A New EDA with Dimension Reduction Technique for Large Scale Many-Objective Optimization

Mingli Shi, Lianbo Ma^(✉), and Guangming Yang^(✉)

Northeastern University, Shenyang, China
305337697@qq.com, {malb, yanggm}@swc.neu.edu.cn

Abstract. The performance of many-objective evolutionary algorithms deteriorates appreciably in solving large-scale many-objective optimization problems (MaOPs) which encompass more than hundreds variables. One of the known rationales is the curse of dimensionality. Estimation of distribution algorithms sample new solutions with a probabilistic model built from the statistics extracting over the existing solutions so as to mitigate the adverse impact of genetic operators. In this paper, an Gaussian Bayesian network-based estimation of distribution algorithm (GBNEDA-DR) is proposed to effectively tackle continued large-scale MaOPs. In the proposed algorithm, dimension reduction technique (i.e. LPP) is employed in the decision space to speed up the estimation search of the proposed algorithm. The experimental results show that the proposed algorithm performs significantly better on many of the problems and for different decision space dimensions, and achieves comparable results on some compared with many existing algorithms.

Keywords: Estimation of distribution algorithm · Gaussian Bayesian network · Dimension reduction

1 Introduction

Many-objective optimization problems (MaOPs) refer to the problems that involve a large number of conflicting objectives to be optimized simultaneously. Due to the complexity and difficulty of MaOPs, it is meaningful to investigate the ways of dealing with a given difficult MaOP. The main difficulty associated with MaOPs is often referred to as the curse of dimensionality. Currently, scalability with respect to the number of objectives has attracted considerable research interests. This is due to the fact that in many-objective optimization, most candidate solutions become nondominated with each other, thus causing

L. Ma and G. Yang—This work is supported by the National Natural Science Foundation of China under Grant No. 61773103, Fundamental Research Funds for the Central Universities No. N180408019 and Huawei HIRP project under Grant No. HO2019085002.

failure of dominance-based selection strategies in traditional MOEAs. To tackle the MaOPs, a number of new multiobjective evolutionary algorithms (MOEAs) have been proposed, such as NSGA-III [6], MOEA/D [22], Tk-MaOEA [12], IBEA [26], and KnEA [24]. However, in spite of the various approaches that are focused on the scalability of MOEAs to the number of objectives, scalability in terms of the number of decision variables remains inadequately explored.

Recently, large-scale optimization has already attracted certain interests in the single-objective optimization problem (SOP). Akin to the large-scale single-objective optimization [14, 19], some authors have attempted to adapt existing techniques for large-scale single-objective optimization to the MaOPs context, such as MOEA/DVA [13], LMEA [23], WOF [25], MOEA/D-RDG [18] and CCGDE3 [2]. The main idea of these methods is divide-and-conquer strategy. Since the target of many-objective optimization is different from that of single-objective optimization [11]. Thus, it is not trivial to generalize such divide-and-conquer strategy proposed in single objective optimization problem (SOP) to solve MaOPs because the objective functions of an MaOP are conflicting with one another.

Deb et al. [4] concluded that the performances of MOEAs are significantly influenced by the genetic operators (i.e. crossover and mutation) which cannot ensure to generate promising offspring. Estimation of distribution algorithms (EDAs) are a relatively new computational paradigm proposed to generate new offspring. EDA generates new solutions by applying probabilistic models, which inferred from a set of selected solutions. These models capture statistics about the values of problem variables and the dependencies among these variables [9].

It has been observed that under mild smoothness conditions, the Pareto set of a continuous MOP is a piecewise continuous $(m-1)$ -dimensional manifold, where m is the number of the objectives. In paper [21], it has shown that reproduction of new trial solutions based on this regularity property can effectively cope with the variable linkages in continuous MOPs. Hence, this characteristics of MOPs can be integrated into EDA to effectively solve large-scale MOPs. This paper proposes a new EDA applied to large-scale MOPs, called GBNEDA-DR. The idea is to combine EDA with dimension reduction methods (i.e. LPP [8]), which are responsible for embedding the solutions used on the probabilistic models in low dimension space.

The rest of this paper is organized as follows. In Sect. 2, we briefly recall some related work on MOEAs for solving large-scale MOPs. Section 3 describes the proposed algorithm GBNEDA/DR. Section 4 illustrates and analyzes the experimental results. Section 5 concludes this paper.

2 Related Works

2.1 Gaussian Bayesian Network

Bayesian networks [15] are multivariate probabilistic graphical models, consisting of two components.

- 1) The structure, represented by a directed acyclic graph (DAG), where the nodes are the problem variables and the arcs are conditional (in)dependencies between twins of variables;
- 2) The parameters, expressing for each variable X_i the conditional probability of each of its values, given different value-settings of its parent variables ($Pa(X_i)$) according to the structure, i.e.

$$p(x_i|pa(X_i)) \quad (1)$$

where $Pa(X_i)$ is a value-setting for the parent variables in $Pa(X_i)$.

In domains with continuous-valued variables, it is usually assumed that the variables follow a Gaussian distribution. The Bayesian network learned for a set of variables, having a multivariate Gaussian distribution $p(x) = N(\mu, \Sigma)$ as their joint probability function, is called a Gaussian Bayesian network (GBN). Here, μ is the mean vector and Σ is the covariance matrix of the distribution.

The structure of a GBN is similar to any other Bayesian network. However, for each node the conditional probability represented by the parameters is a univariate Gaussian distribution, which is determined by the values of the parent variables [7]

$$p(x|pa(X_i)) = N(\mu_i + \sum_{X_j \in Pa(X_i)} \omega_{ij}(x_j - \mu_j), v_i^2) \quad (2)$$

where μ_i is the mean of variable X_i , v_i is the conditional standard deviation of the distribution, and regression coefficients ω_{ij} specify the importance of each of the parents. x_j is the corresponding value of X_j in $Pa(X_i)$. These are the parameters stored in each node of a GBN.

2.2 Estimation of Distribution Algorithm

Traditional genetic operators used for generating new solutions in evolutionary algorithms act almost blindly and are very likely to disrupt the good sub-solutions found so far which will affect the optimization convergence. This disruption is more likely to occur as the correlation between problem variables increases, rendering the algorithm inefficient for such problems. EDAs make use of probabilistic models to replace the genetic operators in order to overcome this shortcoming. A general framework of EDAs is illustrated in Algorithm 1.

Typically, the EDAs-based MOEAs are broadly classified into two categories based on their estimation models. The first category covers the Bayesian network-based EDAs. For example, multiobjective Bayesian optimization algorithm (BOA) [10]. The other category is often known as the mixture probability model-based EDAs. Such as, in [16], the multiobjective hierarchical BOA was designed by the mixture Bayesian network-based probabilistic model for discrete MOPs. It is believed that EDAs are capable of solving MaOPs without suffering the disadvantages of MOEAs with traditional genetic operators.

Algorithm 1. Framework of an EDA

```

1:  $t = 0$ 
2:  $P_t$  = Randomly initialize the population
3: while termination is not satisfied do
4:    $M$  = Built probabilistic models from  $P_t$ 
5:    $t = t + 1$ 
6:    $U_t$  = Generate offspring from  $M$ 
7:    $P_t$  = Select promising solutions from  $U_t \cup P_{t-1}$ 
8: end while
9: return  $P_t$ 

```

2.3 Locality Preserving Projections

Locality Preserving Projection (LPP) [8] is a general method for manifold learning. Though it is still a linear technique, it seems to recover important aspects of the intrinsic nonlinear manifold structure by preserving local structure. In many real-world applications, the local structure is more important. In this section, we give a brief description of LPP. The complete derivation and theoretical justifications of LPP can be traced back to [8]. LPP seeks to preserve the intrinsic geometry of the data and local structure. Given a set $x_1, x_2, \dots, x_m$ in R^n , find a transformation matrix A that maps these m points to a set of points $y_1, y_2, \dots, y_m$ in R^l . The objective function of LPP is as follows:

$$\min \sum_{ij} (y_i - y_j)^2 S_{ij} \quad (3)$$

where the matrix S is a similarity matrix. The following restrictions are imposed on the equation (3):

$$y^T D y = 1 \quad (4)$$

D is a diagonal matrix, its entries are column sum of S , $D_{ii} = \sum_j S_{ij}$. A possible way of defining S is as follows:

$$S_{ij} = \begin{cases} \exp(-\|x_i - x_j\|^2 / t), & \|x_i - x_j\|^2 < \varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Here, ε defines the radius of the local neighborhood and t is a custom parameter.

3 Proposed Algorithm

In this section, the framework of the proposed algorithm, i.e., GBNEDA-DR, is given first. Then elaborate on the four important components in it, i.e., reducing the dimension of decision space, building the probability model, repairing and environmental selection.

3.1 Framework of the Proposed Algorithm

The framework of the proposed algorithm is listed in Algorithm 2. It consists of the following three main steps. First, a population of N candidate solutions is randomly initialized. Second, by constructing the Gauss Bayesian network model, N new individuals are sampled and generated. This step has many sub-steps. Algorithm 2 shows an overview of this step. Next, the main sub-steps are described in detail. Finally, an environmental selection strategy is used to select excellent individuals, thus, a set of solutions with a better quality in convergence and diversity are obtained. The three steps are performed one by one in a limit number of fitness evaluations. In addition, maximum fitness evaluation, population size and respective threshold for dimension reduction and model building need to be made available prior to the proposed algorithm running.

Algorithm 2. Main framework of GBNEDA-DR

Input: N (population size) $nSel$ (the number of Sel_Pop)

Output: P (final population)

```

1:  $P = \text{Initialize}(N)$ 
2: while termination criterion not fulfilled do
3:    $Sel\_Pop = \text{Selection}(P, nSel)$ 
4:    $Sel\_Pop' = \text{LPP}(Sel\_Pop)$ 
5:    $\Phi = \text{Establishing probability model in low dimensional space by } Sel\_Pop'$ 
6:    $offspring = \text{sampling}(\Phi, N)$ 
7:    $offspring' = \text{Inverse transformation of } (offspring)$ 
8:    $offspring'' = \text{repairing}(offspring')$ 
9:    $P = \text{Environmental Selection}(P, offspring'')$ 
10: end while
11: return  $P$ 

```

3.2 Reducing the Dimension of Decision Space

In this paper, LPP [8] is used to reduce the volume of exploration space to speed up the search of sampling new solutions. A set of Pareto solutions (PS) is selected to be the training data and then exploitation is performed in the subspace. In this paper, the LPP is employed because: 1) LPP is based on the inner geometric structure of manifolds, it shows the stability of embedding; 2) LPP algorithm is a linear dimension reduction method.

At the beginning of dimensionality reduction, the Pareto solutions, which are denoted by Sp , are selected from the population. For convenience of the development, a matrix X is used to represent Sp . Specifically, each row in X denotes one solution while the columns refer to the different dimension of decision variables.

In most PCA-based methods, none of solutions sampled from the reduced space needs to be operated in the original space. However, in the proposed

design, the solutions must be transformed back to the original space for fitness evaluation. After the new offspring is projected back to its original space, it can participate in environmental selection.

The contribution should contain no more than four levels of headings. Table 1 gives a summary of all heading levels.

3.3 Building the Probability Model

The probabilistic model used in this paper for model learning is the Gaussian Bayesian network (GBN). A search+score strategy is used in GBNEDA-DR to learn the GBN from the data. In this strategy, a search algorithm is employed to explore the space of possible GBN structures to find a structure that closely matches the data. The quality of different GBN structures obtained in this search process is measured using a scoring metric, usually computed from data. A greedy local search algorithm is used to learn the structure of GBN. The algorithm finally returns the highest scoring network in all these subsearches. The Bayesian information criterion (BIC) [17] is used to score possible GBN structures.

The parameters of this type of GBN are computed from the mean vector and covariance matrix of the Gaussian distribution (GD) estimated for the joint vector of variables and objectives: $N(\hat{\mu}_{\langle 1, n+m \rangle}, \hat{\Sigma}_{\langle n+m, n+m \rangle})$. Usually the maximum likelihood (ML) estimation is used to estimate the parameters of GD (the mean vector and covariance matrix) from the data.

3.4 Repairing Cross-Border Values

After the new offspring is projected back to its original space, values in some dimensions are illegal. Repairing methods are commonly used to guarantee the feasibility of solutions. They modify (repair) a given individual to guarantee that the constraints are satisfied. In GBNEDA-DR the repairing procedure is invoked at every generation. In this paper, there are two methods to fix illegal values. The first method changes the values of each out of range variable to the minimum (respectively maximum) bounds if variables are under (respectively over) the variables ranges. The second method truncates the values of each out of range variable to a random value within the feasible range. Each method is triggered in a random way.

3.5 Environmental Selection

The purpose of environmental selection is for maintaining a size of population with the same number of initialized population. Figure 1 shows an example of the structure of an GBN. The set of arcs in the structure is partitioned into three subsets. The red arc represents the dependency of decision variables, the blue arc represents the dependency of objective variables, and the black arc represents the dependency between the decision variables and the objective variables. An

analysis of the structures learnt by GBNEDA-DR along the evolution path show that the proposed algorithm is able to distinguish between relevant and irrelevant variables. It can also capture stronger dependencies between similar objectives. The dependencies learnt between objectives in the MOP structure can be used to analyze relationships like conflict or redundancy between sets of objectives.

Fig. 1. An example of a Gaussian Bayesian network structure

With an increase of objective, almost the entire population acquires the same-rank of non-domination. This makes the Pareto-dominance based primary selection ineffective. Objective reduction approaches can solve this problem very well. In this paper, with the help of GBN, GBNEDA-DR can remove redundant objectives and make effective use of non-dominated sorting. In other words, if a class node in the GBN has no parent, it is not redundant, whereas if a class node (i.e. a objective) has a parent, it does not participate in the comparison in the non-dominant sort. For example, there are no arrows pointing at f_1 , f_3 and f_4 in Fig. 1. But there are arrows pointing at f_2 , therefore, f_2 is redundant and can be ignored in non-dominant sort. The framework of the environmental selection in GBNEDA-DR is similar to that of VaEA [20]. This method performs a careful elitist preservation strategy and maintains diversity among solutions by putting more emphasis on solutions that have the maximum vector angle to individuals which have already survived (more details can be found in [20]).

4 Experiments

Since the proposed GBNEDA-DR is an EDA-based algorithm for solving large-scale MaOPs, we verify the performance of GBNEDA-DR by empirically comparing it with four state-of-the-art MOEAs covering two categories: 1) traditional MOEAs (MOEA/DVA [13] and NSGA-III [6]) and 2) EDA-based evolutionary algorithm (RM-MEDA [21] and MBN-EDA [9]). In the following sections, the selected benchmark test problems are introduced first. Then, the performance indicators and parameter settings are introduced and declared respectively. Finally, experiments on compared algorithms are performed and their results measured by the selected performance indicators are analyzed.

4.1 Test Problems and Performance Metrics

The experiments are conducted on 14 test problems taken from two widely used test suites, DTLZ and LSMOP. The DTLZ test suite is a class of widely used benchmark problems for testing the performance of MOEAs. The first four test problems are DTZL1-DTLZ4 taken from the DTLZ test suite. DTLZ test problems are considered less challengeable. LSMOP is a recently developed test suite for large-scale multiobjective and many-objective optimization. It can better reflect the challenges in large-scale MaOPs than existing test suits. Hence, the LSMOP1-LSMOP5 problems from the LSMOP test suit are included into the considered benchmark test problems.

One widely used performance metric, inverted generational distance (IGD) [3], which can simultaneously quantify the performance in convergence and diversity of the algorithms, is adopted in these experiments. The definition of the IGD from P^* (a set of uniformly distributed points in the objective space along the PF) to P (an approximation to the PF) is illustrated as follow,

$$IGD(P^*, P) = \frac{\sum_{v \in P^*} d(v, P)}{|P^*|} \quad (6)$$

where $d(v, P)$ represents the minimum Euclidean distance between v and the points in P .

4.2 Parameter Settings

In order to show the best effect of all algorithms, all compared algorithms adopt the recommended parameter values to achieve the best performance. To be specific, the parameter settings for all conducted experiments are as follows.

- 1) Crossover and Mutation: SBX [1] and polynomial mutations [5] are employed by traditional MaOEAs as the crossover operator and mutation operator, respectively. The crossover probability and mutation probability are set to $P_c = 1.0$ and $P_m = 1/D$, respectively, where D denotes the number of decision variables. In addition, the distribution index of NSGA-III is set to be 30 according to the suggestions in [6] while others are set to be 20.
- 2) Population Sizing: The population size of NSGA-III cannot be arbitrarily specified, which is equal to the number of reference vectors, other peer algorithms adopt the same population size for a fair comparison. Furthermore, the two-layer reference vector generation strategy is adopted here. The settings for reference vector and population size are listed in Table 1. $H1$ and $H2$ are the simplex-lattice design factors for generating uniformly distributed reference vectors on the outer boundaries and the inside layers, respectively.
- 3) Other Settings: The number of evaluations is used as the termination criterion for all considered algorithms. The maximum number of function evaluations is set to 100 000 in all experiments. On each test instance, 10 independent runs are performed for each algorithm to obtain statistical results. For MOEA/DVA, the number of interaction analysis and the number of control

property analysis are set to the recommended values, namely, $NIA = 6$ and $NCA = 50$. For RM-MEDA, the number of clusters in local PCA varies in [31 50 50].

Table 1. Settings for reference vectors and population size.

M	# of division		# of reference vectors	Population size
	H1	H2		
3	14	0	120	120
5	5	0	126	128
8	3	2	156	156
10	4	2	265	265

Table 2. Mean and standard deviation results of IGD obtained on DTLZ1-5 (Dimension of decision space is 100).

Problem	DTLZ1		DTLZ2		DTLZ3		DTLZ4	
Obj.	3	5	3	5	3	5	3	5
RM-MEDA	2.69e+1	2.21e+1	2.43e-1	4.62e+0	3.20e+3	3.12e+3	4.64e-1	3.38e+0
	(2.18e-1)	(5.97e-1)	(1.80e-2)	(5.76e-1)	(7.26e+1)	(2.48e+2)	(5.97e-2)	(2.45e-1)
MBN-EDA	5.10e+1	1.58e+1	2.47e-1	3.02e+0	2.39e+2	6.52e+3	2.62e-1	1.42e+0
	(1.88e-2)	(4.20e-2)	(2.05e-3)	(1.63e+0)	(5.03e+1)	(7.36e+2)	(4.48e-4)	(1.37e-1)
GBNEDA-DR	2.17e+1	3.82e+1	5.67e-1	7.58e-1	6.62e+3	1.24e+3	9.70e-1	1.14e+0
	(1.23+0)	(1.02e-2)	(1.17e-2)	(9.80e-3)	(9.13e+2)	(1.18e+3)	(7.14e-3)	(3.54e-3)

4.3 Results

Tables 2 present the results of the IGD metric values of the three compared algorithms on the five DTLZ test problems with 100 decision variables. As can be seen from the tables, the IGD values obtained by GBNEDA-DR on each test problem are consistently good as the number of objectives increases from 3 to 5, which confirms a promising scalability of GBNEDA-DR. The results on the eight LSMOP test problems with 1000 decision variables are given in Table 3, with both the mean and standard deviation of the IGD values averaged over 10 independent runs being listed for the five compared MOEAs, where the best mean among the five compared algorithms is highlighted. It is clearly shown in Table 3 that when the number of decision variables is increased, the proposed algorithm obtains all the best mean IGD results against its competitors over LSMOP1 and LSMOP4 with 5-, and 10-objective. In summary, the dimension reduction technique can significantly improve the performance of the proposed algorithm especially in solving large scale MaOPs.

Table 3. Mean and standard deviation results of IGD obtained on LSMOP1-8 (Dimension of decision space is 1000).

Problem	Obj.	NSGA-III	MOEA/DVA	RM-MEDA	MBN-EDA	GBNEDA-DR
LSMOP1	5	7.5893e+0 (2.64e-1)	9.9607e+0 (3.66e-1)	1.0004e+1 (3.07e-1)	9.9980e+0 (7.19e-1)	4.1242e+0 (1.04e+0)
		8.9901e+0 (2.33e-1)	9.6706e+0 (6.75e-1)	9.4195e+0 (1.74e-1)	9.0837e+0 (1.59e+0)	7.3057e+0 (1.60e-1)
	10	8.9592e-2 (1.97e-5)	1.2071e-1 (7.29e-3)	1.1071e-1 (1.98e-4)	1.0065e-1 (1.08e-3)	1.0832e-1 (1.35e-3)
		3.6174e-1 (1.37e-2)	3.8925e-1 (3.90e-4)	2.8695e-1 (2.29e-3)	2.5539e-1 (9.35e-3)	2.0060e-1 (4.26e-4)
LSMOP2	5	1.4664e+1 (1.05e-1)	7.2671e+1 (4.77e+1)	1.8609e+1 (1.35e-1)	2.0636e+1 (2.26e-1)	2.1846e+1 (6.18e-1)
		2.8425e+1 (1.95e+0)	1.3015e+3 (9.78e+2)	2.4350e+1 (4.12e+0)	2.4913e+1 (5.76e+0)	1.8471e+2 (1.04e+1)
	10	1.7878e-1 (1.31e-3)	1.9876e-1 (1.47e-4)	2.0692e-1 (2.99e-3)	1.9606e-1 (8.35e-5)	1.6982e-1 (3.67e-4)
		3.6897e-1 (7.09e-3)	3.8537e-1 (9.83e-3)	3.0729e-1 (2.11e-3)	2.9669e-1 (9.93e-3)	2.3592e-1 (4.51e-4)
LSMOP3	5	1.2234e+1 (1.75e-1)	8.7623e+1 (5.02e+0)	6.4356e+0 (2.82e+0)	9.5447e+0 (5.39e-2)	1.4250e+1 (3.61e-1)
		1.5426e+1 (9.83e-1)	4.8658e+1 (4.81e+1)	2.1896e+1 (2.10e-1)	1.8065e+1 (5.37e-1)	7.3963e+0 (3.31e-1)
	10	1.2234e+1 (1.75e-1)	8.7623e+1 (5.02e+0)	6.4356e+0 (2.82e+0)	9.5447e+0 (5.39e-2)	1.4250e+1 (3.61e-1)
		1.5426e+1 (9.83e-1)	4.8658e+1 (4.81e+1)	2.1896e+1 (2.10e-1)	1.8065e+1 (5.37e-1)	7.3963e+0 (3.31e-1)

From these results, we can see that GBNEDA-DR outperforms MOEA/DVA, RM-MEDA, MBN-EDA, and NSGA-III on DTLZ and LSMOP test problems in terms of IGD, especially for problems with more than 100 decision variables. Therefore, we can conclude that the proposed GBNEDA-DR is effective to handle large-scale MaOPs.

5 Conclusion

In this paper, we have proposed a Gaussian Bayesian network-based EDA, termed GBNEDA-DR, for solving large-scale MaOPs. GBNEDA-DR, the proposed algorithm, models a promising area in the search space by a probability model, which is used to generate new solutions. This model can capture the relationships between variables like other EDAs. It must be pointed out that the variables here are those in the low-dimensional space after dimensionality reduction. Because dimension reduction technique (i.e. LPP) is utilized to reduce the cost of exploitation and exploration. Through experimental evaluation we showed significant improvements of the performance on various benchmark problems compared to classical optimization methods as well as existing large scale approaches.

This paper demonstrates that the idea of using EDA to generates new solutions for large-scale MaOPs is very promising. In our future research, we would

also like to further improve the computational efficiency of the dimension reduction procedure, which is the main computational cost of the proposed GBNEDA-DR.

References

1. Agrawal, R., Deb, K., Agrawal, R.: Simulated binary crossover for continuous search space. *Complex Syst.* **9**(2), 115–148 (2000)
2. Antonio, L.M., Coello, C.A.C.: Use of cooperative coevolution for solving large scale multiobjective optimization problems. In: 2013 IEEE Congress on Evolutionary Computation, pp. 2758–2765, June 2013
3. Bosman, P.A.N., Thierens, D.: The balance between proximity and diversity in multiobjective evolutionary algorithms. *IEEE Trans. Evol. Comput.* **7**(2), 174–188 (2003)
4. Deb, K., Sinha, A., Kukkonen, S.: Multi-objective test problems, linkages, and evolutionary methodologies. In: *Proceeding of Genetic and Evolutionary Computation Conference*, vol. 2, pp. 1141–1148, January 2006
5. Deb, K., Goyal, M.: A combined genetic adaptive search (GeneAS) for engineering design. *Comput. Sci. Inf.* **26** (1999)
6. Deb, K., Jain, H.: An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. *IEEE Trans. Evol. Comput.* **18**(4), 577–601 (2014)
7. Geiger, D., Heckerman, D.: Learning Gaussian networks. In: *Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence (UAI-1994)*, February 2013
8. He, X., Niyogi, P.: Locality preserving projections (LPP). *IEEE Trans. Reliab. TR* **16** (2002)
9. Karshenas, H., Santana, R., Bielza, C., Larrañaga, P.: Multiobjective estimation of distribution algorithm based on joint modeling of objectives and variables. *IEEE Trans. Evol. Comput.* **18**(4), 519–542 (2014). <https://doi.org/10.1109/TEVC.2013.2281524>
10. Khan, N., Goldberg, D., Pelikan, M.: Multiple-objective Bayesian optimization algorithm, p. 684, January 2002
11. Lianbo, M., Cheng, S., Shi, Y.: Enhancing learning efficiency of brain storm optimization via orthogonal learning design. *IEEE Trans. Syst. Man Cybern. Syst.*, 1–20 (2020). <https://doi.org/10.1109/TSMC.2020.2963943>
12. Ma, L., et al.: A novel many-objective evolutionary algorithm based on transfer matrix with Kriging model. *Inf. Sci.* **509**, 437–456 (2020)
13. Ma, X., et al.: A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables. *IEEE Trans. Evol. Comput.* **20**(2), 275–298 (2016)
14. Mahdavi, S., Shiri, M.E., Rahnamayan, S.: Metaheuristics in large-scale global continues optimization: a survey. *Inf. Sci.* **295**, 407–428 (2015)
15. Pearl, J.: Bayesian networks: a model of self-activated memory for evidential reasoning. In: *Proceedings of the 7th Conference of the Cognitive Science Society* (1985)
16. Pelikan, M., Sastry, K., Goldberg, D.: Multiobjective hBOA, clustering, and scalability. In: *GECCO 2005 - Genetic and Evolutionary Computation Conference*, March 2005

17. Schwarz, G.: Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978)
18. Song, A., Yang, Q., Chen, W., Zhang, J.: A random-based dynamic grouping strategy for large scale multi-objective optimization. In: 2016 IEEE Congress on Evolutionary Computation (CEC), pp. 468–475, July 2016. <https://doi.org/10.1109/CEC.2016.7743831>
19. Tang, K., Li, X., Suganthan, P., Yang, Z., Weise, T.: Benchmark functions for the CEC 2008 special session and competition on large scale global optimization, December 2009
20. Xiang, Y., Zhou, Y., Li, M., Chen, Z.: A vector angle-based evolutionary algorithm for unconstrained many-objective optimization. *IEEE Trans. Evol. Comput.* **21**(1), 131–152 (2017)
21. Zhang, Q., Zhou, A., Jin, Y.: RM-MEDA: a regularity model-based multiobjective estimation of distribution algorithm. *IEEE Trans. Evol. Comput.* **12**(1), 41–63 (2008). <https://doi.org/10.1109/TEVC.2007.894202>
22. Zhang, Q., Hui, L.: MOEA/D: a multiobjective evolutionary algorithm based on decomposition. *IEEE Trans. Evol. Comput.* **11**(6), 712–731 (2008)
23. Zhang, X., Tian, Y., Cheng, R., Jin, Y.: A decision variable clustering-based evolutionary algorithm for large-scale many-objective optimization. *IEEE Trans. Evol. Comput.* **22**, 99 (2016)
24. Zhang, X., Tian, Y., Jin, Y.: A knee point driven evolutionary algorithm for many-objective optimization. *IEEE Trans. Evol. Comput.* **19**(6), 761–776 (2014)
25. Zille, H., Ishibuchi, H., Mostaghim, S., Nojima, Y.: A framework for large-scale multiobjective optimization based on problem transformation. *IEEE Trans. Evol. Comput.* (2017). <https://doi.org/10.1109/TEVC.2017.2704782>
26. Zitzler, E., Künzli, S.: Indicator-based selection in multiobjective search. In: Yao, X., et al. (eds.) PPSN 2004. LNCS, vol. 3242, pp. 832–842. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30217-9_84