ORIGINAL PAPER

An investigation on niching multiple species based on population replacement strategies for multimodal functions optimization

Minqiang Li · Dan Lin · Jisong Kou

Abstract This paper studies the niching mechanism based on population replacement in the process of evolution to solve the multimodal functions optimization (MMFO) problems. In order to niche multiple species for the MMFO tasks, the overlapping population replacement is surely needed because the offspring population most probably does not inherit all of the genetic information contained in its parental population, and the basic procedure for niching genetic algorithms with overlapping population replacement is presented. Then four niching schemes, the nearest neighbors replacement crowding (NNRC), the species conservation technique (SCT), the HFC-I (implicit hierarchical fair competition), and the CPE (clearing procedure with elitist) are investigated. These niching schemes are characterized with regard to different niching strategies and parameterizations, and the corresponding niching procedures are outlined. Finally, experiments are carried out on a suite of test functions to compare different niching strategies regarding niching efficiency and scalability. Experimental results illustrate the intrinsic difference of the four niching schemes. The NNRC and HFC-I have a mechanism of multiple species coevolution via adapting multiple species to different niches, while the SCT and CPE tend to make use of a mandatory mechanism to conserve species just like the grid searching over the solution space based on species distance or clearing radius. All niching methods are able to deal with complex MMFO problems, while the NNRC and HFC-I show a better performance in terms of niching efficiency and scalability, and are more robust regarding the algorithm parameterization.

1 Introduction

Multimodal optimization problems are ubiquitous in the real world. It is often necessary to find multiple optima in the solution space so as to get more information about an optimization problem for supporting decision making tasks in engineering and management. There have been growing interests in multimodal function optimization (MMFO) in the community of the genetic algorithms (GAs) or evolutionary computation (EC) (Bäck et al. 2000; Eiben and Smith 2003).

The multimodality of a multimodal function poses a great difficulty for the GAs and conventional heuristic searching methods (Horn 1997; Li and Kou 2008). The simple genetic algorithms (SGA) are not suited to solve the MMFO problems because the final population tends to converge to a single solution (mostly a local optimum). Since the fitness sharing was proposed by Goldberg and Richardson (Goldberg and Richardson 1987; Goldberg et al. 1992) to address the MMFO problems, various approaches have been invented by aiming at maintaining the diversity of the population for concurrently searching a

M. Li (✉) · D. Lin · J. Kou
School of Management, Tianjin University, 300072 Tianjin,
People's Republic of China
e-mail: mqli@tju.edu.cn

D. Lin
e-mail: dlin@tju.edu.cn

J. Kou
e-mail: jskou@tju.edu.cn

variety of peaks in the solutions space of an optimization problem (Sareni and Krähenbühl 1998; Cioppa et al. 2004).

Niching methods or techniques, called niching GAs [or niching evolutionary algorithms (EAs)], have been proposed to solve MMFO problems. These approaches allow the GAs or EAs to identify and locate multiple optima simultaneous or sequentially by employing specific niching techniques, so that the final population or multiple subpopulations are distributed spatially among all or most of the niches, and only a subset of individuals is allowed to adapt to a niche by niching schemes (Goldberg 1989; Bäck et al. 2000). In other words, a local optimum could only be exploited by a subset of genetically similar individuals (called a species), and the local adaptation makes different species of individuals adapt to different niches on the multimodal landscape. All of the species gradually get to the distribution equilibria proportional to the fitness or the attraction basin size of all niches, and are thus preserved forever (Goldberg 1989; Li and Kou 2005).

Existing niching techniques can be classified into two categories, explicit niching schemes or implicit niching schemes, by whether they utilize explicit divide-and-conquer strategy or not. The explicit niching scheme adopt clustering, restricted mating, or multiple subpopulations to do coevolutionary searching of the multimodal landscape via multiple species. The typical methods of explicit niching include the individual clustering based sharing schemes (Yin and Germany 1993; Lin and Yang 1998; Pelikan and Goldberg 2000; Jelasity et al. 2001; Gan and Warwick 2001), the species conservation GAs (Li et al. 2002) and species conservation particle swarm optimizer (Iwamatsu 2006), the restricted competition and mating (Goldberg 1989; Goldberg et al. 1990; Harik 1995), the sequential niching (Beasley et al. 1993), the island model GAs and the parallel GAs (Siarry et al. 2000, 2002; Gudla and Ganguli 2005), the estimation of distribution algorithm (EDA) (Peña et al. 2005), the parallel hill-climbing and multi-start strategies of hill-climbing in conventional mathematical programming domain. The crowding and preselection (De Jong 1975; Mahfoud 1995; Cedeño and Vemuri 1999), the fitness sharing (Goldberg and Richardson 1987; Goldberg 1989), and recently proposed crowding with nearest neighbors replacement (NNRC) (Li and Kou 2005, 2008), are typical implicit niching ones.

The performance of niching methods in solving MMFO problems is usually measured by efficacy, efficiency, robustness, and pervasive applicability. It is unrealistic to compare and evaluate all of the niching methods over a broad class of MMFO problems. But we can gain a good understanding of different niching mechanisms by investigating representatives of either implicit niching schemes or explicit niching schemes. In addition, we confine our research on the niching strategies working in a single

population and having global exploration capability. Here, the recombination operations of individuals from different species are required. Although the global crossover operation may probably yield low fitness individuals when parents do not come from identical species, which delays the species-level convergence of the population, it is necessary to recombine building blocks for exploring the total solution space without a massive population size. Finally, our research is oriented to real-valued MMFO problems with continuous solution spaces.

Therefore, four specific schemes for niching multiple species in a single population are selected for detailed investigation: the crowding with nearest neighbors replacement NNRC (Li and Kou 2005, 2008), the species conservation technique (SCT) (Li et al. 2002), the hierarchical fair competition model (HFC model) (Hu and Goodman 2004; Hu et al. 2005), the clearing procedure (Pétrowski 1996). They are selected in terms of simple niching mechanism, good niching capability, high efficiency, and pervasive applicability.

In Sect. 2, the niching mechanism is modeled as the process of population replacement in evolution (either generational or steady-state), and multiple species are conserved by niching methods with particular strategies for implementing population replacement. The NNRC and the SCT are introduced in Sects. 3 and 4, and are characterized in terms of particular niching mechanisms. The implicit HFC and the clearing procedure with elitist are introduced in Sect. 5. A suite of multimodal functions are selected and experiments are conducted in Sect. 6. Section 7 concludes this paper and recommends directions in the future work.

2 Niching multiple species with population replacement

A niche represents a local optimum and its attraction basin in the landscape of the MMFO (Goldberg 1989; Mahfoud 1995; Horn 1997; Eiben and Smith 2003). The niching GAs emulates the mechanics of the coevolution of multiple species in nature for finding multimodal optima. This aim is realized by implementing concurrently exploration and exploitation of all or most of the niches in MMFO problems.

Since the original GAs by Holland (1992) assumed the generational replacement paradigm, the population, consisting of N chromosomes, was modified by selection, crossover, and mutation operations (Reeves and Rowe 2003). De Jong (1975) introduced the elitism and overlapping population replacement strategy, where offspring only replaced a part of individuals in the population in each iteration. The evolution strategies (ESs) by Rechenberg and Schwefel (Beyer 2001) operates with the (μ, λ) or $(\mu + \lambda)$

strategies, where μ stands for the population size, λ indicates the number of offspring produced by genetic operators. The $(\mu + \lambda)$ scheme implements the overlapping population replacement, where the selection pool contains both the offspring and parental individuals, and truncation selection is usually implemented.

When the MMFO tasks are considered, the overlapping population replacement is surely needed in the evolution process of GAs. The multiple species should be maintained in the population, which is called the species-level diversity. Since offspring populations most probably do not inherit all of the genetic information contained in parental populations, the species that exist in the proceeding population can only be conserved successfully by merging the two populations and then exerting a specific selection or deletion operation to yield the new one. This manipulation may be called species-level elitism in comparison with the population-level elitist selection in the standard GAs, and the former is aimed at obtaining multimodal optima convergence against the global optimum convergence by the latter.

Suppose that a real-valued MMFO problem is represented as $x^* = \arg\max\{f(x), x \in R^n\}$ The real-coded chromosome is a vector of continuous variables $a = (a_1, a_2, \ldots, a_n)$, $a \in R^n$ In order to niche multiple species in the population, a specific operation should be introduced to complete the population replacement. Based on the idea of overlapping population evolution, and suppose that $P(t)$ and $P_{\mathrm{offspring}}(t)$ represent the parental population and offspring population individually in the tth generation or iteration, we introduce the species niching operation as the final step in one repetition of genetic operations to produce the new population:

$$P(t+1) = \Phi(P(t), P_{\mathrm{offspring}}(t)), \qquad (1)$$

where $P_{\mathrm{offspring}}(t) = S \circ C \circ M(P(t)); P(t+1)$ indicates the population in the $(t+1)$-th generation or iteration; S, C, M are the selection, crossover, and mutation operations exerting on $P(t)$

Usually, we do not assume the size of the offspring population as identical to that of the parental population, and $|P_{\mathrm{offspring}}(t)| \geq 1$ is the feasible lower bound for the niching operation. What is critical to niching multiple species is the niching strategies involving the parental population and the offspring population. Therefore, the conceptual algorithm of the niching GAs (or EAs) based on overlapping population replacement is built as Fig. 1.

There are numerous ways to implement the overlapping population replacement for the purpose of niching multiple species, such as the De Jong's crowding and preselection (De Jong 1975), the deterministic crowding (DC) (Mahfoud 1995), the multi-niche crowding (MNC) (Cedeño and Vemuri 1999), the NNR crowding (Li and Kou 2005,

```
procedure Niching_GAs (P_offspring(t), P(t))
    set: t := 0;
    initialize: P(0);
    compute individual fitness: {f_1(0), f_2(0), ..., f_N(0)};
    while stopping_criteria_not_ met do
        produce offspring by performing selection, crossover, mutation:
            P_offspring(t) := S ∘ C ∘ M(P(t));
        compute individual fitness: {f_1(t), f_2(t), ..., f_{|P_offspring(t)|}(t)};
        produce new population by species niching:
            P(t+1) := Φ(P(t), P_offspring(t));
        set: t := t + 1;
    end while
    return {P(t)}
end procedure
```

Fig. 1 The basic procedure for niching GAs with overlapping population replacement

2008), and the SCT (Li et al. 2002), etc. Three interrelated goals should be attained: (1) the competition among individuals is localized so that genetic drift is prevented; (2) the species-level elitism should be sustained so that species are maintained once they are located; (3) the exploitation of different niches should be done concurrently so that local optima are found. A perfect niching strategy should be able to achieve all objectives.

3 Nearest neighbors replacement crowding (NNRC)

The NNRC was originally proposed in the recombination-replacement paradigm of the steady-state GAs, where random selection was done to take advantage of the global exploration power of the crossover operation. Since more than one nearest neighbors can be used to do local replacement for an offspring, the scheme is also called the q-NNR crowding strategy (Li and Kou 2005, 2008). Since offspring only replace their genotypically most similar ones in the population, the competition among individuals is confined locally in different niches, and thus multiple species can be conserved to adapt to multiple niches on the landscape (Li and Kou 2008).

3.1 NNRC algorithm

The NNRC works in the paradigm of both the steady-state GAs and the generational GAs. For the sake of generality, we consider the generic paradigm of the niching GAs modeled in Sect. 2, where the parental population of the GAs is denoted as $P = \{a_1, a_2, \ldots, a_N\}$, offspring population is denoted as $P_{\mathrm{offspring}} = \{a_1, a_2, \ldots, a_M\}$. Then the crowding scheme based on the q-NNR to yield a new population is presented as below.

The similarity measure for finding the $q(1 \leq q < N)$ number of nearest neighbors is computed based on the Euclidean distance with the real-valued MMFO. For any $\boldsymbol{a}' \in \boldsymbol{P}_{\text{offspring}}$, we have:

$$d(\boldsymbol{a}', \boldsymbol{a}_i) = ||\boldsymbol{a}' - \boldsymbol{a}_i||_2 = \sqrt{\sum_{k=1}^{n} (a'_k - a_{i,k})^2}, \quad \boldsymbol{a}_i \in \boldsymbol{P} \quad (2)$$

so that we get $\{d(\boldsymbol{a}', \boldsymbol{a}_i) | i = 1, 2, \ldots, N\}$, and the q nearest individuals are found as: $\boldsymbol{P}_{\text{NN}} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_q\}(\boldsymbol{P}_{\text{NN}} \subset \boldsymbol{P}, q << N)$. Then the one with the lowest fitness is to be replaced. Suppose that we get $\boldsymbol{a}^* = \arg \min\{f(\boldsymbol{a}), \boldsymbol{a} \in \boldsymbol{P}_{\text{NN}}\}$, and if $f(\boldsymbol{a}') > f(\boldsymbol{a}^*)$, then $\boldsymbol{a}^* \in \boldsymbol{P}$ will be replaced by $\boldsymbol{a}'$.

The value of parameter q is determined by considering both the available resource (as the population size) and the computational efficiency of the niching algorithm. We might as well take $q = 1$ as default. If the bigger q is taken, the replacement will cover a larger region in the solution space, and the wrong replacement will happen more probably, which leads to genetic drift. When $q = N$ (the worst deletion scheme in the standard steady-state GAs), there will appear global competition among individuals in the population replacement, and all individuals in the final population evolve only to one niche.

The basic procedure of the NNRC scheme for producing a new population $\boldsymbol{P}(t+1)$ based on $\boldsymbol{P}_{\text{offspring}}(t)$ and $\boldsymbol{P}(t)$ between succeeding generations or iterations is shown in Fig. 2.

The NNRC seems to be similar to the MNC GA (Cedeño and Vemuri 1999) within the steady-state GAs, but there exists important difference regarding their niching strategies.

The MNC adopted crowding both in the selection step and replacement step for confining the individual

procedure NNR_crowding($\boldsymbol{P}_{\text{offspring}}(t), \boldsymbol{P}(t), q$)

    $\boldsymbol{P}_{\text{temp}} := \boldsymbol{P}(t)$;

    for $\boldsymbol{a}' \in \boldsymbol{P}_{\text{offspring}}(t)$ do

        compute: $\boldsymbol{P}_{\text{NN}} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_q\}$, $\boldsymbol{P}_{\text{NN}} \subset \boldsymbol{P}_{\text{temp}}$;

        find: $\boldsymbol{a}^* = \arg \min\{f(\boldsymbol{a}), \boldsymbol{a} \in \boldsymbol{P}_{\text{NN}}\}$;

        if $f(\boldsymbol{a}') > f(\boldsymbol{a}^*)$ then

            replace: $\boldsymbol{a}^* := \boldsymbol{a}'$, $\boldsymbol{a}^* \in \boldsymbol{P}_{\text{temp}}$;

        end if

    end for

    $\boldsymbol{P}(t+1) := \boldsymbol{P}_{\text{temp}}$;

    return $\{\boldsymbol{P}(t+1)\}$

end procedure

Fig. 2 The basic procedure for the NNRC scheme

competition in the same niche. The crowding selection becomes the random global selection with $s = 1$, and works like the restricted mating when $s = N$. So, it was a hybrid between the global random selection and the restricted mating schemes. The worst among most similar (WAMS) was developed for the crowding replacement in the MNC by specifying the crowding groups (size g) and crowding size (number f). The WAMS becomes the NNRC ($q = 1$) with $f = 1$ and $g = N$ if we do not consider the competition between the offspring and the parental individual. The NNRC ($q = 1$) shares similarity with the MNC with $s = 1$, $f = 1$, and $g = N$. The major difference is that the NNRC takes into consideration of the competition between the offspring and the parental individuals with the lowest fitness, and the winner enters the new population, otherwise the replacement does not happen. This local elitism is good for the convergence of species residing in different niches. For the case $q > 1$ that is able to promote the local searching or exploitation in niches, there is no correspondence between the two schemes. The WAMS with $f > 1$ and $g < N$ raises the possibility of the wrong replacement of an individual in a niche by an offspring in another niche.

The real-parameter GAs for finding multiple optimal solutions in multi-modal optimization by Ballester and Carter (2003) was also an efficient niching scheme based on the population replacement. But the niches with local optima were usually not maintained although they could be found in the evolution process. This is caused by the competition between the global optima and local optima in individual replacement. This is the main reason that the nearest individuals for replacement of an offspring in the NNRC strategy are selected from the total parental population instead of from the randomly preselected group; or in other words, the crowding factor is the population size. The construction of the nearest neighbor set in the NNRC is different from the tournament selection of individuals for replacement. It only considers the distance, either the Euclidean or the Hamming, between the offspring and individuals in the parental population. This makes it possible for a species to be conserved once it is formed on a niche even with lower fitness, especially with $q = 1$.

3.2 Niching capability

With the NNRC scheme, less competitive niches with smaller attraction basins and lower fitness can still be preserved because of the confined local competition among offspring and nearest individuals in the population replacement. The number of individuals adapting to a niche is mainly proportional to the size of its attraction basin in the fitness landscape in contrast to the fitness-proportionate individual's distribution with the fitness sharing method (Horn 1997).

When local optima or global optima are found, will the species adapting to a niche draw individuals from species of neighbor niches? This question is associated with the properties of the multimodal functions. In order to make a descriptive analysis, let's consider an univariate continuous function with two peaks:

$$\max f(x) = \begin{cases} \dfrac{a(c-x)}{c} & 0 \leq x < c \\ \dfrac{b(x-c)}{1-c} & c \leq x \leq 1 \end{cases}, \quad x \in [0,1], \quad (3)$$

where $\{a, b, c\}$ are parameters to control the shape of the univariate function, and $a < b$, $0 < c < 0.5$. It has two niches, $x \in (c, 1]$ for niche A, and $x \in [0, c]$ for niche B (see Fig. 3). The global optimum is $x_A^* = 1, f(x_A^*) = b$ in niche A, and the local optimum is $x_B^* = 1, f(x_B^*) = a$ in niche B.

When there are more than two niches, we only consider two neighborhood niches. Suppose that there are two individuals that adapt to niche A and niche B separately, $x_A = 1, f(x_A) = b$, and $x_B = 0, f(x_B) = a$, called species A and species B. There are following facts about this univariate function.

$$f(x_1^*) = f(x_B) = a, \quad \text{where } x_1^* = \frac{a(1-c)}{b} + c;$$
$$f(x) > f(x_B) = a, \quad x_1^* < x \leq 1;$$

If an offspring x' falls between x_1^* and 0.5, $x_1^* < x' < 0.5$, then x' will replace x_B with the $q = 1$ replacement strategy because $f(x') > f(x_B)$ and then the species B will disappear in the population hereafter. Even if x_B is produced in succeeding generations, it will not make successful replacement because its fitness is not competitive, $f(x_B) < f(x_A)$ In other words, there is not a chance to recover species B in latter generations. In such cases, the NNRC scheme fails to conserve species B no matter how big the population size is. If $x_1^* > \frac{x_B + x_A}{2} = 0.5$ or
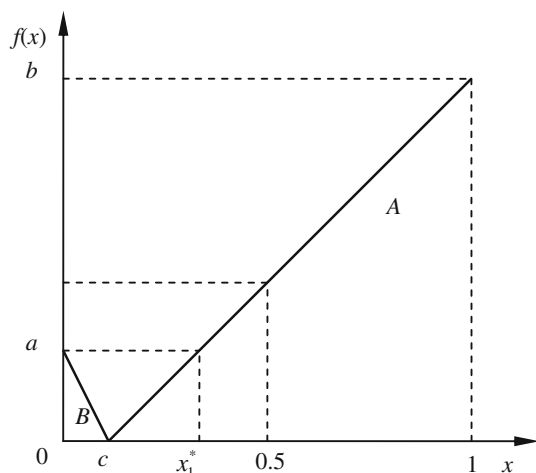

Fig. 3 A two-peak univariate function

$f(x_B) > f\left(\frac{x_B + x_A}{2}\right) = f(0.5)$, both species will be kept to exploit the two niches A and B separately.

Actually, niches with very low fitness and pretty small attraction basins are treated as residing in the attraction basins of neighborhood big ones in real-world situations. They do not contribute extra information to the decision making, and decision makers often discard minor local optima without suffering significant loss.

As to the niching scheme with $q > 1$ and $q << N$, the niching efficiency can be speeded up (Li and Kou 2008). However, if q takes a very big value, replacement errors could probably appear, and species adapting to low fitness niches could be taken over by species for neighborhood ones with bigger fitness. Suppose that $f(x_B) < f(x_A)$, and individuals belonging to species A and species B are numbered as $N_A, N_B(N_A + N_B = N)$, the species B can not be conserved with $q = N$. This happens definitely when the replace-worst strategy is employed. In the case there are multiple niches with niche A and niche B as neighbors, $f(x_B) < f(x_A), N_A + N_B < N$, it is not possible to preserve the species B if $q \geq N_A + N_B$.

Therefore, the q is parameterized by considering a good balance between the species conservation and algorithm efficiency. For this purpose, the Boltzmann scheme for sizing the nearest neighbor set dynamically is designed as:

$$q_t = q_{\min} + \left[\frac{e^{\left(\frac{t}{T_{\max}}\right)} - 1}{e - 1}\right]^{\alpha} \times (q_{\max} - q_{\min}), \quad (4)$$

where q_t is the nearest neighbor set size in tth generation or iteration; $q_{\min}, q_{\max}$ are the minimum value (also the initial value) and the maximum (or final value) for the size of nearest neighbors set in total evolution process. Usually we take $q_{\min} = 1$ as default. $T_{\max}$ is the maximum generations or iterations of the algorithm. α is the scaling factor, and we have $\alpha = 1$ or $\alpha = 2$.

4 Species conserving technique (SCT)

The SCT was proposed by Li et al. (2002) to evolve multiple species in parallel in a single population. It adopted explicit niching scheme to divide the population into species according the similarity of individuals. By defining the species distance threshold σ_s (the niches radius is denoted as $\sigma_s/2$), two individuals are considered to be similar if their distance (genotypic or phenotypic) is smaller than half of the threshold. A species is a subset of individuals in the population.

Given a population, $P = \{a_1, a_2, \ldots, a_N\}$ a partition of the population consists of species, $S_1, S_2, \ldots, S_K$, that are exclusive and complete, $P = \bigcup_{k=1}^{K} S_k, S_{k_1} \cup S_{k_2} = \varnothing, k_1 \neq k_2, k_1, k_2 = 1, 2, \ldots, K$. A dominating individual a_k^* is defined as the best one in the species S_k:

$$f(\boldsymbol{a}_k^*) \geq f(\boldsymbol{a}_k), \quad \boldsymbol{a}_k^*, \boldsymbol{a}_k \in S_k \tag{5}$$

and the species S_k is centered around its dominating individual $\boldsymbol{a}_k^*$:

$$d(\boldsymbol{a}_k^*, \boldsymbol{a}_k) \leq \sigma_s/2, \quad \boldsymbol{a}_k^*, \boldsymbol{a}_k \in S_k \tag{6}$$

Intuitively, individuals with bigger fitness are candidates to be conserved. Besides, the SCT also tends to conserve individuals that do not have big fitness but are different enough to the current best ones which have been conserved. The dominating individuals are also called species seeds, and are identified by the algorithm in Fig. 4.

After the species seeds are identified, the population is easily split into multiple species. But the SCGA still works as the SGA regarding the genetic operations, including selection, crossover, and mutation. However, what are conserved are the species seeds instead of all individuals in species. Once the offspring population is obtained, a species conserving process is conducted to ensure that species are preserved in the new population. In fact, some species may get lost after the genetic operations in the offspring population, thus the species conserving procedure is mandatory as shown in Fig. 5.

The SCT uses the species seeds set X_S instead of the old population $\boldsymbol{P}$ to do overlapping population replacement, where the species seeds set X_S is an approximate cover of the population $\boldsymbol{P}$ with regard to the species distance threshold σ_s. When proper parameterization is done about σ_s with a priori knowledge about a MMFO problem, the X_S

```
procedure Determine_species_seeds (P(t))
    X_S(t) := ∅ ;
    while not (no more unmarked individuals in  P(t) ) do
        search for the best unmarked  a* ∈ P(t) ;
        mark  a*  as processed;
        found:= false;
        for all  a ∈ X_S(t)  do
            if  d(a, a*) ≤ σ_s / 2  then
                found:= true;
                break;
            end if
        end for
        if (not found) then
            X_S(t) := X_S(t) ∪ {a*} ;
        end if
    end while
    return { X_S(t) }
end procedure
```

Fig. 4 The algorithm for determining species seeds

```
procedure Conserve_species (P_offspring(t),  X_S(t))
    P(t + 1) := P_offspring(t) ;
    mark individuals in  P(t + 1)   as unprocessed;
    for all  a' ∈ X_S  do
        select the worst unmarked individuals  a ∈ S(a', σ_s) ;
        if (a  exists) then
            if ( f(a') > f(a)  exists) then
                a := a' ;
            end if
        else
            select the worst unmarked individuals  a ∈ P(t + 1) ;
            a := a' ;
        end if
        mark  a  as processed;
    end for
    return {P(t + 1)}
end procedure
```

Fig. 5 The algorithm for conserving species

is able to represent all niches in the landscape. Thus all species in the old populations are preserved in the new ones through the species conserving algorithm even if some species are not contained in the offspring population $\boldsymbol{P}_{\text{offspring}}$.

However, if σ_s is set as bigger than the actual distance between niches, which usually happens when we do not have good problem-specific knowledge about the MMFO problem, the SCT will most probably fail to conserve some species (Li et al. 2002).

Let's consider the univariate continuous function with two peaks in Fig. 3. With $\sigma_s > 1$, it will be impossible for the species seeds set X_S to contain a seed residing in the niche B, and the species B will be lost forever in the population. Even if the niche A and niche B have equal peak fitness, $f(x_A^*) = f(x_B^*)$, it is still possible that they are not preserved simultaneously in the population. If either is not retained in the offspring population because of the genetic drift, and meanwhile the species seeds set X_S does not have a seed of the species, it will disappear in the next population.

There is another potential risk with the SCT scheme, some species seeds do not actually indicate niches on the fitness landscape when the species distance threshold is much smaller than the actual niche distance, or they indicate fake niches. Since the population size is limited, an individual is determined as a seed just because of its dissimilarity with other seeds, and it could probably not be a local optimum or close to an optimum. Thus, the final set of species seeds yielded after the algorithm terminates has to be checked to make sure whether a seed is a true local optimum or not. This is to be observed in experiments.

5 Niching strategies based on hierarchical fair competition and clearing procedure

There are two niching techniques, the hierarchical fair competition and the clearing procedure, that work in similar mechanisms with the NNRC and the SCT. They are investigated in the paradigm of population replacement for niching multiple species as well.

5.1 Hierarchical fair competition (HFC)

The HFC model was proposed by Hu and Goodman (2004) based on the combination of spatial niching and continuous temporal niching. The HFC algorithm employed multiple subpopulations to form an assembly-line structure with different fitness levels to implement the preservation and sustainable evolution of hierarchical building blocks (Hu et al. 2005). The HFC (or quick HFC, QHFC) is actually a special niching technique that hybridizes the spatial crowding (deterministic crowding) and fitness crowding, and it is able to achieve good efficiency and robustness with a proper division of population in the fitness range.

The main difficulty of this algorithm arises with the determination of the multiple subpopulations structure for specific MMFO problems because the fitness ranges is usually unknown a priori. The inventors proposed the adaptive allocation of subpopulations to fitness levels to address this issue, which in turn introduces more complexity in application. The second difficulty is the migration policy of individuals from lower-level subpopulations to higher-level ones that needs to be finely tuned.

In contrast to the explicit niching scheme of the HFC, we propose to reconstruct the HFC as an implicit niching technique (HFC-I). The division of the total population in the fitness range is done dynamically through restricted selection for reproduction and restricted crowding for replacement in terms of fitness levels.

Suppose the population consists of N individuals, $P = \{a_1, a_2, \ldots, a_N\}$ and the first parent a_1 is selected randomly, then the second parent a_2 is chosen randomly as:

$$a_2 \in \{a | f(a) \in [f(a_1) - \Delta_f, f(a_1) + \Delta_f]\},$$

where Δ_f is a small real number, and $\Delta_f > 0$. The two offspring, a_1', a_2', produced by crossover and mutation operations, are inserted back to the population by replacing the nearest neighbors, $a_1'^*, a_2'^*$, belonging to identical fitness levels separately:

$$a_1'^* = \arg\min\{f(a), a \in P_{NN,1}\},$$
$$P_{NN,1} = \{a | f(a) \in [f(a_1') - \Delta_f, f(a_1') + \Delta_f]\},$$
$$a_2'^* = \arg\min\{f(a), a \in P_{NN,2}\},$$
$$P_{NN,2} = \{a | f(a) \in [f(a_2') - \Delta_f, f(a_2') + \Delta_f]\},$$

which are embedded in the NNRC niching procedure by NNR_crowding $(P_{offspring}(t), P(t), q)$ in the paradigm of the SSGA. If Δ_f takes big values, the HFC-I will perform more closely to the NNRC, where the fitness crowding will downgrade or disappear. A relative parameter is used in experiments as $\Delta_f = \gamma f$, where $0 < \gamma < 1$.

Thus, the mating and replacement of individuals will be confined within certain fitness sub-ranges just like genetic operations in subpopulations in the original HFC. This scheme avoids the global competition of individuals with different fitness, and is good for the preservation of building blocks of various patterns and strength. The NNRC, instead of the DC as in the original HFC, is used in the HFC-I to further enhance the localization of individual competition in identical fitness levels. Both the HFC-I and the HFC yield a multimodally converged population, while the HFC-I does not introduce random individuals during total evolution process. The nearest neighbors set is computed as in the NNRC. The potential risk with the HFC-I and also the original HFC is that the restricted mating blocks the mixing of building blocks of different strength, which might hinder the efficiency of recombination operation.

5.2 Clearing procedure with elitist (CPE)

The clearing procedure was proposed by Pétrowski (1996) via modifying the resources allocation of the fitness sharing among similar individuals to only elitists of species. The numbers of elitists for reproduction was computed based on the fitness of elitists representing multiple species adapting to different niches, and the stochastic universal selection (SUS) was employed to produce the mating pool individuals. The clearing procedure used the clearing radius (σ) to identify niches and the capacity (k) to indicate the number of winners of individuals in a niche. The basic procedure is shown in Fig. 6.

The clearing procedure works quite similar to the SCT algorithm, and the number of winners with nonzero fitness is decided by the clearing radius (σ) and the capacity (k). Since the capacity is usually set as 1, the clearing procedure becomes the SCT for finding species seeds. The SUS sampling is based on $f(a_i) / \sum_{j=1}^{c} f(a_j)$ to produce the mean number of copies of an elitist in the mating pool, where c is number of elitist. Actually, the niching GA by Lee et al. (1999)adopted the restricted competition selection (RCS) to select elitist to be conserved directly in the next generation, which was almost identical to the clearing procedure with elitist. Therefore, the CPE is selected to make comparative investigation, as shown in Fig. 7.

In the elitist conservation procedure, an elitist in P' replaces the worst individual in P'' if it does not have a copy in the offspring population. The CPE scheme resembles

```
procedure Clearing(P ,σ ,k)
    SortFitness(P );
    for i=0 to N-1 do
        if  f(aᵢ) > 0  then
            Num_winners: =1;
            for j=i+1 to N-1 do
                if  f(aⱼ) > 0  and  d(aᵢ,aⱼ) < σ  then
                    if Num_winners<k then
                        Num_winners: = Num_winners+1;
                    else
                        f(aⱼ) :=0;
                    end if
                end if
            end for
        end if
    end for
    return {P }
end procedure
```

Fig. 6 The basic algorithm for clearing procedure

```
procedure CPE (P_offspring(t) ,P(t) )
    set:  t := 0 ;
    initialize:  P(0) ;
    while stopping_criteria_not_met do
        find elitist  P ′ by Clearing(P(t) ,σ ,1);
        generate offspring population by SUS selection, crossover, mutation:
            P″ := S ∘ C ∘ M (P ′);
        compute the fitness of  P″ ;
        produce the new population by elitist conservation:
            P(t+1) := elitist_ preservation(P ',P″);
        set:  t := t + 1 ;
    end while
    return {P(t) }
end procedure
```

Fig. 7 The basic procedure for CPE

greatly the SCT for the purpose of niching multiple species, and the elitist or species seeds are found by specifying radius of niches. But their conservation operations in the paradigm of the generational GAs are quite different. The CPE exerts mandatory conservation by elitism, while the SCT may fail to conserve some points if there remain no unmarked individuals in current population. Thus, once there an individual falling into the attraction basin of a niche during either initialization or reproduction processes, the niche will be conserved if the clearing radius is smaller than half of the distance between niches.

6 Experiments and discussions

In this section, experiments are conducted to evaluate the NNRC, the SCT, the HFC-I, and the CPE on a suite of typical real-valued multimodal functions. Due to the stochastic nature of algorithms, the results are obtained by averaging the experimental values of 30 independent executions of niching methods with specific parameterization unless otherwise indicated.

In order to evaluate the four niching methods based on the overlapping population replacement paradigm, the steady-state GAs (SSGA) or generational GAs (GGA) with constant population sizes is employed. Here, standard genetic operators are used unless otherwise indicated. Since the NNRC works in the context of both the SSGA and the GGA, the corresponding algorithms are denoted as NNRC SSGA (NNRC-SSGA) and NNRC GGA (NNRC-GGA). Originally, the SCT only works in generation with constant population size, and specific procedure has to be designed to make it fitted to the cases where the offspring population size is different from parental population. Thus, the SCT is only embedded in the GGA, and has two cases, the original SCGA as proposed by inventors (Li et al. 2002), and the SCT GGA (SCT-GGA) that takes the form of generational GAs with standard crossover and mutation operators. There is only one form of the HFC-I that works in the steady-state GAs (HFC-I-SSGA). The CPE runs in generational GAs (CPE-GGA). Main parameters and settings are specified in Table 1.

The population size is indicated as it is used in specific context in experiments. Usually, the population size N is denoted as *pop_size* in figures. If there are many niches that are to be located on the landscape of a MMFO problem, the population should be sized as bigger than the number of local optima. Suppose that the number of local optima is *num_local_optima*, we should take the population size as *pop_size* > 2*num_local_optima*.

There are two cases for setting q in the NNRC and HFC-I, $q = 1$, or $q_{max} = pop\_size \times \eta$, where $1 > \eta > 0$. By considering the *num_local_optima*, q_{max} is essentially set as $q_{max} < \lfloor pop\_size/num\_local\_optima \rfloor$. Actually, there are many feasible values for q that are able to promote the performance of the NNRC and the HFC-I for specific MMFO problems. The settings of q_{max} in experiments are only rule-of-thumb values when the NNRC and HFC-I are used to solve MMFO problems. The maximum fitness (function) evaluations (FEs) are specified in experiments as the stopping criterion of algorithms.

6.1 Test functions

1.   Univariate monotone function (UM)

The univariate monotone function is defined as:

$$\max f(x) = x, \quad x \in [0, 1] \tag{7}$$

which is adopted to illustrate the difference of niching schemes regarding niching dynamics about species conservation.

Table 1 Main parameters and settings for niching methods

Niching methods	Parameters and settings		
	Selection	Crossover	Mutation
NNRC-SSGA	Random selection	Blend crossover[a], or two-point crossover, rate: $0.6 \sim 1.0$	Gaussian$(0,0.1)$, rate: $1/2n$, or 0.1
NNRC-GGA	Fitness proportionate selection		
SCGA[b]		$\boldsymbol{a}' = \boldsymbol{a}_1 + u(\boldsymbol{a}_2 - \boldsymbol{a}_1)$,	$\boldsymbol{a}' = \boldsymbol{a}' + r_m \mathbf{R}(\boldsymbol{a}^u - \boldsymbol{a}^l)$
		$\boldsymbol{u} = \mathrm{rand}(0,1)$, rate: $0.6 \sim 1.0$	$\mathbf{R} = \mathrm{rand}(0,1), r_m = 0.15$, $\boldsymbol{a}^u, \boldsymbol{a}^l$ are upper and lower bounds vector of $\boldsymbol{a}$, rate: 0.05
SCT-GGA		Blend crossover, or two-point crossover, rate: $0.6 \sim 1.0$	Gaussian$(0,0.1)$, rate: $1/2n$, or 0.1
HFC-I-SSGA	Hierarchical random selection		
CPE-GGA	Clearing and SUS selection		

[a] BLX-alpha crossover (0.5) (Eshelman and Shaffer, 1993; Herrera et al. 2003)

[b] Parameters are set as the same in literature (Li et al. 2002)

2. Twin Gaussian bells function (TGB)

A bimodal function is defined by combining two Gaussian functions in a variety of ways as:

$$\max f(\boldsymbol{x}) = \max \left\{ \exp\left(-\frac{||\mathbf{x} - \boldsymbol{\mu}_1||^2}{\sigma_1^2} \right), \alpha \exp\left(-\frac{||\boldsymbol{x} - \boldsymbol{\mu}_2||^2}{\sigma_2^2} \right) \right\},$$
$$x_1, x_2 \in [0, 12] \tag{8}$$

where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ and σ_1, σ_2 are the centroid and width parameters respectively that adjust the overlapping of two niches; α changes the comparative heights of two peaks. The two Gaussian bells are called niche A $(\boldsymbol{\mu}_1, \sigma_1)$ and niche B $(\boldsymbol{\mu}_2, \sigma_2)$. The function TGB is similar to the Gaussian landscape generator developed by Yuan and Gallagher (2003) which provides a good platform for testing niching GAs on multimodal optimization problems with randomly spaced Gaussian niches. The TGB is designed to test the species conservation mechanism of niching methods by associating individual distribution among species at equilibrium with niches' height and size of attraction basins.

3. Schaffer function

The Schaffer function is a symmetric multimodal function with a multimodality of ring ranges of different heights. A modified version is as follows:

$$\min f(x_1, x_2) = 0.5 + \frac{\sin^2 \sqrt{x_1^2 + x_2^2} - 0.5}{[1 + \alpha \times (x_1^2 + x_2^2)]^2},$$
$$x_1, x_2 \in [-10, 10] \tag{9}$$

where α is used to regulate the shape of the function, and we take $\alpha = 0.01$ in experiment. The niches consist of a set of torus. The origin (0,0) is the global minimum with $f(0,0) = 0$ and denoted as $niche_0$. The four toruses that locate at $\sqrt{x_1^2 + x_2^2} = \{\pi, 2\pi, 3\pi, 4\pi\}$ are the local minima denoted as $\{niche_\pi, niche_{2\pi}, niche_{3\pi}, niche_{4\pi}\}$. The origin (0,0) and four toruses constitute the five ring niches of the Schaffer function.

4. Two-dimensional Shubert function (2-d Shubert)

The two-dimensional Shubert function is defined by:

$$\min f(x_1, x_2) = \prod_{i=2}^{2} \left(\sum_{j=1}^{5} j \cos[(j+1)x_i + j] \right),$$
$$x_1, x_2 \in [-10, 10] \tag{10}$$

which has 760 local minima, and 18 of them are global minima with an objective function value of -186.73 approximately. These 18 global minima are grouped into 9 clusters that are evenly spaced with coordinate distance of 5.65. Each group has two neighboring global optima with a distance of 0.98. The Shubert function has been used commonly to test the performance of niching methods, including the SCGA (Li et al. 2002), the multi-species particle swarm optimizer (MSPSO) (Iwamatsu 2006). Empirically, the species distance threshold has a big impact on niching results.

5. Multivariate real-valued deceptive function with independency (MRDFi)

$$\max f_{\mathrm{MRDFi}}(\boldsymbol{X}) = \sum_{i=1}^{n} f_{\mathrm{URDF}}(x_i),$$
$$x_i \in [0, 1], \ i = 1, 2, \ldots, n, \tag{11}$$
$$f_{\mathrm{URDF}}(x_i) = \begin{cases} \dfrac{a(c - x_i)}{c} & 0 \le x_i < c \\ \dfrac{b(x_i - c)}{1 - c} & c \le x_i \le 1 \end{cases},$$

where $\{a, b, c\}$ are parameters to control the shape of the univariate real-valued deceptive function (URDF, see Fig. 8), and $a < b, 0.5 < c < 1$. These parameters take
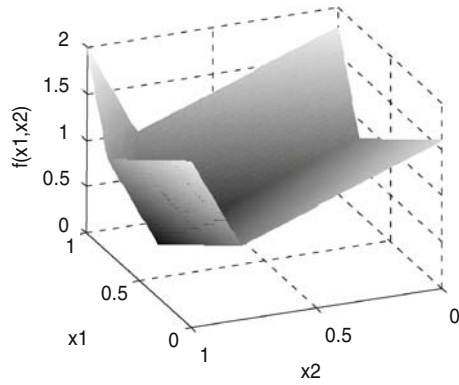
Fig. 8 2-d MRDFi function

Fig. 9 2-d r-Ising model

default values as $\{a = 0.8, b = 1.0, c = 0.8\}$ in experiments. The global maximum for the $f_{\text{URDF}}(x_i)$ is $x_i^* = 1, f(x_i^*) = b$, and the deceptive attractor is $x_{i,\text{deceptive\_attractor}} = 0, f(x_{i,\text{deceptive\_attractor}}) = a$. The global maximum for the $f_{\text{MRDFi}}(\mathbf{X})$ is $\mathbf{X}^* = (1, 1, \ldots, 1), f(\mathbf{X}^*) = nb$, and the globally deceptive attractor is $\mathbf{X}_{\text{deceptive\_attractor}} = (0, 0, \ldots, 0), f(\mathbf{X}_{\text{deceptive\_attractor}}) = na$. There are totally 2^n local optima with the n-d MRDFi, and $2^n - 2$ of them are partially deceptive attractors.

The MRDFi has a rugged fitness landscape, and there is only one global maximum locating at the boundary of the solution space. The global maximum consists of n number of one-dimensional real-valued building blocks, and is characterized as a real-valued needle-in-a-haystack problem. It is extremely difficult for the SGA. In addition, recombination operation is needed definitely to make conjunction of building blocks.

6. Real-valued Ising model (r-Ising)

The real-valued Ising model is defined as:

$$\max f_{r-\text{Ising}}(\mathbf{X}) = \sum_{i=1}^{n} f_{\text{URDF}}(x_i) \times [1 - (x_i - x_{i+1})^2],$$
$$x_i \in [0, 1], \ i = 1, 2, \ldots, n, \tag{12}$$

where, $x_{n+1} = x_1$. The f_{URDF} is identical as in (11), but has parameters set as $\{a = b = 1.0, c = 0.5\}$. The f_{URDF} has two global maxima as $x_0^* = 0, x_1^* = 1, f(x_0^*) = f(x_1^*) = 1$, that have equal attraction basins. The $f_{r-\text{Ising}}(\mathbf{X})$ has two global maxima, $\mathbf{X}_0^* = (0, 0, \ldots, 0), \mathbf{X}_1^* = (1, 1, \ldots, 1)$, and $f(\mathbf{X}_0^*) = f(\mathbf{X}_1^*) = n$.

The r-Ising model is designed by following the binary Ising model (Van Hoyweghen et al. 2002). There exists nonlinear correlation between neighborhood real-valued variables, and a real-valued chromosome is characterized with pinning effects and epistasis on the rugged landscape (see Fig. 9). These features make a simple GA get stuck in a local optimum far away from the global optima. Empirically, the non-inferior building blocks is necessary to find
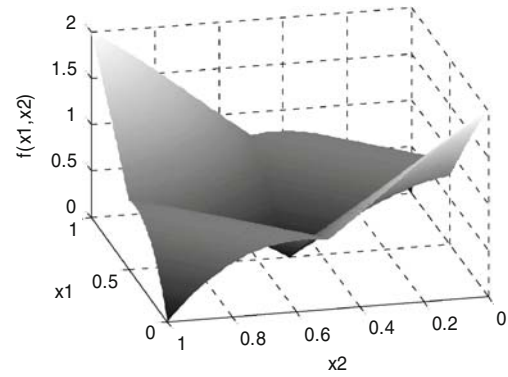
optima for the binary Ising model, which could be only conserved by using niching techniques (Van Hoyweghen et al. 2002; Li and Kou 2008). This is also true for the r-Ising model. Additionally, we do not adopt the f_{URDF} with unequal size of attraction basins for the two global optima because it will make the r-Ising model extremely difficult for niching techniques.

6.2 Experimental results

Experiments are conducted on test functions, and results are evaluated in four aspects.

1. Niching convergence and equilibrium[1]

The niching GAs tends to allocate individuals to niches in the style of species, and the distribution of individuals among species gets to an equilibrium which is kept forever in the niching paradigm of the overlapping population replacement. The experimental results on the UM function is shown in Fig. 10.

The initial populations are identical for all algorithms, and the individual distribution is drawn in Fig. 10a. There is little difference between the context of SSGA and GGA for the NNRC with $q = 1$ or $q_{\max} = 5$ in terms of population convergence on the UM function, but the $q_{\max} = 5$ makes the exploitation of niches much faster with the dynamic NNRC (see Fig. 10b–e, p, q).

The performance of the SCT varies saliently with different species distance regarding both the SCGA and the

---

[1] The niching convergence and equilibrium is measured by population entropy. First, the real-valued variables are discretized into 32 intervals or 32×32 grids in the 1-d or 2-d definition domain of a function. Second, the population entropy is computed by $E_{\text{pop}} = -\sum_{i=0}^{K-1} \sum_{j=0}^{K-1} p_{ij} \log p_{ij}$, where p_{ij} is the approximate proportion of individuals in the $i \times j$ grid, $i, j = 1, 2, \ldots, K$, $K = 31$. If the relative decrease rate of the population entropy (smoothed for algorithms with fitness proportional selection) per 1000 fitness evaluations was smaller than the threshold (5%), the algorithm is taken as reaching niching convergence and equilibrium.

SCT-GGA (see Fig. 10f–i, r, s). When $\sigma_s = 0.1$ is used, the population does not get converged at equilibrium. The species seeds set contains low fitness individuals that are not local optima or peak points of niches (Fig. 10f, h). This is true even if a bigger population size is employed in experiment. The converged population is obtained only when species distance σ_s takes bigger values (Fig. 10g, i). These phenomena are also observed on the TGB function regarding the SCGA and SCT-GGA (see Fig. 11).

The HFC-I with small $\{\gamma, q\}$ ($\gamma = 0.1, q = 1$) is able to keep the diversity of population (see Fig. 10j, t), but individuals in the final population are not converged, which is quite like the SCT-GGA with small species distance. This illustrates the effect of the fitness crowding that is done in the implicit hierarchical fair competition between individuals. However, the HFC-I with bigger $\{\gamma, q\}$ ($\gamma = 0.5, q_{max} = 5$) yields converged population much faster by promoting the competition of individuals in replacement (see Fig. 10k, t).

The CPE evolves the population just like the fitness sharing, and the convergence is not obtained in the final population (see Fig. 10l, u) if the clearing distance is smaller than the niche radius. The clearing procedure and elitist strategy tends to retain individuals with low fitness by using small clearing radius. The final population gets converged asymptotically when the clearing radius is equal to or bigger than the niche radius (see Fig. 10o, u).

The TGB function is parameterized with two equal-height niches that have equal attraction basin size. The NNRC is able to evolve the population into the two species adapting separately to two niches, and the equilibrium is sustained (see Fig. 11b, c, m, n). The SCT fails to obtain the converged populations with a small species distance (see Fig. 11d, e, f, p, q), and a big species distance $(\sigma_s > 8\sqrt{2})$ results in the disability of the niching method to evolve two species in the population. This feature is to be investigated later.

The HFC-I is able to find both global optima, but the final population is not converged because of the restricted selection and replacement based on implicit fitness level (see Fig. 11g, r). However, the bigger γ promotes the convergence by increasing the competition of individuals in replacement, and there are fewer individuals with lower fitness in the final population (see Fig. 11h, r). With the CPE, the population evolves in a dispersed state without convergence by the clearing and elitist strategy (see Fig. 11i–k, s). If the clear radius if bigger than $4\sqrt{2}$, there is only one niche that has individuals converged to the peak point (see Fig. 11l).

By niching equilibrium, we mean the constant individual distribution among species adapting to all niches. The fitness sharing scheme partitions individuals into species in accordance to the peak fitness of niches. The NNRC tends to evolve species based on the size of attractive basins with random selection and global crossover. Since the SCT makes use of an explicit niching based on species seeds identification, the niching equilibrium is quite different. The HFC-I yields results similar to the NNRC, while the CPE reaches niching equilibrium just like the fitness sharing. Statistical results are computed regarding different niching methods on the TGB function as shown in Table 2, where three cases of the TGB function are considered in experiments with different heights and attraction basin sizes of the niche A and B.

The NNRC scheme, in either the context of the SSGA or the GGA, outputs species with nearly equal sizes at equilibrium in case (a), where the replacement error is not significant statistically. The niche A has a smaller height than the niche B in case (b), and less individuals adapt to niche A as compared to the ideal attraction basin size ratio between A and B because of replacement error occurred in the region cross the boundary of the two niches. The GGA yields bigger error than the SSGA. When the niche A has a smaller attraction basin size than the niche B but with equal height as in case (c), there are more individuals in species A as compared to the ideal attraction basin size ratio due to better local competitiveness of the niche A in the region cross the boundary of the two niches. Therefore, the NNRC scheme is mainly regarded as niching species according the attraction basin sizes of the niches.

The SCT method shows quite a different niching mechanism. There is one species that has only a few individuals in either species A or species B in case (a). The ratio of individual distribution among the two niches is biased significantly from the ideal ones for both the 2-d and 5-d TGB. This is caused by the fitness-based proportionate selection in the SCGA. Although both species adapting to niches A and B are equally competitive in the initial evolution stages in case (a), the selection error surely happens just like the SGA, and one species will ultimately dominate the population. When the species distance threshold is smaller than the real distance between two niche peaks, the SCT is able to preserve a few individuals in either the species A or species B in a mandatory way. Although neither species can take over the population, the individual distribution among species at population equilibrium is determined randomly in case (a). In case (b) and (c), the niche A has a smaller height or a smaller attraction basin size, it is not competitive against the niche B, and the species B will surely dominate the population. The SCT-GGA performs similarly with the SCGA. Since the niches distance 2.0 is much smaller, nearly all individuals are found as species seeds with a population size 200 on the 5-d TGB. Thus, the algorithm is disabled to make evolution. This issue can be addressed by using bigger population sizes.
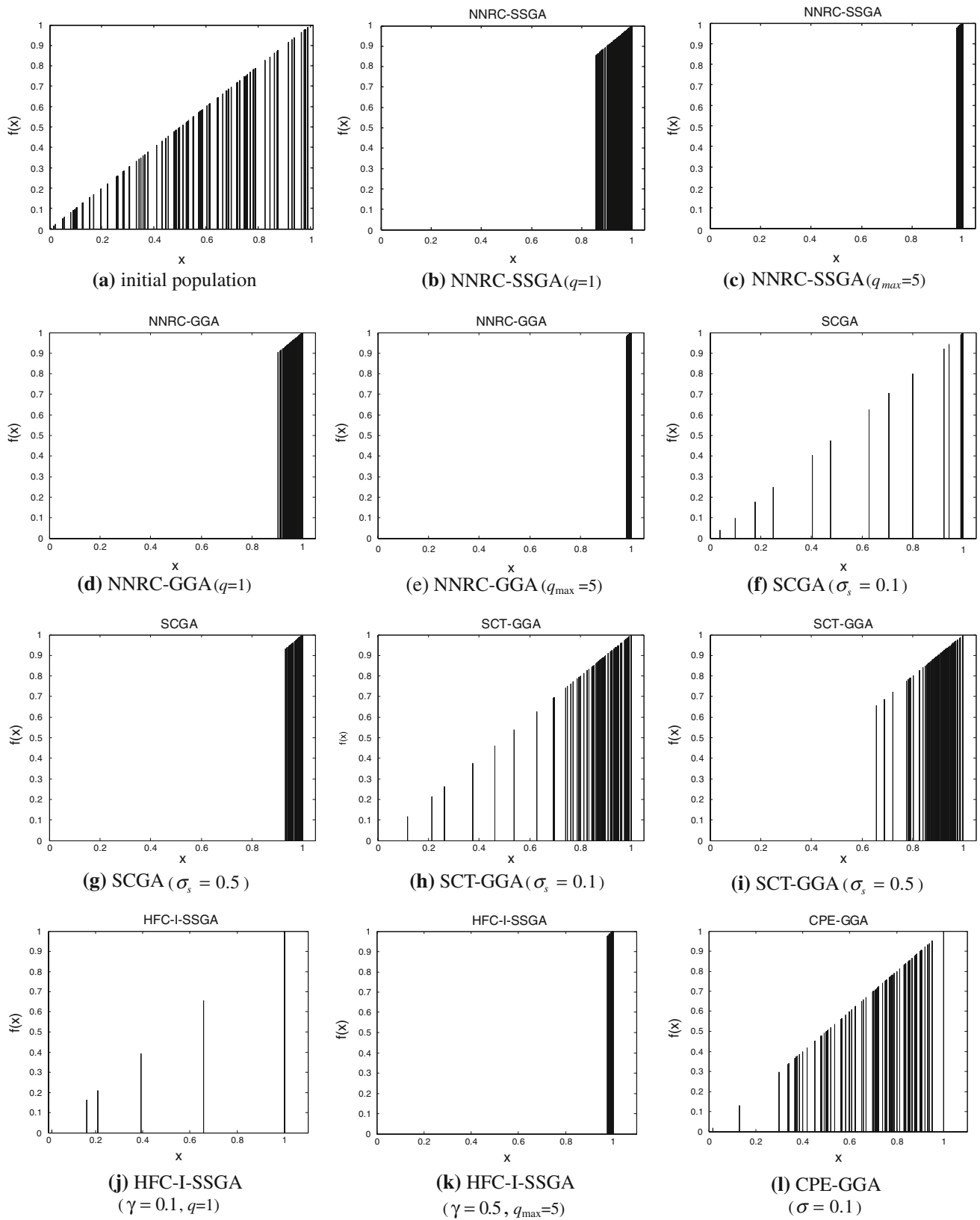
(a) initial population

(b) NNRC-SSGA(q=1)

(c) NNRC-SSGA(q_{max}=5)

(d) NNRC-GGA(q=1)

(e) NNRC-GGA(q_{max} =5)

(f) SCGA($\sigma_s = 0.1$)

(g) SCGA($\sigma_s = 0.5$)

(h) SCT-GGA($\sigma_s = 0.1$)

(i) SCT-GGA($\sigma_s = 0.5$)

(j) HFC-I-SSGA
($\gamma = 0.1$, q=1)

(k) HFC-I-SSGA
($\gamma = 0.5$, q_{max}=5)

(l) CPE-GGA
($\sigma = 0.1$)

Fig. 10 Population convergence with niching methods on UM function (*pop_size* = 100, maximum FEs: 50,000)

(m) CPE-GGA ($\sigma = 0.5$)

(n) CPE-GGA ($\sigma = 0.9$)

(o) CPE-GGA ($\sigma = 1.0$)

(p) NNRC-SSGA

(q) NNRC-GGA

(r) SCGA

(s) SCT-GGA

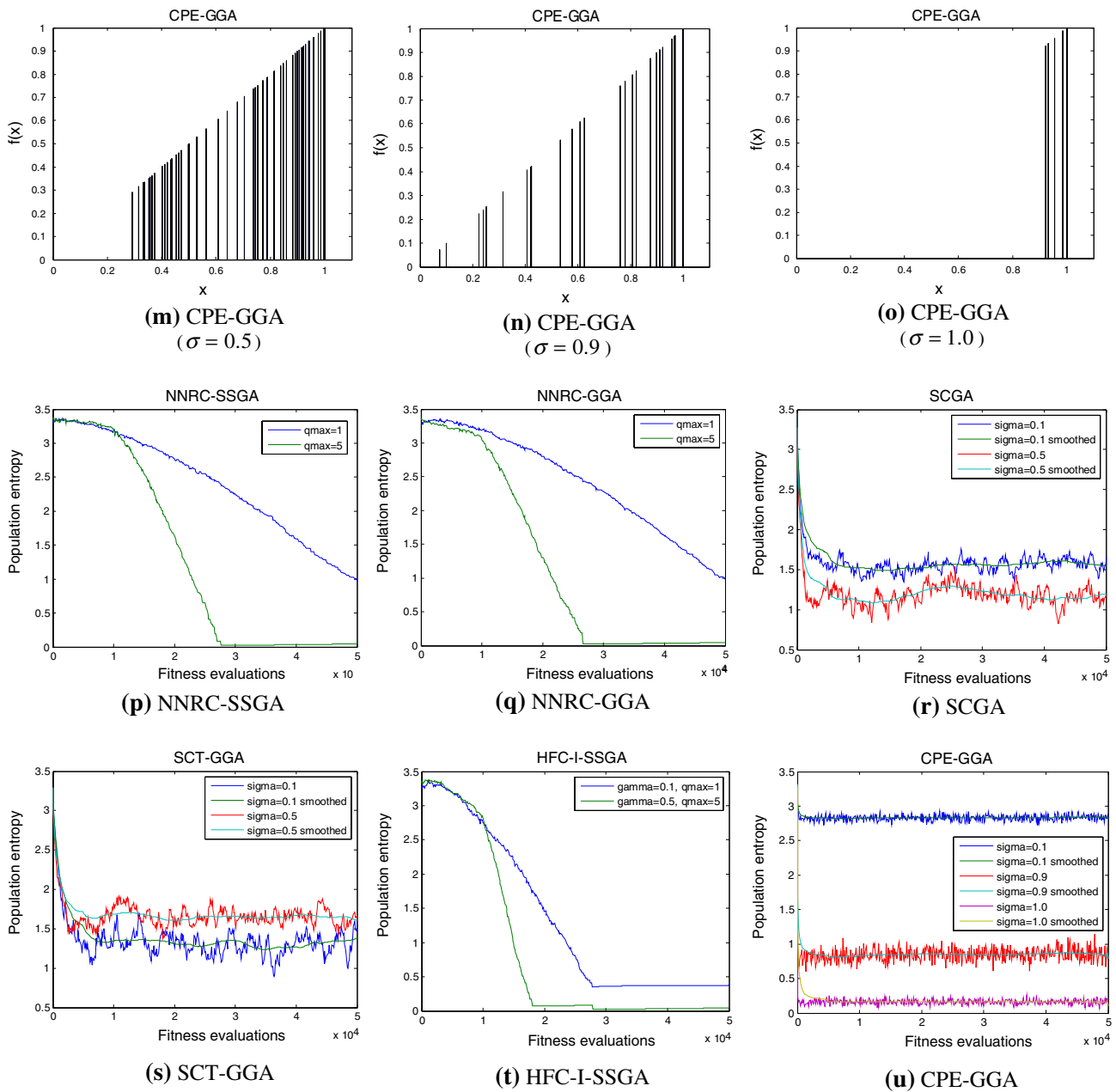**(t)** HFC-I-SSGA

(u) CPE-GGA

Fig. 10 continued

The HFC-I shows similar behavior as the NNRC on the TGB in three cases, but the number of individuals converging to peaks of niches is smaller because there are a few individuals that locate away from the local optima in the final populations. Similar results are obtained by the CPE. Although all niches are found with a good distribution of individuals in case (a), but it fails to get to convergence. Since the niche A of the TGB in case (c) has a much smaller attraction basin with the 5-d TGB, it can not be found definitely in a run of the algorithm. With bigger population sizes, there will be a greater probability to conserve both niches.

2. Niching convergence on non-Gaussian niches

The ring basins with infinite local optima on the Schaffer function constitute non-Gaussian niches. Ideally, the found optima should scatter along the ring basins evenly so as to reveal the true feature of the local multimodality. This is what the NNRC achieves in the context of either SSGA or GGA (see Fig. 12b, c). The population
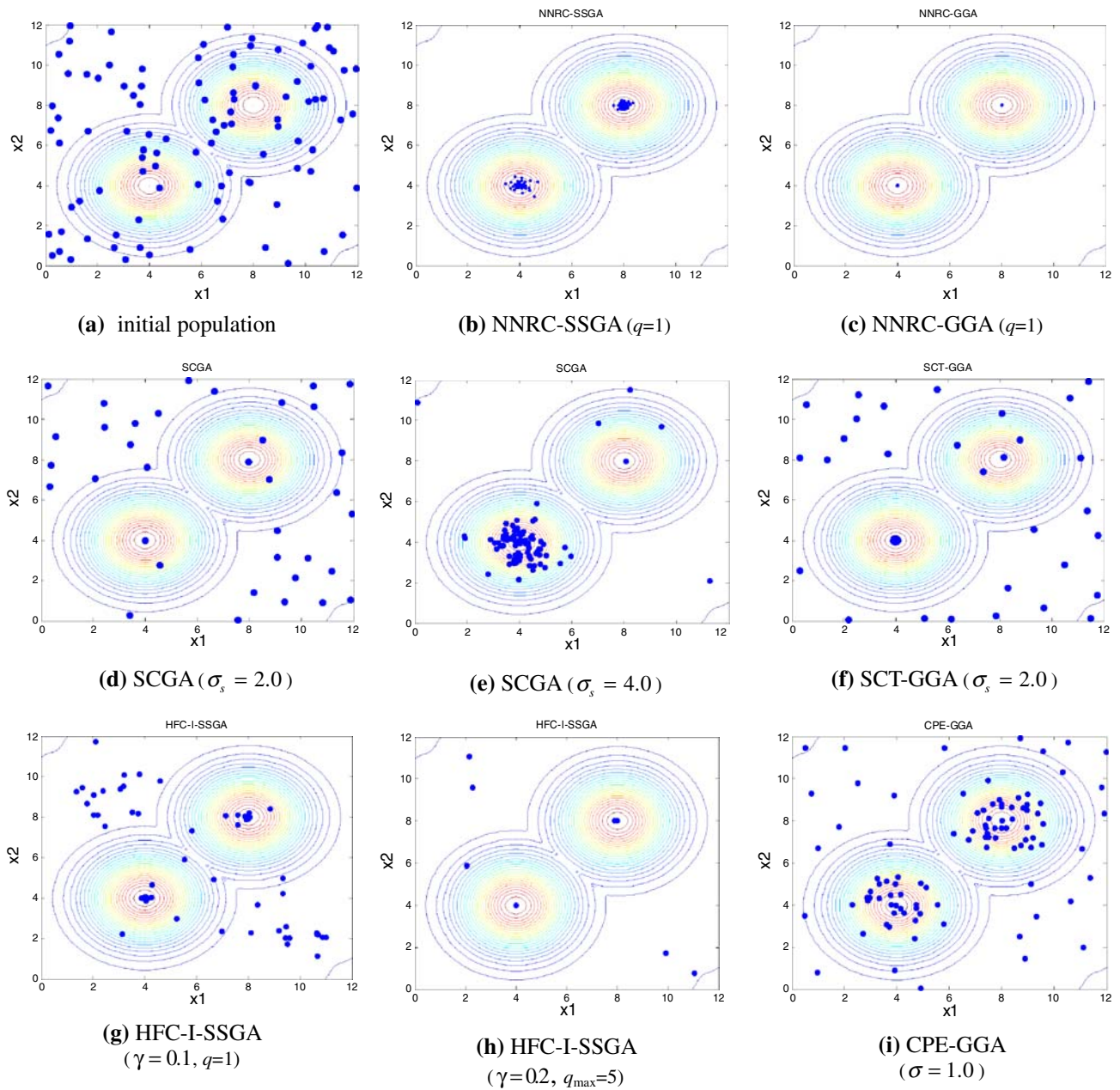
Fig. 11 Population convergence with niching methods on TGB function ($\boldsymbol{\mu}_1 = (4, 4), \boldsymbol{\mu}_2 = (8, 8), \sigma_1 = 2, \sigma_2 = 2, \alpha = 1.0$; pop_size = 100, maximum FEs: 10,000)

convergence depends on the species distance with the SCT, where the optima on the same ring basins are partitioned at least by half of the species distance (see Fig. 12d, e, f), and the final population at convergence is unable to characterize the local multimodality of ring basins when big species distance is used.

The ring basins also constitute a difficulty for the HFC-I. With smaller $\{\gamma, q\}(\gamma = 0.1, q = 1)$, the diversity of population is preserved quite well, but the ring local optima are not found (see Fig. 12g). This is due to the fact

that the recombination operation of individuals with similar fitness but very different phenotypes fails to produce offspring within the attraction basins of local optima. There are not building blocks shared by the optima of different fitness. Thus, the fitness crowding with hierarchical fitness levels confines the global searching capacity of the algorithm. When the hierarchical fitness crowding is loosen ($\gamma = 0.5$), the significant improvement is achieved in terms of individual convergence to ring local optima (see Fig. 12h).
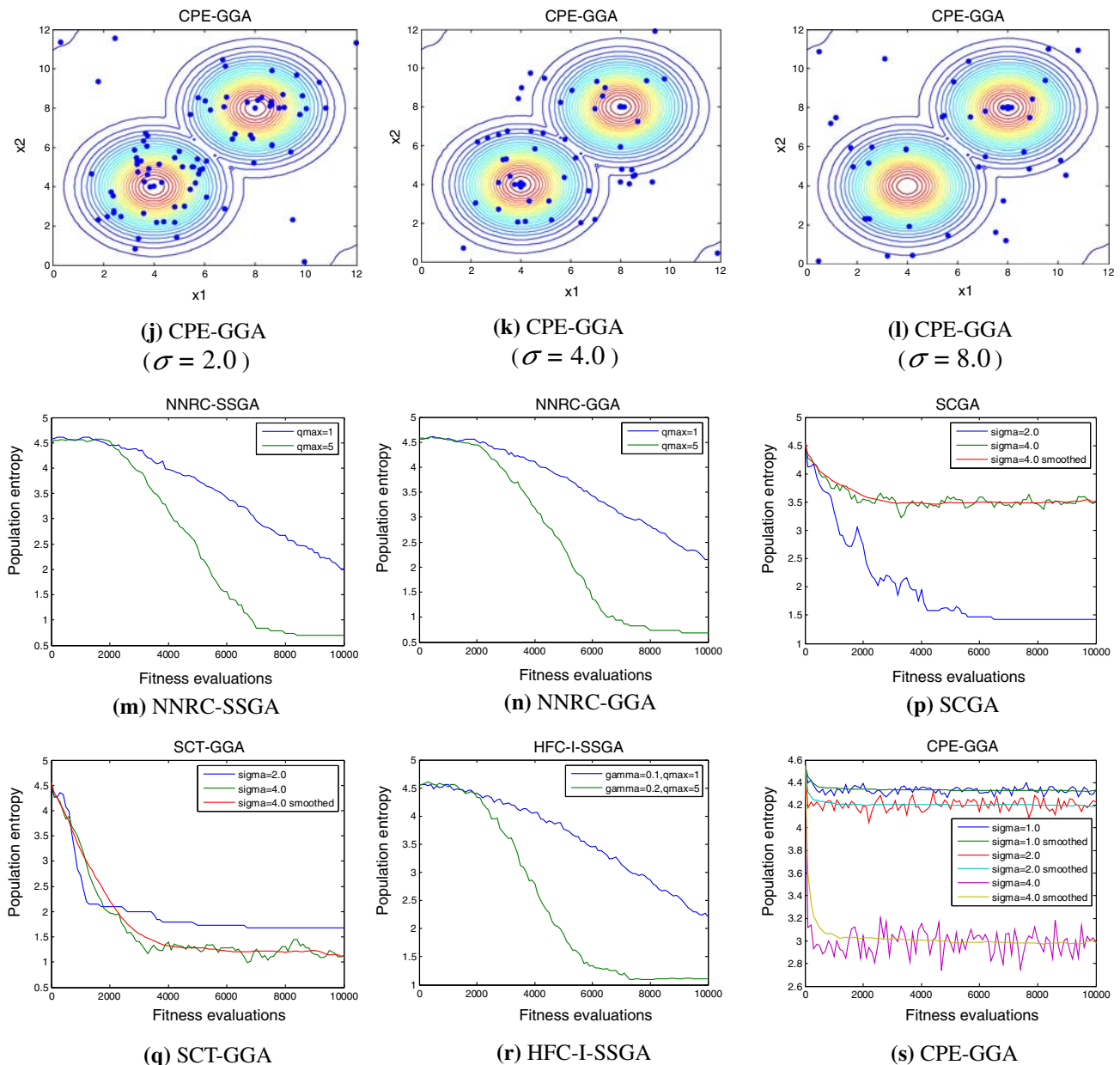
(j) CPE-GGA ($\sigma = 2.0$)

(k) CPE-GGA ($\sigma = 4.0$)

(l) CPE-GGA ($\sigma = 8.0$)

(m) NNRC-SSGA

(n) NNRC-GGA

(p) SCGA

(q) SCT-GGA

(r) HFC-I-SSGA

(s) CPE-GGA

Fig. 11 continued

The CPE finds and maintains successfully all ring local optima (see Fig. 12i–k). The elitist group with clearing radius 1.0 is shown in Fig. 12l, where the winners (or elitists) on the same ring are separated by at least clearing radius just like the distribution of seeds with the SCT. It illustrates the good niching capability with the clearing procedure and elitist strategy.

3. Niching efficiency

The mean number of fitness (function) evaluations (MNFEs) is a nice measure of algorithm efficiency in the field of the GAs and EC. The standard deviation (STD) of

the number of fitness evaluations is computed to indicate the robustness of the algorithms. With the Shubert function, a niche is found if an individual adapts to the niche and gets at least into the neighborhood of the optimum of the niche with a radius of 0.05 (solution precision). Experimental results are reported in Table 3, where all of 18 niches with global optima are found (the success rate of 100% is assumed).

Table 3 shows that the NNRC needs less MNFEs than the SCT for finding all global optima of the Shubert function within specific solution precision, and the difference is significant by the t-statistic with a confidence of

Table 2 Niching equilibrium (mean individual distribution (%)) on TGB function

Niching methods	2-d			5-d		
	Pop_size = 100, maximum FEs: 10,000			*Pop_size* = 200, maximum FEs: 100,000		
	Niche A	Niche B	Difference	Niche A	Niche B	Difference
Case (a): $\mu_1 = (4,4), \mu_2 = (8,8), \sigma_1 = 2, \sigma_2 = 2, \alpha = 1.0$; attraction basin size A:B = 50:50 for both 2-d and 5-d						
NNRC-SSGA ($q = 1$)	50.23	49.77	0.46	49.21	50.79	1.57
NNRC-SSGA ($q_{max} = 5$)	49.07	50.83	1.76	51.10	48.90	2.20
NNRC-GGA ($q = 1$)	50.97	49.03	1.94	48.55	51.45	2.90
NNRC-GGA ($q_{max} = 5$)	48.83	51.17	2.34	51.58	48.42	3.17
SCGA ($\sigma_s = 2.0$)	9.25 or 90.67	9.33 or 90.75	–	–	–	–
SCGA ($\sigma_s = 8.0$)	2.28 or 97.33	2.67 or 97.63	–	19.05 or 81.21	18.79 or 80.95	–
HFC-I-SSGA ($\gamma = 0.1, q = 1$)	51.25	48.75	2.50	50.67	49.33	1.33
HFC-I –SSGA ($\gamma = 0.2, q_{max} = 5$)	48.17	51.83	3.67	51.67	48.33	3.33
CFE-GGA ($\sigma = n$)	51.37	48.63	2.74	48.45	51.55	3.10
Case (b): $\mu_1 = (4,4), \mu_2 = (8,8), \sigma_1 = 2, \sigma_2 = 2, \alpha = 2.0$; attraction basin size A:B $\approx$ 46.8:53.2 for 2-d, 48.5:51.5 for 5-d						
NNRC-SSGA ($q = 1$)	45.37	54.63	9.26	45.42	54.58	9.17
NNRC-SSGA ($q_{max} = 5$)	43.00	57.00	14.00	44.25	55.75	11.50
NNRC-GGA ($q = 1$)	37.83	67.17	29.33	37.00	63.00	26.00
NNRC-GGA ($q_{max} = 5$)	37.00	63.00	26.00	34.83	65.17	30.33
SCGA ($\sigma_s = 2.0$)	14.60	85.40	70.80	–	–	–
SCGA ($\sigma_s = 8.0$)	1.10	98.90	97.80	17.60	82.40	64.80
HFC-I-SSGA ($\gamma = 0.1, q = 1$)	37.17	62.83	25.67	40.42	59.58	19.17
HFC-I-SSGA ($\gamma = 0.2, q_{max} = 5$)	26.50	73.50	47.00	35.17	64.83	29.67
CFE-GGA ($\sigma = n$)	27.84	72.16	44.32	30.25	69.75	39.50
Case (c): $\mu_1 = (4,4), \mu_2 = (8,8), \sigma_1 = 1, \sigma_2 = 3, \alpha = 1.0$, attraction basin size A:B $\approx$ 10.1:89.9 for 2-d, 0.8:99.2 for 5-d						
NNRC-SSGA ($q = 1$)	17.50	82.50	65.00	1.75	98.25	96.50
NNRC-SSGA ($q_{max} = 5$)	18.50	81.50	63.00	1.38	98.63	97.25
NNRC-GGA ($q = 1$)	17.67	82.33	64.67	0.92	99.08	98.17
NNRC-GGA ($q_{max} = 5$)	17.83	82.17	64.33	0.75	99.25	98.50
SCGA ($\sigma_s = 2.0$)	2.20	97.80	95.60	–	–	–
SCGA ($\sigma_s = 8.0$)	1.20	98.80	97.60	0.60	99.40	98.80
HFC-I –SSGA ($\gamma = 0.1, q = 1$)	13.17	86.83	73.67	4.83	95.17	90.33
HFC-I –SSGA ($\gamma = 0.2, q_{max} = 5$)	12.83	87.17	74.33	4.08	95.92	91.83
CFE-GGA ($\sigma = n$)	16.40	83.60	67.20	0.25	99.75	99.50

Attraction basin size A:B ratio is calculated approximately by randomly generating points in the solution space

95%. The GGA makes a better paradigm than the SSGA in terms of MNFEs on the Shubert function regarding the NNRC scheme. This is partly due to the fitness based proportionate selection which promotes the production of competitive individuals with bigger fitness. The HFC-I performs also efficiently in terms of MNFEs because of the hierarchical recombination. This is due to the fact that the all of the global optima share real-value building blocks, and the recombination of individuals with bigger fitness helps find new global optima. The performance of the CPE is close to the HFC-I and NNRC-SSGA by using the clearing procedure and elitist preservation. In addition, the STD values of the NNRC, HFC-I, and CPE are much smaller than that of the SCT, which shows that the former schemes are more robust than the latter one. The MNFEs of the SCGA is nearly doubled when the population size increases from 200 to 500, and the STDs are quite bigger than other schemes in both cases.

However, these niching methods produce quite different sizes of species adapting to different niches at equilibrium. The NNRC and the HFC-I yield equal-size species that have nearly equal number of individuals in species. The species sizes are close to q_{max}. In contrast, the SCT and CPE tend to output species with variant numbers of individuals. There are usually more than one copies of local optimum in a few niches, and there is only one individual
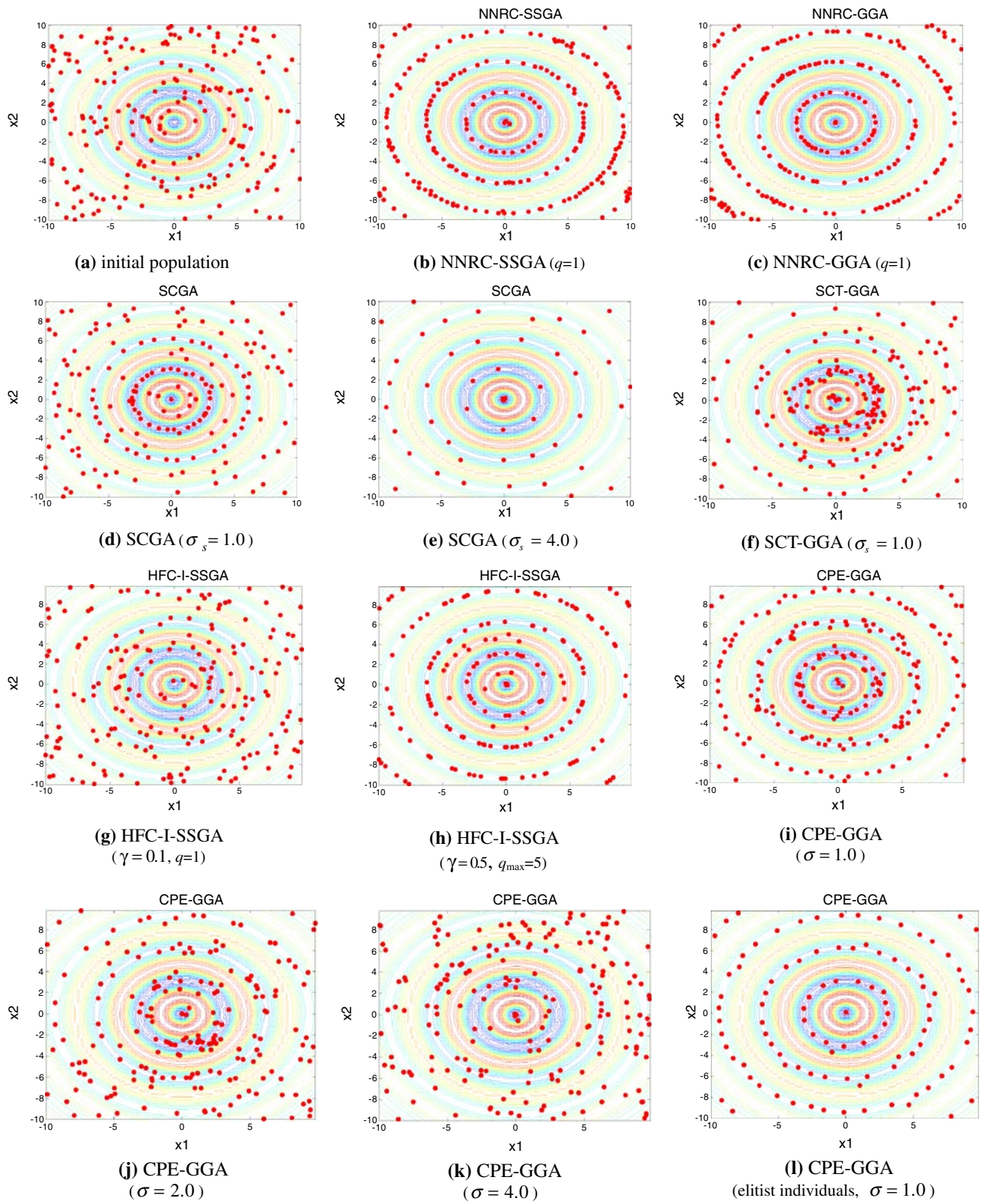
(a) initial population

(b) NNRC-SSGA ($q=1$)

(c) NNRC-GGA ($q=1$)

(d) SCGA ($\sigma_s = 1.0$)

(e) SCGA ($\sigma_s = 4.0$)

(f) SCT-GGA ($\sigma_s = 1.0$)

(g) HFC-I-SSGA ($\gamma = 0.1$, $q=1$)

(h) HFC-I-SSGA ($\gamma = 0.5$, $q_{max}=5$)

(i) CPE-GGA ($\sigma = 1.0$)

(j) CPE-GGA ($\sigma = 2.0$)

(k) CPE-GGA ($\sigma = 4.0$)

(l) CPE-GGA (elitist individuals, $\sigma = 1.0$)

Fig. 12 Population convergence with niching methods on Schaffer function (*pop_size* = 200, maximum FEs: 20,000)

in each of other niches. Fig. 13 illustrates the numbers of individuals in species of the 18 niches in the evolution process.

Table 3 Niching efficiency of niching methods on Shubert function

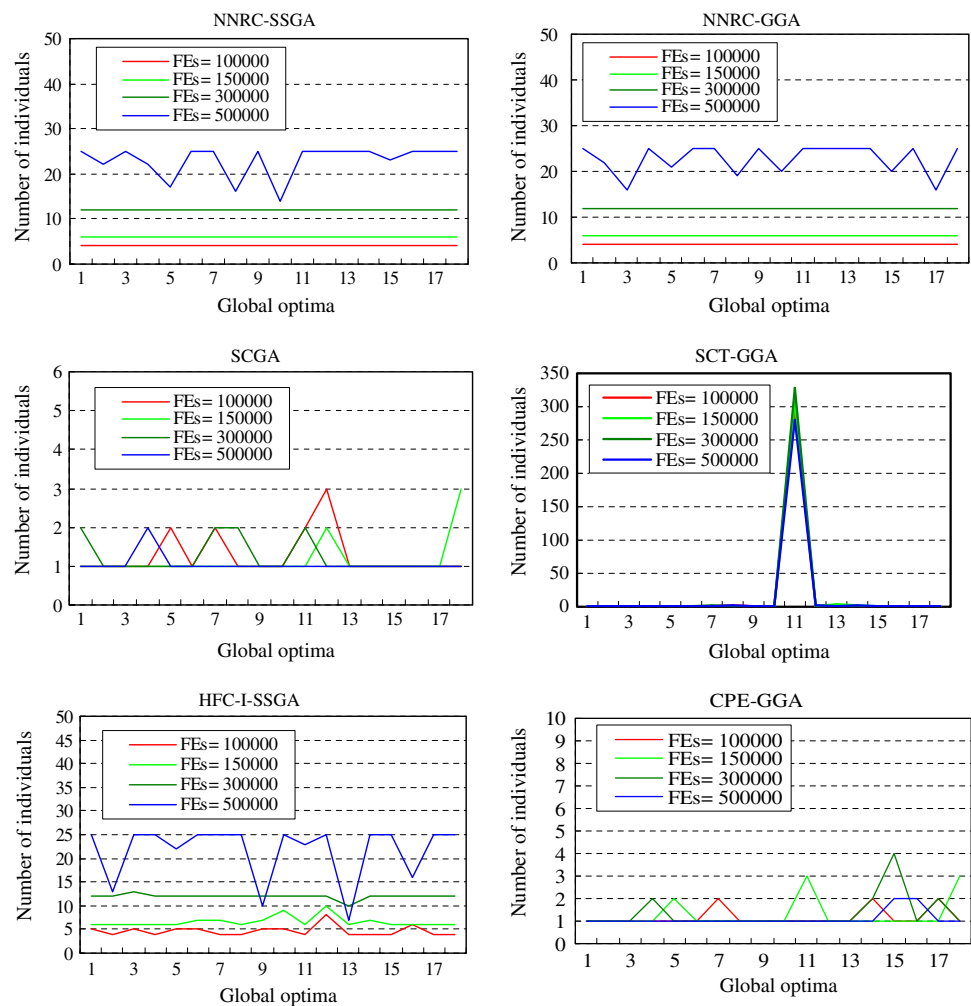| Niching methods	MNFEs	STD
Case (a): *pop_size* = 200, maximum FEs: 100,000		
NNRC-SSGA ($q_{max} = 10$)	34,860	6,177
NNRC-GGA ($q_{max} = 10$)	19,580	2,354
SCGA ($\sigma_s = 1.6$)	54,260	17,532
SCT-GGA ($\sigma_s = 1.6$)	41,440	26,723
HFC-I-SSGA ($\gamma = 0.2, q_{max} = 10$)	30,600	6,512
CPE-GGA ($\sigma = 0.8$)	31,400	8,674
Case (b): *pop_size* = 500, maximum FEs: 500,000		
NNRC-SSGA ($q_{max} = 25$)	43,050	5,500
NNRC-GGA ($q_{max} = 25$)	22,250	3,939
SCGA ($\sigma_s = 1.6$)	90,350	48,261
SCT-GGA ($\sigma_s = 1.6$)	49,800	48,872
HFC-I-SSGA (($\gamma = 0.2, q_{max} = 25$)	37,650	8,975
CPE-GGA ($\sigma = 0.8$)	43,900	9,468

Figure 13 reveals once again the essential difference among niching methods with regard to the niching results at equilibrium. All of the 18 niches have equal height and equal size of attraction basins, there should be an ecological similarity about multiple species adapting to different niches. The NNRC and the HFC-I obtain multiple species with nearly equal number of individuals in species at equilibrium, but the SCT and CPE yield multiple species with varying number of individuals although it can locate the optima of all niches. Thus, the SCT and CPE look more like making a distributed searching rather than niching multiple species via coevolution because there is only one or a few individuals in most of the species at equilibrium. All niching schemes are able to conserve species once they are formed on niches in the evolution process.

4. Niching scalability

The scalability is an important property of searching algorithms. There are many ways to verify the scalability of niching approaches, and the basic idea is to test the computation time of niching methods for finding the global

Fig. 13 Individual number in species conserved in the evolution process on Schubert function (pop_size = 500, maximum FEs: 500,000)

optima in association with the dimensionality or size of the MMFO problems. The parameters of niching methods are set in association with dimension of the problems. A success rate of 95% is assumed for 30 executions of all algorithms. The niching scalability is investigated on the MRDFi and the r-Ising model.

(a) MRDFi

The population size $pop\_size = 20n$ and dynamic nearest neighbor sizing $q_{max} = pop\_size \times 0.05$ are used in the NNRC and HFC-I (with $\gamma = 0.1$), and two-point crossover is adopted to ensure the recombination efficiency of building blocks. Species distance threshold is set as $\sigma_s = 1.0 + 0.15n$ for the SCT, and the clearing radius is set as $\sigma = 1.0$ for $n \leq 10$ and $\sigma = 1.5$ for $n > 10$ with the CPE. The SCT-GGA and CPE-GGA employs the population size as $pop\_size = 30n + 200$ and two-point crossover as well. The SCGA uses bigger population sizes as {350, 380, 510, 840} for the MRDFi with dimensions as {5, 6, 7, 8}, the crossover and mutation operators are the original ones for the SCGA. The SGA is also tested on the MRDFi for comparison, and the population sizes are set as {350, 380, 510, 840, 1,070, 1,300} with dimensions of {5, 6, 7, 8, 9, 10}; the two-point crossover, Gaussian mutation, and elitist selection are used. The global optimum is found if there is at least one individual locating in its neighborhood with fitness greater than $(f^* - 0.025)$. Experimental results are shown in Fig. 14.

The NNRC and HFC-I achieve sub-quadratic scalability on the MRDFi in terms of the MNFEs, and the difference between them is not significant by the t test with a confidence level of 95%. It is verified empirically that the NNRC and HFC-I are able to preserve diverse local optima or building blocks for the specific decomposable problem, and two-point crossover promotes the mixing of lower-order building block for creating high-order ones. The

scalability of the NNRC-SSGA with $q > 1$ is the best of all.

The scalability of the CPE and SCT-GGA on the MRDFi is approximately quadratic. They are able to obtain the global optimum with the two-point crossover when a diverse population is maintained. In contrast, the SCGA can only find the global optimum of the low-dimensional MRDFi, and much bigger population sizes and more generations are required along with the increasing of the problem dimension. The computation time (MNFEs) increases exponentially or in high-order polynomial for locating the global optimum. This is due to the inefficiency of the recombination operator of the SCGA on the MRDFi although the SCT niches a population with good diversity.

The scalability of the SGA is a little better than the SCGA when two-point crossover operator is employed, but it still has a scalability of exponential or high-order polynomial on the MRDFi because it is unable to niche diverse local optima to provide the building blocks for the recombination operation.

When we compare the performance of the NNRC, SCT, HFC-I, and CPE with the SGA, we find that the niching strategy is critical to the scalability of algorithms. The comparison of the SCGA and the SCT-GGA in terms of scalability reveals the importance of the proper recombination operation for the niching methods on the MRDFi, which shows that the particular recombination operation has to be designed for specific multimodal problems so as to achieve good scalability.

(b) R-Ising model

The population size $pop\_size = 10n + 100$ and dynamic nearest neighbor sizing $q_{max} = pop\_size \times 0.02$ are used in the NNRC and HFC-I (with $\gamma = 0.1$). The species distance threshold is set as $\sigma_s = 2.0$ for the SCT. The SCT-GGA uses the same population sizing as the NNRC, but the SCGA employs population size as {150, 200, 250, 400, 750} for the r-Ising model with dimensions as {5, 10, 15, 20, 25}. The CPE-GGA employs {250, 300, 350, 400, 550, 700, 950, 1,300} as specific population sizes for dimensions of {5, 10, 15, 20, 25, 30, 35, 40}, and the clearing radius is set as $\sigma = 1.5$ for $n \leq 15$ and $\sigma = 2.0$ for $n > 15$. The two-point crossover and Gaussian mutation are used for the NNRC-SSGA, NNRC-GGA, SCT-GGA, HFC-I-SSGA, CPE-GGA. The SGA is not able to niche two species corresponding to the two global optima. With the r-Ising model, a global optimum is found if there is at least one individual locating in its neighborhood with a radius of 0.5 in the Euclidean distance. Experimental results are plotted in Fig. 15.

The NNRC and HFC-I achieve sub-quadratic scalability on the r-Ising model in terms of the MNFEs due to the nearest neighbors replacement strategy, and the t test
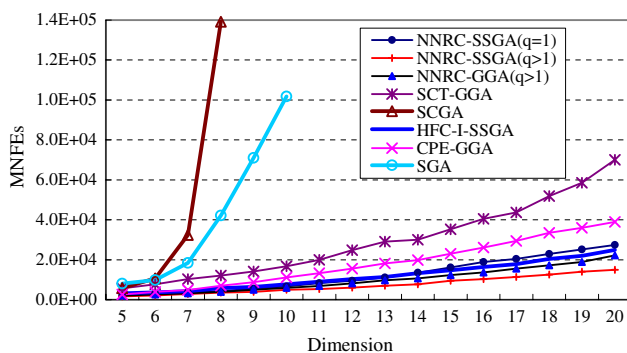

Fig. 14 Scalability of niching methods on MRDFi (Polynomial orders of scalability are estimated as NNRC-SSGA($q > 1$): 1.58, NNRC-GGA($q > 1$): 1.69, SCT-GGA: 2.18, SCGA: 5.69, HFC-I-SSGA:1.72, CPE-GGA: 1.91, SGA: 3.98)
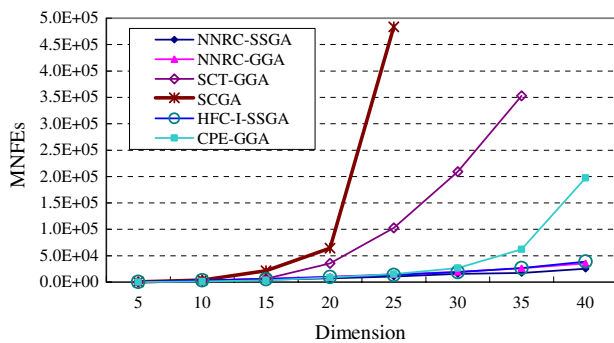
Fig. 15 Scalability of niching methods on r-Ising model (Polynomial orders of scalability are estimated as NNRC-SSGA: 1.68, NNRC-GGA: 1.79, SCT-GGA: 2.89, SCGA: 5.44, HFC-I-SSGA:1.80, CPE-GGA: 2.56)

indicates that the difference is not significant with a confidence level of 95%. The scalability of the SCT-GGA and the CPE falls between the quadratic and trinomial. The SCGA has a high order or near exponential scalability. The proportionate selection in these schemes often lead to the mating of two individuals with genotypically different building blocks, and thus the recombination operation fails to produce more competitive offspring because of the pinning effects on chromosomes. With the SCGA and the SCT-GGA, the specific crossover operator in the former algorithm is not as efficient as the two-point crossover in the latter on the epitasis landscape of the r-Ising model.

Empirically, the NNRC and HFC-I have a good capability to preserve building blocks of various length on this non-decomposable function, and the random selection and two-point crossover are able to accommodate the mixing of building blocks for creating high order ones under specific niching strategy with population replacement.

7 Discussions

Experimental results illustrate the intrinsic characteristics of the four niching schemes with overlapping population replacement. The NNRC and HFC-I work with a mechanism of multiple species coevolution, where species adapt to different niches on the landscape with the number of individuals approximately proportional to the attraction basin sizes of niches. The SCT and CPE make use of a mandatory mechanism to conserve species, which works more like the grid searching over the solution space that depends strictly on the thresholds of species distance or clearing radius. Although, each of the niching methods is able to deal with complex MMFO problems, the NNRC shows a better performance in terms of niching efficiency and scalability than others on specific test problems. In addition, the NNRC and the HFC-I are more robust

regarding the algorithm parameterization than the SCT and the CPE.

8 Conclusions

The main contribution of this paper is the introduction and evaluation of the niching mechanism based on overlapping population replacement for conserving multiple species. Four niching schemes, the NNRC, the SCT, the HFC-I, and the CPE are analyzed with respect to particular strategies for implementing population replacement. A suite of multimodal optimization functions are selected to test the performance of niching schemes regarding niching efficiency and scalability. All niching schemes perform competitively in dealing with complex MMFO problems and achieve sub-quadratic and close to quadratic scalability on decomposable MMFO problems although they adopt different niching strategies. As to the non-decomposable r-Ising model, the NNRC and the HFC-I schemes also achieve sub-quadratic polynomial computation efficiency in terms of the MNFEs.

These niching schemes need to be further investigated on more MMFO problems with different particularities. Since the overlapping population replacement is a good platform for niching species, there are possibilities to invent new niching methods on this platform that are more efficient, robust, and have better scalability in solving complex multimodal optimization problems in the real world.

References

Bäck T, Fogel DB, Michalewicz Z (2000) Evolutionary Computation. Institute of Physics Publishing, Bristol and Philadelphia

Ballester P, Carter J (2003) Real-parameter genetic algorithms for finding multiple optimal solutions in multi-modal optimization. In: Erick Cantú-Paz et al (eds) Proceedings of the 2003 genetic and evolutionary computation conference (GECCO 2003), Lecture Notes in Computer Science 2, Springer, Berlin, pp 706–717

Beasley D, Bull DR, Martin RR (1993) A sequential niche technique for multimodal function optimization. Evolut Comput 1(2):10–125

Beyer H-G (2001) The theory of evolution strategies. Springer, Berlin

Cedeño W, Vemuri V (1999) Analysis of speciation and niching in multi-niche crowding genetic algorithms. Theor Comput Sci 229(1–2):177–197

Cioppa AD, De Stefano C, Marcelli A (2004) On the role of population size and niche radius in fitness sharing. IEEE Trans Evolut Comput 8(6):580–592

De Jong KA (1975) An Analysis of The Behavior of A Class of Genetic Adaptive Systems. Dissertation, University of Michigan, Ann Arbor. (University Microfilms No. 76-9381)

Eiben AE, Smith JE (2003) Introduction to Evolutionary Computing. Springer, Berlin

Eshelman LJ, Shaffer JD (1993) Real coded genetic algorithms and interval schemata. In: Whitley LD (ed) Foundations of genetic algorithms II. Morgan Kaufmann, San Mateo, pp 187–202

Gan J, Warwick K (2001) Dynamic niche clustering: a fuzzy variable radius niching technique for multimodal optimization in GAs. In: Proceedings of the 2001 Congress on Evolutionary Computation (CEC 2001), vol 1. 27–30 May 2001, pp 215–222

Goldberg DE (1989) Genetic algorithms in search, optimization and machine learning. Addison-Wesley Publishing Company, New York

Goldberg DE, Richardson JJ (1987) Genetic algorithms with sharing for multimodal function optimization. In: Proceedings of the second international conference on genetic algorithms (ICGA 2), pp 41–49

Goldberg DE, Deb K, Korb B (1990) Messy genetic algorithms revisited: studies in mixed size and scale. Complex Syst 4(4):415–444

Goldberg DE, Deb K, Horn J (1992) Massive multimodality, deception, and genetic algorithms. Parallel Problem Solving from Nature, 2, pp 37–46. (Also IlliGAL Report No. 92007)

Gudla PK, Ganguli R (2005) An automated hybrid genetic-conjugate gradient algorithm for multimodal optimization problems. Applied Mathematics and Computation 167(2):1457–1474

Harik GR (1995) Finding multimodal solutions using restricted tournament selection. In: Proceedings of the sixth international conference on genetic algorithms (ICGA 6), pp 24–31 (Also IlliGAL Report No. 94002)

Herrera F, Lozano M, Sànchez AM (2003) A taxonomy for the crossover operator for real-coded genetic algorithms: an experimental study. Int J Intell Syst 18(3):309–338

Holland JH (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, 2nd edn. MIT Press, Cambridge

Horn J (1997) The Nature of niching: genetic algorithms and the evolution of optimal, cooperative populations. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL61801. (Also IlliGAL Report No. 97008)

Hu J, Goodman E (2004) Robust and efficient genetic algorithms with hierarchical niching and sustainable evolutionary computation model. In: Deb K et al (eds) Proceedings of genetic and evolutionary computation conference (GECCO 2004), Seattle, Washington, USA, June 2004; Part 1, Lecture Notes in Computer Science, vol 3,102, Springer, Berlin, pp 1220–1232

Hu J, Goodman ED, Seo K, Fan Z, Rosenberg R (2005) The hierarchical fair competition (HFC) framework for sustainable evolutionary algorithms. Evolut Comput 13(2):241–277

Iwamatsu M (2006) Multi-species particle swarm optimizer for multimodal function optimization. IEICE Trans Inf Syst E89-D(3):1181–1187

Jelasity M, Ortigosa PM, Garcia I (2001) UEGO, an abstract clustering technique for multimodal global optimization. J Heuristics 7(3):215–233

Lee C-G, Cho D-H, Jung H-K (1999) Niching genetic algorithm with restricted competition selection for multimodal function optimization. IEEE Transactions on Magnetics 35(3), Part 1, pp 1722–1725

Li M, Kou J (2005) A novel type of niching methods based on steady-state genetic algorithm. In: Wang L, Chen K, Ong YS (eds) Advances in natural computation: first international conference (ICNC 2005), lecture notes in computer science, vol 3612/2005, Springer, Berlin, pp 37–47

Li M, Kou J (2008) Crowding with nearest neighbors replacement for multiple species niching and building blocks preservation in binary multimodal functions optimization. J Heuristics 14(3):243–270

Li J-P, Balazs ME, Parks GT, Clarkson PJ (2002) A species conserving genetic algorithm for multimodal function optimization. Evolut Comput 10(3):207–234

Lin C-Y, Yang Y-J (1998) Cluster identification techniques in genetic algorithms for multimodal optimization. Comput Aided Civil Infrastruct Eng 13(1):53–62(10)

Mahfoud SW (1995) Niching methods for genetic algorithms. Dissertation, University of Illinois at Urbana-Champaign, Urbana, IL61801 (Also IlliGAL Report No. 95001)

Pelikan M, Goldberg DE (2000) Genetic algorithms, clustering, and the breaking of symmetry. In: Proceedings of parallel problem solving from Nature VI, Springer, Berlin, pp 385–394 (Also IlliGAL Report No. 2000013)

Peña JM, Lozano JA, Larrañaga P (2005) Globally multimodal problem optimization via an estimation of distribution algorithm based on unsupervised learning of Bayesian networks. Evolut Comput 13(1):43–66

Pétrowski A (1996) A clearing procedure as a niching method for genetic algorithms. In: Proceedings of the IEEE international conference evolutionary computation, Nagoya, Japan, pp 798–803

Reeves CR, Rowe JE (2003) Genetic algorithms: principles and perspectives—a guide to GA theory. Kluwer Academic Publishers, Boston

Sareni B, Krähenbühl L (1998) Fitness sharing and niching methods revisited. IEEE Trans Evolut Comput 2(3):97–106

Siarry P, Pétrowski A, Bessaou M (2000) Island model cooperating with speciation for multimodal optimization. In: Schoenauer M et al (eds) Proceedings of 6th international conference on parallel problem solving from nature (PPSN-VI}), parallel problem solving from nature, Paris, France. Springer, Berlin, pp 437–446

Siarry P, Pétrowski A, Bessaou M (2002) A multipopulation genetic algorithm aimed at multimodal optimization. Adv Eng Softw 33(4):207–213

Van Hoyweghen C, Naudts B, Goldberg DE (2002) Spin-flip symmetry and synchronization. Evolut Comput 10(4):317–344

Yin X, Germany N (1993) A fast algorithm with sharing scheme using cluster analysis methods in multimodal function optimization. In: Albrecht RF, Reeves CR, Steel NC (eds) Proceedings of the international conference on artificial neural nets and genetic algorithms. Springer, Berlin, pp 450–457

Yuan B, Gallagher M (2003) On building a principled framework for evaluating and testing evolutionary algorithms: a continuous landscape generator. In: Sarkar R et al (ed) Proceedings of the 2003 congress on evolutionary computation (CEC 2003), vol 1, 8–12 Dec 2003, pp 451–458