

## Research Article

# Multiobjective Optimization Using Cross-Entropy Approach

Karim Sebaa,<sup>1</sup> Abdelhalim Tlemçani,<sup>1</sup> Mounir Bouhedda,<sup>2</sup> and Noureddine Henini<sup>1</sup>

<sup>1</sup> Research Laboratory Electrical Engineering and Automation, University of Médéa, Quartier Ain Dheb, 26000 Medea, Algeria

<sup>2</sup> Laboratory of Advanced Electronic Systems, University of Médéa, Quartier Ain Dheb, 26000 Medea, Algeria

Correspondence should be addressed to Karim Sebaa; karim.sebaa@gmail.com

Received 5 June 2013; Revised 12 September 2013; Accepted 13 September 2013

Academic Editor: Ling Wang

Copyright © 2013 Karim Sebaa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new approach for multiobjective optimization is proposed in this paper. The method based on the cross-entropy method for single objective optimization (SO) is adapted to MO optimization by defining an adequate sorting criterion for selecting the best candidates samples. The selection is made by the nondominated sorting concept and crowding distance operator. The effectiveness of the approach is tested on several academic problems (e.g., Schaffer, Fonseca, Fleming, etc.). Its performances are compared with those of other multiobjective algorithms. Simulation results and comparisons based on several performance metrics demonstrate the effectiveness of the proposed method.

## 1. Introduction

Optimization is a basic tool for several decision making in engineering area. In these fields, we have many conflicting objectives to satisfy. Usually, we convert all objectives to one, single objective (SO) function. The goal is to find out the maximum/minimum of this SO function subject to maintain the physical constraints of the system. The results reflect a compromise between all objectives. The idea is to formulate the function to achieve this desired compromise.

By aggregating all objectives in a weighted function, or by transforming all objectives into only one objective, and retaining one objective which will be added to constraints, the conversion from MO to SO is done. But the weakness and the limitation of the aggregation are as follows.

- (1) The requirement of a prior knowledge about the relative importance of the objectives, and the limits on the objectives that are converted into constraints.
- (2) The aggregated function leads to only one solution.
- (3) Trade-offs between objectives cannot be easily evaluated.
- (4) The solution may not be attainable unless the search space is convex.

The aggregation is not recommended for the systems with conflicting objectives. Also, we need to know all possible

solutions of all objectives simultaneously. In the business word it is called “the trade-off analysis”. There are several areas in engineering where the performing of the trade-off analysis is necessary, such as the following.

- (1) The design of controllers while reducing the cost, which are two conflicting objectives.
- (2) The placement of more functional blocks on chip while minimizing the chip area and/or power dissipation.
- (3) The finding of the vehicle which has the highest range while at the same time consuming minimum amount of energy.
- (4) The minimization of the operation cost while maintaining a stable work force [1–3].

MO problems are more difficult to solve compared to the SO problems. In case of SO optimization solution is unique. However, in the MO optimization there is a set of acceptable trade-off optimal solutions which is called the Paretofront. In fact, MO optimization is considered as analytical stage of the multicriteria decision making (MCDM) procedure; it consists of the determination of all solutions for the MO problem, which are the optimal in Pareto logic [4]. The preferred solution, desired by the designer or decision maker, (DM) was selected from it.

Pareto set has several advantages. It contains the solutions that are optimum from an “overall” standpoint and allows the DM to make an informed decision by seeing a wide range of compromise (trade-off). This characteristic is helpful, because it provides better knowledge of the system according to the consequences of decisions [5].

There are many metaheuristic methods which use the Pareto ranking in order to determine the probability of replication of an individual. The basic idea is to find the set of nondominated individuals in the population. There are several meta-heuristic methods such as Niched Pareto Genetic Algorithm (NPGA) [6], Nondominated Sorting Genetic Algorithm (NSGA) [7], Strength Pareto Evolutionary Algorithm (SPEA) [8], the Nondominated Sorting Genetic Algorithm-II (NSGAII) [9], and, recently, Multiobjective Particle Swarm Optimization (MOPSO) [10–12].

In the determinist methods there is only one method, the Normal boundary Intersection (NBI) based [13] to generate *Pareto optimal solutions* for a general nonlinear multiobjective optimization.

Since 1997, the cross-entropy (CE) (or Kullback-Leibler CE) was proposed by Rubinstein for solving rare event simulation problems [14] and was afterward extended to solving combinatorial problems and continuous mono objective optimization. Several randomized optimization algorithms based on CE method have been proposed in the literature and have been shown to lead to good performances on many optimization problems, often outperforming other randomized algorithms [15]. In this decade, CE method has been applied in several engineering applications (see, e.g., [16–18]).

In [19] the authors have extended the negative-log-likelihood (NLL) approach on combination with a multiobjective evolutionary algorithms (MOEA) to resolve an Multiobjective Optimization, and in [20] the CE method was extended to multiobjective optimization in general and to multiobjective water distribution systems design in particular; this was done by converting the MO optimization to the SO optimization via a weighting method. Hence, the Pareto set will be generated by using different weighting of objectives. Recently, in [21], an extension of the CE method to MO optimization where the CE’s parameters are tuning based on information collected from clustered nondominated solution based on fuzzy c-means algorithm (FCM), and in [22] a cross-entropy used with Pareto ranking is developed.

Motivated by the satisfaction given through the applications of the CE method for the resolution of the SO optimization, in this paper an original use of the CE to resolve an MO optimization was presented. A novel approach for computing the *Pareto optimal* front based on rare event simulation, the nondominated sorting ranking, and a crowding operator [9] to select the elitist solutions was proposed. It is simple compared to the algorithm proposed in [21].

This document is divided into three parts. In the first part, the basics such as the rare event simulation (RES), the Nondomination, the crowded distance computation and the crowded distance operator are defined. In the second one, the proposed algorithm is presented, and, at the last, the testing

application of the proposed algorithm via very known test problem was done, a comparison with other meta-heuristic methods, and concurrent MOCE approach was done.

## 2. The Multiobjective Optimization via CE Method

The CE method has been designed to solve rare event simulation problems and extended to optimize SO problems. This section proposes an new extension to MO optimization. The first part is devoted to the presentation of CE method for SO optimization and is inspired by [16].

**2.1. From Rare Event Simulation to SO Optimization.** The CE method is a method to evaluate the probability that a rare event occurs. Let  $X$  be a random variable taking its values in some space  $\mathcal{X}$  with a probability density function (pdf)  $f(\cdot)$  and  $S(\cdot)$  be a real value function defined on  $\mathcal{X}$ . The problem is to evaluate  $l = \mathbb{P}_{X \sim f}(S(x) \geq \gamma)$ , where  $\gamma$  is a given threshold. An estimator of this probability is given by (1) where  $I(S(X_i) \geq \gamma)$  is the indicator function of the event and  $N$  is the number of samples:

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N I(S(X_i) \geq \gamma). \quad (1)$$

(The indicator function  $I(S(X_i) \geq \gamma)$  is equal to 1 when  $\{S(X_i) \geq \gamma\} = \text{true}$  and 0 otherwise.)

If this probability is very low, this estimator requires a great number of samples. For example, in order to estimate  $l = 10^{-6}$ , with a relative error  $\kappa = 0.01$ , a sample size of  $N \approx 1/\kappa^2 l = 10^{10}$  is required. The basic idea of CE is to introduce an importance sampling pdf  $g(\cdot)$  to reduce the required number of samples. The new estimator given by (2) where the samples are drawn from the density  $g(\cdot)$  can then be used:

$$\hat{l} = \frac{1}{N} \sum_{i=1}^N I(S(X_i) \geq \gamma) \frac{f(X_i)}{g(X_i)}. \quad (2)$$

Obviously the best estimator would be based on the ideal importance sampling pdf given by (3) since  $l$  is constant using this “ideal” importance sampling (3) would lead to an estimator (2) having a zero variance

$$g^*(X) = \frac{I(S(X_i) \geq \gamma)}{l} f(X). \quad (3)$$

The main idea of the CE method for rare event simulation is to find inside a prior given set  $\mathcal{G}$  of pdfs, defined on  $\mathcal{X}$ , the element  $g(\cdot)$  such that its distance with the “ideal” sampling distribution is minimal.

A particularly convenient measure of distance between two pdfs  $a(\cdot)$  and  $b(\cdot)$  on  $\mathcal{X}$  is the Kullback-Leibler distance, which is also termed the cross-entropy between  $a(\cdot)$  and  $b(\cdot)$ . The Kullback-Leibler distance, which is not a “distance” in the formal sense since it is, for example, not symmetric, is defined as follows:

$$\mathcal{D}(a, b) = E_{X \sim a(\cdot)} \left[ \ln \frac{a(X)}{b(X)} \right]. \quad (4)$$

The CE method reduces the problem of finding an appropriate importance sampling pdf to the optimization problem given by

$$\arg \min_{g \in \mathcal{G}} \mathcal{D}(g^*, g). \quad (5)$$

One can show through simple mathematical derivations that solving (5) is equivalent to solve (6) which does not explicitly depend on  $l$  anymore:

$$\arg \max_{g \in \mathcal{G}} E_{X \sim f(\cdot)} [I(S(X) \geq \gamma) \ln g(X)]. \quad (6)$$

Using a given importance sampling with a pdf  $h(\cdot)$  ( $h \in \mathcal{G}$ ) it is possible to get an estimation of the solution of (6) by solving its stochastic counterpart (7) where the set of samples  $X_1, X_2, \dots, X_M$  is drawn according to  $h(\cdot)$  and  $W(x) = f(x)/h(x)$  is the likelihood ratio between  $f(\cdot)$  and  $h(\cdot)$ :

$$\arg \max_{g \in \mathcal{G}} \sum_{j=1}^M I(S(X_j) \geq \gamma) \ln g(X_j) W(X_j). \quad (7)$$

Under specific assumptions on  $\mathcal{X}$ ,  $f(\cdot)$ , and  $\mathcal{G}$  it is possible to analytically solve (7). For example, when  $\mathcal{X} \subset \mathbb{R}$ ,  $\mathcal{G}$  is the set of  $n$ -dimensional exponential distributions with independent components specified by (8), where  $u$  is the vector of parameters and  $X[i]$  the  $i$ th component of  $X$ , and  $f$  also belongs to this set; the parameters of the solution of (7) are given by (9)

$$\text{Exp}_n(X, u) = \exp \left( -\sum_{i=1}^n \frac{X[i]}{u[i]} \right) \prod_{i=1}^n \frac{1}{u[i]}, \quad (8)$$

$$v[i] = \frac{\sum_{j=1}^M I(S(X) \geq \gamma) X_j[i] W(X_j)}{\sum_{j=1}^M I(S(X) \geq \gamma) W(X_j)}. \quad (9)$$

Of course the solution of (7) is a better estimation of the solution of (6) when the number of samples such as  $S(X) \geq \gamma$  is high. It is then necessary to adopt an iterative approach to compute the pdf  $h$  to favor the occurrence of the desired event. If  $l$  is very low it is even necessary to introduce an iteration that increases the value of  $\gamma$ .

This approach evaluates the probability that the function  $S(X)$  is greater than or equal to a given value by drawing samples with a pdf that evolves in such a way that the number of event increases, that is, such that the value of  $S(X)$  is great. It is then attractive to use the approach to evaluate the maximal value of the function.

Let us consider the single objective optimization problem specified by (10) where  $\mathcal{X} \subset \mathbb{R}^n$  and  $S(X)$  is a real value objective function. The basic idea of CE for SO optimization is to use a stochastic approach based on importance sampling in order to get a good evaluation of  $x^*$ :

$$x^* = \arg \max_{x \in \mathcal{X}} (S(x)). \quad (10)$$

Starting from a prior pdf to draw samples, the method iteratively computes series of pdf that increase the probability

to draw a sample near the global optimal solution. With respect to the previous problem the main difference is that the event that is used to iteratively compute the pdf is not given by the problem but has to be chosen. Generally, this is done by choosing a given number  $k$  and considering that the relevant event is that the sample belongs to the  $k$  better samples according to the objective function  $S(x)$ . At each step the new pdf is computed according to (11), where Elite is the set of the  $k$  best samples that can be analytically computed in some cases. For example, when  $\mathcal{X} \subset \mathbb{R}$  and  $n$ -dimensional exponential distributions with independent components specified by (8) are chosen the parameters of the solution of (11) are given by a formula close to (9)

$$\arg \max_{g \in \mathcal{G}} \sum_{j=1}^M I(X_j \in \text{Elite}) \ln g(X_j) W(X_j). \quad (11)$$

The rare event simulation is used to maximize an objective function. However, the use of this method in MO optimization amounts to resolve (11). Hence, this approach can be used in minimization via the modification of the belonging to the Elite's set.

**2.2. Multiobjective Optimization.** In MO problems a set of real value objective functions  $(S_i(x), i = 1 \dots K)$  are considered. The aim is to find a set of better trade-offs between the various objectives called *Pareto optimal solutions*  $\Omega$  [7], when the meaning of better is considered according to the nondominated criterion [23]. The dominance operator ( $\prec$ ) that is used in this criterion is defined by (12), that is, a solution  $x$  is said to dominate the other solution  $y$  if  $x$  is better than  $y$  for at least one objective  $S_i(\cdot)$  and is not worse for any other  $S_j(\cdot)$ . The *Pareto optimal solutions* are then the elements that are not dominated by any other element:

$$x \prec y \iff \forall j \quad S_j(x) \leq S_j(y), \quad \exists i \text{ s.a. } S_i(x) < S_i(y). \quad (12)$$

The Pareto front is then defined as the subset of  $\mathbb{R}^K$  image of the *Pareto optimal solutions* by the objective functions (13). The aim of this part is to adapt the SO optimization procedure to find points of this Pareto front for a given MO problem:

$$\mathcal{P} = \{(S_1(x), \dots, S_K(x)) \mid x \in \Omega\}. \quad (13)$$

In order to apply the CE approach to MO optimization the same approach as for single objective is used. At each step a random sample is generated according to a pdf, and a new pdf is computed as the solution of (11) with respect to a subset Elite.

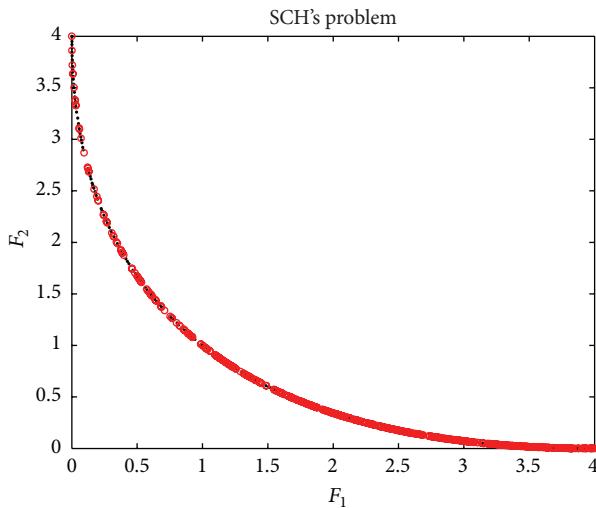
The nondominated criterion is used to build the Elite set. From the initial sample  $X^{(1)}$  it is possible to compute the *Pareto optimal solutions*  $\Omega_1$  and then a new set  $X^{(2)}$  by removing the element of  $\Omega_1$  from  $X^{(1)}$  ( $X^{(2)} = X^{(1)} - \Omega_1$ ). The concept of *Pareto optimal solutions* within  $X^{(2)}$  leads to a new set  $\Omega_2$ . By iteration, it is possible to compute sets  $\Omega_j$  such that each set  $\Omega_{j-1}$  dominates the set  $\Omega_j$ . According to the number of elements that are to select in the Elite set

**Input:** A number of performances functions  $M$ , a performances functions  $S_1, \dots, S_M: \mathcal{U} \rightarrow \mathbb{R}$  where  $\mathcal{U}$  is the set of the feasible solutions, the starting CE's parameter  $v_0$  and two parameters:  $C$  and  $\varrho$ .  
**Output:** An element  $v_{\text{output}} \in \mathbb{R}^+$  and the Pareto front  $P_{\text{nbElite}}$ .

**Algorithm:**

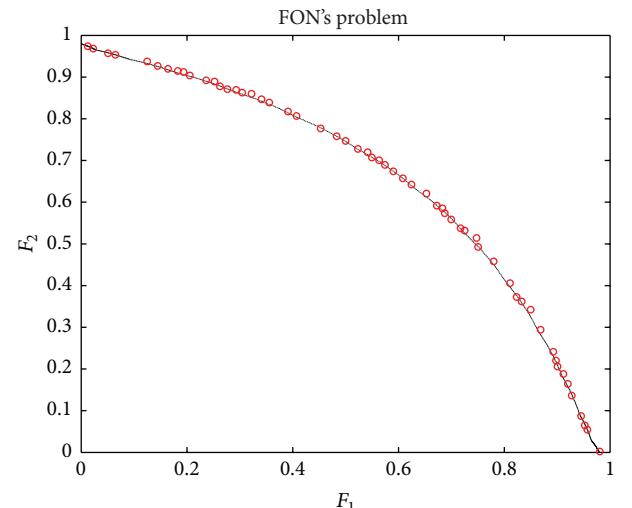
- Step 1. Set  $t$  equal to 1 and the components  $v_t[i]$  of the  $n$ -dimensional vector  $v_t$  equal to  $v_0$ .  
 Set nbElite equal to the largest integer inferior or equal to  $\varrho \times C \times n$ .
- Step 2. Set  $X_t$  equal to an empty set.
- Step 3. Draw independently  $C \times n$  elements according to the Exponential pmf  $\text{Exp}_n(\cdot, v_t)$  and set them in  $X_t$ .
- Step 4. Search all possible non-dominated solutions of  $X_t$  ( $\Omega = \Omega_1, \Omega_2, \dots$ ).
- Step 5. Set  $i$  to 1 and the elite set of sample  $P_{\text{nbElite}}$  to the empty set.
- Step 6. Computation of  $P_{\text{nbElite}}$  Set based on non dominance and crowding operator.
- Step 6 a. Until  $|P_{\text{nbElite}}| + |F_i| \leq \text{nbElite}$ .
- Step 6 b. Include  $i$ th non-dominated front in the elite set  $P_{\text{nbElite}}$  ( $P_{\text{nbElite}} = P_{\text{nbElite}} \cup F_i$ ).
- Step 6 c. Check the next front for inclusion ( $i \leftarrow i + 1$ ).
- Step 7. Calculate crowding distance in  $F_i$  and order the elements according to the  $\prec_n$  operator.
- Step 8. Choose the first ( $\text{nbElite} - |P_{\text{nbElite}}|$ ) elements of  $F_i$  and added its to the elite set  
 $(P_{\text{nbElite}} = P_{\text{nbElite}} \cup F_i(1: (\text{nbElite} - |P_{\text{nbElite}}|)))$ .
- Step 9. If stopping conditions are met, then return  $v_{\text{output}} = v_t$  and stop. Otherwise go to Step 10.
- Step 10. Set  $v_t[i] = \frac{\sum_{k=1}^{C \times n} I_{\{(S_1(X_k), \dots, S_M(X_k)) \in P_{\text{nbElite}}\}} X_k}{\sum_{k=1}^{C \times n} I_{\{(S_1(X_k), \dots, S_M(X_k)) \in P_{\text{nbElite}}\}} W(X_k; u, v_{t-1}[i])}$  for  $i = 1, 2, \dots, n$  and  $t \leftarrow t + 1$ . Go to Step 2.

ALGORITHM 1: A cross-entropy approach for the computation of the nondominated solutions.



- True Pareto front
- Front obtained by CE

FIGURE 1: Optimal front for SCH's study.



- True Pareto front
- Front obtained by CE

FIGURE 2: Optimal front for FON's study.

the first sets  $\Omega_i$  will be included. If the addition of the last set cannot be done (exceeding the required number of samples), The nondomination operator can not be used any more. The selection is then based on the crowding operator  $\prec_n$  [9].

### 3. CE for Multiobjective Optimization Algorithm

The main algorithm is very simple; it is based on rare event simulation with an original event as shown in Algorithm 1.

The selection of the Elite solutions is based on two steps. At the first one, The nondominated sorting is applied to

generate all possible fronts the best fronts are totally added to the Elite set  $P_{\text{nbElite}}$ . At the second step, the remaining elements to reach the total size of  $P_{\text{nbElite}}$  are chosen from the next front based on crowding operator. More detail about these operators (crowding distance, nondominated sorting) can be founded in [9].

The complexity of this method is the overall of the below algorithm complexity, which is  $O(Mn^2)$ .

- (1) The nondominated sorting (with the crowding distance and the sorting on  $\prec_n$ )  $O(M(\varrho \times C \times n)^2)$ .
- (2) The cross entropy (empirical result)  $O(n \log n)$  [24].

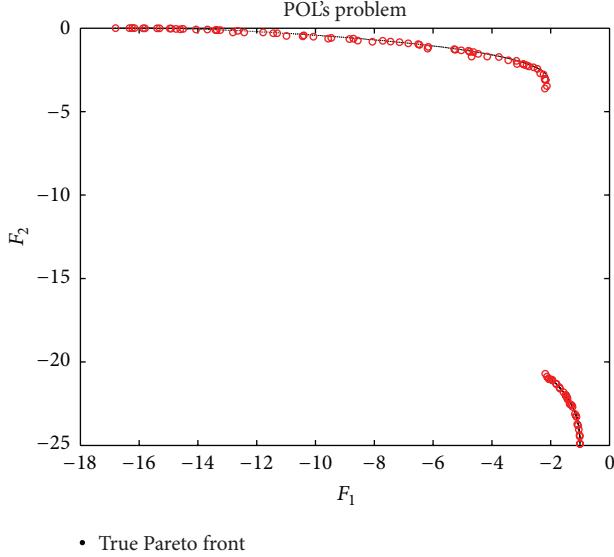


FIGURE 3: Optimal front for POL's study.

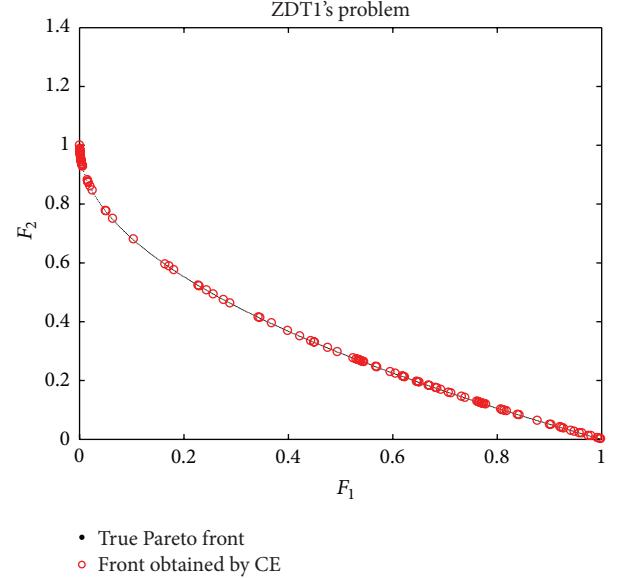


FIGURE 5: Optimal front for ZDT1's study.

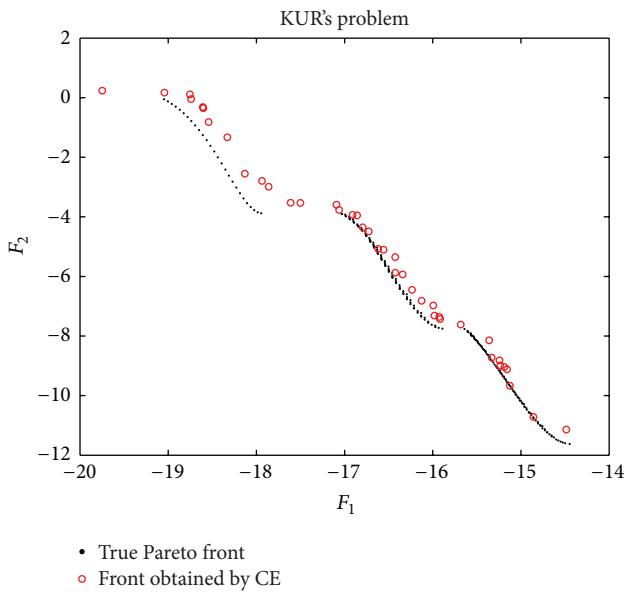


FIGURE 4: Optimal front for KUR's study.

## 4. Simulation Results

In this section, we first describe the cases studies used to compare the CE approach to compute the Pareto front with the method based on CE proposed by [21] and the other meta-heuristic methods such as NSGA-II, the Pareto Archived Evolution Strategy (PAES), and Strength Pareto Evolutionary Algorithm (SPEA). The CE with RES approach was implemented in two steps. First, the goal is to tune the CE's parameters; these parameters are obtained by trial and error. The obtained best parameters are  $C \times n = 1000$  and  $\varrho = 0.01$ . In the second step, we use  $C \times n = 10000$  and  $\varrho = 0.01$ , where the goal is the building of a significant nondominated solutions curves. All runs are repeated 10 times and the

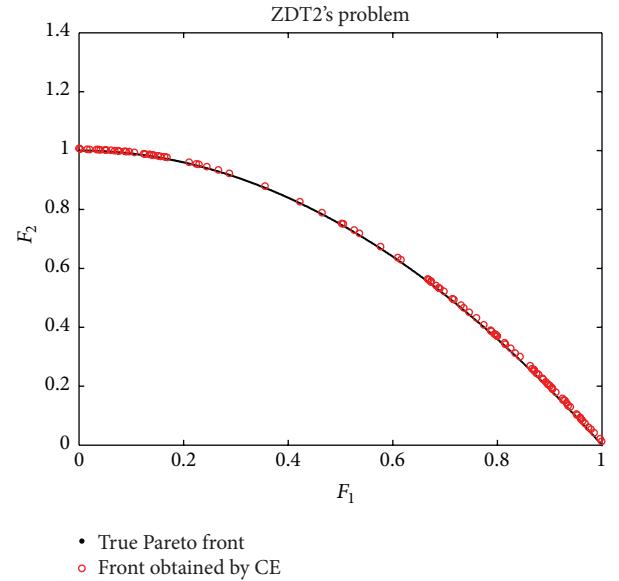


FIGURE 6: Optimal front for ZDT2's study.

starting points are changed. The outcomes correspond to the best solutions. For the problems which have constraints or boundaries, the acceptance-rejection method is used.

**4.1. Cases Studies.** In this section the examples are described. We propose to review the ten tests problems (nine without constraints): the Schaffer's study (SCH), the Fonseca and Fleming's study (FON), the Poloni's study (POL), Kursawe's study (KUR), Zitzler's studies (ZDT1, ZDT2, ZDT3, ZDT4, and ZDT6), and one constrained study (CONSTR or DEB) as in [9].

**4.2. Performance Metrics.** Since our approach is a Pareto-based approach, which will obtain a set of nondominated

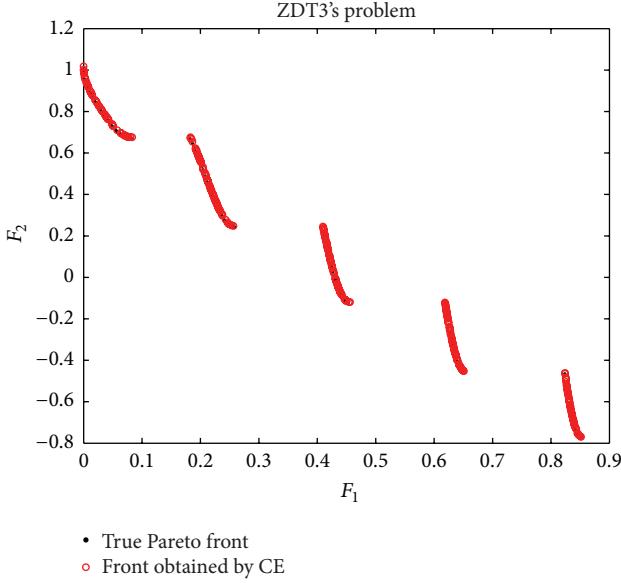


FIGURE 7: Optimal front for ZDT3's study.

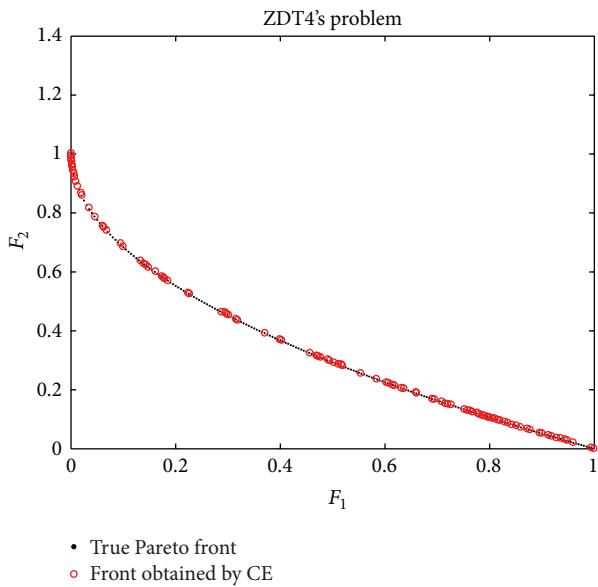


FIGURE 8: Optimal front for ZDT4's study.

solutions  $\Omega$ . Here, we introduce several metrics based on the obtained nondominated solutions to measure the searching quality of the algorithm, which can be used for comparison between different algorithms [25].

**4.2.1. Overall Nondominated Vector Generation (ONVG).** For an obtained nondominated solution set  $\Omega$ , the metric ONVG is defined as  $|\Omega|$  which is the number of distinct nondominated solutions.

**4.2.2. C Metric (CM).** A CM is used to compare two nondominated solution sets  $\Omega$  and  $\Omega'$  obtained by two algorithms, which maps the ordered pair  $(\Omega, \Omega')$  to the interval  $[0, 1]$  to

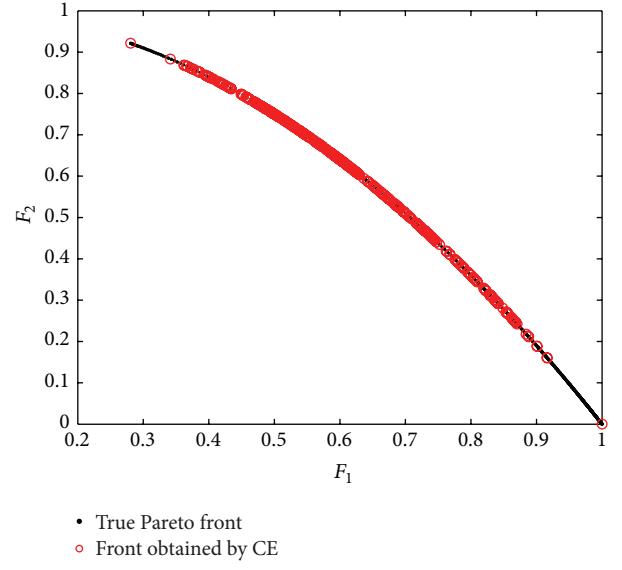


FIGURE 9: Optimal front for ZDT6's study.

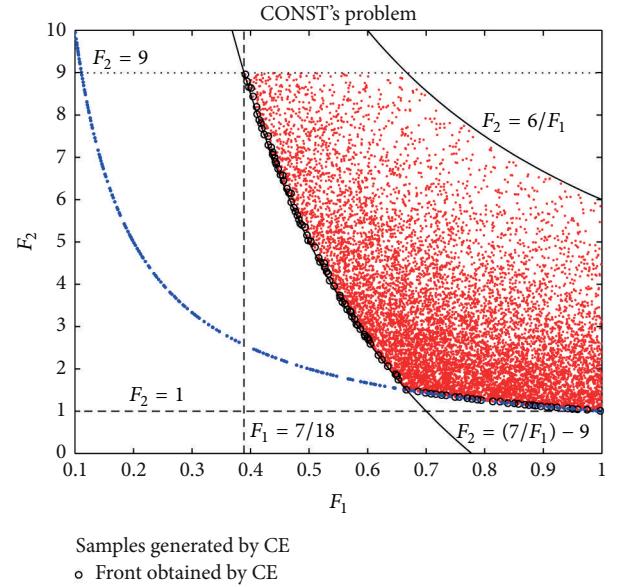


FIGURE 10: Optimal front for CONSTR's study.

reflect the dominance relationship between solutions in the two sets

$$C(\Omega, \Omega') = \frac{|\{x' \in \Omega' \mid \exists : x \in \Omega, x' \prec x\}|}{|\Omega'|}. \quad (14)$$

If all the solutions in  $\Omega'$  are dominated by those in  $\Omega$ , then  $C(\Omega, \Omega') = 1$ . Conversely, if all the solutions in  $\Omega$  are dominated by those in  $\Omega'$ , then  $C(\Omega, \Omega') = 0$ . It should be noticed that the sum of  $C(\Omega, \Omega')$  and  $C(\Omega', \Omega)$  is not always equal to 1, because there may be some solutions in  $\Omega$  and  $\Omega'$  that are not dominated to each other.

**4.2.3. Distance Metrics ( $D_{av}$  and  $D_{max}$ ).** Two distance metrics were used to measure the performance of nondominated

TABLE 1: Other metric comparison based.

Algorithm	Method	SCH	FON	POL	KUR	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
ONVG	CE method	1	1	1	1	1	1	1	1	1
	NAGA-II	1	1	1	1	1	1	1	1	1
CM	CE method	1.000	0.915	0.96	1	0.37	1	0.81	0.01	0.9200
	NAGA-II	0.996	0.89	0.57	0.2	1	0.75	0.975	1	1
$D_{av}$	CE method	0.0022	0.0014	-0.0001	0.0154	0.0069	0.0092	0.0020	0.0066	0.0229
	NAGA-II	0.0012	0.0009	0.0009	0.0067	0.0106	0.0053	0.0022	0.3292	0.0025
$D_{max}$	CE method	0.0111	0.0116	0.0079	0.0453	0.0368	0.0316	0.0055	0.0317	0.1045
	NAGA-II	0.0073	0.0043	0.0046	0.0213	0.0221	0.0143	0.0067	0.3736	0.0119
TS	CE method	0.0017	0.0030	0.0034	0.1908	0.0126	0.0020	0.0032	0.0090	0.3012
	NAGA-II	0.0005	0.0003	0.0019	0.0582	0.0022	0.0042	0.0045	0.0022	0.0018
MS	CE method	1.0005	0.9856	0.9976	0.9631	0.9982	0.9967	1.0038	1.0042	1.0000
	NAGA-II	1.0000	1.0028	0.9990	1.0000	1.0055	0.9976	0.9996	1.3387	1.6902
AQ	CE method	0.5235	0.3346	-1.2381	-5.8113	0.2061	0.3272	0.1317	0.2047	0.085469
	NAGA-II	0.5271	0.3294	-1.2440	-6.0626	0.2039	0.3256	0.1318	0.5111	0.006664
$t$	CE method	13.01548	13.776714	24.715435	13.708412	28.320879	41.6873	16.4662	13.344223	11.692149
	NAGA-II	67.06799	70.4706	38.285153	35.7219	98.879344	101.7208	101.95494	99.9021	108.873080

TABLE 2: Mean (first row) and variance (second row) of the convergence metric  $\Upsilon$ .

Algorithm	SCH	FON	POL	KUR	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
CE method with RES	0.000140	0.002704	0.057898	0.162699	0.001282	0.005001	0.003321	0.002771	0.009426
	0	0.000005	0.005125	0.011727	0	0	0.000007	0.000007	0.000001
CE method with FCM		0.0026			0.0434	0.1002	0.0245	1.2056	0.0029
NSGA-II	0.003391	0.001931	0.015553	0.028964	0.033482	0.072391	0.114500	0.513053	0.296564
Real coded	0	0	0.000001	0.000018	0.004750	0.031689	0.007940	0.118460	0.013135
NSGA-II	0.002833	0.002571	0.017029	0.028951	0.000894	0.000824	0.043411	3.227636	7.806798
Binary coded	0.000001	0	0.000003	0.000016	0	0	0.000042	7.307630	0.001667
SPEA	0.003403	0.125692	0.037812	0.045617	0.001799	0.001339	0.047517	7.340299	0.221138
	0	0.000038	0.000088	0.00005	0.000001	0	0.000047	6.572516	0.000449
PEAS	0.001313	0.151263	0.030864	0.082085	0.126276	0.023872	0.023872	0.854816	0.085469
	0.000003	0.000905	0.000431	0.011989	0.008679	0.036877	0.000001	0.527238	0.006664

solution set  $\Omega$  relative to a reference set or a true Pareto front ( $R$ )

$$D_{av} = \frac{\sum_{x_R \in R} \min_{x \in \Omega} d(x, x_R)}{|R|}, \quad (15)$$

$$D_{max} = \max_{x_R \in R} \left\{ \min_{x \in \Omega} d(x, x_R) \right\},$$

where  $d(x, x_R) = \max_{j=1, \dots, n} \{(f_j(x) - f_j(x_R)) / \Delta_j\}$ ,  $x \in \Omega$ ,  $x_R \in R$ . More details can be found in [25].

4.2.4. *Tans Spacing* (TS). The following spacing metric was used to measure how evenly the solutions are distributed:

$$TS = \sqrt{\frac{(1/|\Omega|) \sum_{i=1}^{|\Omega|} (D_i - \bar{D})^2}{\bar{D}}}, \quad (16)$$

where  $\bar{D} = \sum_{i=1}^{|\Omega|} D_i / |\Omega|$ .

4.2.5. *Maximum Spread* (MS). This metric is developed to measure how well the true Pareto front is covered by the obtained nondominated solutions in the set  $\Omega$  through the hyperboxes formed by the extreme function values observed in the optimal Pareto front and  $\Omega$ :

$$MS = \sqrt{\frac{1}{n} \sum_{j=1}^n \left( \frac{\max_{i=1}^{|\Omega|} f_j(x_i) - \min_{i=1}^{|\Omega|} f_j(x_i)}{F_j^{\max} - F_j^{\min}} \right)}, \quad (17)$$

where  $F_j^{\max}$  and  $F_j^{\min}$  are the maximum and minimum values, respectively, of the  $j$ th objective in the true Pareto front.

4.2.6. *Average Quality* (AQ). This metric was proposed to measure the quality of the solution set, which was originally expressed in the form of weighted Tchebycheff function  $s_{\infty}$ . But that function may hide certain aspects about the quality of solution set because poor performance with respect to proximity could be compensated by good performance in the distribution of solutions. Therefore, diversity indicators of spread and space are added to the formulation to overcome

TABLE 3: Mean (first row) and variance (second row) of the diversity metric  $\Delta$ .

Algorithm	SCH	FON	POL	KUR	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
CE method with RES	1.130705	0.427151	0.984515	0.630363	0.867578	0.754868	0.733369	0.736887	0.767926
CE method with FCM		0.2504			0.0921	0.1777	0.1686	0.5697	1.0459
NSGA-II	0.477899	0.378065	0.452150	0.411477	0.390307	0.430776	0.738540	0.702612	0.668025
Real coded	0.003471	0.000639	0.002868	0.000992	0.001876	0.004721	0.019706	0.064648	0.009923
NSGA-II	0.449265	0.395131	0.503721	0.442195	0.463292	0.435112	0.575606	0.479475	0.644477
Binary coded	0.002062	0.001314	0.004656	0.001498	0.041622	0.024607	0.005078	0.009841	0.035042
SPEA	1.021110	0.792352	0.972783	0.852990	0.784525	0.755148	0.672938	0.798463	0.849389
	0.004372	0.005546	0.008475	0.002619	0.004440	0.004521	0.003587	0.014616	0.002713
PEAS	1.063288	1.162528	1.020007	1.079838	1.229794	1.165942	0.789920	0.870458	1.153052
	0.002868	0.008945	0	0.013772	0.004839	0.007682	0.001653	0.101399	0.003916

the limitation, and a metric is given by the average value of a scalarizing function over a representative sample of weight vectors as follows:

$$AQ = \frac{\sum_{\lambda \in \Lambda} s_a(f, z^0, \lambda, \rho)}{|\Lambda|}, \quad (18)$$

where

$$\begin{aligned} s_a(f, z^0, \lambda, \rho) &= \min_i \left\{ \max_j \{ \lambda_j (f_j(x_i) - z_j^0) \} \right. \\ &\quad \left. + \rho \sum_{j=1}^n \lambda_j (f_j(x_i) - z_j^0) \right\}, \quad (19) \\ \Lambda &= \left\{ \Lambda \in (\lambda_1, \dots, \lambda_n) \mid \lambda_j \in \left\{ 0, \frac{1}{r}, \frac{2}{r}, \dots, 1 \right\}, \right. \\ &\quad \left. \sum_{j=1}^n \lambda_j = 1 \right\}. \end{aligned}$$

**4.2.7. Running Time ( $t$ ).** To reflect the efficiency of a multiobjective optimization algorithm, running time is also used as a metric.

**4.3. Computation Results.** To evaluate the performance of the results, we use the convergence metric and the diversity metric parameters as described in [9].

Table 2 shows the convergence metric  $\Upsilon$  obtained using six algorithms, CE with RES, CE with FCM, NSGA-II, SPEA, and PAES. The data for CE with FCM is provided from [21] and the data of the four meta-heuristic algorithm are provided from [9]. CE with RES is able to converge better in the problems SCH, ZDT1, ZDT4, and ZDT6.

Table 3 shows the diversity metric  $\Delta$  using the five algorithms. It is deduced that the diversity of the proposed method is small compared to the diversity of other methods because it uses an exponential pdf. The better diversity is

obtained by the CE based on FCM which is assured using clustering. We project in the future works the use of other distribution function as normal distribution.

In Table 1, a comparison of our method with the NSGA-II method is shown. To properly evaluate both algorithms and highlight the contribution of our approach, the parameters of the previous subsection are calculated (we choose the algorithm NSGA-II for comparison, given its obvious success). We conclude that both approaches have almost identical performance. However, in computation time, our method shows its efficiency (calculated with Inter processor Core i3 CPU M370 @ 2.40 GHz). The comparison of the proposed method with the true Pareto front is yield in Figures 1–10. Our method gives solutions which are closes to the true Pareto fronts.

## 5. Conclusion

In this paper a simple original algorithm for solving a multiobjective optimization based on the cross-entropy method (particularly the rare event simulation) is developed and tested through several cases. This method is fast and very simple to implement in economic and engineering fields.

## References

- [1] W. Stadler, *Multicriteria Optimization in Engineering and in the Sciences*, Plenum Press, New York, NY, USA, 1998.
- [2] M. T. Tabucanon, *Multiple Criteria Decision Making Industry*, Elsevier Science, New York, NY, USA, 1988.
- [3] C. A. Coello Coello, D. A. van Veldhuizen, and G. B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems*, Kluwer Academic Publishers, New York, NY, USA, 2002.
- [4] T. P. Bagchi, *Multiobjective Scheduling by Genetic Algorithms*, Kluwer Academic Publisher, New York, NY, USA, 1999.
- [5] A. Berizzi, C. Bovo, M. Innorta, and P. Marannino, “Multiobjective optimization techniques applied to modern power systems,” in *Proceedings of the IEEE Power Engineering Society Winter Meeting*, vol. 3, pp. 1503–1508, Columbus, Ohio, USA, February 2001.
- [6] J. Horn, N. Nafpliotis, and D. E. Goldberg, “A Niched Pareto genetic algorithm for multiobjective optimization,” in *Proceedings of the 1st IEEE Conference on Evolutionary Computation*,

- IEEE World Congress on Computational Intelligence*, vol. 1, pp. 82–87, Orlando, Fla, USA, June 1994.
- [7] N. Srinivas and K. Deb, “Multiobjective function optimization using nondominated sorting genetic algorithms,” *Evolutionary Computation*, vol. 2, no. 3, pp. 221–248, 1994.
  - [8] E. Zitzler and L. Thiele, “An evolutionary algorithm for multiobjective optimization: the strength pareto approach,” Tech. Rep. 43, TIK, Swiss Federal Institute of Technology (ETH), 1998.
  - [9] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.
  - [10] J. Kennedy and R. C. Eberhart, *Swarm Intelligence*, Morgan Kauffman, Burlington, Mass, USA, 2001.
  - [11] M.-P. Song and G.-C. Gu, “Research on particle swarm optimization: a review,” in *Proceedings of the International Conference on Machine Learning and Cybernetics*, vol. 4, pp. 2236–2241, August 2004.
  - [12] C. A. Coello Coello, G. T. Pulido, and M. S. Lechuga, “Handling multiple objectives with particle swarm optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 256–279, 2004.
  - [13] I. Das and J. E. Dennis, “Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multi-criteria optimization problems,” *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 631–657, 1998.
  - [14] R. Y. Rubinstein, “Optimization of computer simulation models with rare events,” *European Journal of Operational Research*, vol. 99, no. 1, pp. 89–112, 1997.
  - [15] R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning*, Information Science and Statistics, Springer, New York, NY, USA, 2004.
  - [16] D. Ernst, M. Glavic, G.-B. Stan, S. Mannor, and L. Wehenkel, “The cross-entropy method for power system combinatorial optimization problems,” in *Proceedings of the 7th IEEE Power Engineering Society (PowerTech '07)*, pp. 13–17, Lausanne, Switzerland, July 2007.
  - [17] P. T. de Boer, *Analysis and efficient simulation of queueing models of telecommunication systems [Ph.D. thesis]*, University of Twente, Enschede, The Netherlands, 2000.
  - [18] B. E. Helvik and O. Wittner, “Using the cross-entropy method to guide/govern mobile agent’s path finding in networks,” in *Proceedings of the 3rd International Workshop on Mobile Agents for Telecommunication Applications (MATA '01)*, pp. 255–268, 2001.
  - [19] A. V. Babayan, D. A. Savic, and G. A. Walters, “Multiobjective optimization for the least-cost design of water distribution systems under correlated uncertain parameters,” in *Proceedings of the World Water and Environmental Resources Congress*, pp. 1–11, Anchorage, Alaska, USA, May 2005.
  - [20] M. Carney and P. Cunningham, “Calibrating probability density forecasts with multi-objective search,” TCD-CS-2006-07, Department of Computer Science, The University of Dublin, Dublin, Ireland, 2006.
  - [21] A. Ünveren and A. Acan, “Multi-objective optimization with cross entropy method: stochastic learning with clustered pareto fronts,” in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '07)*, pp. 3065–3071, Singapore, September 2007.
  - [22] J. Bekker and C. Aldrich, “The cross-entropy method in multi-objective optimisation: an assessment,” *European Journal of Operational Research*, vol. 211, no. 1, pp. 112–121, 2011.
  - [23] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Boston, Mass, USA, 1989.
  - [24] J. Wu and A. C. S. Chung, “Cross entropy: a new solver for Markov random field modeling and applications to medical image segmentation,” in *Proceedings of the on Medical Image Computing and Computer-Assisted Intervention (MICCAI '05)*, vol. 3749 of *Lecture Notes in Computer Science*, pp. 229–237, October 2005.
  - [25] B.-B. Li and L. Wang, “A hybrid quantum-inspired genetic algorithm for multiobjective flow shop scheduling,” *IEEE Transactions on Systems, Man, and Cybernetics, Cybernetics B*, vol. 37, no. 3, pp. 576–591, 2007.

