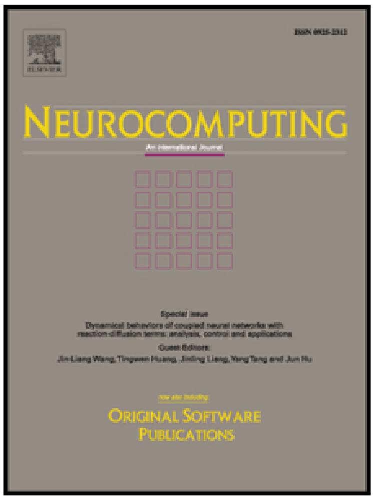
Accepted Manuscript



Objective Reduction Particle Swarm Optimizer based on Maximal Information Coefficient for Many-Objective Problems

Yi Liang , Wangli He , Weimin Zhong , Feng Qian

PII: S0925-2312(17)31550-3
DOI: 10.1016/j.neucom.2017.09.033
Reference: NEUCOM 18913

To appear in: *Neurocomputing*

Received date: 9 May 2017
Revised date: 21 August 2017
Accepted date: 11 September 2017

Objective Reduction Particle Swarm Optimizer based on Maximal Information Coefficient for Many-Objective Problems

Yi Liang, Wangli He, Weimin Zhong*, Feng Qian*

Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China

Abstract It is challenging to solve reducible many-objective problems due to difficulties caused by the unknown number of non-conflicting objectives. Objective reduction method is one of promising and efficient solutions in which two fundamental problems should be addressed: how to find the redundant objectives and which objectives should be selected or omitted. A novel objective reduction algorithm is proposed in this paper, named Maximal Information Coefficient based Multi-Objective Particle Swarm Optimizer (MIC-MOPSO). By a powerful MIC indicator, the algorithm could find hidden linear or nonlinear relationships between two objectives. Another indicator, the change rate of non-dominated population, is used to judge whether there exist non-conflicting objectives or not. An effective way to rapidly select the retained objectives is also developed based on these two indicators. Tested by a series of benchmark experiments and a real industrial optimization problem, the results show that our approach significantly improve the performance on both reducible and irreducible many-objective problems.
Keywords maximal information coefficient, particle swarm optimization, objective reduction, many-objective problems

1. Introduction

Many industrial problems that involve simultaneous optimization of several objective functions could be treated mathematically as multi-objective optimization problems (MOPs). It is well known that multi-objective evolutionary algorithms (MOEAs), such as Non-dominated Sorting Genetic Algorithm II(NSGA-II) [1], Strength Pareto Evolutionary Algorithm (SPEA) [2], perform well in MOPs with two objectives. However, they often fail in the MOPs when the objective number is larger than three. Generally speaking, a MOP with more than three objectives is commonly termed as a Many-Objective Problem (MaOP).

Ishibuchi et al. [3] argued that three main reasons may lead traditional multi-objective algorithms failed in MaOPs. These distinct challenges made MaOPs become much more difficult optimization problems than MOPs.

- When the number of objectives increases, almost all solutions in each population become non-dominated. This severely weakens the Pareto dominance-based selection pressure toward the Pareto front. That means the convergence property of heuristic algorithms is severely deteriorated.
- The requirement of the number of solutions is exponential increased for approximating the entire Pareto front. Since the Pareto front is a hyper-surface in the objective space, the number of solutions required for its approximation exponentially increases with the dimensionality of the objective space, which may need thousands of non-dominated solutions to approximate the entire Pareto front of a many-objective problem.
- The visualization of the entire Pareto front becomes increasingly difficult.

Over the past decades, many valuable efforts have been dedicated to overcome the problems mentioned above. These improvements could be briefly classified into five directions:

- Pareto dominance modification methods: Sato et al. [4] addressed that using a modified dominance could increase the selection pressure toward the Pareto front and improve the performance of NSGA-II for MaOPs. Instead of the standard Pareto dominance, a more relaxing dominance relation was adopted and some non-dominated solutions became dominated by others. Similar improvements were also proposed such as dominance relaxing incorporated with user preference [5], ϵ-dominance [6], k-optimality [7], objective subspace partition [8] and the control of dominance area of solutions (CDAS) [9, 10].
- Aggregation-based methods: The idea behind this kind of algorithms is to transfer MaOPs to a group of weighted sub-problems. For example, MOEA/D [9] optimized a number of single-objectives in parallel which were defined by the weighted sum or weighted Tchebycheff function.
- Indicator-based methods: Zitzler and Künzli [10] proposed a framework of indicator-based evolutionary algorithms (IBEAs), while preference based approaches could effectively limite heuristic algorithm to a sub-region that do not need many candidate solutions.
- Preference information methods: Fleming et al. [11] proposed a method by preference articulation to

* Corresponding authors
E-mail addresses: yliangnwpu@163.com (Yi Liang), {wanglihe, wmzhong, fqian}@ecust.edu.cn (Wangli He, Weimin Zhong, Feng Qian)

many-objective problems where the focused region gradually became smaller during MOEAs. Reference point-based preference methods is also widely used. Deb and Sundar [12] incorporated reference point-based preference information into NSGA-II. Britto and Pozo [13] proposed a reference point based MOPSO to update the external archive. Thiele et al. [14] used reference point-based information in IBEA.

- Objective reduction methods: For some MaOPs, the dimension of the true Pareto front might be less than the number of objectives, which means some objectives are redundant and do not contribute for the Pareto front. Objective reduction approaches aim to identify the redundant objectives and reduce the dimension of problem and then existing Pareto-based algorithms for MOPs could be used. Refs. [15-21,31-32] proposed different kinds of reduction based methods. A more thorough review could be seen in the following.

The objective reduction method is a promising choice to solve MaOPs as it could transfer reducible many-objective to multi-objective problem and the dimensional difficulties will dramatically decrease. The main goal is to remove unnecessary objectives while maintaining the shape of the Pareto front in the reduced objective space. The existing objective reduction-based approaches include the following directions [21]:

- Dominance relation preservation-based objective reduction: Brockhoff and Zitzler [15] adopted a measure of variation of the dominance structure and a general notion of minimum objective subsets (MOSS) problems based on an error term δ. The algorithms in [15] intended to find a δ-minimal objective set which cannnot be further reduced without changing the dominance structure with an error of at most δ.
- Pareto corner search: Singh et al. [16] proposed the Pareto corner search evolutionary algorithm (PCSEA) to identify the relation among objectives and got a conflicting objective subset by using the information of the corner of Pareto fronts.
- Machine learning-based objective reduction: This kind of algorithms treat objective reduction as feature selection. The critical aim is to find a proper feature to distinguish potential non-conflicting objectives from conflicting ones. Techniques in machine learning, such as correlation matrix [17], principal component analysis (PCA) [18] and maximum variance unfolding (MVU) [19, 20], were used in the literature.
- Mutual information-based objective reduction: Most of the relationship among objectives is non-linear, and mutual information is a useful technique to identify non-linear correlation between variables. Wang and Yao [21] adopted a non-linear correlation information entropy (NCIE) to detect redundant objectives.

The objective reduction methods mentioned above have their own limitation in practice. Some could not be used in the complex non-linear relationship, some might misjudge the conflicting objectives. In this paper, a novel objective reduction algorithm is proposed, named Maximal Information Coefficient based Multi-Objective Particle Swarm Optimizer (MIC-MOPSO). Maximal Information Coefficient (MIC) [22] is a method which could effectively identify variable correlation between two objectives. Using MIC, the algorithm could find latent correlation between objectives, which may be ignored by other correlation method such as correlation coefficient method. Also in MIC-MOPSO, a more effective approach, the change rate of non-dominated population, is proposed to judge whether there is non-conflicting objective or not, and it could greatly accelerate objective selection process without performance loss.

This paper is organized as follows. In Section 2, the conflicting relationships are defined and different cases of redundant objectives in MaOPs are shown. In Section 3, the MIC-MOPSO algorithm is proposed. Section 4 presents the experimental results and some discussions are also given. Then MIC-MOPSO is used to solve a real four objective optimization problem of entrained flow coal gasification process with one redundant objective. Finally, Section 5 gives the conclusion.

2. Reducible many-objective problem

2.1 Conflicting relationship

Intuitively, the conflicting relationship between two objectives means that, for example in a minimum optimization problem, when one objective reaches the minimal value, the other could not get its minimum, even maybe reach its maximal point. In other words, the values of these two objectives could not increase or decrease simultaneously. On the contrary, if two objectives are not conflicting, they are non-conflicting or harmonious [23]. Without loss of generality, in this paper we consider a minimization problem with M-objective functions.

Purshouse et al. [23] proposed formal definitions of such relationships, as shown in *Definition 1-Defintion 3* below. Let i and j be indices to particle objective: $i, j \in [1, \cdots, M]$,where M is the objective number. Let a and b be indices to objective value vector instances: $a, b \in [1, \cdots, |X|]: x_a, x_b \in X$, where X is a particular region of all feasible values. Also let (a, b) denote a pair of instances when $a \neq b$.

Definition 1. Conflicting relationship: objectives i and j exhibit evidence of conflicting relationship if the condition $(x_a^i < x_b^i) \wedge (x_a^j > x_b^j)$ is satisfied. If $\exists(a, b)$ such that the condition holds $\forall(a, b)$, then there is total conflict [23].

Definition 2. Harmonious or non-conflicting relationship: objectives i and j exhibit evidence of harmonious or non-conflicting relationship if the condition $(x_a^i < x_b^i) \wedge (x_a^j < x_b^j)$ is satisfied. If $\exists(a, b)$ such that the condition holds $\forall(a, b)$, then there is total harmony [23].
Definition 3. Weak harmonious relationship: objectives i and j exhibit evidence of weak harmonious relationship if the condition $[(x_a^i < x_b^i) \wedge (x_a^j = x_b^j)] \vee [(x_a^i = x_b^i) \wedge (x_a^j < x_b^j)]$ is satisfied. If $\exists(a, b)$ such that the condition holds $\forall(a, b)$, then there is total weak harmony [23].

2.2 Reducible many-objective problems

Many-objective problems do not always only have objectives conflicting with each other, some of them may be non-conflicting or latent non-conflicting. If we omit some non-conflicting ones in the optimization process, the sequence of non-dominance will not change, and will not influence the final Pareto front. Those non-conflicting objectives are named redundant objectives. MaOPs which contain redundant objectives are reducible many-objective problems. Let us take functions in table 1 as an example. In case 1, $f_1(x)$ and $f_2(x)$ are conflicting, but $f_2(x)$ and $f_3(x)$ are non-conflicting. Assumed that in the objective space, there are solutions $x_1 = (0,1,0)$, $x_2 = (1,0,2)$, $x_3 = (2, -1,4)$, and all of them are non-dominant solutions. If we omit the third element of solutions, the dominant relation will not change. This could also be seen on its parallel coordinate graph on Fig. 1, in which crossing lines referring to the objectives are conflicting and parallel lines are non-conflicting with each other. $f_2(x)$ and $f_3(x)$ have no crossing line. Therefore, the objective $f_3(x)$ will not influence the final Pareto optimal results.

Table 1

Examples of minimal MaOPs with redundant objectives, where $x \in [0,1]$

Case 1	$f_1 = x$
	$f_2 = 1 - x$
	$f_3 = 1 - 2x$
Case 2	$f_1 = x$
	$f_2 = 1 - x$
	$f_3 = 1$
Case 3	$f_1 = x$
	$f_2 = 1 - x$
	$f_3 = \begin{cases} -0.8x + 0.9, x \in [0,0.4) \\ -0.4x + 0.7, x \in [0.4,0.5) \\ x - 0.5, x \in [0.5,1] \end{cases}$
Case 4	$f_1 = -6x + 4$
	$f_2 = 6x + 1$
	$f_3 = -(6x - 4)^2 + 10$

Fig. 1 represents some conflicting relationships on the parallel coordinate graph of functions in table 1. Each line represents a solution while x-axis refers to the objective and y-axis represents the values of objective functions. It could be seen from these graphs how the conflict varies and also the conflict could be concentrated at certain values. In case 2, $f_1(x)$ and $f_2(x)$ are conflicting, but $f_3(x)$ is not conflicting with $f_2(x)$ because of $f_3(x)$ is a constant value. Case 3 and case 4 are MaOPs with local conflicts. In case 3, $f_3(x)$ is a piecewise function, it conflicts with $f_2(x)$ for low values. Case 4 is on the opposite side which conflicts for high values. The existing objective reduction methods often focused on problems such as case 1 and 2. However, they could not effectively handle with problems similar to case 3 or 4. In this paper, a novel method is proposed to solve MaOPs with local conflicts.

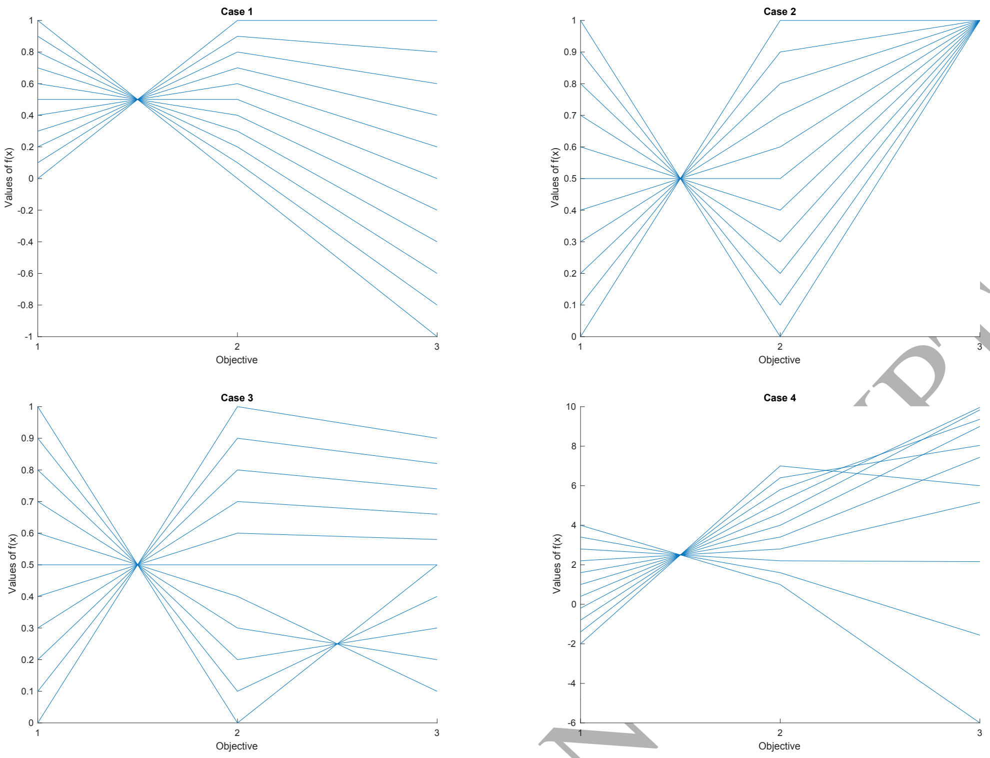


Fig. 1. Parallel coordinate graphs of cases 1-4

3. Maximal information coefficient based multi-particle swarm optimizer

In order to solve MaOPs with reducible objectives, one of the critical questions is how to identify the redundant part from all objectives. Intuitively, if two objectives are non-conflicting, their values should increase or decrease simultaneously and might exist some certain relationships between each other. Therefore, finding a method to identify the relationship and estimate whether the relation is conflicting or not will answer the question above. In this paper, a new algorithm hybridizing with MOPSO is proposed to solve objective reduction problem. By computing in parallel maximal information coefficient and the change rate of non-dominated population, each objective will be given a score based on these two indicators. The higher the score is, the more possibility of redundancy the objective is. The non-conflicting objectives will be omitted and most conflicting will be retained. And this objective selection process could find potential non-conflicting relationship. After selection, MOPSO [24] will optimize the reduced MaOPs. The details of the proposed algorithm, MIC-MOPSO, will be discussed in the next sub-section and the flowchart is shown in Fig. 2.

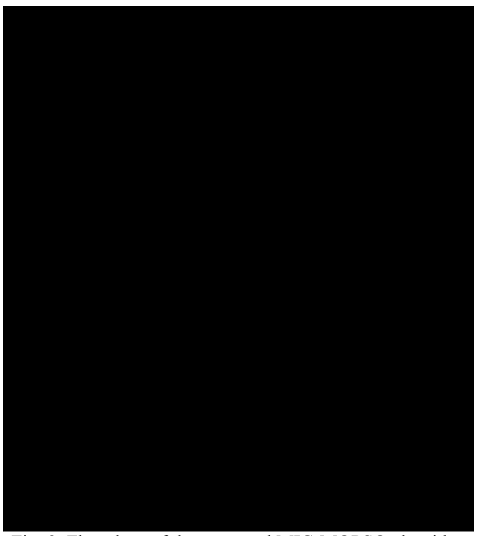


Fig. 2. Flowchart of the proposed MIC-MOPSO algorithm

3.1 Maximal Information Coefficient

When it comes to explore relationship in a data set, several kinds of methods could be considered as candidate measurements, such as method based on correlation coefficient [17] is able to estimate linear relationship among the objectives, method based on mutual information entropy [21] could find either linear or nonlinear relationship. Practically, nonlinear relationship between two objectives is more common and more difficult to discover in many real optimization problems. Hence, a method to find potential nonlinear relationship is one of the core techniques in objective reduction. Maximal Information Coefficient (MIC) [22] is a powerful approach to detect various relation of variables, which belongs to a large class of maximal information-based non-parametric exploration statistics for identifying and classifying relationships. MIC captures a wide range of associations both functional and not, which is useful for data-driven optimization problems. For functional relationship, it provides a score that roughly equals the coefficient of determination of the data relative to the regression function. Fig. 3 shows a bar graph of the ability of MIC, compared with Pearson correlation coefficient and Nonlinear Correlation Information Entropy (NCIE) [21] which is a kind of mutual information methods, to find various functional relationship, including strong nonlinear functions. For comparison, the absolute value of Pearson correlation coefficient is used.

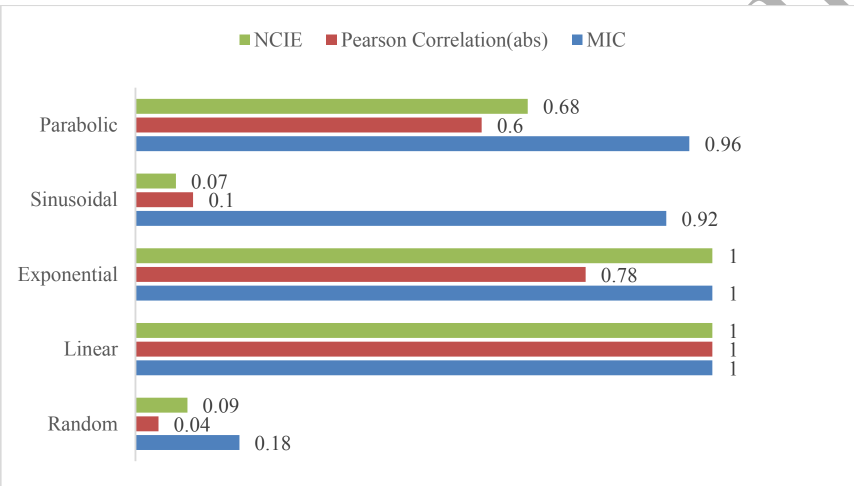


Fig. 3. Comparison of MIC to Pearson Correlation Coefficient (absolute value) and NCIE in various noiseless functional relationships.

MIC is based on the idea that if there exists certain relationship between two variables, a grid could be drawn on the scatter plot of two variables to encapsulate this relationship and to get the maximal information entropy simultaneously. The detail procedure of MIC method is as follows:

Step 1: Partition scatter plot of variables with a resolution pair (x, y), where $(x, y) \in \mathbb{N}$

Step 2: Compute the largest possible mutual information I_{xy} by (x, y) grid

Step 3: Repeat Step 1 and 2 with different pairs (x, y) each time, until reach the maximal grid resolution which depends on the sample size

Step 4: Normalize all of maximal mutual information values $I_{xy} \in [0,1]$

Step 5: Generate the characteristic matrix M_{xy}

$$M_{xy} = \frac{I_{xy}^*}{log\{min\{x,y\}\}} \tag{1}$$

where I_{xy}^* is the normalized maximal mutual information

Step 6: Compute $\mathrm{MIC} = max\{M_{xy}\}$

3.2 Change Rate of non-dominated population

By means of MIC matrix, it could measure whether a statistical significant correlation exists or not between two objectives. However, it is not sufficient for objective reduction because that it is still do not know whether such objectives are conflicting or not. Roughly speaking, it is difficult to directly observe the non-conflicting evidence among the objectives, especially for those with potential non-conflicting relationship. Take $DTLZ5(I, M)$ as an example [25], and the corresponding Eq. (2) is shown below, where I denotes the conflicting number, and is also the dimensionality of the Pareto-optimal front and M is the number of objectives in the problem. For $I = 2$, a minimum of two objectives will be enough to represent the true Pareto-optimal front. When $g(x_M)=0$, there is a relationship $f_1(x)=c_1f_2(x)=c_2f_3(x)=\cdots=c_{M-2}f_{M-1}(x)$, and $f_1(x), f_2(x),\cdots,f_{M-1}(x)$ become non-conflicting with each other. While the increasing of the number of objectives, it will be more difficult to reach the condition $g(x_M)=0$ and also more difficult to find reducible objective efficiently. In this sub-section, a new indicator, change rate of non-dominated population σ, is proposed to find reducible objective in advance, which could be also decrease the computational cost.

$$\begin{cases} \text{Min} \quad f_1(x)=\big(1+100g(x_M)\big)\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{M-2})\cos(\theta_{M-1}), \\ \text{Min} \quad f_2(x)=\big(1+100g(x_M)\big)\cos(\theta_1)\cos(\theta_2)\cdots\cos(\theta_{M-2})\sin(\theta_{M-1}), \\ \qquad \vdots \\ \text{Min} \quad f_{M-1}(x)=\big(1+100g(x_M)\big)\cos(\theta_1)\sin(\theta_2), \\ \text{Min} \quad f_M(x)=(1+100g(x_M))\sin(\theta_1), \\ \text{where} \quad g(x_M)=\sum_{x_i\in x_M}(x_i-0.5)^2, \\ \qquad \theta_i=\begin{cases} \frac{\pi}{2}x_i, & \text{for } i=1,\cdots(I-1), \\ \frac{\pi}{4(1+g(x_M))}(1+2g(x_M)x_i), & \text{for } i=I,\cdots,(M-1) \end{cases}, \\ \qquad 0\le x_i\le 1, \quad \text{for } i=1,2,\cdots,n. \end{cases} \tag{2}$$

When two objectives are harmony or non-conflicting, it actually means that the values of two objectives increase or decrease simultaneously, and only one non-dominated point, namely the optimal value, rules all other points. On the contrary, if there is only one non-dominated exists, two objectives are non-conflicting relationship. For those reducible objectives, the non-dominated set will shrink as optimization is going and finally converge to one point. Consider $DTLZ5(2,5)$ again, Fig. 4 (a)-Fig. 4 (c) represent the distribution change of MOPSO population on $f_1(x)$ and $f_2(x)$, where x in red refers to non-dominated particles. It is observed that non-dominated particles become fewer and finally converge to one particle. Fig. 4 (d) is the average distance of non-dominated population along with generation. Such convergence trend could also be seen from Fig. 4 (d). By the change rate of average distance and number of non-dominated population, we could analyze whether two objectives are conflicting.

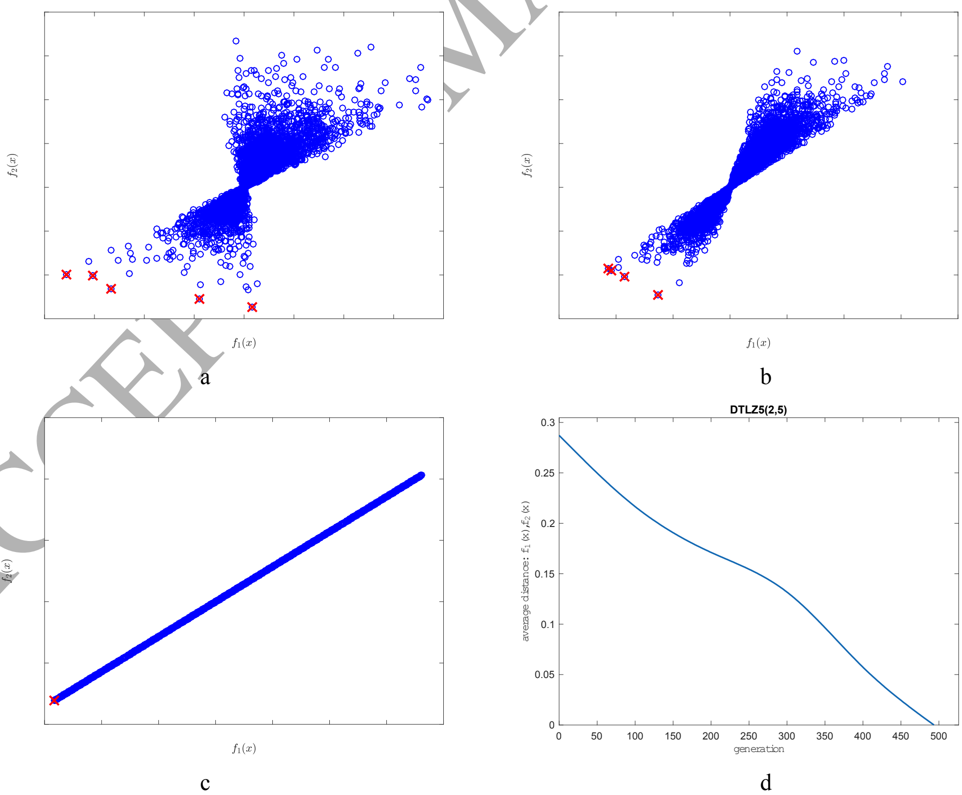


Fig. 4. Distribution change of MOPSO population on $f_1(x)$ and $f_2(x)$ of DTLZ5(2,5).

Firstly, we will discuss the change rate of average distance of non-dominated population. To compute it, the most direct method is to calculate the distance change of two consecutive generations, as is shown in (3), where $\Delta d_i=d_{i+1}-d_i$, d_{i+1} denotes the i_{th} average distance of non-dominated population, i is the i_{th} generation. The average distance of two consecutive generations might vary in a large range, and for computational simplicity, set $d\in[-1,1]$. Positive value means the decrease of d and negative value means the increase of d. The larger d is, the more possibility two objectives are harmony. Therefore, Eq. (3) is adapted to Eq. (4). However, there are still two important situations should be considered:

1. If two change rates of average distances are the same, for example one is $d_i = 2 \to d_{i+1} = 1$ and the other is $d_i = 0.2 \to d_{i+1} = 0.1$, it is obvious that the latter has more possibility of non-conflicting than the former.
2. If two change rates of average distances are equal to 1, for example one is $d_i = 5 \to d_{i+1} = 5$ and the other is $d_i = 0.1 \to d_{i+1} = 0.1$, the latter has more possibility of non-conflicting than the former.

$$d_1 = \frac{\Delta d_i}{d_{i+1}} \tag{3}$$

$$d_1 = \begin{cases} 1 & \Delta d_i = 0 \\ \frac{\Delta d_i}{d_{i+1}} & \Delta d_i > 0 \\ \frac{\Delta d_i}{d_i} & \Delta d_i < 0 \end{cases} \tag{4}$$

$$d_2 = 1 - \frac{d_{i+1}}{\sum d_{i+1}} \tag{5}$$

$$d = d_1 \times d_2 \tag{6}$$

To handle such problems, we should consider the contribution of d_{i+1} itself and the proportion of d_{i+1} in all of the i_{th} average distance $\sum d_{i+1}$. For instance, in DTLZ5(2,3), we have the average distances $d_i^{1,2}$=10 for objective pair $(f_1(x), f_2(x))$, $d_i^{1,3}$=1 for $(f_1(x), f_3(x))$, and $d_i^{2,3}$=0.1 for $(f_2(x), f_3(x))$ on the i_{th} generation. It is observed that $(f_2(x), f_3(x))$ has the most possibility of non-conflicting, and the proportion of $d_i^{2,3}$ is 0.009. If larger value means more harmonious, Eq. (5) could be got. Therefore Eq. (4) is revised to Eq. (6).

Practically, the change rate of average distance sometimes misleads the relationship analysis. Therefore, we combine the change rate of the number of non-dominated population to remedy the mistake. Because the proposed algorithm is based on MOPSO and MOPSO has fixed eternal archive to retain the non-dominated population, it is easy to use the following equation to calculate the change rate,

$$n = \frac{n_i}{N} \tag{7}$$

where n_i represents the i_{th} number of non-dominate population and N is the total archive number of MOPSO.

$$\alpha = d \times n \tag{8}$$

Finally, the change rate of non-dominated population is computed according to Eq. (8), and usually for computational convenience, a change rate matrix of all objective pairs A is used, which has the same size of MIC.

3.3 Objective selection

With MIC and the change rate of non-dominated population, the information of conflicting objective relationship is given. The next step is to determine which objectives should be omitted to achieve the goal of objective reduction. In this paper, a simple way is adopted to just multiply MIC and Change Rate of α, which is in the range of [-1,1]. Then to sum all the scores of each objective, we get a final score of all objectives. The higher the score is, the more possibility to be omitted. The idea behind this method is that the smaller negative objective means it has more conflicting correlation with other objectives, and to omit the objective with less conflicting correlation. Sometimes, the final score is positive, but quite small such as 0.1, we consider this situation as the objective has few positive relationship and will be not omitted. A threshold *T* is set, only when *Score*>*T* the objective will be omitted. The details are shown in Algorithm 1, where S_t is the objective set, S_r is the selected set or retained set.

Algorithm 1. Pseudo code of objective selection of MIC-MOPSO

Initialization: MIC matrix and the change rate matrix of non-dominated population $\boldsymbol{A}$ of the current generation of all objective pairs, the number of objectives M, $\boldsymbol{S_t}$=[1:M], $\boldsymbol{S_r}$=∅, and threshold $\boldsymbol{T}$
$\boldsymbol{S = MIC \times A}$
For *i*=1:M
  $\boldsymbol{Score_i = \sum S_i}$
  IF $\boldsymbol{Score_i < T}$
    $\boldsymbol{S_r = \{S_r, i\}}$
  Else
    $\boldsymbol{S_r = S_r}$
End For
End

3.4 Computational complexity

For an M objective problem with a decision variables set of size N and a fixed population size K, the MIC matrix has complexity $O(M^2K)$, and the change rate of non-dominated population is $O(MK^2N)$. Therefore, the total complexity of MIC-MOPSO is $O(M^2K^2N)$ at each generation.

4. Experimental analysis

4.1. Test problems and settings

The experiments adopt the well-known MaOPs benchmark problems, the $DTLZ$ test suite [26]. Among these test problems, $DTLZ1-DTLZ4, DTLZ6-7$ are irreducible MOPs, and the reducible MaOP $DTLZ5(I,M)$ is used instead of original $DTLZ5$, where I represents the dimensionality of the Pareto front and M is the number of objectives in the problem. All of comparative algorithms in the following are repeated for 50 independent runs and population size is set to 1000. The testing machine is a Personal Computer with 2.7 GHz Intel® i5 CPU and 8G RAM.

4.2 IGD performance on irreducible DTLZ problems

Firstly, the optimization performance of irreducible MaOPs (DTLZ1-7 except for DTLZ5) are tested between MIC-MOPSO and other well-known algorithms solving MaOPs. IGD [27, 28] is chosen to quantify the performance of algorithms. For the IGD metric, the true Pareto fronts of the test instances are required as the referenced fronts to measure the performance metric. The more the samples of a true Pareto front are available, the better the IGD metric will be for an MaOP problem, yet the higher the computational cost will be incurred.

Table 2

IGD values of the different algorithms on DTLZ1-4 by Mann-Whitney *U* test in the form of "*mean(std.) #*", where # stands for one of the significance symbol (+, = or -) of U-test.

Function	M	MIC-MOPSO	MOPSO	NSGA-II	PCA-NSGA-II	NSGA-II (NCIE)	IBEA	IBEA（NCIE）	MOEA/D
DTLZ1	3	**3.31E-2 (3.80E-2)**	1.10E+0 (4.69E+0) +	4.97E+0 (1.136E+0) +	3.78E+0 (8.24E+0) +	3.52E+0 (5.41E+0) +	2.82E+0 (8.24E+0) +	1.81E+0 (3.14E+0) +	7.21E-2 (3.41E-2) +
	5	**3.01E-1 (1.59E-2)**	3.16E+0 (8.44E+0) +	2.91E+1 (3.51E+1) +	3.41E+0 (9.43E+0) +	3.14E+0 (1.45E+0) +	1.23E+0 (9.14E-1) +	4.13E-1 (3.95E-1) +	4.71E-1 (4.35E-1) +
	15	**9.35E-1 (9.68E-1)**	2.11E+1 (5.90E+1) +	3.65E+1 (2.31E+1) +	2.98E+1 (1.39E+1) +	1.58E+1 (1.71E+1) +	7.52E+0 (1.88E+0) +	2.91E+0 (4.13E-1) +	5.80E+0 (2.59E+0) +
DTLZ2	3	1.10E-1 (1.01E-1)	8.51E-1 (4.12E-1) +	1.04E+0 (6.52E-2) +	1.01E-1 (5.31E-2) =	**4.13E-2 (9.10E-2) -**	8.71E-2 (2.4E-2) -	1.14E-1 (1.01E-1) =	5.86E-2 (2.60E-3) -
	5	1.22E+0 (1.31E-1)	2.28E+0 (2.76E-1) +	1.28E+0 (3.87E-1) +	1.38E+0 (2.50E-1) +	5.18E+0 (2.10E-1) +	1.87E-1 (4.00E-3) -	**1.25E-1 (3.15E-3) -**	4.71E-1 (2.51E-3) -
	15	9.41E-1 (5.09E-1)	3.51E+1 (5.30E+1) +	2.71E+0 (1.81E-1) +	1.39E+0 (3.18E-2) +	2.49E+0 (3.19E-1) +	2.66E-1 (4.13E-2) -	**1.52E-1 (9.01E-3) -**	2.83E-2 (5.12E-3) -
DTLZ3	3	**3.93E-2 (7.27E-2)**	6.90E+0 (2.08E+0) +	9.69E+0 (2.14E+0) +	2.44E+0 (3.49E-1) +	1.29E+0 (6.22E-1) +	4.96E+0 (9.52E+0) +	4.64E-1 (1.14E-1) +	5.05E-1 (1.42E-1) +
	5	**3.37E-1 (6.43E-1)**	2.32E+1 (3.56E+1) +	2.09E+2 (2.01E+1) +	2.18E+1 (3.03E+1) +	1.29E+2 (6.20E+1) +	6.32E+0 (4.21E+0) +	4.39E+1 (3.72E+1) +	4.21E+0 (9.44E+0) +
	15	**3.91E-1 (2.04E-1)**	2.77E+2 (1.06E+0) +	2.32E+2 (1.35E+1) +	1.19E+2 (2.57E+1) +	2.16E+2 (1.53E+1) +	4.13E+0 (9.86E+0) +	3.93E+1 (1.64E+1) +	5.95E+1 (7.13E+1) +
DTLZ4	3	**5.99E-2 (4.00E-2)**	6.11E-1 (0E+0) +	1.41E+0 (7.50E-2) +	2.11E-1 (4.00E-2) +	1.34E-1 (6.02E-1) +	1.09E-1 (1.00E-1) +	1.36E-1 (4.20E-3) +	3.30E-1 (1.81E-1) +
	5	**2.11E-1 (9.40E-2)**	3.61E+0 (2.14E-2) +	1.34E+0 (3.99E-2) +	4.66E-1 (5.62E-1) +	7.89E-1 (3.02E-1) +	2.96E+0 (1.40E-1) +	1.16E+0 (2.00E-4) +	3.34E-1 (7.23E-2) =
	15	**1.04E-1 (4.20E-3)**	1.77E+0 (4.50E+0) +	1.16E+0 (1.14E-2) +	1.00E+0 (9.13E+0) +	1.29E+0 (6.34E-2) +	1.33E+0 (0E+0) +	1.33E+0 (0E+0) +	1.02E+0 (6.89E+0) +
DTLZ6	3	**6.11E-3 (4.13E-3)**	7.28E-1 (1.00E-2) +	9.91E-1 (2.22E-2) +	4.34E-2 (0E+0) +	2.66E-2 (6.12E-2) +	1.61E-1 (9.80E-3) +	1.58E-2 (4.00E-2) +	1.45E-2 (4.72E-4) +
	5	2.34E-1 (4.37E-2)	3.87E+0 (1.01E+0) +	6.22E+0 (4.50E+0) +	4.11E+0 (9.19E+0) +	7.03E+0 (3.11E-1) +	5.31E-1 (3.40E-2) =	4.98E-1 (1.00E-2) =	**7.24E-2 (2.34E-2) -**
	15	3.10E+1 (0E+0)	9.90E+1 (2.81E+0) +	1.20E+2 (1.31E-1) +	3.72E+1 (2.50E-1) =	5.69E+1 (3.01E+0) =	1.88E+0 (2.10E-1) -	**9.01E-1 (3.01E-2) -**	3.91E+0 (2.23E-1) -
DTLZ7	3	**4.92E-2 (3.12E-2)**	1.31E+0 (4.00E-2) +	1.49E+0 (1.10E+0) +	1.81E-1 (2.58E-2) +	1.33E+0 (2.22E-2) +	2.04E-1 (1.45E-2) +	9.82E-2 (8.31E-2) +	1.49E-1 (4.12E-2) +
	5	**2.50E-1 (8.12E-2)**	1.10E+0 (9.00E-2) +	2.00E+0 (0E+0) +	1.52E+0 (3.72E-2) +	9.10E-1 (0E+0) +	9.93E-1 (1.20E-1) +	3.49E-1 (2.13E-3) =	3.79E-1 (2.00E-2) =
	15	**9.13E-1 (1.31E-1)**	1.59E+1 (3.11E+1) +	7.41E+1 (3.10E+0) +	8.72E+0 (4.13E+0) +	8.99E+0 (3.14E-1) +	7.09E+0 (2.44E-1) +	1.33E+0 (6.56E-2) +	1.43E+0 (4.56E-2) +
Better (+)			18	18	16	16	13	12	11
Same (=)			0	0	2	1	1	3	2
Worse (-)			0	0	0	1	4	3	5

To show the performance of our approach clearly, MIC-MOPSO is compared with MOPSO [24], NSGA-II [1], PCA-NSGA-II [18], MOEA based on NCIE [21], IBEA[10], IBEA based on NCIE[21] and MOEA/D [9] which could be seen in Table 2. In order to provide the statistical quantifications on IGD performance metric, the nonparametric statistical hypothesis test, the Mann-Whitney-Wilcoxon rank-sum test (U-test) [30] is applied and after each value one of the three significance symbols (+, = or -) is marked in Table 2. With the ability of objective reduction, the MIC-MOPSO has a better performance on most benchmark problems than other algorithms, especially for those with large number of objectives. However, the IGD values are not greatly improved by our approach on the hard-to-converge problem DTLZ2.

4.3 Parameter sensitivity of T-threshold

The performance of MIC-MOPSO varies with different Ts due to the value of threshold T effects the final selected objective set. Fig. 5 shows the sensitivity of threshold T with reducible (left) and irreducible (right) MaOPs over 200 generations in 50 independent runs. Comparing with the two sub-figures, we find that the value of T affects the behavior of our approach on reducible problems more than irreducible problems, especially for problem with large number of objective. And for the globally irreducible problem DTLZ2, the proposed MIC-MOPSO still could find some locally redundant objectives. If T is set to be too large, some redundant objectives would not be removed, which would waste the computational expense. If T is too small, the conflicting objectives would be abandoned. Therefore, a robust T is very important to the algorithm, and in this paper, T is set to be 0.2.

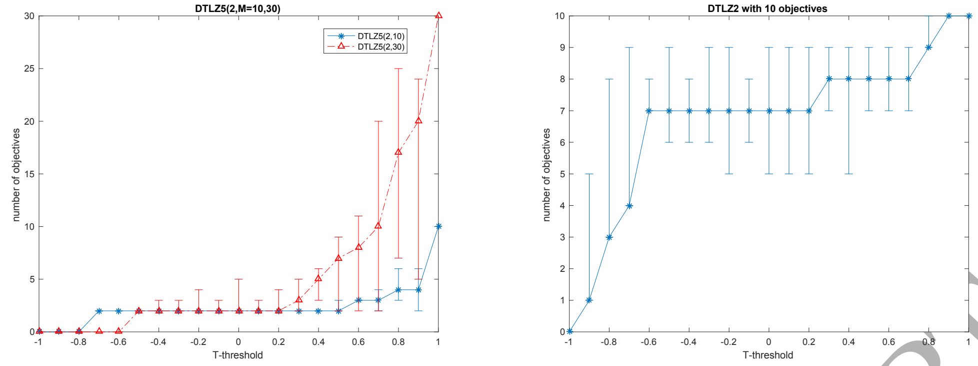


Fig. 5. Median number of objectives after reduction with different *T* s on DTLZ5(2, M=10, 30) (left) and DTLZ2 with 10 objectives (right).

4.4 Comparison of objective selection methods on DTLZ5(2,M)

Besides the proposed reduction technique MIC + Change Rate, there are other objective reduction approaches, such as PCA and NCIE. Fig. 6 shows the median number of objectives after reduction over 200 generation and 50 times repetition, DTLZ5(2, M=5,10,20,30,50) is chosen as the test problem in this subsection, PCA and NCIE are embedded in NSGA-II. As it is shown in Fig. 6, the proposed approach could effectively find the reducible objectives except DTLZ5(2,50). For the 50-objective problem, the performance of MIC + Change Rate approach shows unstable situation compared with NCIE, which sometimes only find 3 redundant objectives.

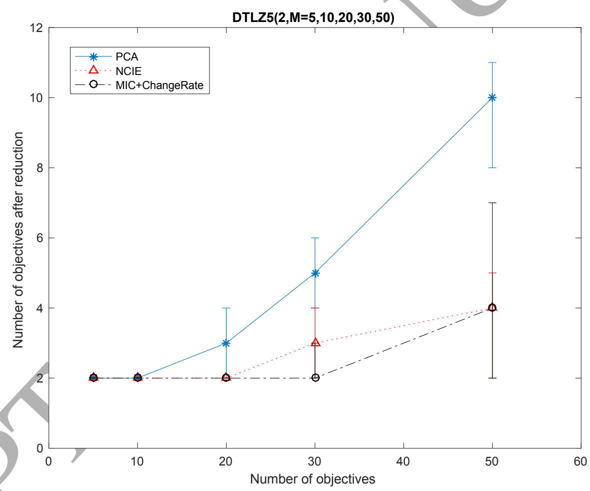


Fig. 6. Median number of objectives after reduction by MIC-MOPSO, PCA and NCIE based NSGA-II as objective selection methods on DTLZ5(2, M)

4.5 Comparison of computational speed on DTLZ5(2,M)

Fig. 7 shows the execution time of MIC-MOPSO with PCA and NCIE based NSGA-II on DTLZ5(2, M). With the increasing number of objectives, both approaches increase their execution time. It is shown that the increasing speed of MIC-MOPSO is much slower than other approaches in the lower dimension. However, when the number of objectives is larger than 20, the execution time of MIC-MOPSO has a rapid increase. The reason might be the computation of average distance costs lots of time with large dimension.

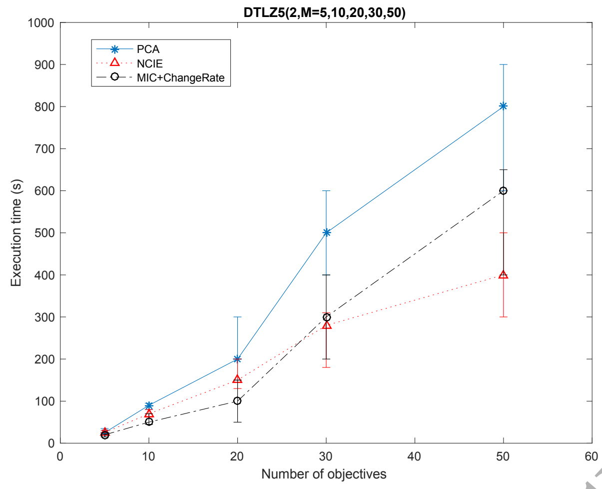


Fig. 7. Execution time of MIC-MOPSO, PCA-NSGA-II and NCIE based NSGA-II on DTLZ5(2, M)

4.6 The application of MIC-MOPSO on gasification process operation optimization

In this sub-section, we apply the proposed algorithm, MIC-MOPSO, to a real industrial many-objective optimization problem, gasification process operation optimization. The gasification process is the basis of the modern coal-based energy and chemical technology. The main purpose of gasification process optimization is to solve gasification model by many-objective optimization algorithms, seek the optimal operation conditions under the constraints of process requirements.

Effective gas production rate is one of the most important performance indicator of gasification system, which could evaluate the operation conditions and also is a key economic indicator of coal gasification. CO, CO_2 and H_2 are principal components of synthesis gas. Based on different process, hydrogen and ammonia generation for example, maximizing hydrogen gas production is needed because the final product is H_2. Meanwhile, the gasification process needs to decrease oxygen consumption in generating synthesis gas, which means oxygen consumption rate is another important optimization value. Effective gas production rate η, H_2 gas production F_{H_2}, CO gas production F_{CO} and oxygen consumption rate α are often contradictory and could not get the optimization condition in the same time, this is a typical many-objective problem in gasification field. In this paper, a Texaco gasification model based on Aspen (Fig. 8) is used, which is firstly built by Kong, et. al. [29]. The parameters $\eta, F_{H_2}, F_{CO}, \alpha$ are chosen as objectives. For calculation convenience, reciprocal of η and F_{H_2} are used.

$$F(x) = \left[\frac{1}{\eta}, \frac{1}{F_{H_2}}, \frac{1}{F_{CO}}, \alpha\right]^T \tag{7}$$

Under the conditions of operation safety, the temperature of gasification reaction should be controlled in 1180-1380°C. And conversion rate of carbonization should be up to a certain value. Therefore, the constraint conditions are:

$$\begin{cases} 1180 \le T \le 1380 \\ \eta_{conv} \ge 0.98 \end{cases} \tag{8}$$

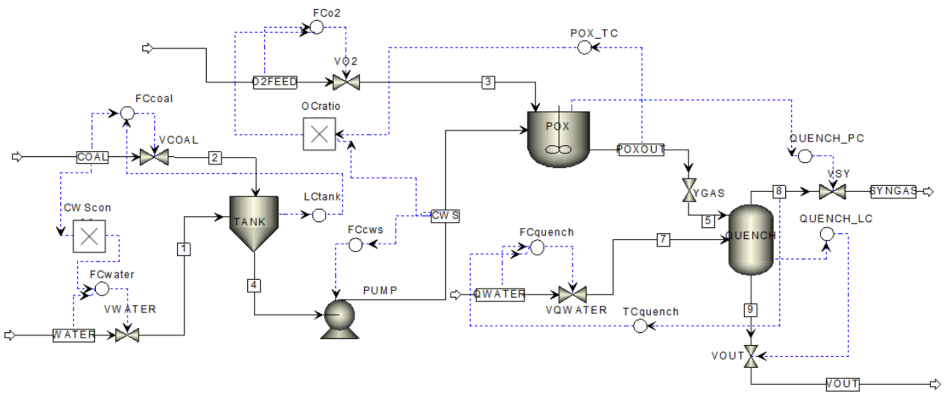


Fig. 8. Flow diagram of the gasification model [29]

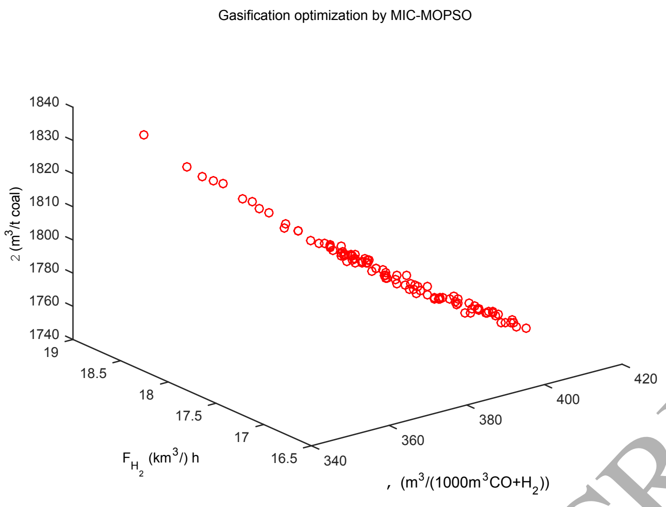


Fig. 9. Approximation of Pareto Font of MIC-MOPSO of coal gasification optimization process

In this paper, MIC-MOPSO is used to optimize many-objective operation problem in gasification process compared with NSGA-II. The parameters setting is the same as the benchmark test above. In Fig. 9, it is shown that the approximation of Pareto front by MIC-MOPSO. And we find that the objective CO gas production F_{CO} is reducible and does not have much contribution to the Pareto front. In order to observe the optimization impact to the real operation plant, the operation sets in Pareto optimal solution are selected in this paper. To keep the coal-water-slurry's concentration constant, the data of oxygen-coal ratio are compared as shown in Table 3. It is revealed that although hydrogen production and oxygen consumption rate decrease, the effective gas production rate is dramatically higher than before.

Table 3
Performance indicators comparisons after optimization

	coal-water-slurry's concentration	oxygen-coal ratio	gas production rate (m^3/kg)	hydrogen production (m^3/h)	oxygen consumption rate (m^3/1000m^3)
Before optimization	61.5	1.011	1596	17854	398
After optimization	61.5	0.973	1645	17240	380

5. Conclusions

In this paper, a novel objective reduction particle swarm optimization algorithm is proposed, named Maximal Information Coefficient based Multi-Objective Particle Swarm Optimizer (MIC-MOPSO). MIC could discover hidden linear or non-linear relationship between two objectives. While significant correlation of objectives is detected, a fast and effective technique, the change rate of non-dominated population, will be adopted to judge whether objectives are conflicting or not. For non-conflicting objectives, MIC-MOPSO decides which objective should be selected or omitted by comparing MIC values of non-conflicting objective pairs. Then MOPSO is used to solve multi-objective problem without redundant objectives. Based on the experiments that are carried out on a set of benchmark functions in comparison with other multi-objective algorithms from the literature, it can be shown that our proposed algorithm is an efficient optimizer for MaOPs, especially for higher dimension ones.

For the future work, it is straightforward to examine the effect of the proposed schemes with other MOEAs improvements, such as the aggregation based methods or preference based methods, on the performance of MIC-MOPSO for MaOPs.

Acknowledge

This work was supported by the National Natural Science Foundation of China (61422303, 61773163), Fundamental Research Funds for the Central Universities (222201717006, 22A201514046), Young Elite Scientists Sponsorship Program by CAST(2016QNRC001), Natural Science Foundation of Shanghai (17ZR1444600) and Shanghai talent development funding.

References

[1] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE Trans. Evol. Comput. 6 (2002) 182-197.
[2] E. Zitzler, L. Thiele, Multiobjective evolutionary algorithms: a comparative case study and the strength pareto

approach, IEEE Trans. Evol. Comput. 3 (4) (1999) 257-271.
[3] H. Ishibuchi, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization: a short review, in: IEEE Congress on Evolutionary Computation, 2008, pp. 2419-2426.
[4] H. Sato, H.E. Aguirre, K. Tanaka, Controlling dominance area of solutions and its impact on the performance of MOEAs, in: Evolutionary Multi-Criterion Optimization, Springer, 2007, pp. 5-20.
[5] J. Branke, T. Kaubler, H. Schemck, Guidance in evolutionary multiobjective optimization, Adv. Eng. Software, 32(6) (2001) 499-507.
[6] M. Laumanns, L. Thiele, K. Deb, E. Zitzler, Combining convergence and diversity in evolutionary multiobjective optimization, Evol. Comput. 10(3) (2002) 263-282.
[7] M. Farina, P. Amato, A fuzzy definition of 'optimality' for many-criteria optimization problems, IEEE Trans. Syst. Man, Cbyern., Part A: Syst. Hum. 34(3) (2004) 315-326.
[8] A.L. Jaimes, C.A. Coello Coello, Study of preference relations in many-objective optimization, in: Proceedings of 11th Annual Conference on Evolutionary Multi-Criterion Optimization, 2011, pp. 151-165.
[9] A.B. de Carvalho, A. Pozo, Measuring the convergence and diversity of CDAS multi-objective particle swarm optimization algorithms: a study of many-objective problems, Neurocomputing, 75 (2012) 43-51.
[10] C. Dai, Y. Wang, A new uniform evolutionary algorithm based on decomposition and CDAS for many-objective optimization, Knowledge-Based Systems, 85 (2015) 131-142.
[9] Q. Zhang, H. Li, MOEA/D: a multiobjective evolutionary algorithm based on decomposition, IEEE Trans. Evol. Comput. 11 (6) (2007) 712-731.
[10] E. Zitzler, S. Künzli, Indicator-based selection in multiobjective search, in: Parallel Problem Solving from Nature - PPSN VIII, Springer, 2004, pp. 832-842.
[11] P.J. Fleming, R.C. Purshouse, R.J. Lygoe, Many-objective optimization: an engineering design perspective, in: Evolutionary Multi-Criterion Optimization, Springer, 2005, pp. 14-32.
[12] K. Deb, J. Sundar, Preference point based multi-objective optimization using evolutionary algorithms, in: Proceeding of 2006 Genetic and Evolutionary Computation Conference, 2006, pp. 635-642.
[13] A. Britto, A. Pozo, Using reference points to update the archive of MOPSO algorithms in many-objective optimization, Neurocomputing, 127 (2014) 78-87.
[14] L. Thiele, K. Miettinen, P.J. Korhonen, J. Molina, A preference-based interactive evolutionary algorithm for multiobjective optimization, Working Paper, Helsinki School of Economics, W-412, 2007.
[15] D. Brockhoff, E. Zitzler, Are all objectives necessary? on dimensionality reduction in evolutionary multiobjective optimization. in: Parallel Problem Solving from Nature - PPSN IX, Springer, 2006, pp. 533-542.
[16] H. Singh, A. Isaacs, T. Ray, A pareto corner search evolutionary algorithm and dimensionality reduction in many-objective optimization problems, IEEE Trans. Evol. Comput. 15(4) (2011) 539-556.
[17] A.L. Jaimes, C.A. Coello Coello, D. Chakraborty, Objective reduction using a feature selection technique, in: Proceedings of 2008 Genetic and Evolutionary Computation Conference, 2008, pp. 673-680.
[18] K. Deb, D. Saxena, On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems, Kangal report 11.
[19] D.K. Saxena, J.A. Duro, A. Tiwari, K. Deb, Q. Zhang, Objective reduction in many-objective optimization: linear and nonlinear algorithms, IEEE Trans. Evol. Comput. 17 (1) (2013) 77-99.
[20] J.A. Duro, D.K. Saxena, K. Deb, Q.F. Zhang, Machine learning based decision support for many-objective optimization problems, Neurocomputing, 146 (2014) 30-47.
[21] H. Wang, X. Yao, Objective reduction based on nonlinear correlation information entropy, Soft Comput. 20 (2016) 2393-2407.
[22] D.N. Reshef, Y.A. Reshef, H.K. Finucane, S.R. Grossman, G. McVean, P.J. Turnbaugh, E.S. Lander, M. Mitzenmacher, P.C. Sabeti, Detecting novel associations in large data sets, Science, 334 (2011) 1518-1524.
[23] R.C. Purshouse, P.J. Fleming, Conflict, harmony, and independence: relationships in evolutionary multi-criterion optimisation, Lecture Notes in Computer Science, 2632 (2003) 16-30.
[24] C.A.C Coello, G.T. Pulido, M.S. Lechuga, Handling multiple objective with particle swarm optimization, IEEE Trans. Evol. Comput. 8(3) (2004) 256-279.
[25] K. Deb, D. Saxena, On finding pareto-optimal solutions through dimensionality reduction for certain large-dimensional multi-objective optimization problems. Kangal report 11.
[26] K. Deb, L. Thiele, M. Laumanns, E. Zitzler, Scalable multi-objective optimization test problems, in: IEEE Congress on Evolutionary Computation, 2002, pp. 825-830.
[27] D.A. Van Veldhuizen, G.B. Lamont, Multiobjective evolutionary algorithm research: a history and analysis, Tech. rep. TR-98-03, 1998.
[28] D.A. Van Veldhuizen, G.B. Lamont, On measuring multiobjective evolutionary algorithm performance, in: IEEE Congress on Evolutionary Computation, 2000, pp. 204-211.
[29] X.D. Kong, Research on modeling, control and optimization for industrial entrained flow coal gasification process,

PhD thesis, 2014, 69-78.
[30] T. Wagner, N. Beume, B. Naujoks, Pareto-, aggregation-, and indicator-based methods in many-objective optimization, in: Evolutionary Multi-Criterion Optimization, Springer, 2007, pp. 742-756.
[31] Z. He, G.G. Yen, Many-objective evolutionary algorithm: objective space reduction and diversity improvement, IEEE Trans. Evol. Comput. 20(1) (2016) 145-160。
[32] Y. Cheung, F. Gu, H. Liu, Objective extraction for many-objective optimizzation problems: algorithm and test problems, IEEE Trans. Evol. Comput. 20(5) (2016) 755-772.

Yi Liang received the B.S. degree in electronic engineering from Northwest Polytechnical University, Xi'an, China in 2008 and M.S. degree in process control and optimization from East China University of Science and Technology, Shanghai, China in 2011.

He is currently working toward the Ph.D. degree at Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China. His current research interests include evolutionary computation and process optimization.

Dr. Wangli He received the B.S. degree in information and computing science and Ph.D. degree in applied mathematics from Southeast University, Nanjing, China, in 2005 and 2010, respectively.

From January 2010 to November 2011, she was a Post-Doctoral Research Fellow at East China University of Science and Technology, China. From July 2010 to July 2011, she was a Visiting Postdoctoral Research Fellow at Central Queensland University, Australia. From 2013 to 2017, she held several visiting positions with the Department of Mechanical Engineering, the University of Hong Kong, the Centre for Chaos and Complex Networks at the City University of Hong Kong, the Potsdam Institute for Climate Research Institute, Germany and Tokyo Metropolitan University, Japan.

She is currently an Associate Professor with the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. Her current research interests include cooperative control of multi-agent systems, security control of CPS and networked systems.

Prof. Weimin Zhong received the B.S. degree in automation and Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China in 1998 and 2006, respectively.

From 2006 to 2008, he was a Post-Doctoral Research Fellow at East China University of Science and Technology, China. From September 2013 to August 2014, he was a Visiting Research Fellow with Department of Chemical Engineering, Lehigh University, Bethlehem, PA, United States.

He is currently a Professor, Vice Dean of the School of Information Science and Engineering, East China University of Science and Technology, Shanghai, China. His current research interests modeling, control, optimization and integration of industrial process. He has hosted more than 10 projects including National Science Foundation of China, "863" Project, The National Key Technology R&D Program. He was a recipient of Natural Science Fund for Distinguished Young Scholar and has published more than 80 papers.

Professor Feng Qian received the B.Sc. degree in chemical automation and meters from Nanjing Institute of Chemical Technology, Nanjing, China, in 1982, and the M.S. and Ph.D. degrees in automation from East China Institute of Chemical Technology, Shanghai, China, in 1988 and 1995, respectively.

From 1999 to 2001, he was the Director of Automation Institute, East China University of Science and Technology, China, where he served as the head of Scientific and Technical Department from 2001 to 2006. He is the current Vice President of East China University of Science and Technology, the Director of Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, China, and the Director of Process System Engineering Research Centre, Ministry of Education, China.

Professor Qian's research interests include modeling, control, optimization and integration of petrochemical complex industrial process and the industrial applications, neural networks theory, real-time intelligent control technology and their applications on the ethylene, PTA, PET and refining industries. He is now the member of China Instrument and Control Society, Chinese Association of Higher Education and China's PTA Industry Association. He is a Fellow of Academy of Engineering of China.