEA is used to locate the sparsity level of the knee solution on the PF found by **ChainThresholdingSearch** in MOEA/D-GMC.

TABLE 3. The sparsity error values found by four different strategies, i.e., the mixture gaussian clustering MOEA/D-GMC, the weighted sum method in MOEA/D-1/2, the thresholding method in SPLS/HALF, and the spline-based method in StMOEA.

Noise	Instance	MOEA/D-GMC	MOEA/D-L1/2	SPLS/HALF	SPLINE
$\sigma = 0.01$	P1	0.053846	0.053846	**0.042308**	0.415385
	P2	**0.053846**	0.115385	0.057692	0.430769
	P3	0.053846	0.053846	**0.046154**	0.423077
	P4	**0.030769**	0.063077	0.036923	0.407308
	P5	0.030769	0.063077	**0.028846**	0.419615
	P6	0.030769	0.063077	**0.030000**	0.427692
$\sigma = 0.02$	P1	0.053846	0.053846	**0.038462**	0.415385
	P2	0.053846	0.146154	**0.046154**	0.438462
	P3	0.053846	0.053846	**0.030769**	0.423077
	P4	**0.030769**	0.063077	0.302308	0.409615
	P5	**0.030769**	0.063077	0.266154	0.418846
	P6	**0.016154**	0.063077	0.186154	0.425769
$\sigma = 0.05$	P1	0.115385	**0.076923**	0.284615	0.415385
	P2	**0.053846**	0.207692	0.269231	0.434615
	P3	0.115385	**0.053846**	0.238462	0.423077
	P4	**0.063077**	0.065385	0.596154	0.409615
	P5	0.063077	**0.033846**	0.466154	0.417692
	P6	**0.130000**	0.146154	0.384231	0.429615
$\sigma = 0.1$	P1	**0.115385**	0.338462	0.530769	0.423077
	P2	**0.015385**	0.207692	0.330769	0.438462
	P3	**0.115385**	0.276923	0.423077	0.423077
	P4	**0.095385**	0.323077	0.766923	0.416538
	P5	**0.033846**	0.226923	0.591923	0.422308
	P6	**0.162308**	0.178462	0.441538	0.434231

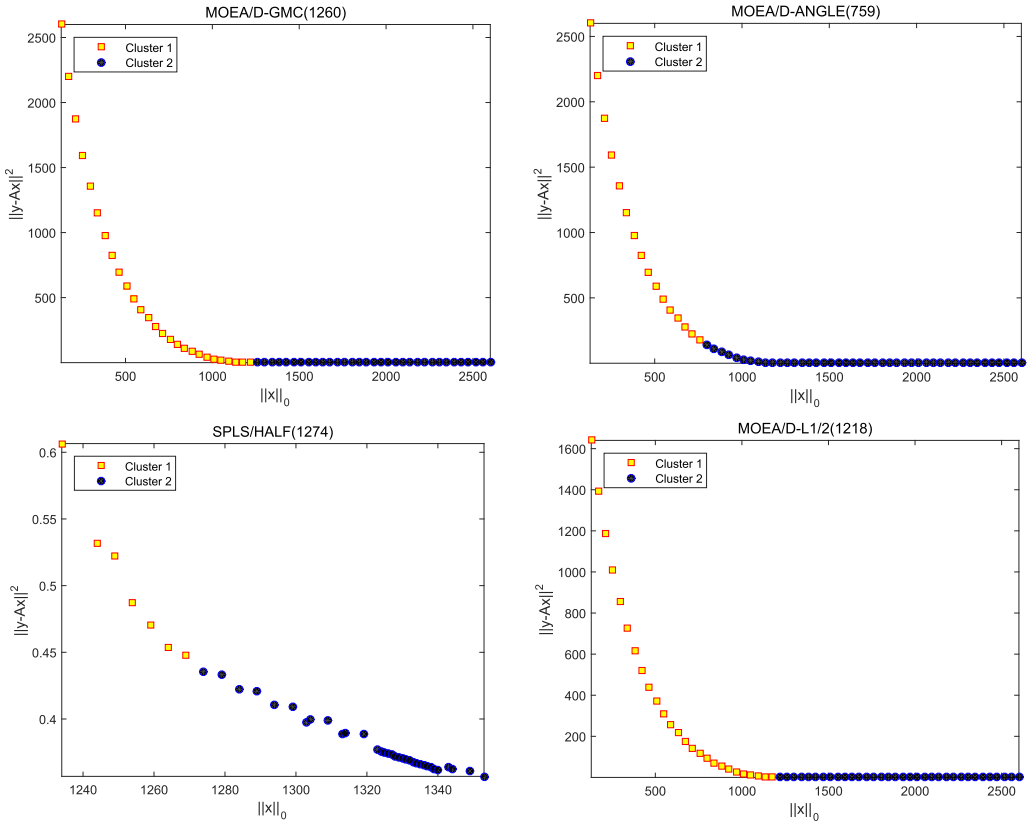


FIGURE 14. The plots of final solutions in two clusters obtained by four strategies in MOEA/D-GMC, ANGLE, SPLS, and MOEA/D-L1/2 on P4 with $\sigma = 0.01$.

SPLS/HALF find smaller RSE values than MOEA/D-L1/2 and SPLINE. In the case of $\sigma = 0.01$, these two algorithms found the RSE values less than 0.06 while the RSE value obtained by SPLINE is larger than 0.4. On the three long-length instance sets P4-P6 with $\sigma = 0.02$, the RSE values obtained by SPLS/HALF are clearly greater than those found by MOEA/D-GMC.

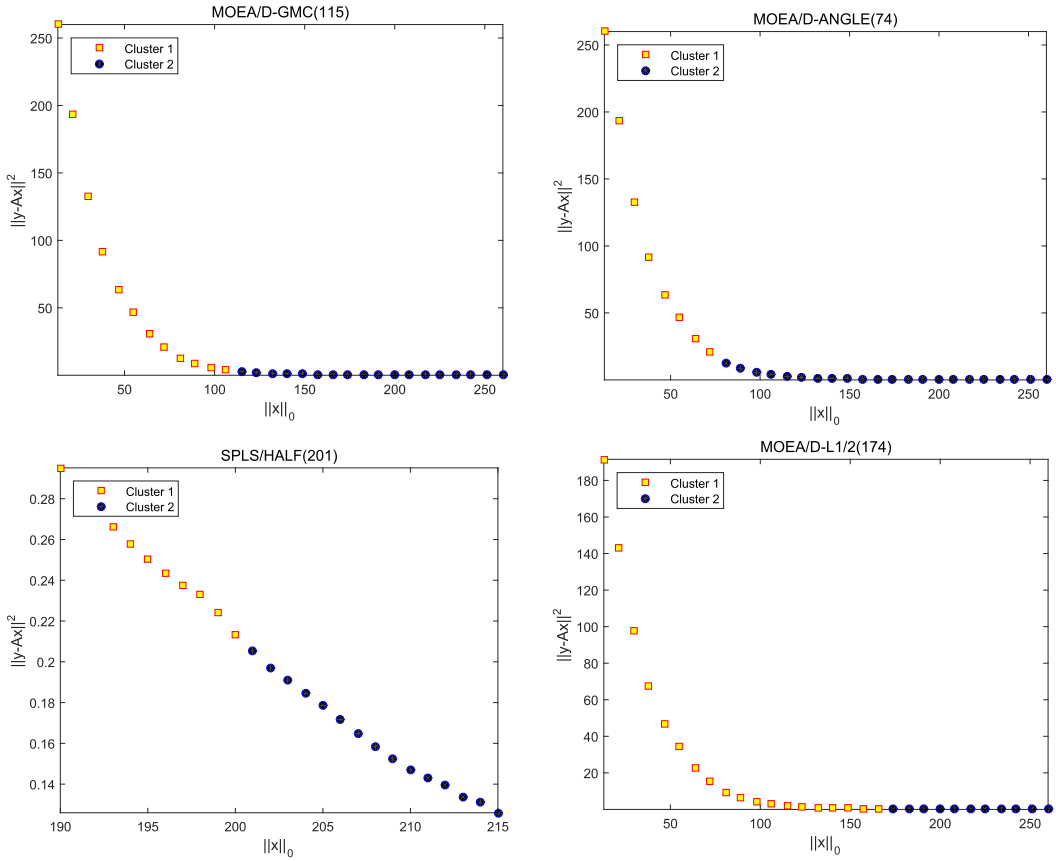


FIGURE 15. The plots of final solutions in two clusters obtained by four strategies in MOEA/D-GMC, ANGLE, SPLS, and MOEA/D-L1/2 on P1 with $\sigma = 0.1$.

- When six instance sets P1-P6 are involved with large noise levels $\sigma = 0.05$ and 0.1, both MOEA/D-GMC and MOEA/D-L1/2 perform better than the other two algorithms in terms of the mean RSE values. In case of the largest noise level $\sigma = 0.1$, the RSE values found by MOEA/D-GMC are smaller than those found by MOEA/D-L1/2 on P1-P6. Also, the RSE values found by SPLS/HALF and SPLINE are much greater than those found by MOEA/D-GMC.

Figs.13-16 show the final solutions in two clusters obtained by the finding-knee strategies in MOEA/D-GMC, ANGLE, SPLS/HALF, and MOEA/D-L1/2 on one small instance set P1 and one large instance set P4 at two noise levels, i.e., $\sigma = 0.01, 0.1$. In case of small noise level $\sigma = 0.01$, MOEA/D-GMC, SPLS/HALF, and MOEA/D-L1/2 found the knee sparsity levels that are very close to 130 of P1 and 1300 of P4. In contrast, the knee sparsity level found by ANGLE is far from the true sparsity level on both instance sets. In case of large noise level $\sigma = 0.1$, the knee sparsity levels found by MOEA/D-GMC are 115 and 1176, which are still close to 130 and 1300 respectively. The other three algorithms failed to find the knee sparsity levels close to the true sparsity levels of P1 and P4 with $\sigma = 0.1$. Among them, both SPLS and MOEA/D found the quite large knee sparsity levels while ANGLE found the small sparsity levels.

TABLE 4. The percentages (%) of the correct location of nonzero entries by MOEA/D-GMC, MOEA/D-L1/2, and SPLS/HALF.

Instance	MOEA/D-GMC	MOEA/D-L1/2	SPLS/HALF
P1(0.01)	**100**	94.62	**100**
P1(0.05)	**90**	87.69	88.46
P4(0.01)	**99.23**	93.69	99
P4(0.05)	**95.69**	93.77	95.23

Compared with the other three algorithms, MOEA/D-GMC has a better ability to detect the sparsity levels of knee solution, particularly when the noise level is relatively large. Moreover, the poor performance of SPLINE for locating knee solutions is due to the coarse approximation of the PF found by the multiobjective algorithms.

To show the performance of MOEA/D-GMC for locating percentage of nonzero entries in sparse signals, we tested three EMO algorithms on P1 and P4 with either $\sigma = 0.01$ or $\sigma = 0.05$. From the results shown in Table 4, it is clear that MOEA/D-GMC performs better than the other two EMO algorithms regarding the percentages of the correct location of nonzeros entries.

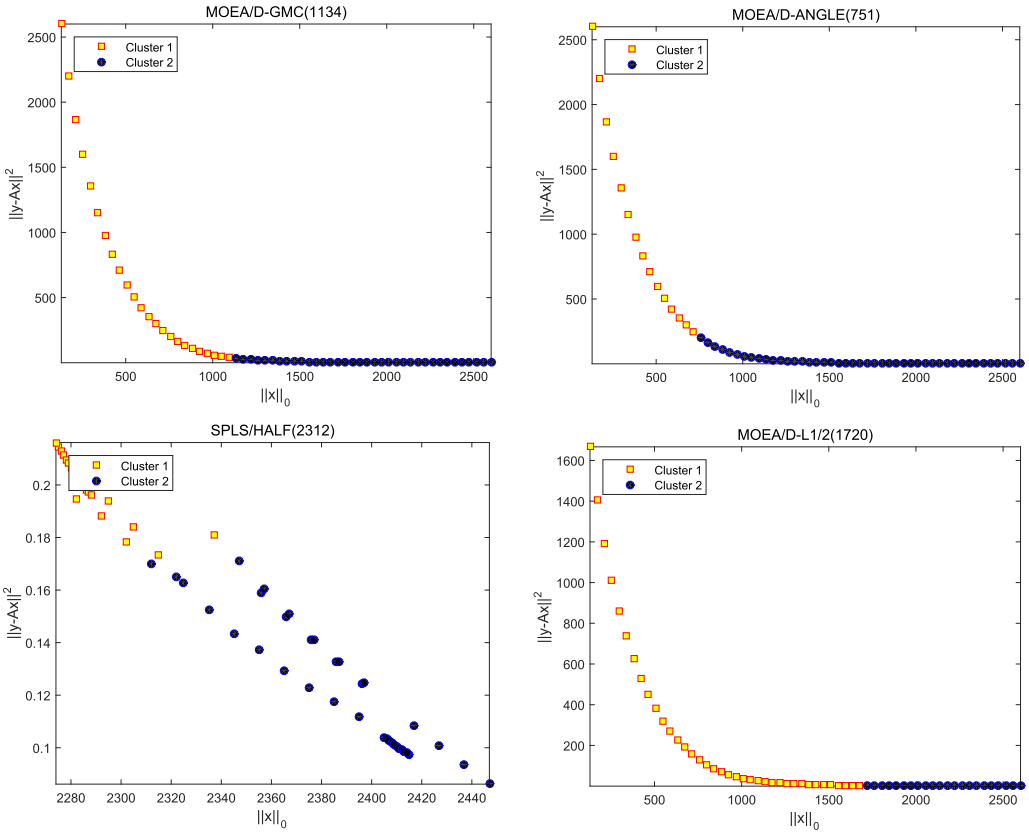


FIGURE 16. The plots of final solutions in two clusters obtained by four strategies in MOEA/D-GMC, ANGLE, SPLS, and MOEA/D-L1/2 on P4 with $\sigma = 0.1$.

V. CONCLUSION

In this paper, we proposed an improved MOEA/D based on mixture gaussian clustering, denoted by MOEA/D-GMC. In our proposed algorithm, the subproblems defined via the same objective function and various sparsity constraints are first optimized in a chain manner. The knee region is then detected by mixture gaussian clustering. Unlike other sparsity detection in MOEA/D-L1/2 and SPLS/HALF, no extra parameter is involved in mixture gaussian clustering. Our experimental results on six instance sets with four different noise levels indicated that MOEA/D-GMC is clearly superior to three other representative sparse algorithms in minimizing the median MSE values and the RSE values for sparsity detection. Moreover, MOEA/D-GMC is also suitable to detect the knee position when the sparse signal is noiseless. In our future work, we will investigate the applications of MOEA/D-GMC on some real world sparse optimization problems.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[2] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY, USA: Springer, 2010.

[3] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[4] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[5] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.

[6] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.

[7] C. A. C. Coello, "Constraint-handling using an evolutionary multiobjective optimization technique," *Civil Eng. Environ. Syst.*, vol. 17, no. 4, pp. 319–346, 2000.

[8] Z. Cai and Y. Wang, "A multiobjective optimization-based evolutionary algorithm for constrained optimization," *IEEE Trans. Evol. Comput.*, vol. 10, no. 6, pp. 658–675, Dec. 2006.

[9] Y. Wang, Z. Cai, Y. Zhou, and W. Zeng, "An adaptive tradeoff model for constrained evolutionary optimization," *IEEE Trans. Evol. Comput.*, vol. 12, no. 1, pp. 80–92, Feb. 2008.

[10] H. Li, X. L. Su, Z. B. Xu, and Q. Zhang, "MOEA/D with iterative thresholding algorithm for sparse optimization problems," in *Proc. 12th Int. Conf. Parallel Problem Solving Nature (PPSN)*, 2011, pp. 93–101.

[11] L. Li, X. Yao, R. Stolkin, M. G. Gong, and S. He, "An evolutionary multiobjective approach to sparse reconstruction," *IEEE Trans. Evol. Comput.*, vol. 18, no. 6, pp. 827–845, Dec. 2014.

[12] H. Li, Y. Y. Fan, Q. Zhang, Z. B. Xu, and J. D. Deng, "A multi-phase multiobjective approach based on decomposition for sparse reconstruction," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2016, pp. 601–608.

[13] Y. Zhou, K. Sam, G. Hainan, Z. Xiao, and Q. Zhang, "A two-phase evolutionary approach for compressive sensing reconstruction," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2651–2663, Sep. 2017.

[14] H. Li, Q. Zhang, J. D. Deng, and Z.-B. Xu, "A preference-based multiobjective evolutionary approach for sparse optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1716–1731, May 2018.

[15] B. Yan, Q. Zhao, Z. H. Wang, and J. A. Zhang, "Adaptive decomposition-based evolutionary approach for multiobjective sparse reconstruction," *Inf. Sci.*, vol. 462, pp. 141–159, Sep. 2018.

[16] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, Apr. 2002.

[17] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.

[18] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2007.

[19] P. Bühlmann and S. van de Geer, *Statistics for High-Dimensional Data: Methods, Theory and Applications* (Springer Series in Statistics). Berlin, Germany: Springer, 2011.

[20] B. K. Natraajan, "Sparse approximation to linear systems," *SIAM J. Comput.*, vol. 24, no. 2, pp. 227–234, 1995.

[21] J. Zeng, J. Fang, and Z. Xu, "Sparse SAR imaging based on $L_{1/2}$ regularization," *Sci. China Inf. Sci.*, vol. 55, no. 18, pp. 1755–1775, 2012.

[22] Y. Liang *et al.*, "Sparse logistic regression with a $L_{1/2}$ penalty for gene selection in cancer classification," *BMC Bioinf.*, vol. 14, no. 1, pp. 1–198, 2013.

[23] J. Zeng, S. Lin, Y. Wang, and Z. Xu, "$L_{1/2}$ regularization: Convergence of iterative half thresholding algorithm," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2317–2329, May 2014.

[24] P. H. Yin, Y. F. Lou, Q. He, and J. Xin, "Minimization of $\ell_{1/2}$ for compressed sensing," *SIAM J. Sci. Comput.*, vol. 37, no. 1, pp. 536–563, 2015.

[25] L. Wu, Z. Sun, and D. H. Li, "A Barzilai-Borwein-like iterative half thresholding algorithm for the $L_{1/2}$ regularized problem," *J. Sci. Comput.*, vol. 67, no. 2, pp. 581–601, 2016.

[26] Y. Li *et al.*, "MUSAI-$L_{0.5}$: Multiple sub-wavelet-dictionaries-based adaptively-weighted iterative half thresholding algorithm for compressive imaging," *IEEE ACCESS*, vol. 6, pp. 16795–16805, 2018.

[27] Z. Li, T. Hayashi, S. Ding, and Y. Li, "Dictionary learning with the $\ell_{1/2}$-regularizer and the coherence penalty and its convergence analysis," *Int. J. Mach. Learn. Cybern.*, vol. 9, no. 8, pp. 1351–1364, 2018.

[28] J. Branke, K. Deb, and M. Osswald, "Finding knees in multi-objective optimization," in *Proc. 8th Conf. Parallel Problem Solving Nature (PPSN VIII)*, in Lecture Notes in Computer Science, vol. 3242. Birmingham, U.K.: Springer, 2004, pp. 722–731.

[29] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

[30] H. Li and Q. Zhang, "Multiobjective optimization problems with complicated Pareto sets, MOEA/D and NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 13, no. 4, pp. 284–302, Apr. 2009.

[31] Q. Zhang, W. Liu, and H. Li, "The performance of a new version of MOEA/D on CEC09 unconstrained MOP test instances," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, May 2009, pp. 203–208.

[32] H. Li and D. Landa-Silva, "An adaptive evolutionary multi-objective approach based on simulated annealing," in *Evolutionary Computation*, vol. 19, no. 4. Cambridge, MA, USA, 2011, pp. 561–595.

[33] L. Ke, Q. Zhang, and R. Battiti, "MOEA/D-ACO: A multiobjective evolutionary algorithm using decomposition and antcolony," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1845–1859, Dec. 2013.

[34] K. Li, Á. Fialho, S. Kwong, and Q. Zhang, "Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 18, no. 1, pp. 114–130, Feb. 2014.

[35] H.-L. Liu, F. Gu, and Q. Zhang, "Decomposition of a multiobjective optimization problem into a number of simple multiobjective subproblems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 450–455, Jun. 2014.

[36] M. Asafuddoula, T. Ray, and R. Sarker, "A decomposition-based evolutionary algorithm for many objective optimization," *IEEE Trans. Evol. Comput.*, vol. 19, no. 3, pp. 445–460, Jun. 2015.

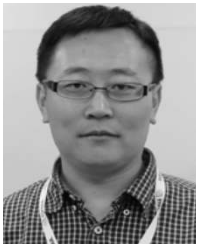
[37] Y. Yuan, H. Xu, B. Wang, B. Zhang, and X. Yao, "Balancing convergence and diversity in decomposition-based many-objective optimizers," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 180–198, Apr. 2016.

[38] H. Li, Q. Zhang, and J. D. Deng, "Biased multiobjective optimization and decomposition algorithm," *IEEE Trans. Cybern.*, vol. 47, no. 1, pp. 52–66, Jan. 2017.

[39] L. Paquete and T. Stützle, "A two-phase local search for the biobjective traveling salesman problem," in *Proc. 2nd Int. Conf. Evol. Multi-Criterion Optim.* Springer, 2003, pp. 479–493.

HUI LI (M'16–SM'17) received the B.Sc. and M.Sc. degrees in applied mathematics from the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China, in 1999 and 2002, respectively, and the Ph.D. degree in computer science from the University of Essex, Colchester, U.K., in 2008. From 2007 to 2010, he was a Postdoctoral Research Associate with the School of Computer Science, University of Nottingham, Nottingham, U.K. He is currently an Associate Professor with the School of Mathematics and Statistics, Xi'an Jiaotong University. His current research interests include evolutionary computation, multiobjective optimization, and machine learning. He was a recipient of the 2010 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award as one of the inventors for MOEA/D.

JIANYONG SUN (M'15–SM'18) received the B.Sc. degree in computational mathematics from Xi'an Jiaotong University, Shaanxi, China, in 1997, and the M.Sc. and Ph.D. degrees in applied mathematics and computer science from the University of Essex, Colchester, U.K., in 1999 and 2005, respectively. He was a Senior Lecturer with the Faculty of Engineering and Science, University of Greenwich, Greenwich, U.K. He is currently a Professor with Xi'an Jiaotong University. He has published over 30 peer-reviewed papers on evolutionary algorithms and statistical machine learning in prestigious journals, such as the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON COMPUTATIONAL BIOLOGY AND BIOINFORMATICS, and PNAS. His research interests include evolutionary computation and statistical machine learning, and their applications in computational biology, bioinformatics, image processing, and big data analytics.



DEYU MENG received the B.Sc., M.Sc., and Ph.D. degrees from Xi'an Jiaotong University, Xi'an, China, in 2001, 2004, and 2008, respectively, where he is currently a Professor with the Institute for Information and System Sciences. He was a Visiting Scholar with Carnegie Mellon University, Pittsburgh, PA, USA, from 2012 to 2014. His current research interests include self-paced learning, noise modeling, and tensor sparsity.

QINGFU ZHANG (M'01–SM'06–F'17) received the B.Sc. degree in mathematics from Shanxi University, Taiyuan, China, in 1984, and the M.Sc. degree in applied mathematics and the Ph.D. degree in information engineering from Xidian University, Xi'an, China, in 1991 and 1994, respectively. He is currently a Professor with the Department of Computer Science, City University of Hong Kong, Hong Kong, and a Changjiang Visiting Chair Professor with Xidian University, China. He is also leading the Metaheuristic Optimization Research Group, City University of Hong Kong. MOEA/D, a multiobjective optimization algorithmic framework, developed in his group, is one of the most widely used and researched multiobjective evolutionary algorithms. His current research interests include evolutionary computation, optimization, neural networks, data analysis, and their applications. He is an Editorial Board Member of three other international journals. He was a recipient of the 2010 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award. He was a 2016 Web of Science highly cited researcher in computer science. He is also an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and the IEEE TRANSACTIONS ON CYBERNETICS.