A Parameterless-Niching-Assisted Bi-objective Approach to Multimodal Optimization

Sunith Bandaru and Kalyanmoy Deb
Kanpur Genetic Algorithms Laboratory
Indian Institute of Technology Kanpur
Kanpur 208016, India
Email: {sunithb,deb}@iitk.ac.in

KanGAL Report No. 2013009

***Abstract*—Evolutionary algorithms are becoming increasingly popular for multimodal and multi-objective optimization. Their population based nature allows them to be modified in a way so as to locate and preserve multiple optimal solutions (referred to as Pareto-optimal solutions in multi-objective optimization). These modifications are called niching methods, particularly in the context of multimodal optimization. In evolutionary multi-objective optimization, the concept of dominance and diversity preservation inherently causes niching. This paper proposes an approach to multimodal optimization which combines this power of dominance with traditional variable-space niching. The approach is implemented within the NSGA-II framework and its performance is studied on 20 benchmark problems. The simplicity of the approach and the absence of any special niching parameters are the hallmarks of this study.**

I. Introduction

Multimodality in single-objective optimization refers to the existence of more than one optimal solution. While a single solution is usually sought for implementation, the knowledge of multiple optima can be beneficial in two ways. Firstly, it gives the user an option to switch from one optimal solution to another. This may be required when user preferences cannot be incorporated in the search process or when system dynamics cause a favorable solution to become unfavorable. Secondly, data-mining of multiple optima can reveal properties which make them optimal. This is similar in principle to the concept of *innovization* [1] defined for multi-objective problems. Multimodal optimization deals with the location of multiple optima. Optima can be local or global depending on whether they are optimal only in their immediate neighborhood or in the entire search space. The distinction is important because local optima can have different function values whereas all global optima have the same value for the objective function.

Evolutionary algorithms (EAs) use a population-based approach to optimization. Due to the nature of the selection scheme, they typically converge to a single solution, most likely to a global optimum. Early attempts at multimodal optimization involved sequential location of multiple optima where a traditional EA is run several times with the hope of finding a different optima each time. Such methods are called temporal niching methods [2], [3]. On the other hand spacial niching methods involve modifications to the basic framework of an EA so as to allow the maintenance of sub-populations which ultimately leads to simultaneous location of multiple optima. Popular spacial niching methods are fitness sharing [4], [5], crowding [6], [7], clearing [8], clustering [9] and restricted mating [10]. A survey of niching methods in evolutionary algorithms can be found in [11], [12], [13].

Evolutionary algorithms are especially popular in the field of multi-objective optimization. As in the case of multimodal optimization, multi-objective optimization also requires obtaining multiple optimal solutions, largely referred to as Pareto-optimal solutions. Consider a multi-objective optimization problem of the form

$$\begin{array}{ll} \text{Minimize} & \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_M(\mathbf{x})\} \\ \text{Subject to} & \mathbf{x} \in S \end{array} \qquad (1)$$

where $f_i : \mathbb{R}^D \rightarrow \mathbb{R}$ are $M(> 2)$ conflicting objectives that have to be simultaneously minimized and the variable (or decision) vector $\mathbf{x} = \{x_1, x_2, \ldots, x_D)\}$ belongs to the non-empty feasible region $S \subset \mathbb{R}^D$. The feasible region is formed by the constraint functions and the variable bounds. A variable vector $\mathbf{x_1}$ is said to dominate $\mathbf{x_2}$ and denoted as $\mathbf{x_1} \succ \mathbf{x_2}$ if

$$(f_m(\mathbf{x_1}) \leq f_m(\mathbf{x_2}) \ \forall m) \wedge (\exists n | \ f_n(\mathbf{x_1}) < f_n(\mathbf{x_2})). \qquad (2)$$

In the absence of such an n, $\mathbf{x_1}$ and $\mathbf{x_2}$ are said to be non-dominated with respect to each other.

The consideration of multiple objectives combined with the concept of dominance leads to inherent niching and preserves multiple solutions in the population that are said to be non-dominated. Additionally, diversity preservation is usually employed to obtain a good spread of solutions. The latter also serves as a secondary measure for selecting individuals.

In this paper, we deal with single objective multimodal optimization problems and our goal is to locate all the global optima. We convert the original problem into a bi-objective problem, a process known as *multiobjectivization* [14], by taking the original objective as the first objective and a suitable second objective. Thereafter, the bi-objective problem is solved using NSGA-II [15] with a modified dominance criterion. An adaptive constraint is also added to the bi-objective problem in order to avoid local optima. Without this constraint, the algorithm can also locate and preserve the local optima. Minimal changes are required in the original NSGA-II code [16]. Moreover, no parameters are introduced in the process, which makes this approach very easy to use.

The paper is organized as follow: In Section II, we briefly discuss the original NSGA-II algorithm. Section III discusses the modifications required for the proposed approach. The test problems used in this work are summarized in Section IV. The performance measures and experimental setup are described in Section V and Section VI respectively. Finally, the results are presented in Section VII.

II. Original NSGA-II

Algorithm 1 shows the basic pseudocode of NSGA-II. It takes the initial random population P_0, population size POP and the number of generations $NGEN$ as input parameters along with standard genetic algorithm (GA) parameters like crossover and mutation probabilities (p_c and p_m, respectively). After $NGEN$ generations, it generates the final population Q_{NGEN} containing the trade-off (non-dominated) objective and variable values. The main strength of NSGA-II comes from its fast non-dominated sorting procedure (step 4 in Algorithm 1) and the use of a diversity preserving operator ($\prec_c$, also called the crowded-comparison operator) for filling the child population Q_t (steps 6 to 12 in Algorithm 1). The standard sub-routines for performing non-dominated sorting of the population members `nondominated-sort()`, and calculating their crowding distance measures `crowding-distance()`, can be found in [15]. The `select-cross-mutate()` routine performs tournament selection, crossover and mutation of the population members. The original NSGA-II implementation uses Simulated Binary Crossover (SBX) [17] and polynomial mutation [11] for real-variables which require the specification of distribution indices η_c and η_m respectively.

Algorithm 1 Elitist Non-Dominated Sorting Genetic Algorithm (NSGA-II).

Require: P_0, POP, $NGEN$, p_c, p_m, η_c, η_m
Ensure: Q_{NGEN}
1: **set** $t \leftarrow 0$, $Q_0 =$ `select-cross-mutate(`P_0`)`
2: **while** $t < NGEN$ **do**
3: $\quad R_t = P_t \cup Q_t$
4: $\quad \mathcal{F} =$ `nondominated-sort(`R_t`)`
5: $\quad P_{t+1} = \Phi$ and $i \leftarrow 1$
6: $\quad$ **while** $|P_{t+1}| + |\mathcal{F}_i| \leq POP$ **do**
7: $\quad\quad$ `crowding-distance(`$\mathcal{F}_i$`)`
8: $\quad\quad P_{t+1} = P_{t+1} \cup \mathcal{F}_i$
9: $\quad\quad i \leftarrow i + 1$
10: $\quad$ **end while**
11: $\quad$ Sort($\mathcal{F}_i, \prec_c$)
12: $\quad P_{t+1} = P_{t+1} \cup \mathcal{F}_i[1 : (POP - |P_{t+1}|)]$
13: $\quad Q_{t+1} =$ `select-cross-mutate(`P_{t+1}`)`
14: $\quad t \leftarrow t + 1$
15: **end while**

The dominance definition in Eq. 2 is valid when both variable vectors belong to the feasible space S. However, in the presence of constraints, random populations may contain infeasible individuals (variable vectors). Therefore, NSGA-II's non-dominated sorting procedure uses the following criterion for establishing dominance between two individuals:

1) If one is feasible and the other is not, then the feasible individual dominates the other.
2) If both are feasible, then use Eq. 2.
3) If both are infeasible, then the individual with lower constraint violation dominates the other.

NSGA-II uses crowding distance as a measure for the diversity preservation operator $\prec_c$ in step 11. An individual with a higher value for crowding distance is more isolated in the objective space and therefore will be preferred over another individual with the same rank but with a lower crowding distance value. This establishes a partial order in selecting individuals for forming the mating pool. The NSGA-II algorithm is, by nature, elitist and hence the best individuals (in terms of rank and crowding distance) from any generation will be maintained till the end of all generations.

III. Proposed Approach

Consider the multimodal function optimization of $F(\mathbf{x})$ in the form,

$$\begin{array}{ll} \text{Minimize} & F(\mathbf{x}) \\ \text{Subject to} & x_d^{(L)} \leq x_d \leq x_d^{(U)} \ \forall\, d = \{1, 2, \ldots, D\}. \end{array} \tag{3}$$

In the following sections, we discuss the methodology for obtaining multiple global optima of $F(\mathbf{x})$.

A. Multiobjectivization

The first step involves multiobjectivization of the original problem with two objectives. This approach to multimodal optimization was proposed recently [18], although here we use a different secondary objective. The primary objective remains the minimization of the original function, i.e.

$$f_1(\mathbf{x}) = F(\mathbf{x}). \tag{4}$$

Locating all global optima requires that the population have as much diversity as possible. This can be achieved by considering a secondary objective which maximizes diversity among the individuals. Specifically, for the i-th population member we consider the function,

$$f_2(\mathbf{x_i}) = \frac{1}{\sum_{j=1}^{POP} \|\mathbf{x_i} - \mathbf{x_j}\|^2}. \tag{5}$$

The minimization of f_2 (or maximization of the quantity in the denominator) causes the population members to spread out in the search space, thus enabling exploration.

The resulting bi-objective problem can be solved using NSGA-II. However, the two objectives considered above need not be conflicting and hence the population may still eventually converge to a single optimum. In order prevent this, we assist NSGA-II with variable space niching in the following section.

B. Parameterless-Niching-Assisted NSGA-II

As discussed in SectionII, NSGA-II comes with a diversity preserving operator that uses crowding distances calculated in the objective space. However, when the objectives are non-conflicting, such objective space niching will be ineffective. We therefore introduce a scheme for niching in the decision space. Most niching methods introduce special parameters whose values can drastically affect the performance of the algorithm. The proposed parameterless-niching approach uses available parameters to tune itself.

Consider performing this parameterless-niching on a population size of POP in a D-dimensional decision space. Theoretically, the number of global optima that can be obtained using this population is POP. In the ideal case, the decision space can be divided into POP equal sub-regions and each population member can locate an optima in each sub-region. For obtaining equal sub-regions, we divide each of the D dimensions into T divisions. An appropriate value for T can be obtained as follows:

$$T^D = POP \Rightarrow T \approx \lfloor e^{\frac{\ln POP}{D}} \rfloor. \tag{6}$$

This ensures that each sub-region can theoretically be represented by at least one population member thus covering the extreme case when the multimodal function has POP global optima, one in each sub-region.

A niching distance can now be defined for each dimension as:

$$\nu_d = \frac{x_d^{(U)} - x_d^{(L)}}{T} \; \forall \; d = \{1, 2, \ldots, D\}. \tag{7}$$

Most niching methods use Euclidean distance metric (L_2 norm) as a measure of proximity. It has been shown elsewhere [19] that the Manhattan distance metric (L_1 norm) is consistently more preferable than higher norms in high-dimensional datasets. The niching distance defined above enables us to use the L_1 norm in each dimension.

Next, we define the proximity of two population members $\mathbf{x_1}$ and $\mathbf{x_2}$ based on the niching distances. $\mathbf{x_1}$ and $\mathbf{x_2}$ are said to be proximate and denoted by proximate($\mathbf{x_1}$,$\mathbf{x_2}$) $= 1$ if

$$|(\mathbf{x_1} - \mathbf{x_2})_d| \le \nu_d \; \forall \; d = \{1, 2, \ldots, D\}. \tag{8}$$

Otherwise, $\mathbf{x_1}$ and $\mathbf{x_2}$ are non-proximate, i.e. proximate($\mathbf{x_1}$,$\mathbf{x_2}$) $= 0$. Niching can now be introduced by using this proximity measure to modify the dominance criterion in Eq. 2 as follows:

$\mathbf{x_1} \succ \mathbf{x_2}$ if

1) $(f_m(\mathbf{x_1}) \le f_m(\mathbf{x_2}) \; \forall m) \wedge (\exists n | \; f_n(\mathbf{x_1}) < f_n(\mathbf{x_2}))$,
2) and proximate($\mathbf{x_1}, \mathbf{x_2}$) $= 1$.

Otherwise $\mathbf{x_1}$ and $\mathbf{x_2}$ are non-dominated with respect to each other. Note that we have not restricted the mating of individuals from different sub-regions.

C. Adaptive Constraint

The proposed niching-assisted NSGA-II, in addition to locating global optima, also maintains niches at the local optima. When local optima are undesired, a constraint based on the current-best function value can be added to the problem to make the corresponding variable vectors infeasible. Since the dominance relations between population members are easily influenced by their feasibility (see Section II), we adopt an adaptive constraint that is insignificant at the start of the generations, but becomes increasingly difficult to satisfy as the generations progress.

Let ϵ be the desired level of accuracy for the global optima and let F^{best} be the least value of $F(\mathbf{x})$ among all evaluated solutions till the current generation gen. Then the adaptive constraint at gen is given by,

$$f_1(\mathbf{x}) \le F^{best} + f_{gen} \cdot \epsilon. \tag{9}$$

The factor f_{gen} decays exponentially from a large value (10^{14}) at $gen = 1$ to 2.0 at the final generation. For a fixed number of function evaluations $MaxFEs$, the final generation is $\left(\dfrac{MaxFEs}{POP} - 1\right)$. For $f_{gen} = a \; e^{b \cdot gen}$, a and b can be calculated using,

$$a = 10^{14}/e^b \tag{10}$$

$$b = \left(\ln 2 - \ln 10^{14}\right) \Big/ \left(\frac{MaxFEs}{POP} - 1\right) \tag{11}$$

$$= -31.543 \Big/ \left(\frac{MaxFEs}{POP} - 1\right). \tag{12}$$

The exponentially decaying factor is used to give sufficient number of generations for the population in the later generations to locate and crowd the global optima. After all local optima become infeasible, the population members close to them start crowding near the global optima.

The two objectives and the adaptive constraint can be directly coded in the original NSGA-II implementation. The only modification concerns the domination criterion discussed above and is incorporated in the `nondominated-sort()` routine.

IV. Summary of Benchmark Functions [20]

The benchmark set includes the following 20 multimodal test functions:

- F_1: Five-Uneven-Peak Trap (1D)
- F_2: Equal Maxima (1D)
- F_3: Uneven Decreasing Maxima (1D)
- F_4: Himmelblau (2D)
- F_5: Six-Hump Camel Back (2D)
- F_6: Shubert (2D, 3D)
- F_7: Vincent (2D, 3D)
- F_8: Modified Rastrigin - All Global Optima (2D)
- F_9: Composition Function 1 (2D)
- F_{10}: Composition Function 2 (2D)
- F_{11}: Composition Function 3 (2D, 3D, 5D, 10D)
- F_{12}: Composition Function 4 (3D, 5D, 10D, 20D)

The problem definitions can be found in [20].

These multimodal test functions have following properties:

1) All test functions are formulated as maximization problems. For minimization in the context of the proposed approach, we use $-F(\mathbf{x})$.
2) F_1, F_2 and F_3 are simple 1D multimodal functions.
3) F_4 and F_5 are simple 2D multimodal functions. These functions are not scalable.
4) F_6 to F_8 are scalable multimodal functions. The number of global optima for F_6 and F_7 are determined by the dimension D. However, for F_8, the number of global optima is independent from D, therefore can be controlled by the user.

TABLE I. PARAMETERS USED FOR PERFORMANCE MEASUREMENT.

Function	r	F^*	No. of global optima (NKP)
F_1 (1D)	0.01	200.0	2
F_2 (1D)	0.01	1.0	5
F_3 (1D)	0.01	1.0	1
F_4 (2D)	0.01	200.0	4
F_5 (2D)	0.5	1.03163	2
F_6 (2D)	0.5	186.731	18
F_7 (2D)	0.2	1.0	36
F_6 (3D)	0.5	2709.0935	81
F_7 (3D)	0.2	1.0	216
F_8 (2D)	0.01	-2.0	12
F_9 (2D)	0.01	0.0	6
F_{10} (2D)	0.01	0.0	8
F_{11} (2D)	0.01	0.0	6
F_{11} (3D)	0.01	0.0	6
F_{12} (3D)	0.01	0.0	8
F_{11} (5D)	0.01	0.0	6
F_{12} (5D)	0.01	0.0	8
F_{11} (10D)	0.01	0.0	6
F_{12} (10D)	0.01	0.0	6
F_{12} (20D)	0.01	0.0	8

5) F_9 to F_{12} are scalable multimodal functions constructed by several basic functions with different properties. F_9 and F_{10} are separable, and non-symmetric, while F_{11} and F_{12} are non-separable, non-symmetric complex multimodal functions. The number of global optima in all composition functions is independent from D, and therefore can be controlled by the user.

The following information about the test functions is used for assessing the performance of the proposed algorithm:

1) The number of global optima.
2) The fitness of the global optima (or peak height) F^*, which is known or can be estimated.
3) A niche radius value r that can sufficiently distinguish two closest global optima.

Table I shows the above information for all test problems.

V. PERFORMANCE MEASURES [20]

In order to study the capability of the proposed approach in locating *all* global optima, we first specify a level of accuracy ($0 < \epsilon < 1$), a threshold value under which we would consider that a global optimum is found. The decision vectors $\mathbf{x}$ from the final population Q_{NGEN} of niching-assisted NSGA-II is provided as input to an algorithm which counts the number of global optima in the population. The involved steps are as follows:

Step 1: Evaluate $F(\mathbf{x})$ for all individuals in Q_{NGEN}.
Step 2: Sort Q_{NGEN} in decreasing value of $F(\mathbf{x})$ (since the original function is to be maximised) to form the sorted list of individuals L_{sorted}.
Step 3: SET $k \leftarrow 1$ and $S = \Phi$.
Step 4: IF $k > |L_{sorted}|$, STOP.
Step 5: FOR the k-th individual in L_{sorted}, set $already\_found \leftarrow$ FALSE.
Step 6: IF $F^* - F(\mathbf{x_k}) > \epsilon$, $k \leftarrow k + 1$ and GOTO Step 4.
Step 7: SET $t \leftarrow 1$.
Step 8: FOR the t-th individual in S, IF $\|\mathbf{x_t} - \mathbf{x_k}\| \leq r$, set $already\_found \leftarrow$ TRUE, GOTO Step 10.
Step 9: $t \leftarrow t + 1$ and GOTO Step 8.
Step 10: IF $already\_found =$ FALSE, $S = S \cup \{\mathbf{x_k}\}$.
Step 11: $k \leftarrow k + 1$ and GOTO Step 4.

The above steps result in a solution list S containing all the distinct global optima found by the proposed approach. The niche radius value (r) for each test function in Table I is set to a value less than the distance between two closest global optima. This ensures that the individuals on two found global optima would be treated as two different solutions in Step 8. The number of global optima found is given by $NPF = |S|$.

The performance of the proposed approach is measured in terms of the *peak ratio*, *success rate*, and *averaged number of evaluations* required to achieve a given accuracy ϵ for locating all global optima over multiple runs [21]. These measures are calculated over multiple runs of the proposed approach.

A. Peak Ratio (PR)

For a given number of maximum number of function evaluations ($MaxFEs$) and a required accuracy level ϵ, PR measures the average percentage of all known global optima found over NR multiple runs and is given by,

$$PR = \frac{\sum_{run=1}^{NR} NPF_i}{NKP \times NR}, \tag{13}$$

where NPF_i denotes the number of global optima found at the end of the i-th run, NKP is the number of known global optima from Table I.

B. Success Ratio (SR)

The success ratio measures the percentage of successful runs, NSR (a successful run is defined as a run where all known global optima are found), out of all NR runs. It is given by,

$$SR = \frac{NSR}{NR}. \tag{14}$$

C. Convergence Speed

The convergence speed of the proposed approach is measured by counting the number of function evaluations (FEs) required to locate all known global optima, at a specified accuracy level ϵ. The function evaluations are averaged over multiple runs as,

$$AveFEs = \frac{\sum_{run=1}^{NR} FEs_i}{NR}, \tag{15}$$

where FE_i denotes the number of evaluations used in the i-th run. If all the global optima are not found with $MaxFEs$ evaluations, then $MaxFEs$ is used when calculating the average value.

VI. EXPERIMENTAL SETUP [20]

The proposed algorithm is executed for $NR = 50$ runs at each of the five levels of accuracy in $\epsilon = \{1.0E-01, 1.0E-02, 1.0E-03, 1.0E-04, 1.0E-05\}$. The maximum allowable function evaluations used for the test problems are specified in Table II. A random initial population is generated for each run by uniformly varying the seed for the random number

TABLE II. $MaxFEs$ USED FOR 3 RANGES OF TEST FUNCTIONS.

Range of functions	$MaxFEs$
F_1 to F_5 (1D or 2D)	$5.0E+04$
F_6 to F_{11} (2D)	$2.0E+05$
F_6 to F_{12} (3D or higher)	$4.0E+05$

TABLE III. PARAMETERS OF NICHING-ASSISTED NSGA-II.

Parameter	Value
Population Size (POP)	$100D$
No. of generations ($NGEN$)	$\frac{MaxFEs}{POP} - 1$
SBX Crossover Prob. (p_c)	0.9
Polynomial Mutation Prob. (p_m)	0.05
SBX Distribution Index (η_c)	10
Mutation Distribution Index (η_m)	50

generator. Each run is terminated upon reaching $MaxFEs$ evaluations.

The parameters for the niching-assisted NSGA-II are shown in Table III. Note again that no new parameters have been introduced by niching. The population size is kept relatively high, since the stricter dominance criterion in the proposed approach can cause too many non-dominated solutions in turn leading to a stagnant population.

Convergence speed is calculated for the accuracy level $\epsilon = 1.0E-04$. The maximum allowable function evaluations and other parameters are kept the same as in Table II and Table I, respectively.

VII. RESULTS

The PR and SR values obtained using the proposed niching-assisted NSGA-II are tabulated in Table V. The values are significantly better, especially for $D < 5$, when compared to the *untuned* baseline model results provided in [20] obtained using two algorithms [22], [21] based on differential evolution. The algorithm is able to locate all global optima at all accuracy levels in all 50 runs for 5 out of the 20 test problems (namely F_1 (1D), F_2 (1D), F_3 (1D), F_5 (2D) and F_8 (2D)). For the rest of the problems (except F_{12} (20D) where none of the optima were found) Figures 1, 2, 3, 4 and 5 show the number of distinct global minima found at different levels of accuracy.

Table IV shows (at each accuracy level) the number of test problems for which the proposed approach could locate all global optima in at least one of the 50 runs.

Table VI shows the convergence speeds for the proposed approach. In Figures 6, 7 and 8, the mean number of global optima found is shown against generations for problems where all the optima were found (with $\epsilon = 1.0E-04$) in at least one of the runs before reaching $MaxFEs$ function evaluations.

For the sake of comparison, we also present, in Table VII, the peak ratios and success ratios obtained using a recently proposed bi-objective optimization based multimodal optimization algorithm [18].

TABLE IV. NUMBER OF TEST PROBLEMS (OUT OF 20) FOR WHICH ALL GLOBAL OPTIMA COULD BE LOCATED AT LEAST ONCE.

Accuracy (ϵ)	No. of problems
$1.0E-01$	13
$1.0E-02$	6
$1.0E-03$	6
$1.0E-04$	6
$1.0E-05$	6

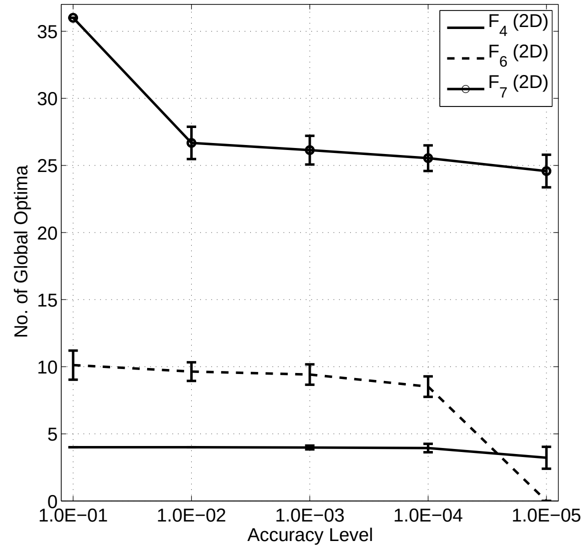

Fig. 1. Number of distinct global minima for problems F_4 (2D), F_6 (2D) and F_7 (2D).

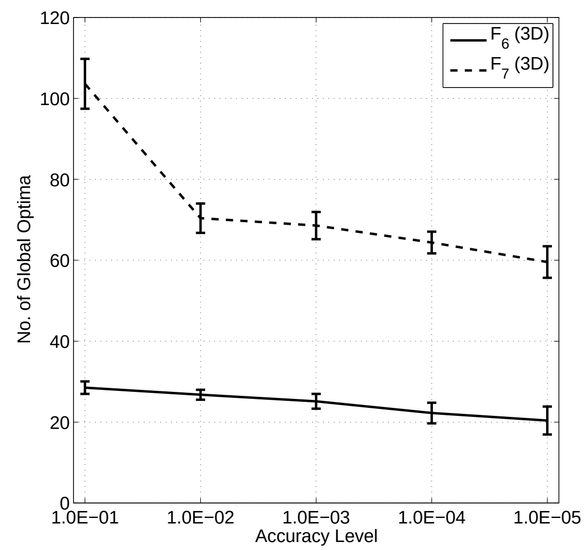

Fig. 2. Number of distinct global minima for problems F_6 (3D) and F_7 (3D).

A. Improving Performance in Higher Dimensions ($D \geq 5$)

It is observed that simply by discarding the secondary objective in Eq. 5 and using the niching-assisted NSGA-II for the sole optimization of the multimodal function improves the performance of the algorithm in terms of PR and SR for problems with $D \geq 5$. The improved results are shown in Table VIII.

VIII. CONCLUSIONS

In this paper, we developed a niching-assisted bi-objective approach to multimodal optimization. The original multimodal problem is first converted into a bi-objective problem through multiobjectivization. A secondary objective is

TABLE V. PEAK RATIOS AND SUCCESS RATES OF THE PROPOSED APPROACH

Accuracy (ϵ)	$F_1(1D)$		$F_2(1D)$		$F_3(1D)$		$F_4(2D)$		$F_5(2D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$1.0E-02$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$1.0E-03$	1.000	1.000	1.000	1.000	1.000	1.000	0.995	0.980	1.000	1.000
$1.0E-04$	1.000	1.000	1.000	1.000	1.000	1.000	0.985	0.960	1.000	1.000
$1.0E-05$	1.000	1.000	1.000	1.000	1.000	1.000	0.805	0.420	1.000	1.000
Accuracy (ϵ)	$F_6(2D)$		$F_7(2D)$		$F_6(3D)$		$F_7(3D)$		$F_8(2D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.562	0.000	1.000	1.000	0.352	0.000	0.480	0.000	1.000	1.000
$1.0E-02$	0.536	0.000	0.741	0.000	0.330	0.000	0.326	0.000	1.000	1.000
$1.0E-03$	0.523	0.000	0.726	0.000	0.310	0.000	0.318	0.000	1.000	1.000
$1.0E-04$	0.473	0.000	0.709	0.000	0.275	0.000	0.298	0.000	1.000	1.000
$1.0E-05$	0.000	0.000	0.683	0.000	0.252	0.000	0.276	0.000	1.000	1.000
Accuracy (ϵ)	$F_9(2D)$		$F_{10}(2D)$		$F_{11}(2D)$		$F_{11}(3D)$		$F_{12}(3D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.877	0.420	0.752	0.000	0.697	0.000	0.933	0.700	0.665	0.060
$1.0E-02$	0.677	0.000	0.715	0.000	0.667	0.000	0.667	0.000	0.495	0.000
$1.0E-03$	0.670	0.000	0.672	0.000	0.667	0.000	0.667	0.000	0.485	0.000
$1.0E-04$	0.680	0.000	0.642	0.000	0.663	0.000	0.663	0.000	0.470	0.000
$1.0E-05$	0.663	0.000	0.573	0.000	0.623	0.000	0.610	0.000	0.443	0.000
Accuracy (ϵ)	$F_{11}(5D)$		$F_{12}(5D)$		$F_{11}(10D)$		$F_{12}(10D)$		$F_{12}(20D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	1.000	1.000	0.917	0.640	0.640	0.640	0.020	0.000	0.000	0.000
$1.0E-02$	0.523	0.000	0.347	0.000	0.117	0.000	0.020	0.000	0.000	0.000
$1.0E-03$	0.523	0.000	0.338	0.000	0.113	0.000	0.043	0.000	0.000	0.000
$1.0E-04$	0.417	0.000	0.300	0.000	0.110	0.000	0.017	0.000	0.000	0.000
$1.0E-05$	0.323	0.000	0.245	0.000	0.093	0.000	0.010	0.000	0.000	0.000

TABLE VI. CONVERGENCE SPEEDS OF THE PROPOSED APPROACH (WITH ACCURACY LEVEL $\epsilon = 1.0E-04$)

Function	$F_1(1D)$	$F_2(1D)$	$F_3(1D)$	$F_4(2D)$	$F_5(2D)$
Mean	32164.0	38888.0	1246.0	44224.0	38064.0
St. D.	593.764	771.069	889.900	3082.972	2953.639
Function	$F_6(2D)$	$F_7(2D)$	$F_6(3D)$	$F_7(3D)$	$F_8(2D)$
Mean	200000.0	200000.0	400000.0	400000.0	161216.0
St. D.	0.000	0.000	0.000	0.000	5772.255
Function	$F_9(2D)$	$F_{10}(2D)$	$F_{11}(2D)$	$F_{11}(3D)$	$F_{12}(3D)$
Mean	200000.0	200000.0	200000.0	400000.0	400000.0
St. D.	0.000	0.000	0.000	0.000	0.000
Function	$F_{11}(5D)$	$F_{12}(5D)$	$F_{11}(10D)$	$F_{12}(10D)$	$F_{12}(20D)$
Mean	400000.0	400000.0	400000.0	400000.0	400000.0
St. D.	0.000	0.000	0.000	0.000	0.000

TABLE VII. PEAK RATIOS AND SUCCESS RATES OF ALGORITHM PROPOSED IN [18].

Accuracy (ϵ)	$F_1(1D)$		$F_2(1D)$		$F_3(1D)$		$F_4(2D)$		$F_5(2D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.930	0.860	1.000	1.000	1.000	1.000	0.720	0.360	0.960	0.920
$1.0E-02$	0.930	0.860	1.000	1.000	1.000	1.000	0.650	0.280	0.950	0.900
$1.0E-03$	0.930	0.860	1.000	1.000	1.000	1.000	0.470	0.020	0.940	0.880
$1.0E-04$	0.930	0.860	1.000	1.000	1.000	1.000	0.320	0.000	0.900	0.800
$1.0E-05$	0.900	0.800	1.000	1.000	1.000	1.000	0.155	0.000	0.680	0.380
Accuracy (ϵ)	$F_6(2D)$		$F_7(2D)$		$F_8(2D)$		$F_6(3D)$		$F_7(3D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.274	0.000	0.739	0.000	0.005	0.000	0.520	0.000	1.000	1.000
$1.0E-02$	0.140	0.000	0.727	0.000	0.000	0.000	0.489	0.000	1.000	1.000
$1.0E-03$	0.041	0.000	0.668	0.000	0.000	0.000	0.345	0.000	0.998	0.980
$1.0E-04$	0.001	0.000	0.509	0.000	0.000	0.000	0.140	0.000	0.953	0.580
$1.0E-05$	0.000	0.000	0.346	0.000	0.000	0.000	0.038	0.000	0.695	0.060
Accuracy (ϵ)	$F_9(2D)$		$F_{10}(2D)$		$F_{11}(2D)$		$F_{11}(3D)$		$F_{12}(3D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.337	0.000	0.225	0.000	0.220	0.000	0.120	0.000	0.102	0.000
$1.0E-02$	0.237	0.000	0.155	0.000	0.113	0.000	0.003	0.000	0.020	0.000
$1.0E-03$	0.100	0.000	0.048	0.000	0.027	0.000	0.000	0.000	0.000	0.000
$1.0E-04$	0.033	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1.0E-05$	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Accuracy (ϵ)	$F_{11}(5D)$		$F_{12}(5D)$		$F_{11}(10D)$		$F_{12}(10D)$		$F_{12}(20D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1.0E-02$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1.0E-03$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1.0E-04$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$1.0E-05$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

TABLE VIII. IMPROVED PEAK RATIOS AND SUCCESS RATES OF NICHING-ASSISTED NSGA-II FOR PROBLEMS WITH $D \geq 5$.

Accuracy (ϵ)	$F_{11}(5D)$		$F_{12}(5D)$		$F_{11}(10D)$		$F_{12}(10D)$		$F_{12}(20D)$	
	PR	SR	PR	SR	PR	SR	PR	SR	PR	SR
$1.0E-01$	0.617	0.000	0.427	0.000	0.880	0.780	0.277	0.160	0.000	0.000
$1.0E-02$	0.617	0.000	0.427	0.000	0.290	0.000	0.147	0.000	0.000	0.000
$1.0E-03$	0.623	0.000	0.430	0.000	0.260	0.000	0.130	0.000	0.000	0.000
$1.0E-04$	0.610	0.000	0.367	0.000	0.223	0.000	0.130	0.000	0.000	0.000
$1.0E-05$	0.423	0.000	0.258	0.000	0.160	0.000	0.123	0.000	0.000	0.000

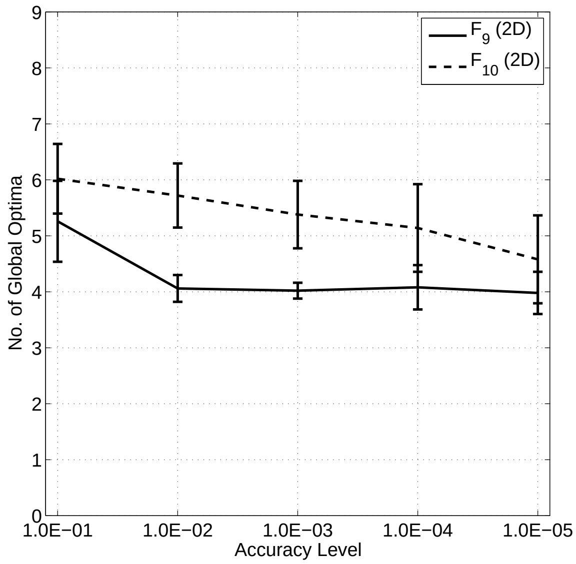

Fig. 3. Number of distinct global minima for problems F_9 (2D) and F_{10} (2D).

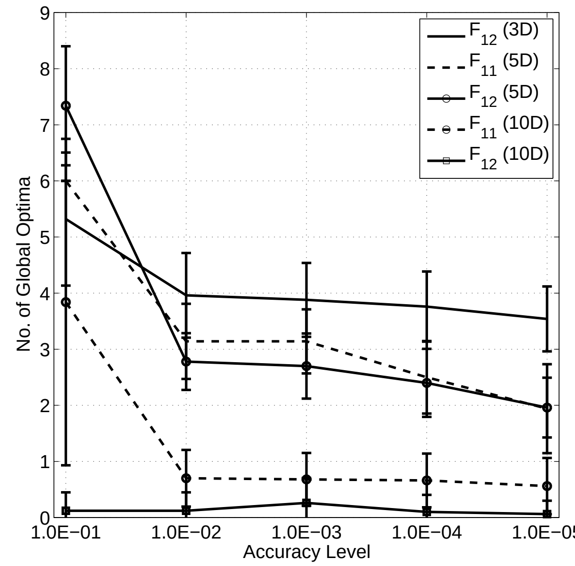

Fig. 5. Number of distinct global minima for problems F_{12} (3D), F_{11} (5D), F_{12} (5D), F_{11} (10D) and F_{12} (10D).

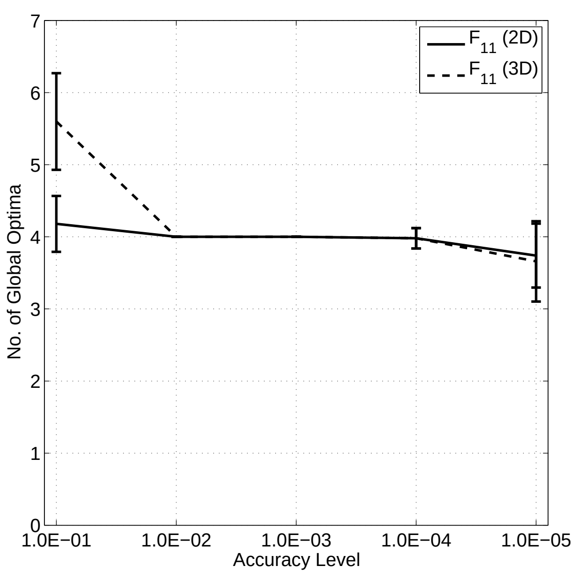

Fig. 4. Number of distinct global minima for problems F_{11} (2D) and F_{11} (3D).

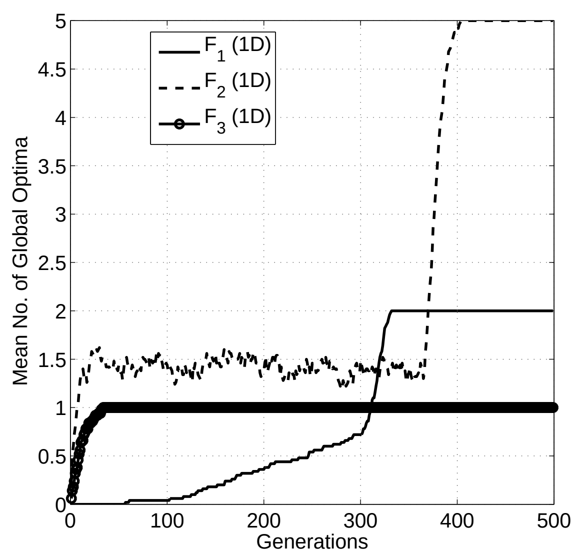

Fig. 6. Mean number of global minima found during the 50 runs for problems F_1 (1D), F_2 (1D) and F_3 (1D) for $\epsilon = 1.0E-04$.

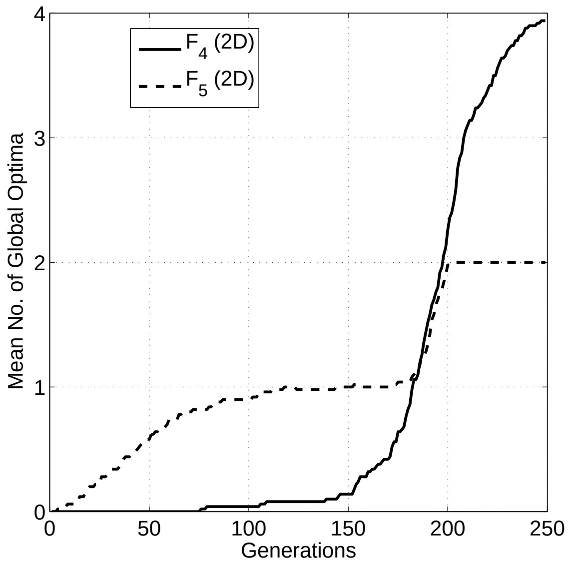

Fig. 7. Mean number of global minima found during the 50 runs for problems F_4 (2D) and F_5 (2D) for $\epsilon = 1.0E - 04$.

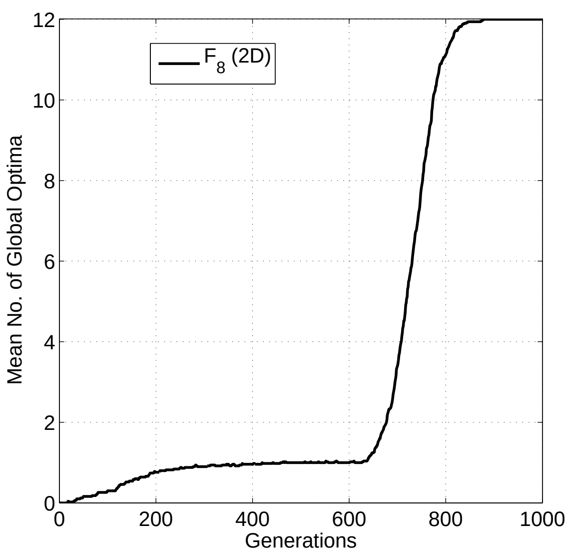

Fig. 8. Mean number of global minima found during the 50 runs for problem F_8 (2D) for $\epsilon = 1.0E - 04$.

introduced which aims at maximizing the diversity of the population. However, it may not always be conflicting with the original multimodal function. A parameterless-niching is introduced which modifies the definition of dominance so as to allow domination of one individual over another only when they are proximate in all dimensions. The niching distance in each dimension is automatically defined from the population size and number of dimensions. The proposed approach has been applied to 20 benchmark problems and various performance measures are obtained.

References

[1] K. Deb and A. Srinivasan, "Innovization: Innovating design principles through optimization," in *Proceedings of the 8th annual conference on genetic and evolutionary computation, GECCO '06*. New York: ACM, 2006, pp. 1629–1636.

[2] S. Mahfoud, "Niching methods for genetic algorithms," Ph.D. dissertation, University of Illinois at Urbana-Champaign, 1995.

[3] S. W. Mahfoud, "A comparison of parallel and sequential niching methods," in *Proceedings of 6th International Conference on Genetic Algorithms*, 1995, pp. 136–143.

[4] D. E. Goldberg and J. Richardson, "Genetic algorithms with sharing for multimodal function optimization," in *Proceedings of the Second International Conference on Genetic Algorithms on Genetic algorithms and their application*. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1987, pp. 41–49.

[5] B. L. Miller and M. J. Shaw, "Genetic algorithms with dynamic niche sharing for multimodal function optimization," in *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*. IEEE, 1996, pp. 786–791.

[6] K. A. De Jong, "An analysis of the behavior of a class of genetic adaptive systems." Ph.D. dissertation, Ann Arbor, MI, USA, 1975, aAI7609381.

[7] O. J. Mengshoel and D. E. Goldberg, "Probabilistic crowding: Deterministic crowding with probabilistic replacement," in *Proc. of the Genetic and Evolutionary Computation Conference (GECCO-99)*, 1999, pp. 409–416.

[8] A. Petrowski, "A clearing procedure as a niching method for genetic algorithms," in *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*. IEEE, 1996, pp. 798–803.

[9] X. Yin and N. Germay, "A Fast Genetic Algorithm with Sharing Scheme Using Cluster Analysis Methods in Multimodal Function Optimization," in *Artificial Neural Nets and Genetic Algorithms*, R. F. Albrecht, N. C. Steele, and C. R. Reeves, Eds. Springer Verlag, 1993, pp. 450–457.

[10] K. Deb, "Genetic algorithms in multi-modal function optimization," Master's thesis, Tuscaloosa, AL, USA, 1989.

[11] ——, *Multi-objective optimization using evolutionary algorithms*. New York: Wiley, 2001.

[12] S. Das, S. Maity, B.-Y. Qu, and P. Suganthan, "Real-parameter evolutionary multimodal optimizationa survey of the state-of-the-art," *Swarm and Evolutionary Computation*, vol. 1, no. 2, pp. 71–88, 2011.

[13] Y. Liu, X. Ling, Z. Shi, M. Lv, J. Fang, and L. Zhang, "A survey on particle swarm optimization algorithms for multimodal function optimization," *Journal of Software*, vol. 6, no. 12, pp. 2449–2455, 2011.

[14] J. Knowles, R. Watson, and D. Corne, "Reducing local optima in single-objective problems by multi-objectivization," in *Evolutionary Multi-Criterion Optimization*. Springer, 2001, pp. 269–283.

[15] K. Deb, S. Agarwal, A. Pratap, and T. Meyarivan, "A fast and elitist multi-objective genetic algorithm: Nsga-ii," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[16] ——, "Multi-objective nsga-ii code in c," 2005, available from: http://www.iitk.ac.in/kangal/codes.shtml.

[17] K. Deb and R. B. Agrawal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, no. 2, pp. 115–148, 1995.

[18] K. Deb and A. Saha, "Multimodal optimization using a bi-objective evolutionary algorithm," *Evolutionary Computation*, vol. 20, no. 1, pp. 27–62, 2012.

[19] C. Aggarwal, A. Hinneburg, and D. Keim, "On the surprising behavior of distance metrics in high dimensional space," *Database TheoryICDT 2001*, pp. 420–434, 2001.

[20] M. G. E. Xiaodong Li, Andries Engelbrecht, "Benchmark Functions for CEC2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization," Royal Melbourne Institute of Technology, Tech. Rep., March, 2013, http://goanna.cs.rmit.edu.au/ xiaodong/cec13-niching/.

[21] M. G. Epitropakis, V. P. Plagianakos, and M. N. Vrahatis, "Finding multiple global optima exploiting differential evolution's niching capability," in *Differential Evolution (SDE), 2011 IEEE Symposium on*. IEEE, 2011, pp. 1–8.

[22] R. Thomsen, "Multimodal optimization using crowding-based differential evolution," in *Evolutionary Computation, 2004. CEC2004. Congress on*, vol. 2. IEEE, 2004, pp. 1382–1389.