HMOEDA_LLE: a hybrid multi-objective estimation of distribution algorithm combining locally linear embedding

Yuzhen Zhang, Guangming Dai*, Lei Peng, Maocai Wang
School of Computer Science
China University of Geosciences
Wuhan, China

Abstract—**Based on the regularity that: the Pareto set of a continuous m-objectives problem is a piecewise continuous (m-1)-dimensional manifold, a novel hybrid multi-objective optimization algorithm is proposed in this paper. In the early evolutionary stage, traditional crossover and mutation operations are used to produce offspring, in addition, the locally linear embedding (LLE) with small neighbor parameter approach is introduced to learn the local geometry of the manifold. When certain regularity in population's distribution is detected, new offspring are sampled from the probability models created by the statistical distribution information. An entropy-based criterion is imported to determine the switching time of the two different phases of evolutionary search. The proposed hybrid multi-objective estimation of distribution algorithm combining locally linear embedding (HMOEDA_LLE) adopts several widely used test problems to conduct the comparison experiments with two state-of-the-art multi-objective evolutionary algorithms NSGA-II and RM-MEDA. The simulated results show the effectiveness of the entropy-based criterion and the proposed algorithm has better optimization performance.**

Keywords—regularity; multi-objective optimization; locally linear embedding; entropy-based criterion

I. INTRODUCTION

Multi-Objective Problems (MOPs) arise in many scientific research and engineering applications. Generally, these objectives conflict with each other, that is to say, we can't optimize these objectives simultaneously. Thus, the optimal solution to a MOP is not a single one but a collection of solutions called Pareto set. Due to the population based attribute of approximating the Pareto set in a single run, evolutionary algorithm has been proved to be suited for solving MOPs.

Since 1985, researchers have done a lot of work in this field, formulating the first generation and the second generation MOEAs which emphasize on simplicity and efficiency respectively [1]. In recent years, some new approaches have been proposed. They can be grouped into two categories: newly proposed dominance mechanisms and the application of different heuristic mechanisms. Due to the ineffectiveness of

Pareto dominance in solving many-objective optimization, several new dominance mechanisms were proposed to enhance the selection pressure toward the Pareto front. e.g., a relaxed Pareto dominance based on the relationship among different objectives [2], dominance area control [3], ε-dominance [4], K-optimality [5], partial dominance [6], Pareto-adaptive-dominance [7], fuzzy Pareto dominance [8], grid-dominance [9]. Meanwhile, different heuristic mechanisms have been introduced in multi-objective evolutionary optimization including particle swarm optimization, ant colony algorithm, artificial immune systems, estimation of distribution algorithm, memetic algorithm, etc. MOPSO [10] proposed by Coello et al, non-dominated neighbor immune algorithm [11], Zhang's MOEA/D [12] and RM-MEDA [13] are some notable algorithms among them. Instead of using traditional crossover and mutation operators to generate offspring individuals, RM-MEDA builds a probability distribution model of promising solutions, and new individuals are then sampled from the model [14] by applying the regularity that the PF and PS of a continuous MOP are piecewise continuous (m-1)-dimensional manifold under mild conditions [15]. A basic idea behind RM-MEDA is to use globally statistical information to guide the search of EDA, thus the local information of each individual seems to be considered inadequately.

Locally linear embedding (LLE) [16] is a representative manifold learning method. It can exploit the local linear patches of the manifold. Due to the simplicity and few parameters, LLE gains much attention and has been used to solve various problems in machine learning [17].

In the early stage of evolutionary algorithm, the distribution of population may lack certain regularity. Difference exists between search direction of sampled individuals and the promising direction. By mixing local exploitation with traditional crossover and mutation operation, the regularity can be found more quickly and effectively, the performance of EDA gets enhanced consequently.

In this paper, a hybrid multi-objective estimation of distribution algorithm combining locally linear embedding (HMOEDA_LLE) is proposed. At the early evolutionary stage, LLE-based method is introduced to generate partial offspring

and the rest population is produced by traditional genetic operators in non-dominated sorting genetic algorithm [18]. After some generations, estimation of distribution modelling and sampling method is used to learn and sample new solutions. An entropy based criterion helps to determine the switching time of the two phases of different evolutionary search. Experiments have been conducted on several state-of-the-art multi-objective problems to compare HMOEDA_LLE with two representative algorithms NSGA-II and RM-MEDA.

The rest of the paper is organized as follows. Section II briefly describes multi-objective problem and LLE algorithm. In Section III, the proposed algorithm is presented in detail. The experimental studies and results analysis are given in section IV. Finally, section V is devoted to conclusion and future work.

II. PRELIMINARY

A. Multi-objective Optimization Problems (MOPs)

Without losing generality, we consider the following continuous multi-objective optimization problem in this paper:

$$\begin{aligned} minimize \quad & F(x) = (f_1(x), f_2(x), \cdots, f_m(x))^T \\ subject\ to \quad & x \in \Omega \end{aligned} \quad (1)$$

Where Ω is the decision space and $x = (x_1, x_2, \cdots, x_n)^T$ is the decision vector. $F(x)$ consists of m objective functions, $f_i : \Omega \to R, (i = 1, 2, \cdots, m)$, R^m is the objective space.

Definition (Pareto Dominance): Let

$$\begin{aligned} & \forall i \in (1, 2, \cdots, m) : f_i(x) \leq f_i(y) \land \\ & \exists j \in (1, 2, \cdots, m) : f_j(x) < f_j(y) \end{aligned} \quad (2)$$

A vector $x^* \in \Omega$ is Pareto optimal if there is no other $x \in \Omega$ such that $x \prec x^*$, the set of all Pareto optimal points is called Pareto set (PS). The corresponding images of Pareto set in objective space form the Pareto front (PF).

The goal for all the MOEAs is to minimize the distance of solutions to the Pareto front and make them distribute along the Pareto front as diverse as possible.

B. Locally linear embedding

Manifold learning is a kind of nonlinear dimensionality reduction algorithm. The basic idea of the algorithm is to find the low dimensional manifold embedded in a high dimensional space. Since the publication of three classical manifold learning papers [16, 19-20], it has attracted much attention and gained rapidly development. Although many new methods such as Laplacian eigenmap (LE) [21], Hessian eigenmap (HE) [22], Locality tangent space alignment (LTSA) [23] etc. arise, Locally linear embedding (LLE) is still a very attractive method due to its easy implementation and few parameters (only two parameters: the number of nearest neighbors K and the dimension d in lower space).

The main idea of locally linear embedding can be concluded into three steps: 1) Find K nearest neighbors for each point X_i; 2) Reconstruct X_i with K neighbors by linear weights; 3) Compute the low-dimensional embedding vectors through the obtained weights. Details of this algorithm are described in Algorithm1.

Algorithm1 Locally linear embedding

Inputs: N vectors: $X = \{X_1, X_2, \cdots, X_N\}, X_i \in R^D$;
             dimension of embedding manifold: d
             number of nearest neighbors: K;
 Output: low-dimensional embedding vectors:
             $Y = \{Y_1, Y_2, \cdots, Y_N\}, Y_i \in R^d \ (i = 1, \cdots, N\}$

Step1: Find K nearest neighbors $X_{i1}, X_{i2}, \cdots, X_{iK}$ for each
            X_i by Euclidean distance;
 Step2: Reconstruct X_i with its neighbors , compute the
            linear coefficients by minimizing reconstruction
            error:
 $$\varepsilon(W) = \sum_i \left| X_i - \sum_{j=1}^K W_j^{(i)} X_{ij} \right| \quad (3)$$
 Step3: Compute the low-dimensional embedding vector
            Y_i through the linear coefficients W_{ij} :
 $$\varepsilon(Y) = \sum_i \left| Y_i - \sum_{j=1}^K W_{ij} Y_j \right| \quad (4)$$
     It can be converted to:
 $$\varepsilon(Y) = tr(YMY^T) \quad (5)$$
 Where $M = (I - W)^T (I - W)$, then:
 $$Y = \sqrt{N}(e_1, e_2, \cdots, e_d)^T * \quad (6)$$

* $e_1, e_2, \cdots, e_d$ is the corresponding eigenvectors of d eigenvalues of M sorted by ascending order.

LLE is a nonlinear unsupervised learning algorithm of dimension reduction. By using locally linear reconstruction, the neighborhood relationship of the points in high dimensional space can be mapped into low dimensional manifold space, thus the global nonlinear manifold structure is featured based on local geometry.

III. HMOEDA_LLE

A. Entropy-based Criterion

The concept of entropy was proposed by Clausius, Boltzmann developed the theory and introduced the statistical definition of entropy. In 1948, Shannon adopted the theoretic entropy to measure the stochastic process of information, establishing the field of information theory. The application of entropy permeated various fields including multi-objective optimization [24-30]. In the process of evolution, the distribution of population undergoes the change from chaotic to the presentation of certain regularity. Entropy is introduced to describe the population's distribution, on the basis of this, an entropy-based criterion is proposed in this paper.

Definition (The entropy of population): It's a measure of the distribution of all individuals in a population. Suppose the population size is N, the dimension of the decision/objective variables is n, dividing each dimensionality into G parts, the whole grid number of the decision/objective space is G^n, thus the entropy of the population can be defined as:
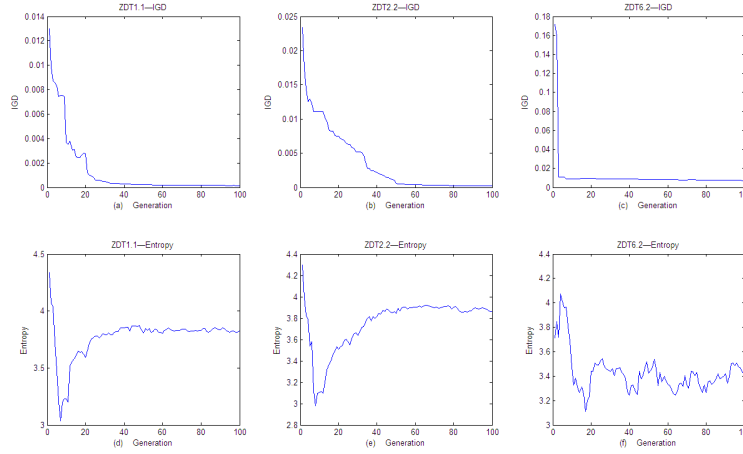
Fig. 1. The IGD((a)(b)(c)) and entropy((d)(e)(f)) curves of ZDT1.1, ZDT2.2, ZDT6.2

$$E = -\sum_{i=1}^{G^n} p_i \log p_i \, (p_i \neq 0) \qquad (7)$$

Where $p_i = counter[i] / N$, $counter[i]$ is the size of individuals in the i-th grid.

Fig. 1 shows the performance curves of three test problems, Fig. 1(a)(b)(c) are the IGD curves of ZDT1.1, ZDT 2.2 and ZDT6.2, Fig. 1(d)(e)(f) are the entropy curves of the above problems. As shown in Fig. 1(a)(b), along with the evolutionary process the population converges. The entropy curves in Fig. 1(d)(e) become stable as the convergence of the population. Whereas, Fig. 1(c) indicates that the IGD performance of ZDT6.2 is not very good, the entropy curve in Fig. 1(f) oscillates accordingly. As it can be seen in Fig. 1, the entropy varies severely in the initial stage, with the evolution progress it gradually decreases. After some generations, entropy changes within a certain range. So entropy can be regarded as a criterion to the distribution performance of the evolutionary population.

Definition (Entropy-based criterion): The entropy value difference of j generations is:

$$\sigma(t) = \sqrt{\frac{1}{j-1} \sum_{g=t-j+1}^{t} \left(E(g) - \overline{E} \right)^2}, \left(\overline{E} = \frac{1}{j} \sum_{g=t-j+1}^{t} E(g) \right) \quad (8)$$

If $\sigma(t) < \varepsilon$ (ε is a predefined constant), certain regularity for the distribution of the evolutionary population can be considered formed.

B. LLE-based approach

According to Karush-Kuhn-Tucker condition: the PF and PS of a continuous MOP are piecewise continuous (m-1)-dimensional manifold for the first class MOPs [31]. That's to say, the PS is a curve/surface for bi-objective/three objective optimization problems with n-dimensional variables. Therefore, LLE can be introduced to excavate the low dimensional manifold of the population's distribution in decision space. Yang et al. [32] propose a hybrid multi-objective optimization algorithm which combines the immune inspired algorithm with the estimation of distribution algorithm based on LLE.

Different from Yang's approach, the LLE-based modeling method in this paper is only adopted at the early evolutionary stage, and a smaller neighbor parameter K makes LLE focus more on the local patches of the manifold, exploitation ability gets enhanced correspondingly. Details of this approach are shown in Algorithm 2.

Algorithm2 LLE-based approach

Inputs: population: P ;
              number of nearest neighbors: K;
              dimension of embedding manifold: d
Outputs: offspring population: Z

Step1: Compute the low-dimensional embedding: Let $X = P$, Algorithm1 is adopted to get the low dimension embedding vectors Y in decision space; Step2: Find the extreme vector in Y : $$a_i = \min(Y_j^i), b_i = \max(Y_j^i) \qquad (9)$$ Where $(i = 1,\cdots,d \, ; \, j = 1,\cdots,N)$; Step3: Orthogonal sampling: Get N low dimensional vectors $Y_1^{'}, Y_2^{'}, \cdots, Y_N^{'}$ by orthogonal sampling* from the subspace (a,b) obtained in Step2; Step4: Reconstruct $Y_i^{'}$ with its neighbors in Y : Compute the linear coefficients by solving the LS problem: $$\varepsilon(W) = \sum_i \left| Y_i^{'} - \sum_{j=1}^{K} W_j^{(i)} Y_{ij} \right| \qquad (10)$$ Step5: Generate offspring: Produce N offspring through linear reconstruction by W_{ij} in Step4: $$Z = \left\{ Z_i \in R^n \middle\| Z_i = \sum_{j=1}^{K} W_{ij} X_{ij} + N(0,\sigma_i) \right\} \quad (11)$$

*More details of the orthogonal sampling method can be found in [33]

C. Framework of the proposed algorithm

The evolutionary process of the proposed algorithm can be divided into two stages. At the first stage, offspring population is produced partially by LLE-based sampling approach and the rest is generated by traditional crossover and mutation operations. At the second stage, reproduction is replaced with

probability model of population [13]. Conversion timing of them is decided by the entropy-based criterion. Framework of the proposed hybrid algorithm is presented in Algorithm3.

Algorithm3 HMOEDA_LLE

Step1: Initialization: t=0, generate the initial population $P(t)$, set the value of K and ε ;

Step2: Reproduction:

  Step2.1 Entropy calculation: Calculate the entropy $E(t)$ of the current population $P(t)$ according to Eq.(7) (8)，if $t > j$ and $\sigma(t) < \varepsilon$ are satisfied, go to Step 2.5, else go to Step 2.2;

  Step2.2 LLE-based approach: Let $P = P(t)$, d=(m-1), adopting Algorithm2 to generate N_1 offspring;

  Step2.3 Genetic operation: Using non-dominated sorting approach in [18] to produce N_2 offspring;

  Step2.4 Combination: Merge the solutions produced by Step 2.2 and Step 2.3 to get the offspring population $Q(t)$, go to step 3;

  Step2.5 Sampling population: Create H probability models according to the method in [13], sampling offspring population $Q(t)$ randomly;

Step3: Selection: Select N individuals from $P(t) \cup Q(t)$ and replace all the individuals in $P(t)$;

Step4: Stopping criterion: If stopping criterion is met, stop and return $P(t)$, otherwise go to Step2.

The estimation of distribution algorithm is proved to be very efficient in solving complex multi-objective optimization problems with variable linkages. However, at the early evolutionary stage, the population may not sufficiently converged, the probability model established by EDA lack of accuracy as a consequence. By introducing non-dominated sorting genetic method, the traditional genetic operators guide the search to promising areas. In addition, locally linear embedding based approach extracts the local characteristics of the underlying manifold with small neighbor parameter K. The linear reconstruction individuals help searching the nearby region, thus exploitation ability gets enhanced. The hybrid of the two means leads the identifying of the regularity of the distribution of PS in decision space more quickly and efficiently. The principal component directions of the probability model built on regularity meet more with the promising solutions.

IV. EXPERIMENT AND RESULT

In order to verify the performance of HMOEDA_LLE, the proposed algorithm is tested on eight test problems which can be classified into two categories: problem with linear variable linkages and problem with nonlinear variable linkages. The results are compared with two state-of-the-art algorithms NSGA-II [18] and RM-MEDA [13].

The eight test instances are from [13], six of them are variants of ZDT test suit [34] and the other two problems are variants of DTLZ test problem [35]. Among them ZDT1.1, ZDT2.1, ZDT6.1, DTLZ2.1 have linear variable linkages and the rest ones have nonlinear variable linkages. The inverted generational distance (IGD) is used to measure both the convergence and diversity of an approximate solution set to the true PF.

A. General Experimental Setting

The proposed HMOEDA_LLE is implemented with C++, the test environment is: CPU: Intel Pentium G630, 2.7GHz; Memory: 4GB; OS: Windows7; Programming environment: Microsoft Visual Studio 2008.

1) Population size: For all the bi-objective problems, the population size is set to 100, and 200 for three objective problems.

2) Parameter setting in HMOEDA_LLE: The value of N_1 in Algorithm3 is:

$$r * |Pop| * e^{-\frac{t}{T}} \qquad (12)$$

The value of N_2 is $|Pop| - N_1$ accordingly, here r is set to 0.7. The neighbor parameter K is 2 for all the bi-objective problems and 3 for all the three objective problems. ε is set to 0.1 for all the test problems, where t is the current generation and T is the maximum generation.

3) Parameter setting in NSGA-II: SBX crossover and Polynomial mutation are adopted with both distribution indexes 20 (i.e. $\eta_c = 20, \eta_m = 20$). The mutation rate $p_m = 1/n$ (where n is the number of decision variables) and crossover rate $p_c = 0.9$.

4) Parameter setting in RM-MEDA: For all the test problems, the number of clusters H is set to 5.

5) Number of runs and stopping condition: Each algorithm runs 30 times independently on each test problem. The stopping criterion is maximal number of function evaluations, which is set to be 10 000 for bi-objective problems and 20 000 for three objective problems.

B. Performance comparison

Under the same conditions, a smaller IGD value indicates a better approximation to the true PF and a broader distribution along the front. Table 1-2 shows the mean and standard deviation over 30 independent runs of NSGA-II, RM-MEDA and the proposed algorithm in this paper, where the best mean of each test instance is displayed with gray background.

As it can be seen in Table I-II, the proposed HMOEDA_LLE obtains the best IGD values on all the eight test problems with linear and nonlinear variable linkages, RM-MEDA obtains better IGD values on ZDT1.1, ZDT2.1, DTLZ2.1, ZDT1.2, ZDT2.2, DTLZ2.2, whereas NSGA-II outperforms RM-MEDA on ZDT6.1 and ZDT6.2. As we all know, RM-MEDA is efficient to solve problems with variable linkages for the use of covariance matrix in RM-MEDA can identify the dependences between variables. The reason why RM-MEDA can't converge to the true PF on ZDT6.1 and ZDT6.2 is that they are bias problems, which means solutions

TABLE I. IGD metric. Mean (first line) and standard deviation (second line)

Algorithm	ZDT1. 1	ZDT2. 1	ZDT6. 1	DTLZ2. 1
NSGA-II	0. 00783383±	0. 0227759±	0. 00949334±	0. 00119925±
	2. 2229e-006	2. 90709e-006	2. 91219e-008	1. 68302e-008
RM-MEDA	0.000159403±	0.000177517±	0. 0936774±	0.000516313±
	8. 37546e-011	7. 80016e-010	5. 21269e-005	3. 19663e-010
HMOEDA_LLE	0.000148039±	0.000150462±	0. 00439391±	0.000501305±
	5. 85535e-012	3. 04897e-012	9. 93824e-008	1. 20358e-010

TABLE II. IGD metric. Mean (first line) and standard deviation (second line)

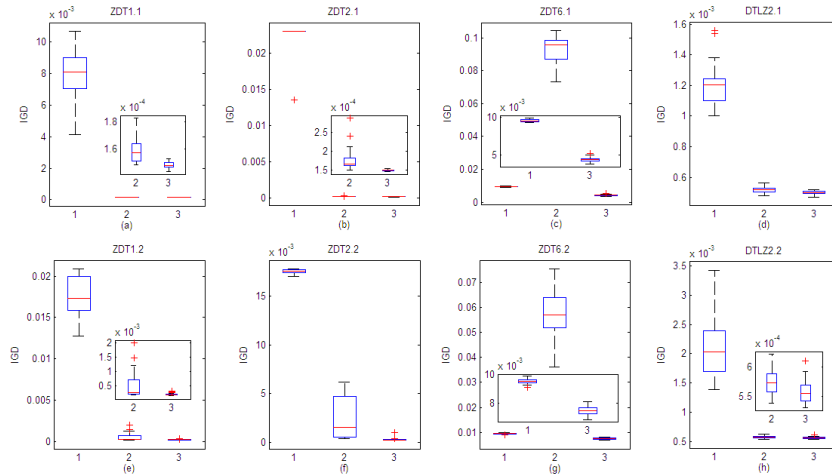| Algorithm	ZDT1. 2	ZDT2. 2	ZDT6. 2	DTLZ2. 2
NSGA-II	0. 0174777±	0. 0175527±	0. 00953686±	0. 00211867±
	5. 06057e-006	4. 96669e-008	2. 98307e-008	2. 82809e-007
RM-MEDA	0.000492489±	0. 00240114±	0. 057498±	0.000574903±
	1. 91828e-007	4. 10512e-006	8. 35169e-005	4. 68283e-010
HMOEDA_LLE	0.000206575±	0.000258601±	0. 00745946±	0.000558326±
	1. 18616e-009	2. 01932e-008	1. 11356e-007	3. 8473e-010

Fig. 2. Comparisons of the IGD boxplots of the three algorithms on all the test problems,1, 2, 3 represent NSGA-II, RM-MEDA and HMOEDA_LLE

of the two problems are not evenly distributed in objective space. The objective vectors are sparser toward the Pareto optimal front for ZDT6.1, ZDT6.2, so the randomly sampling method used in RM-MEDA may not escape from local optimum. NSGA-II is a classical multi-objective evolutionary algorithm. Its optimization ability has been proved in solving various MOPs. Nevertheless traditional crossover and mutation can't capture the dependencies between different variables, so the performance of NSGA-II is not that good when facing problems with variable linkages in decision space. The exploration and exploitation abilities of the proposed hybrid multi-objective estimation of distribution algorithm are adequately utilized by combining locally linear embedding in the early evolutionary stage with traditional crossover and mutation. In addition, it attaches equal importance to the population's distribution regularity by the use of modelling and sampling method in the later evolutionary process. The IGD values in Table I-II clearly show the effectiveness of the proposed HMOEDA_LLE.

To further indicates the optimization ability of the three different algorithms, Fig. 2 gives the boxplot of IGD performance indicator on 30 independent runs. Boxplot is an important tool in economics, comparing to mean and variance, it is more intuitive to analyze the statistical distribution of the data. The bottom and top of the box are the first and third quartiles. The band inside the box is the median. The ends of the whiskers are the minimum and maximum values. "+" represents outliers. 1, 2, 3 stand for NSGA-II, RM-MEDA and HMOEDA_LLE respectively in Fig. 2. It can be seen from Fig. 2(a)-(h) that the optimization ability and robustness of HMOEDA_LLE perform the best, RM-MEDA follows, and the performance of NSGA-II is relatively poor. ZDT6.1 and ZDT6.2 are exceptions as it shows in Fig. 2(c), (g) for NSGA-II is not easy to fall into the local optimal and converges prematurely. The proposed algorithm takes the advantage of both NSGA-II and RM-MEDA. It's not only suitable for solving MOPs with variable linkages but also not easy to fall into local optimum.

In order to demonstrate the evolutionary process of the three different algorithms, Fig. 3 plots the mean IGD trajectories of 30 independent runs on each test problems.
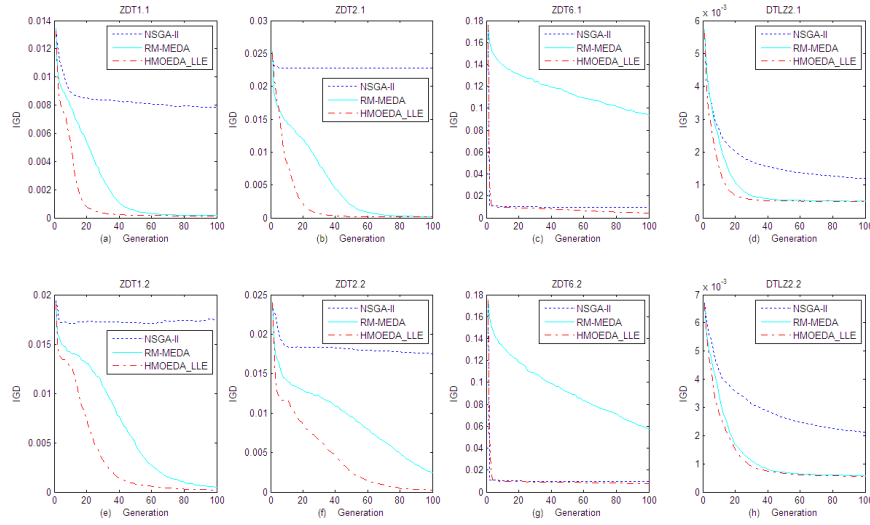
Fig. 3 Comparisons of the IGD mean curves of 30 independent runs of the three algorithms
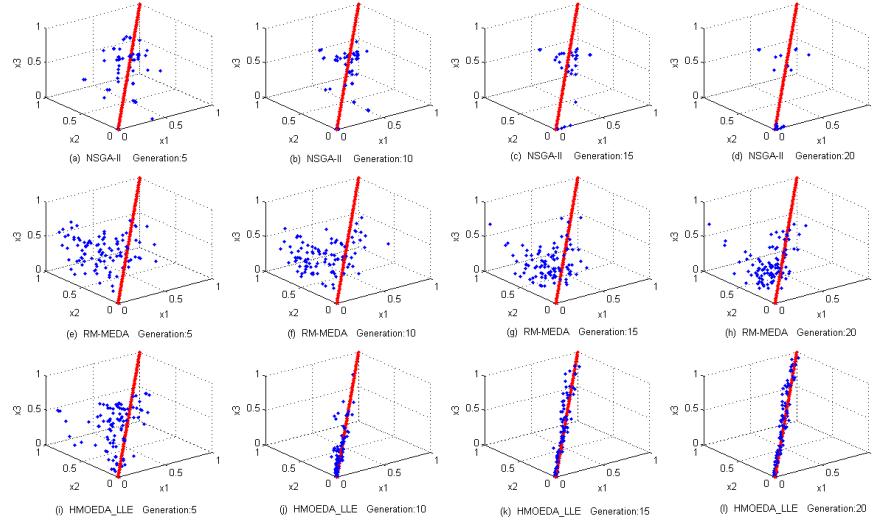

Fig. 4 The distribution of the first three dimensional of PS obtained by NSGA-II, RM-MEDA and HMOEDA_LLE for ZDT1.1 test problem

HMOEDA_LLE converges more quickly than the other two algorithms on ZDT1.1, ZDT2.1, ZDT1.2, ZDT2.2 obviously as it can be seen in Fig. 3 (a)(b)(e)(f). For problems ZDT6.1, ZDT6.2 which are a little hard for RM-MEDA to converge (show in Fig. 3(c)(g)) and DTLZ2.1, DTLZ2.2 which are difficult for NSGA-II to converge (show in Fig. 3(d)(h)), the proposed algorithm still performs very well.

To better analyze the impact of the introduction of LLE-based approach, Fig. 4-5 give the distribution of PS obtained by NSGA-II, RM-MEDA and HMOEDA_LLE for problem with linear variable linkages ZDT1.1 and problem with nonlinear variable linkages ZDT1.2. For the limitation of visualization, we only draw the population's distribution on the first three dimensions. Fig. 4(a)-(d), Fig. 5(a)-(d) give the distribution of populations found by NSGA-II at generation 5,10,15,20. They show that with the evolutionary process NSGA-II can only find a small part of the optimal solutions, thus the mapped solutions in the objective space will distribute along only a small part of the Pareto front, which means the lack of population's diversity. Fig. 4(e)-(h), Fig. 5(e)-(h) show the distribution of populations of ZDT1.1 and ZDT1.2 found by RM-MEDA at generation 5,10,15,20 respectively. Although RM-MEDA has proved to be very efficient to solve problems with variable linkages, but it can be seen from Fig. 4(e)-(h), Fig. 5(e)-(h) that in the initial evolutionary stage, it's hard for RM-MEDA to find the correct direction of the principal component because the statistical information is not very sufficient, whereas the populations obtained by HMOEDA_LLE for ZDT1.1 and ZDT1.2 find the right direction of the principal component and distribute on the PS diversely with the evolutionary process as showed in Fig. 4(i)-(l), Fig. 5(i)-(l). As the description in Section III (B), locally linear embedding based approach extracts the local characteristics of the underlying manifold with small neighbor parameter K by linear reconstruction and maps the neighborhood relationship of individuals in high decision space to underlying lower
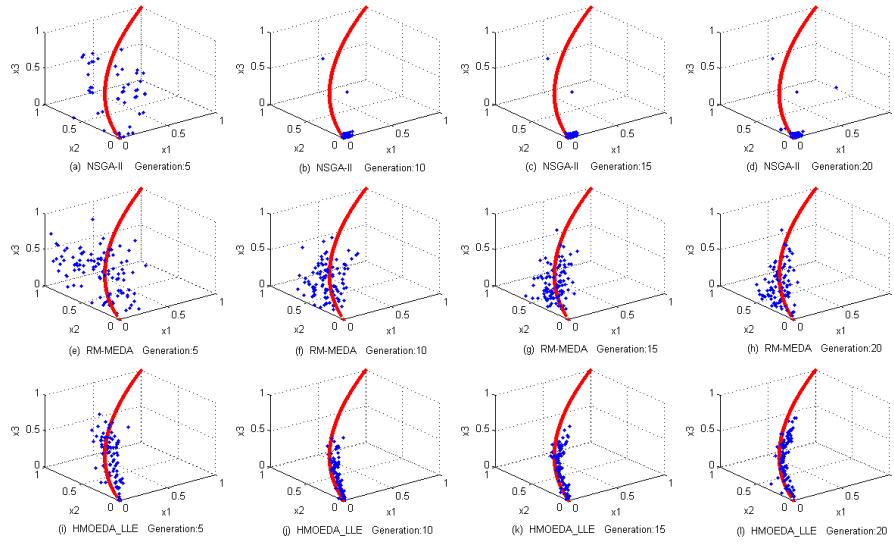
Fig. 5 The distribution of the first three dimensional of PS obtained by NSGA-II, RM-MEDA and HMOEDA_LLE for ZDT1. 2 test problem.

TABLE III. IGD metric. Mean(first line) and standard deviation(second line)

Algorithm	ZDT1.1	ZDT2.1	ZDT6.1	DTLZ2.1
HMOEDA	0.00014978±	0.0223277±	0.0050595±	0.000499892±
	7.35941e-012	1.69563e-005	1.3681e-007	1.67279e-010
HMOEDA_LLE	0.000148092±	0.000150084±	0.00435625±	0.000499022±
	9.33956e-012	3.82886e-012	1.12063e-007	1.97575e-010

TABLE IV. IGD metric. Mean(first line) and standard deviation(second line)

Algorithm	ZDT1.2	ZDT2.2	ZDT6.2	DTLZ2.2
HMOEDA	0.00386678±	0.0143667±	0.0076994±	0.000551672±
	1.36429e-005	2.1145e-006	2.48522e-007	5.23962e-010
HMOEDA_LLE	0.000198611±	0.000259849±	0.00734217±	0.000559813±
	1.02547e-009	8.34037e-009	1.10173e-007	3.70403e-010

dimensional manifold space. The exploitation ability of proposed algorithm in this paper gets enhanced. More importantly, the distribution regularity of PS can be found more quickly and efficiently which helps the modelling and sampling in the later evolutionary stage.

To further illustrate the validity of LLE-based approach, we omit the LLE-based approach in HMOEDA_LLE, and then in the first evolutionary stage, all new individuals are generated by crossover and mutation operation. The second stage is just the same as in HMOEDA_LLE. Entropy-based criterion still performs as the switching judgment of the two stages. We might call this method as HMOEDA. Table III-IV show the mean and standard deviation over 30 independent runs of HMOEDA and HMOEDA_LLE, where the best mean of each test instance is displayed with gray background.

It can be seen from Table III-IV that HMOEDA_LLE performs better than HMOEDA on seven test problems (ZDT1.1, ZDT1.2, ZDT6.1, DTLZ2.1, ZDT1.2, ZDT2.2, and ZDT6.2) and it gets the comparative result on DTLZ2.2. This demonstrates the effectiveness of the introduction of LLE method adequately.

Overall, the introduction of locally linear embedding method into MOEA is beneficial to the optimization process. It can enhance the exploitation ability and feature the local

geometry of the population's distribution. The hybrid algorithm HMOEDA_LLE takes advantages of the genetic operations and probabilistic modelling and sampling method, besides the entropy-based criteria is helpful for the mixture of the two ways of evolutionary search.

V. CONCLUSION

In this study, a novel hybrid multi-objective estimation of distribution algorithm combing locally linear embedding (HMOEDA_LLE) is proposed. The evolutionary process of HMOEDA_LLE can be divided into two stages. In the first stage, LLE-based approach and traditional genetic operation: crossover and mutation are mixed together to produce offspring. After certain regularity in population distribution being detected, the estimation of distribution modelling and sampling method is adopted to generate new population. An entropy based criterion is introduced to guide the switch of the two different stages. To verify the performance of HMOEDA_LLE, eight representative multi-objective test problems with linear and nonlinear variable linkages are employed. Experiment results show that HMOEDA_LLE is superior to two state-of-the-art algorithms NSGA-II and RM-MEDA.

In this study, a fixed value of neighbor parameter K is adopted. In fact, for different problems the optimal value of K is not the same, how to choose a proper value adaptively is our

future work, and the best reproduction ratio of LLE-based approach and traditional genetic operations also deserves further study.

REFERENCES

[1] Coello Coello C A. Evolutionary multi-objective optimization: a historical view of the field [J]. Computational Intelligence Magazine, IEEE, 2006, 1(1): 28-36.

[2] Drechsler N, Drechsler R, Becker B. Multi-objective optimisation based on relation favour [C]//Evolutionary multi-criterion optimization. Springer Berlin Heidelberg, 2001: 154-166.

[3] Branke J, Kaußler T, Schmeck H. Guidance in evolutionary multi-objective optimization [J]. Advances in Engineering Software, 2001, 32(6): 499-507.

[4] Laumanns M, Thiele L, Deb K, et al. Combining convergence and diversity in evolutionary multiobjective optimization [J]. Evolutionary computation, 2002, 10(3): 263-282.

[5] Farina M, Amato P. A fuzzy definition of 'Optimality' for many-criteria optimization problems [J]. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2004, 34(3): 315-326.

[6] Brockhoff D, Zitzler E. Are all objectives necessary? On dimensionality reduction in evolutionary multiobjective optimization [M]//Parallel Problem Solving from Nature-PPSN IX. Springer Berlin Heidelberg, 2006: 533-542.

[7] Hernández-Díaz A G, Santana-Quintero L V, Coello Coello C A, et al. Pareto-adaptive ε-dominance [J]. Evolutionary Computation, 2007, 15(4): 493-517.

[8] Köppen M, Yoshida K. Substitute distance assignments in NSGA-II for handling many-objective optimization problems [C]//Evolutionary Multi-Criterion Optimization. Springer Berlin Heidelberg, 2007: 727-741.

[9] Yang S, Li M, Liu X, et al. A grid-based evolutionary algorithm for many-objective optimization [J]. 2013.

[10] Coello C A C, Pulido G T, Lechuga M S. Handling multiple objectives with particle swarm optimization [J]. Evolutionary Computation, IEEE Transactions on, 2004, 8(3): 256-279.

[11] Gong M, Jiao L, Du H, et al. Multiobjective immune algorithm with non-dominated neighbor-based selection [J]. Evolutionary Computation, 2008, 16(2): 225-255.

[12] Zhang Q, Li H. MOEA/D: A multiobjective evolutionary algorithm based on decomposition [J]. Evolutionary Computation, IEEE Transactions on, 2007, 11(6): 712-731.

[13] Zhang Q, Zhou A, Jin Y. RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm [J]. Evolutionary Computation, IEEE Transactions on, 2008, 12(1): 41-63.

[14] Zhou A, Qu B Y, Li H, et al. Multiobjective evolutionary algorithms: A survey of the state of the art [J]. Swarm and Evolutionary Computation, 2011, 1(1): 32-49.

[15] Miettinen K. Nonlinear multiobjective optimization [M]. Springer, 1999.

[16] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290(5500): 2323-2326.

[17] Chen J, Liu Y. Locally linear embedding: a survey [J]. Artificial Intelligence Review, 2011, 36(1): 29-48.

[18] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. Evolutionary Computation, IEEE Transactions on, 2002, 6(2): 182-197.

[19] Seung H S, Lee D D. The manifold ways of perception [J]. Science, 2000, 290(5500): 2268-2269.

[20] Tenenbaum J B, De Silva V, Langford J C. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290(5500): 2319-2323.

[21] Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation [J]. Neural computation, 2003, 15(6): 1373-1396.

[22] Donoho D L, Grimes C. Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data [J]. Proceedings of the National Academy of Sciences, 2003, 100(10): 5591-5596.

[23] Zhang Z, Zha H. Principal Manifolds and Nonlinear Dimensionality Reduction via Tangent Space Alignment [J]. SIAM Journal on Scientific Computing, 2005, 26(1): 313-338.

[24] Farhang-Mehr A, Azarm S. Diversity assessment of Pareto optimal solution sets: an entropy approach[C]//Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on. IEEE, 2002, 1: 723-728.

[25] Gunawan S, Farhang-Mehr A, Azarm S. Multi-level multi-objective genetic algorithm using entropy to preserve diversity[C]//Evolutionary Multi-Criterion Optimization. Springer Berlin Heidelberg, 2003: 148-161.

[26] Ocenasek J. Entropy-based convergence measurement in discrete estimation of distribution algorithms[M]//Towards a New Evolutionary Computation. Springer Berlin Heidelberg, 2006: 39-50.

[27] Wang Y N, Wu L H, Yuan X F. Multi-objective self-adaptive differential evolution with elitist archive and crowding entropy-based diversity measure[J]. Soft Computing, 2010, 14(3): 193-209.

[28] Wang Y N, Wu L H, Yuan X F. Multi-objective self-adaptive differential evolution with elitist archive and crowding entropy-based diversity measure[J]. Soft Computing, 2010, 14(3): 193-209.

[29] Qin Y, Ji J, Liu C. An entropy-based multiobjective evolutionary algorithm with an enhanced elite mechanism[J]. Applied Computational Intelligence and Soft Computing, 2012, 2012: 17.

[30] LinLin W, Yunfang C. Diversity Based on Entropy: A Novel Evaluation Criterion in Multi-objective Optimization Algorithm[J]. International Journal of Intelligent Systems and Applications (IJISA), 2012, 4(10): 113.

[31] Zhou A, Zhang Q, Jin Y. Approximating the set of Pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm[J]. Evolutionary Computation, IEEE Transactions on, 2009, 13(5): 1167-1189.

[32] Yang D, Jiao L, Gong M, et al. Hybrid multiobjective estimation of distribution algorithm by local linear embedding and an immune inspired algorithm[C]//Evolutionary Computation, 2009. CEC'09. IEEE Congress on. IEEE, 2009: 463-470.

[33] Leung Y W, Wang Y. An orthogonal genetic algorithm with quantization for global numerical optimization[J]. Evolutionary Computation, IEEE Transactions on, 2001, 5(1): 41-53.

[34] Zitzler E, Deb K, Thiele L. Comparison of multiobjective evolutionary algorithms: Empirical results[J]. Evolutionary computation, 2000, 8(2): 173-195.

[35] Deb K, Thiele L, Laumanns M, et al. Scalable multi-objective optimization test problems[C]//Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA). Proceedings of the Congress on Evolutionary Computation (CEC-2002),(Honolulu, USA), 2002: 825-830.