

Maintaining Healthy Population Diversity Using Adaptive Crossover, Mutation, and Selection

Brian Mc Ginley, John Maher, Colm O’Riordan, and Fearghal Morgan

Abstract—This paper presents ACROMUSE, a novel genetic algorithm (GA) which adapts crossover, mutation, and selection parameters. ACROMUSE’s objective is to create and maintain a diverse population of highly-fit (healthy) individuals, capable of adapting quickly to fitness landscape change and well-suited to the efficient optimization of multimodal fitness landscapes. A new methodology is introduced for determining standard population diversity (SPD) and an original measure of healthy population diversity (HPD) is proposed. The SPD measure is employed to adapt crossover and mutation, while selection pressure is controlled by adapting tournament size according to HPD. In addition to selection pressure control, ACROMUSE tournament selection selects individuals according to healthy diversity contribution rather than fitness. This proposed selection mechanism simultaneously promotes diversity and fitness within the population. The performance of ACROMUSE is evaluated using various multimodal benchmark functions. Statistically significant results are presented comparing ACROMUSE’s fitness and diversity performance to that of several other GAs. By maintaining a diverse population of healthy individuals, ACROMUSE responds to fitness landscape change by restoring better fitness scores faster than other GAs. Analysis of the adaptive operators illustrates that the key benefit of ACROMUSE is the synergy of the operators working together to achieve an effective balance between exploration and exploitation.

Index Terms—Genetic algorithm parameter adaptation, healthy population diversity.

I. INTRODUCTION

PREMATURE population convergence [1] about a local optimum is a common problem for traditional genetic algorithms (GAs) [2]. It is a result of individuals hastily congregating within a small region of the search space. Common reasons for this convergence include the following.

- 1) *Incorrect application of selection pressure*: Whereby a “super-performer” in the population dominates the selection procedure [3], eliminating much of the population’s diversity.

Manuscript received February 23, 2009; revised July 9, 2009 and October 13, 2009; accepted January 10, 2010. Date of current version September 30, 2011. This work was supported by the Irish Research Council for Science, Engineering, and Technology (IRCSET) Embark Initiative.

B. Mc Ginley, J. Maher, and F. Morgan are with the Bio-Inspired and Reconfigurable Computing Research Group, National University of Ireland (NUI), Galway, Ireland (e-mail: brianmccginley@gmail.com, john.maher@nuigalway.ie, fearghal.morgan@nuigalway.ie).

C. O’Riordan is with the Computational Intelligence Research Group, National University of Ireland (NUI), Galway, Ireland (e-mail: colm.oriordan@nuigalway.ie).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TEVC.2010.2046173

- 2) *Too low a mutation rate*: Diversity lost through selection and crossover is not recovered via mutation.

- 3) *Loss of crossover efficacy*: As a population becomes more converged, recombination of similar individuals creates similar offspring (inbreeding).

Eiben *et al.* [4] demonstrate that GA parameter choice strongly influences GA performance and that optimal parameter settings vary during the evolutionary process.

This paper presents ACROMUSE, a novel GA which adapts crossover, mutation, and selection parameters. ACROMUSE’s objective is to create and maintain a diverse population of highly-fit (healthy) individuals, capable of adapting quickly to fitness landscape change and well-suited to the efficient optimization of multimodal fitness landscapes (i.e., those typically present in real-world problems).

ACROMUSE employs two measures of population diversity to adapt crossover, mutation, and selection parameters. The first measure: Standard population diversity (SPD) describes a population’s solution space diversity. SPD is similar to other diversity measures described in the literature [5]. The second measure (proposed in this paper): Healthy population diversity (HPD) describes a population’s solution space diversity from a fitness perspective, i.e., a measure of the diversity of healthy individuals. A healthy population refers to a high-fitness population. However, a healthy population may be converged from a solution space perspective. SPD refers purely to solution space diversity with no regard to the health/fitness of individuals. ACROMUSE’s objective is to create a diversity of health in the population [i.e., a (solution space) diversity of high-fitness individuals]. This diversity of health is quantified by the HPD measure.

Other SPD measures reported in the literature [5]–[7] describe the genetic diversity of individuals in a population, without considering the fitness of each solution. Random initialization and disruptive mutation of individuals can easily generate a highly-diverse, randomly scattered population (of low-fitness individuals). This is because highly mutated and randomly initialized individuals are usually unfit. The proposed HPD measure differs in that it quantifies the diversity of highly-fit (healthy) individuals. By employing this HPD measure, ACROMUSE targets healthy diversity maintenance rather than SPD maintenance.

ACROMUSE employs the SPD measure to control mutation and crossover rates, while HPD is used to regulate selection pressure. Crossover employs the SPD measure to divide the population into an exploration section and an exploitation

section. The relative size of each section is controlled by the SPD measure. Mutation is employed in both sections. Mutation is applied adaptively with higher probability in the exploration section to explore potentially unvisited areas of the search space. Mutation is also employed as a local-search mechanism (with low probability) in the exploitation section of the population.

Selection pressure (tournament size) is adapted according to the proposed HPD measure. Tournament size is reduced when HPD is low (converged population) permitting lower-fitness outliers to reproduce, thereby protecting innovation. When HPD is high, tournament size is increased to promote “survival of the fittest.” ACROMUSE tournament selection chooses the individual with the largest healthy diversity contribution as the tournament winner. This differs from traditional tournament selection [8] which selects the individual with the best fitness score. Healthy diversity contribution is calculated as a function of both an individual’s diversity contribution and its fitness. This mechanism simultaneously promotes both diversity and fitness.

Maintaining a diverse population is very important for GA search. Not only does high-genetic diversity increase the population’s search coverage but it also endows the population with a degree of robustness when facing environmental (fitness landscape) change. A GA that demonstrates good fitness performance may be brittle in the face of fitness landscape variation, if it is not equipped to maintain or introduce diversity. Thus, for certain applications (e.g., real-world robotics controllers), it may be argued that diversity in a population is as important as high fitness. Most GA experiments prioritize solution quality (fitness) or algorithm running time/speed [9], [10], neglecting the valuable diversity component of the population. In this paper, the HPD measure is proposed as an additional means of assessing a GAs performance.

ACROMUSE has been evaluated on a group of multimodal function optimization problems [11], [12] and compared to several well-established GAs [2], [11], [13]–[15]. Statistically significant results demonstrate that for the multimodal functions employed, ACROMUSE eclipses other algorithms both in fitness and diversity performance. In addition, ACROMUSE responds to fitness landscape change by restoring higher fitness scores, faster than other GAs investigated.

A detailed critique, from a fitness and diversity perspective, is presented for each GA analyzed in this paper. The SPD and HPD analysis performed in the paper, coupled with best and average population fitness results, provides novel insights into the workings of these algorithms. Analysis of the adaptive operators illustrates that the key benefit of ACROMUSE is the synergy of the operators working together in offering the best fitness and diversity performance.

The remainder of this paper is structured as follows. Section II reviews related work in the field of GA parameter control and proposes the ACROMUSE adaptive operators. Section III details the methodology for calculating SPD and introduces HPD as a novel measure of a population’s diversity of health. Section IV describes how SPD and HPD measures

are employed to adapt ACROMUSE’s crossover, mutation and selection operators. Section V introduces several well-established GAs whose performances are compared to that of ACROMUSE. Section VI introduces the multimodal function optimization benchmarks used to evaluate GA performance. Section VII presents and compares results from the application of ACROMUSE and reference GAs to static multimodal functions. Section VIII presents and discusses the results of GA behavior following fitness landscape change. Section IX analyses ACROMUSE’s adaptive operators. Section X concludes this paper and lists its principal contributions.

II. RELATED WORK

Eiben *et al.* [4] present a thorough review of GA parameter control and demonstrate that optimal GA parameter values vary during the course of evolution. This can be explained through a GA’s twin goals of exploitation and exploration. A delicate balance must be struck between these two goals to prevent either premature convergence (too much exploitation) or blind random search (too much exploration). This section reviews related work in the field of GA parameter control and proposes ACROMUSE’s methods for adapting the mutation, crossover, and selection operators.

A. Mutation Parameter Control

Control of mutation rates is a widely researched mechanism of parameter control in evolutionary algorithm theory, with its roots dating back to Rechenberg’s adaptive “1/5 success rule” of evolutionary strategies [16]. Some approaches to adaptive mutation control [17]–[20] employ parent fitness in determining mutation probability. If selected, highly fit individuals undergo low levels of mutation (minimal disruption), while low-fitness individuals are subjected to large rates of disruptive mutation. A measure of population diversity is employed by [5] and [21] in adapting mutation probabilities. Mutation according to diversity is used to introduce novel diversity to a converged population. In a similar vein, Zhang *et al.* [22] adapt crossover and mutation according to parameters extracted from a *K*-means clustering algorithm. This clustering algorithm is employed as a means of monitoring diversity and the optimization state of the population. In [23], Yuen *et al.* propose an archiving method that prevents revisits by remembering every point visited in the fitness landscape. This approach has merit for applications with time-consuming fitness evaluations. The binary space partitioning archival method employed constitutes a parameter-less adaptive mutation operator that always locates a previously unvisited neighbor in the search space. Srinivas and Patnaik [24] calculate the net applied mutation rate as an average of: 1) parent-fitness determined mutation, and 2) a mutation rate, determined from a population diversity measure.

ACROMUSE’s adaptive mutation, similar to [24], is controlled by an average rate determined from two mechanisms: 1) SPD controlled mutation rate, and 2) parent fitness controlled mutation rate. A drawback with Srinivas and Patnaik’s method [24] is that population convergence is detected according to a fitness-based measure. The degree of diversity loss is

calculated as $f_{\max} - f_{\text{ave}}$, where f_{\max} is the maximum fitness value in the population, while f_{ave} is the average population fitness. Srinivas and Patnaik hypothesize that the closer f_{ave} is to f_{\max} the more converged the population is. In multimodal fitness landscapes, however, many different genotypes can share the same fitness score, so although an average fitness value may be identical to the best fitness value, the population may be scattered widely. ACROMUSE differs from [24] as it employs diversity measures calculated from the genetic diversity rather than the fitness diversity within a population.

B. Crossover Parameter Control

Hagras *et al.* [21], Srinivas and Patnaik [24], and Zhu *et al.* [5] employ a measure of diversity in adapting the crossover rate. Liu *et al.* [25] employ a rank-based approach for varying crossover probabilities where GA parameter settings for an individual depend on the performance of its parents; highly-fit members are protected while low-fitness individuals are subjected to high-crossover probabilities.

ACROMUSE employs adaptive crossover which varies crossover probability (P_c) according to SPD. P_c is reduced on population convergence (low SPD) and increased on high SPD. Individuals that are not subject to crossover have a higher rate of mutation applied instead. This effectively divides the ACROMUSE population into an exploitation section (using crossover and a low rate of mutation for local search), and an exploration section (where mutation is employed adaptively with higher probability to scout potentially unvisited areas of the search space). ACROMUSE differs from previous GA implementations in that crossover is employed to partition the population into separate sections (one dedicated to exploitation, the other to exploration).

C. Selection Parameter Control

Disruptive selection [25], [26] is a classic example of an adaptive mechanism that counters premature convergence through a nonmonotonic fitness function, which favors the strongest and weakest individuals in a population. Affenzeller and Wagner [27] present an adaptive selection mechanism, where selection pressure is varied by adjusting the proportion of individuals within the offspring population which must outperform their parents.

Self-adaptive selection [28], [29] incorporates individual voting within the tournament selection process. A hybrid self-adaptive tournament selection mechanism is also introduced which outperforms the purely self-adaptive algorithm. Vajda *et al.* [30] present a thorough overview of known parameter control methods for GA parent selection operators, noting that “on-the-fly control of selection operators has received little attention” and that “on-the-fly adjustment of the parameter(s) regulating selection is superior to using a constant value.”

Other reported mechanisms which do not specifically adapt the selection operator but yet control population diversity levels include: Fitness sharing (FS) [15], [31], crowding [32], restricted mating [33], diversity maintaining replacement strategies [34], population re-initialization strategies [35], [14], and random offspring generation [36]. In the field of multiob-

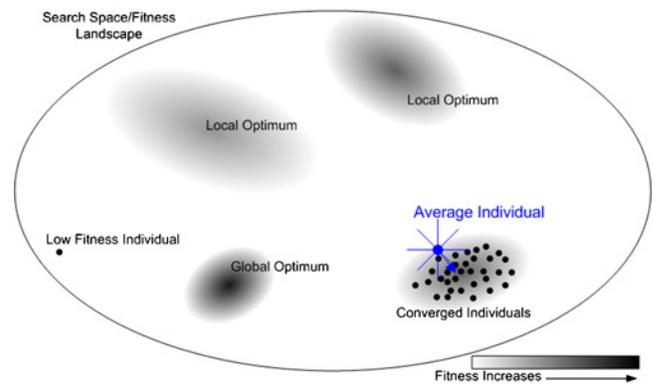


Fig. 1. Typical fitness landscape illustrating a converged population, local optima, global optimum, and the position of the average individual.

jective optimization, [37] provides a detailed analysis of different selection strategies for the promotion and maintenance of a wide diversity of pareto-optimal solutions.

While many of these techniques seek to promote diversity, the proposed ACROMUSE algorithm employs an adaptive selection operator which encourages promotion of HPD. The ACROMUSE adaptive tournament selection operator differs from traditional static tournament selection as ACROMUSE adapts tournament size according to HPD. In addition, ACROMUSE determines the tournament selection winner as the individual with the largest healthy diversity contribution rather than the traditional method of selecting according to best fitness scores. This adaptive selection mechanism simultaneously promotes both diversity and fitness. To the authors’ knowledge, no other adaptive selection operator exists that adapts selection pressure to allow survival of outlying individuals on population convergence, while explicitly rewarding both fitness and diversity contribution in the selection process.

III. SPD AND HPD

This section defines the proposed methods for calculating genotypic SPD and HPD. HPD is presented as a novel measure of diversity that quantifies the diversity of health in a population.

A. Standard Population Diversity (SPD)

SPD [5], [6], [38], and [39] describes the level of variation in a population. Genetic diversity is a very important component of evolutionary exploration since a GA can only search the space offered to it by the genes present in the population. Fig. 1 illustrates a typical fitness landscape containing a converged population, a number of local optima and the global optimum. The location of the hypothetical average individual is also illustrated.

With traditional GAs (TGAs), population convergence, as illustrated in Fig. 1, can occur very quickly [3]. This is caused by individuals congregating at a relatively fit area of the fitness landscape through a process of repeated selection of the same individuals, crossover of similar (converged) individuals and low-mutation rates. This locality of converged solutions may only be at a local optimum, potentially distant

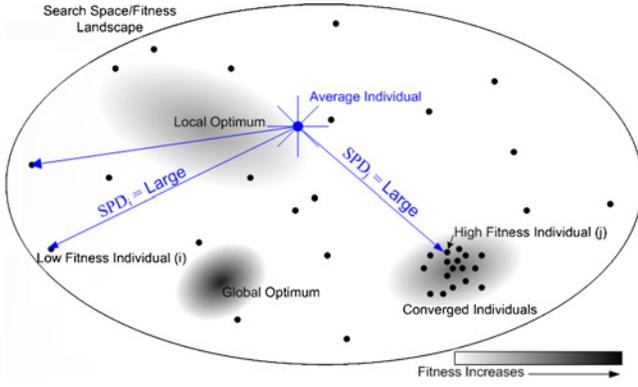


Fig. 2. Fitness landscape illustrating a diverged (highly mutated) population. SPD_i refers to individual i 's contribution to SPD.

from higher peaks or the global optimum. Since the population is converged, little opportunity is afforded to allow individuals explore other solution possibilities. Therefore, the population is trapped within a narrow region of the search space.

Fig. 2 illustrates a fitness landscape where a high-mutation rate has scattered individuals throughout the search space. ACROMUSE implements such an adaptive mutation operator, which scatters individuals to introduce novel genetic diversity. This innovation promotes a broader exploration of the global solution space. Fig. 2 also illustrates that high-mutation rates are generally destructive from a fitness perspective. Note that a partially converged sub-population also exists in the region of a local optimum.

B. Calculating SPD

SPD is calculated by finding the position of the average individual within the population (1) and summing the gene-wise Euclidean distances from this average point to the location of each individual (3).

ACROMUSE's population consists of P individuals (G_1 to G_P) where each individual consists of N genes. $G_{i,n}$ refers to the n th gene of individual i ; $G_i = (G_{i,1}, G_{i,2}, \dots, G_{i,N})$. The average individual in the population is referred to as G^{ave} and is computed as the gene-wise average over all P individuals. G_n^{ave} refers to the average of all n th genes in the population. Note that for certain encodings (e.g., binary valued genomes), an averaging may not produce a realizable individual. This does not affect ACROMUSE, however, as the average individual is purely a hypothetical individual, exclusively employed for diversity calculation

$$G_n^{ave} = \frac{1}{P} \sum_{i=1}^P G_{i,n}. \quad (1)$$

SPD_i refers to individual i 's contribution to SPD. It is calculated as the Euclidean distance between individual i and G^{ave}

$$SPD_i = \sqrt{\sum_{n=1}^N (G_{i,n} - G_n^{ave})^2}. \quad (2)$$

A simple summation of SPD_i could be employed to describe SPD. This approach is not used, however, because the summation cannot be normalized according to the mean (G^{ave}). If SPD is not normalized, SPD measures vary immensely for different problems and populations. To compute a normalized measure of SPD, a gene-wise standard deviation (measure of spread from the average individual) of the population is performed (3). This measure is calculated as the gene-wise standard deviation ($\sigma(G_n^{ave})$) over all P individuals

$$\sigma(G_n^{ave}) = \sqrt{\frac{1}{P} \sum_{i=1}^P (G_{i,n} - G_n^{ave})^2}. \quad (3)$$

Because of the gene-wise approach employed in (3), the standard deviation can be expressed relative to the mean (G^{ave}) as a coefficient of variation (4). $C_v(G^{ave})$ denotes the coefficient of variation of the average individual. This coefficient of variation (4) is used as the measure of SPD, which is employed to determine the adaptive mutation and crossover rates

$$SPD = C_v(G^{ave}) = \frac{1}{N} \sum_{j=1}^N \left(\frac{\sigma(G_j^{ave})}{G_j^{ave}} \right). \quad (4)$$

C. Healthy Population Diversity (HPD)

While the SPD measure of diversity outlined in Section III-B provides a good indication of population spread across the solution space, it does not account for the position of individuals in the fitness space. Fig. 2 illustrates that, although a population may appear diverged, most of the successful highly-fit individuals in the population lie within a converged localized area. At the same time, many of the scattered individuals that contribute to diversity are found in low-fitness areas of the search landscape. A central goal of this research is not only to maintain a good spread of individuals but to maintain a good spread of healthy individuals. Healthy individuals are population members that achieve good fitness scores.

D. Calculating HPD

To deal with the challenge of describing the diversity of population health, a fitness-weighted measure of population diversity (HPD) is introduced. HPD differs from SPD in that each individual's contribution to diversity in the solution space is influenced according to its fitness. A problem's fitness space and solution space are inextricably linked. A fitness landscape is a mapping over a solution space whereby every point has a fitness associated with it. SPD exclusively considers diversity in the solution space. HPD, in contrast, merges both fitness and solution spaces by weighting each individual's distance contribution in the solution space according to its score in the fitness space. Therefore, HPD describes fitness-weighted (fitness landscape) diversity while SPD solely describes solution space diversity.

HPD is calculated by finding the position of the weighted average individual within the population (6) and summing

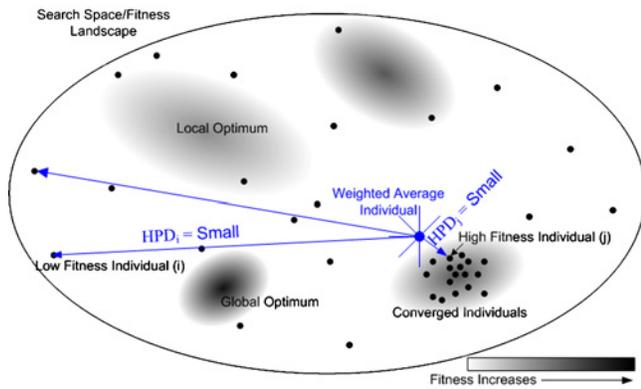


Fig. 3. Fitness landscape illustrating the position of the weighted average individual and the mechanism for calculating HPD. HPD_i refers to individual i 's contribution to healthy diversity.

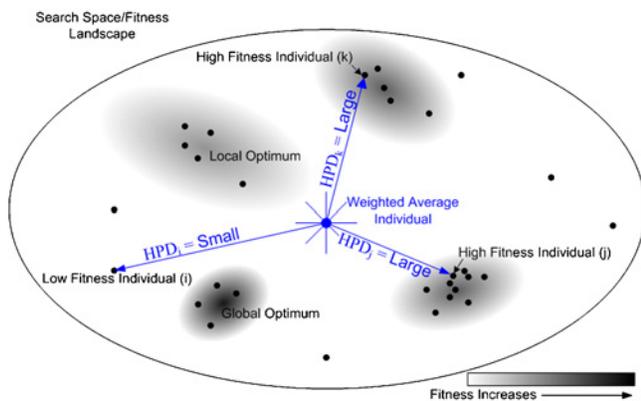


Fig. 4. Fitness landscape illustrating a population with high HPD.

the gene-wise fitness-weighted distances from this weighted average point to the location of each individual (8).

To calculate the position of the weighted average individual ($G^{W.ave}$), each population member's influence on the weighted average individual position is adjusted according to its fitness. An individual's w_i is defined as the individual's fitness expressed as a proportion of total fitness (5).

Fig. 3 illustrates how fitness-weighting each individual's influence on the weighted average has shifted the position of the weighted average individual closer to the converged high-fitness area. Although the population illustrated in Fig. 3 has a large measure of SPD, it has a small HPD measure. This is because all high-fitness individuals are located in a small region of the fitness landscape.

In contrast, Fig. 4 illustrates a population with a large measure of HPD. High-fitness (healthy) individuals attract the weighted average individual away from the cluster of converged individuals (illustrated in Fig. 3). Healthy (high-fitness) individuals are now far from the fitness-weighted average individual, resulting in a large measure of HPD

$$w_i = \frac{f_i}{\sum_{k=1}^P f_k}. \quad (5)$$

Equation (6) describes the method for calculating the weighted average individual ($G^{W.ave}$). $G^{W.ave}$ is computed as

the fitness-weighted gene-wise average across all P individuals

$$G^{W.ave}_n = \sum_{i=1}^P w_i G_{i,n}. \quad (6)$$

HPD_i refers to individual i 's contribution to healthy diversity. Fig. 3 illustrates that HPD_i is small. This is because individual i (although located far from the weighted average individual) is of low-fitness. Similarly, HPD_j (in Fig. 3) is small. This is because, although j is of high-fitness, j is located in close proximity to the weighted average individual.

To calculate HPD_i , individual i 's Euclidean distance to the weighted average individual is weighted according to its fitness (7). This approach means that an individual which is far from the weighted average will only have a large HPD_i measure if the individual is highly-fit

$$HPD_i = w_i \sqrt{\sum_{n=1}^N (G_{i,n} - G_n^{W.ave})^2}. \quad (7)$$

For computation of a normalized HPD, each individual's weighted gene-wise distance to the average individual is summed over the entire population to calculate the population's gene-wise fitness-weighted standard deviation [defined as $\sigma(G^{W.ave}_n)$]

$$\sigma(G^{W.ave}_n) = \sqrt{\sum_{i=1}^P w_i (G_{i,n} - G_n^{W.ave})^2}. \quad (8)$$

The weighted coefficient of variation ($C_v(G^{W.ave})$) is calculated to normalize the value of HPD (by relating the weighted standard deviation to the mean)

$$HPD = C_v(G^{W.ave}) = \frac{1}{N} \sum_{j=1}^N \left(\frac{\sigma(G_j^{W.ave})}{G_j^{ave}} \right). \quad (9)$$

This fitness-adjusted measure of diversity (fitness-weighted coefficient of variation) is used as the measure of HPD. HPD is employed to adapt the tournament selection operator.

IV. ACROMUSE ADAPTIVE OPERATOR IMPLEMENTATION

This section details the implementation of the ACROMUSE adaptive genetic operators and describes how SPD and HPD measures are employed to balance evolution's driving force of "survival of the fittest," with a random exploration of the solution space. ACROMUSE adaptive crossover, adaptive mutation and adaptive tournament selection are introduced. Fig. 5 illustrates the ACROMUSE GA process for producing an offspring population.

Regarding parameter sensitivity, the performance of a TGA is highly dependent on the choice of fixed parameters. In contrast, adaptive GAs vary parameters according to population statistics. For example, with a TGA, an incorrectly chosen selection pressure or mutation rate may result in premature convergence. However, an adaptive GA (such as ACROMUSE) avoids premature convergence by modifying crossover, mutation, and selection according to the population's diversity

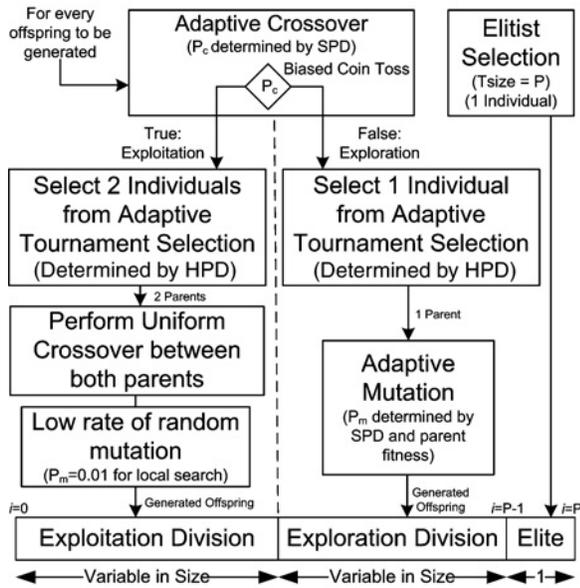


Fig. 5. ACROMUSE GA process for producing an offspring population. The adaptive crossover operator determines the relative sizes of the exploitation/exploration divisions. Individuals that are not chosen for crossover are instead used for exploration (via adaptive random mutation applied with higher probability). All selected individuals are chosen by the adaptive tournament selection operator (except the single elite individual), therefore, ensuring a good diversity of healthy individuals. During tournament selection, the individual with the largest fitness and diversity (HPD_i) contribution is selected.

status. As an adaptive GA's parameters specify upper and lower ranges, they are far less sensitive than a static/fixd value, since the actual parameter settings can assume any value within this range.

A. ACROMUSE Adaptive Crossover

Crossover generally converges a population since no additional diversity is added through its implementation; only recombination of existing genetic material is facilitated. Over time, as individuals become more similar through "survival of the fittest," crossover of similar selected individuals essentially becomes a local search (in-breeding). ACROMUSE adaptively mutates (with higher probability) individuals that do not undergo crossover. This technique essentially divides the population into following two sections (Fig. 5):

- 1) exploitation section (crossover and low mutation);
- 2) exploration section (high-probability adaptive mutation).

The size of each division is determined by the SPD measure. A highly converged (low SPD) population will result in reduced crossover probability to increase the size of the exploration division in the population. A highly diverged population (high SPD) employs more crossover and local search mutation ($P_m = 0.01$) to encourage exploitation. Equation (10) defines the proposed mechanism for calculating crossover probability (P_c). Uniform crossover at gene level is employed by ACROMUSE

$$P_c = \left[\left(\frac{SPD}{SPD_{max}} * (K_2 - K_1) \right) + K_1 \right]. \quad (10)$$

In this paper, P_c varies from 0.4 (K_1) to 0.8 (K_2), based on population diversity ($0 < SPD \leq SPD_{max}$). As P_c controls the relative sizes of the exploration and exploitation sections, K_1 is set to 0.4. This ensures that a minimum (on total population convergence) of 40% of individuals are protected from large-scale disruption in the exploitation division. K_2 is set to 0.8 so that on maximal population diversity, 20% of individuals are dedicated to continued exploration while the majority (80%) of the population works to further exploit the good genes already present. SPD_{max} is set to 0.4. This is because, in practice, the ratio between the population's standard deviation and the mean (4) does not exceed 0.4.

B. ACROMUSE Adaptive Mutation

This section describes the adaptive mutation implemented in the exploration division of the population (Fig. 5). The adaptive mutation rate (P_m) employed is an average of two mechanisms: SPD controlled mutation ($P_m^{Diversity}$) and parent fitness controlled mutation ($P_m^{Fitness}$). Single parent selection (no crossover) (Fig. 5) and random re-initialization mutation (at gene level) are employed.

Equation (11) defines the proposed mutation rate determined from population diversity ($P_m^{Diversity}$). A highly converged population is mutated more to introduce diversity. Grefenstette *et al.* [40] state that employing a P_m above 0.1 essentially corresponds to random search while the CHC random-restart strategy employs a P_m of 0.35 [14]. For the ACROMUSE algorithm, K is set to 0.5 and corresponds to the upper bound of P_m . The value of P_m is equal to K when the population is maximally diverged ($SPD = SPD_{max}$) and when the selected parent is the worst individual in the population. A P_m equal to 0.5 amounts to almost complete disruption of the individual. The selected range of P_m [0–0.5] provides a good resolution of possible mutation rates

$$P_m^{Diversity} = \frac{SPD_{max} - SPD}{SPD_{max}} * K. \quad (11)$$

The rationale for parent fitness-based mutation ($P_m^{Fitness}$) is to preserve a healthy individual's genetic material, while seriously disrupting individuals that are in low-lying fitness areas, to encourage further exploration. Equation (12) defines the proposed method for calculating $P_m^{Fitness}$; where f corresponds to parent fitness, f_{max} and f_{min} correspond to the best and worst fitness individuals in the population respectively. The selected value of K is 0.5

$$P_m^{Fitness} = K * \left(\frac{f_{max} - f}{f_{max} - f_{min}} \right). \quad (12)$$

Equation (13) defines the proposed method for calculating the net adaptive mutation rate (P_m) applied to each individual in the exploration section

$$P_m = \frac{P_m^{Fitness} + P_m^{Diversity}}{2}. \quad (13)$$

C. ACROMUSE Adaptive Selection

This section discusses the limitations of adaptive crossover and mutation. The rationale for including a novel adaptive

selection mechanism is presented and the proposed adaptive selection operator is described.

1) *Adaptive Crossover and Mutation—Limitations*: Observations on a group of randomly generated or highly mutated individuals indicate that most mutations have a negative impact on fitness. This is because, for most problems, high-fitness areas are sparse in the fitness landscape.

In our initial adaptive crossover and mutation experiments, a nonadaptive selection operator continued to choose from the same converged cluster of highly fit individuals (exploitation segment of the population), neglecting the less-fit individuals scattered throughout the fitness landscape by the exploration segment. This resembles the premature convergence problems caused by super-performers in TGAs. As a consequence of strong selection pressure, the same group of (exploitation) individuals continued to reappear in successive populations, albeit undergoing large (mainly destructive) rates of random mutation in the exploration section. No protection or opportunity is afforded to outlying exploration individuals to reproduce and perform a further search of their locality.

2) *Rationale for Adaptive Selection*: From a selection perspective, a diverged parent population does not automatically result in diverged offspring. A novel adaptive selection mechanism is proposed here to overcome this limitation. The ability to tune selection pressure is a product of employing tournament selection. A tournament size equal to the population size equates to elitism, while a tournament size of 1 equates to random selection. Therefore, decreasing tournament size in a converged population provides outlying individuals with a chance of selection. Similarly, if a population is diverged, tournament size is increased to promote “survival of the fittest.”

However, the traditional measure of diversity (SPD) is not suitable for controlling selection pressure. Fig. 2 illustrates that a population may appear highly diverged in the solution space. However, as illustrated, many of the high-fitness individuals lie close to a localized optimum, with diverged or mutated individuals scattered across the low-fitness plains. This means that the population possesses high SPD but low HPD. In this instance, by adapting selection according to the SPD measure, selection pressure will increase (as a consequence of high-solution space diversity) and selection will continuously return to the converged high-fitness section of the population. Therefore, although a population may appear diverged, it is in fact, still highly converged from the perspective of the selection operator.

3) *Proposed Adaptive Selection Operator*: In order to obtain a good spread of healthy individuals, the proposed adaptive selection operator employs the novel HPD measure to control tournament size. When the population is converged from a fitness-based perspective, low-fitness outliers are offered a greater chance of selection, thereby protecting innovation. When diverged, tournament size is increased, thus promoting “survival of the fittest.”

Equation (14) defines the proposed adaptive selection implementation. $Tsize$ refers to the tournament size used by the operator. $Tsize_{max}$ is set to $populationSize/6$ and corresponds to the maximum selection pressure that can be applied to

the population. Empirical results indicate that any tournament size greater than $populationSize/6$ leads to almost instant population diversity loss. This value of $Tsize_{max}$ gives good resolution to possible selection pressures, ranging from random selection ($Tsize = 1$) to very strong selection pressure ($Tsize = populationSize/6$). HPD_{max} refers to the maximum HPD attainable. HPD_{max} is set to 0.3. This is because, in practice, the ratio of the population’s weighted standard deviation to the weighted mean does not exceed 0.3

$$Tsize = \left\lceil \frac{HPD}{HPD_{max}} * Tsize_{max} \right\rceil. \quad (14)$$

4) *Traditional Tournament Selection*: Tournament selection is a well-established GA operator with applications found widely throughout the literature [8]. Tournament selection involves selecting a number ($Tsize$) of individuals randomly from the population, with the best individual from this group being selected as a parent. The definition of “best individual,” however, needs to be examined. To the authors’ knowledge, all existing tournament selection implementations select the individual with the highest fitness score as the best individual (winner) of the tournament. In this paper, a new approach is proposed: Selection according to healthy diversity contribution.

5) *ACROMUSE Tournament Selection—Selection According to Healthy Diversity Contribution (HPD_i)*: Rather than the traditional method of identifying a tournament’s winner as the individual with the best fitness, this paper proposes that to maintain healthy diversity, the individual with the largest healthy diversity contribution is selected. Selecting according to an individual’s standard diversity contribution (SPD_i) would not be appropriate, as GA search would tend toward random selection, negating the very function of the selection operator (“survival of the fittest”). Instead, the HPD_i (fitness-weighted diversity contribution) (7) measure is employed as a means of selecting the best individual. This ensures that selection recognizes both diversity contribution and fitness in choosing the best individuals for reproduction. To the authors’ knowledge, no other adaptive selection operator exists that adapts selection pressure to allow survival of outlying individuals on population convergence, while explicitly rewarding both fitness and diversity contribution in the selection process.

In addition to adaptive tournament selection, ACROMUSE employs elitist selection once ($Tsize = population size$) without crossover and mutation to preserve the best individual between generations (Fig. 5).

V. GA REVIEW

This section describes several well-established reference GAs whose behaviors are compared to that of ACROMUSE. The GAs investigated include; a TGA [2], a GA with Boltzmann selection [13], the CHC GA [14], deterministic crowding (DC) [11], and FS [15]. Table I lists the parameters employed for each GA. A population size of 40 is chosen for all GA experiments. Random re-initialization mutation, uniform crossover, and tournament selection (where appropriate) are employed in order to consistently compare the

TABLE I
BENCHMARK GA PARAMETERS

	ACROMUSE	ACROMUSE Without HPD _i Selection	Traditional GA	Fitness Sharing	Deterministic Crowding	CHC	Boltzmann Roulette Selection
Mutation (P_m)	0–0.5 (according to SPD)	0–0.5 (according to SPD)	0.01 [40, 41]	0.01	0.01	0 (0.35 for soft-restart)	0.01
Crossover (P_c)	0.4–0.8 (according to SPD)	0.4–0.8 (according to SPD)	0.9 [40]	0.9	0.9	1 (if individuals are sufficiently different)	0.9
Selection	$Tsize$ ranges from $1 - popSize/6$ according to HPD. Selection according to fitness and diversity	$Tsize$ ranges from $1 - popSize/6$ according to HPD. Selection according to fitness	$Tsize = 2$ [42]	$Tsize = 2$ [43]	Cross-generational truncation	Random parent and cross-generational survivor	Boltzmann initial temp = 4, final temp = 0.5 [13]
Pop-Size (P)	40						

various diversity maintenance schemes. The maximum number of generations allowed for evolution corresponds to the upper x -limit illustrated on each of the graphs in Section VI.

A. Traditional Genetic Algorithm (TGA)

The TGA does not employ adaptive operators. The TGA's parameters are listed in Table I. The mutation and crossover parameters chosen match the original parameters chosen by De Jong in his seminal work [12]. De Jong's choice of parameters was justified by Grefenstette *et al.* [40] when a meta-GA was employed to evolve these parameters. Following a detailed study of GA control parameters, Schaffer *et al.* [41] found that $P_m = 0.05$ – 0.01 and $P_c = 0.75$ – 0.95 perform best for binary coded genomes.

Many papers [44]–[47] have adopted the above parameters following the research of [12], [40], and [41]. Although these parameters are optimal in many cases, this paper reveals (in Section VII) the low-diversity measures resulting from their application. These low SPD and HPD measures present a significant risk of premature population convergence, demonstrating that for highly-multimodal landscapes, large scale exploration mutation is required. Results presented in this paper focus on stressing the consequences of employing a widely-used mutation rate which is set too low, rather than performing a-priori parameter tuning for best performance.

B. Fitness Sharing (FS)

FS as proposed by Goldberg and Richardson [15] implements a clever mechanism analogous to nature for restricting the number of individuals occupying a niche in the fitness landscape. Individuals within a certain “share-radius” of each other have their fitness penalized according to a sharing function. This limits the number of individuals that can successfully occupy a niche and, therefore, encourages exploration of other potential high-fitness areas.

A consideration of applying FS concerns choosing an appropriate share-radius, since this value is very specific to the environment. A constant share radius assumes that niches are spread uniformly throughout the landscape. This issue has been addressed in recent work relating to FS [31]. FS also

suffers from the limitation that an incorrectly sized population may not provide sufficient coverage for all niches, especially if niches are of unequal fitness.

C. Deterministic Crowding (DC)

DC, proposed by Mahfoud [11], is a novel variation of an earlier crowding method proposed by De Jong [12]. DC is an elegant, elitist algorithm which insists that an offspring must be better than its most genotypically-similar parent in order to progress to the next generation. This process means that offspring, if fitter, will replace adults most similar to themselves, therefore, largely retaining current diversity.

A limitation of DC is that individuals may migrate and converge to a single dominant peak, although much more slowly than with a TGA. Another limitation of DC is that it can only maintain diversity of the pre-existing mixture. There is no mechanism for introducing new diversity if the population becomes converged, or in the event of a fitness landscape change.

D. CHC GA

The CHC GA proposed by Eshelman [14], counters premature convergence by employing random parent selection and an incest prevention mechanism. Individuals are randomly selected for crossover, therefore, giving every individual an equal chance of reproduction. However, crossover is only allowed if individuals are sufficiently different genotypically (according to a hamming distance threshold). If no crossover is possible at the current threshold, the threshold is decremented and crossover is re-attempted. Once the threshold reaches zero, the population is randomly re-initialized (soft-restart) by replacing the population with copies of the best individual and then mutating the population with a large mutation rate (typically $P_m = 0.35$). Selection pressure is added by employing cross-generational truncation selection, where the entire population is replaced by the best individuals from the union of parents and their offspring. No mutation is employed by the CHC algorithm other than during soft-restart.

The CHC algorithm is the only chosen benchmark algorithm that includes a mechanism for large-scale diversity introduc-

tion when the population becomes converged. A negative aspect is that the soft-restart introduces random diversity (around the best individual), rather than maintaining healthy diversity throughout the fitness landscape.

E. Boltzmann Roulette Selection (BRS) GA

The BRS mechanism [13] begins search with a low-selection pressure and increases the pressure steadily according to a predefined schedule. This approach affords greater exploration in the early stages of search, with increased exploitation later (as promising optima emerge). This methodology contrasts with ACROMUSE, where a randomly diverged population at initialization generally lends itself to stronger selection pressure, until a stable SPD/HPD balance is reached. ACROMUSE selection pressure is then reduced to maintain healthy diversity and to perform a continued search of the fitness landscape. BRS performs very well in static fitness landscapes, though it performs poorly when long-term diversity maintenance and continued exploration over time are required.

F. ACROMUSE Without HPD_i Selection (AWHS)

For comparative purposes, a variant of ACROMUSE termed AWHS has been implemented. AWHS is identical to ACROMUSE except that, like traditional tournament selection, AWHS tournament winners are determined by fitness while ACROMUSE tournament winners are determined by an individual's contribution to healthy diversity (HPD_i). It is important to note that both ACROMUSE and AWHS employ identical crossover and mutation operators. Both also employ identical means for calculating tournament size. The only difference is in the tournament selection process:

- 1) AWHS: Best fitness individual wins;
- 2) ACROMUSE: Best diversity and fitness contribution wins.

The AWHS variant is included in the results to highlight the contribution that the HPD_i tournament selection operator makes to ACROMUSE's diversity and fitness performance.

VI. BENCHMARK FUNCTIONS

This section introduces a range of multimodal function optimization benchmarks to which the ACROMUSE and reference GAs are applied. Benchmark functions include Deb *et al.*'s M1, M2, and M4 [48], De Jong's M6 function [12], Goldberg *et al.*'s M7 [43] and Horn and Goldberg's M9 [49]. These benchmarks have been widely and recently employed in assessing the performance of niching GAs [31], [50]–[53].

The functions M1, M2, and M4 [48] are employed to investigate the diversity creation and maintenance properties of each algorithm. These benchmarks are 1-D (encoded as 30 bits), five-peaked sinusoidal functions. Although these functions are simple to optimize fitness-wise, many GAs (including niching GAs) have great difficulty in maintaining more than one solution [11]. These benchmark functions provoke interest because of their multimodality and have consequently been widely employed [31], [54] for the investigation and

comparison of niching and diversity-maintaining evolutionary algorithms. The M6 (also known as Shekel's Foxholes) is a more complex 34-bit encoded, 2-D, 25-peaked function which has been widely used to assess niching GA performance [11].

For the purpose of investigating each GA's response to fitness landscape change, a group of additional functions have been created. The M2a (Fig. 28), M4a (Fig. 39), and M6a functions are reversals of the M2, M4, and M6 functions respectively. These reversed functions are employed to explore how diversity maintenance leads to a more robust search.

The M7 and M9 problems are difficult and deceptive functions with high dimensions and thousands of global optima. These functions are designed to stress test the fitness and diversity performance of the algorithms. All genomes are binary coded. For all problems, the goal is to maximize fitness.

A phenotypic distance is employed by FS, DC (as in [11]), and CHC in experiments M1–M6. This is because (for M1–M6), each individual's binary genotype maps onto a real-valued phenotype. Inter-individual (M1–M6) phenotypic distance measures are more relevant than genotypic distances in these cases. To illustrate, a 1-bit genotypic difference between two individuals could be insignificant or substantial (from a solution space perspective), depending on which bit is flipped. If the least significant bit (of 30 bits) is altered, there is a tiny phenotypic difference. Conversely, if the most significant bit is changed, the difference is much larger. The relative significance of each bit on the genome is not considered by the genotypic distance measure, therefore, a phenotypic (solution space) representation is employed.

A genotypic hamming distance is used by FS, DC (as in [11]), and CHC for the M7 and M9 problems. The functions M7–M9 do not have a real-valued phenotype associated with their binary genotypic representations. Therefore, in these cases a genotypic hamming-distance measure is employed.

ACROMUSE employs a Euclidean inter-individual distance measure for both phenotypic (real-valued) and genotypic (binary-valued) representations. This is because an average of binary genotypes will generally not result in a binary individual, and the hamming distance between a binary individual and the nonbinary average individual is ill-defined. This (individual to average individual) distance measure is necessary for ACROMUSE SPD and HPD calculations.

The definitions and fitness landscapes for each benchmark function are introduced in the following sections alongside the corresponding results for each algorithm.

- 1) Function optimization results:
 - a) M1 function;
 - b) M7 and M9 functions.
- 2) Fitness landscape change results:
 - a) M2–M2a transition;
 - b) M6–M6a transition.
- 3) ACROMUSE operator analysis (M4–M4a transition).

For every experiment, the best individual and average population fitness scores for each GA are illustrated and discussed. SPD and HPD trends are also presented for each algorithm. Each result presented is the mean of 100 independent runs for all GAs except for the CHC algorithm where only 1 GA run is

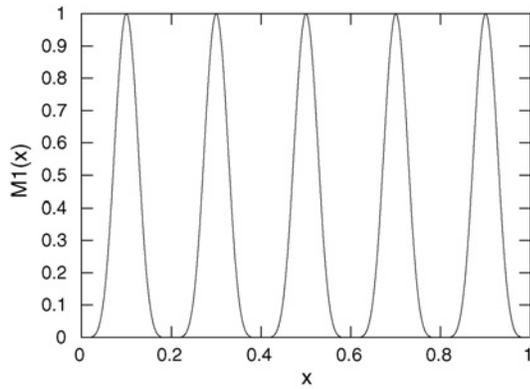


Fig. 6. M1 fitness landscape.

illustrated. This is to preserve the signature CHC “converge-restart” population cycle from averaging effects.

To verify statistical significance of the results, an independent group comparison has been performed on the final generation of each experiment. The final generation is chosen to allow each algorithm diverge sufficiently. A one-way analysis of variance (ANOVA) [55] is employed for all cases except the M7 and M9 best and average fitness results (as residuals of these experiments are not normally distributed). For these non-normally distributed populations, a nonparametric Kruskal–Wallis [56] test is performed instead. A sample size of 100 GA runs is employed. All independent group comparison tests returned significant results ($p < 0.01$), rejecting the null hypothesis that all samples are derived from the same population. Boxplots presented from the M9 experiment illustrate the distribution of best fitness, average fitness, SPD and HPD scores for each algorithm. ANOVA and Kruskal–Wallis statistical analysis results for each experiment are available at [57].

VII. MULTIMODAL FUNCTION OPTIMIZATION RESULTS

A. M1 Function

Fig. 6 illustrates the M1 fitness landscape while (15) defines the M1 function, where x is restricted to the real-valued range $[0, 1]$ and is encoded using 30 bits

$$M1(x) = \sin^6(5\pi x). \tag{15}$$

1) *M1 Best Fitness*: Fig. 7 illustrates best-fitness performance for each of the seven reference GAs applied to the M1 benchmark function. This graph reveals that a global optimum is quickly located by all algorithms. The M1 benchmark is widely employed, not to examine a GA’s fitness performance but to investigate how well each GA can maintain diversity across the multimodal fitness landscape [11]. This diversity maintenance is particularly difficult for TGAs where even moderate selection pressure will result in convergence to a single optimum.

2) *M1 Average Population Fitness*: The average population fitness scores of each algorithm applied to the M1 benchmark are illustrated in Fig. 8. The ACROMUSE and AWHS algorithms maintain the lowest average fitness. This is the cost of

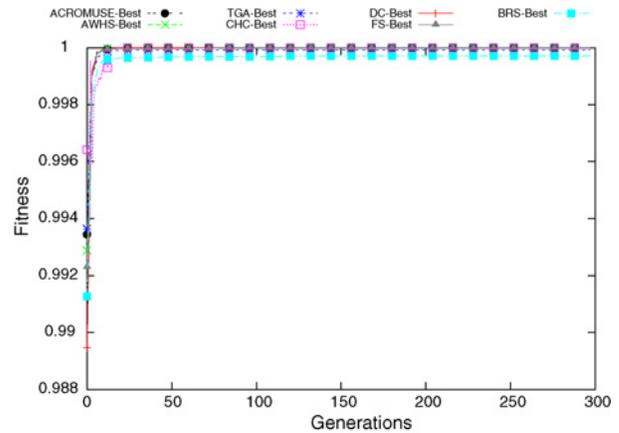


Fig. 7. M1 best fitness.

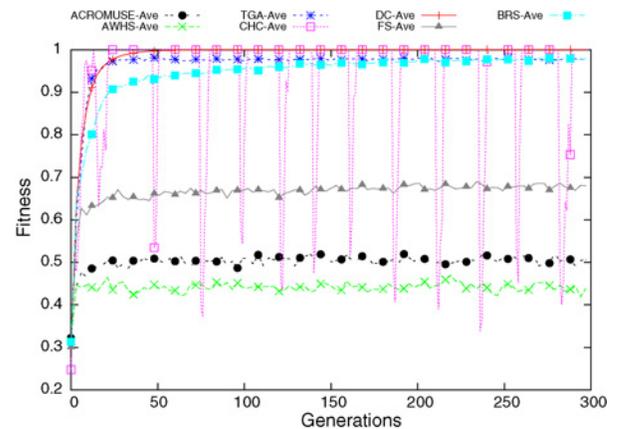


Fig. 8. M1 average population fitness.

applying large mutation rates to the exploration sub-section of the population. ACROMUSE sacrifices low-average fitness in exploration to achieve better fitness and HPD scores.

FS also exhibits a relatively low-average fitness trend, reflecting its tendency to select outlying individuals removed from the dominating optimum. These may be individuals scattered by mutation. These individuals are given a stronger chance to reproduce, as they lie outside the share-radii of individuals located at other optima.

BRS, as expected, demonstrates a slowly rising average fitness curve, as individuals are drawn toward a single optimum by steadily rising selection pressure. The TGA converges quickly to a high-average fitness value, primarily as a result of static selection pressure and a low-mutation rate.

The CHCs average fitness graph reveals the dynamics of the algorithm’s soft-restart. The high-mutation rate employed by this re-initialization causes a sudden dramatic fall in average fitness scores. This drop is quickly remedied by the algorithm’s strong cross-generational survivor selection mechanism. DCs highly elitist cross-generational truncation selection ensures a monotonically rising average fitness curve (in a static environment), as mutated individuals only progress if they are superior to their parents.

3) *M1 SPD*: Fig. 9 illustrates the SPD trend of each GA applied to the M1 benchmark. The FS, ACROMUSE

and AWHs algorithms maintain high levels of SPD through different mechanisms. FS encourages diversity by penalizing individuals which lie within a share radius of each other. These fitness penalties encourage the selection of individuals that lie in other remote areas of the fitness landscape. ACROMUSE and AWHs not only maintain healthy diversity through adapting selection pressure, but also introduce new diversity through the exploration section of their populations. This adaptive exploration section introduces fresh diversity and searches for previously undiscovered optima in the fitness landscape.

ACROMUSE demonstrates the best SPD performance of all the algorithms tested. Although ACROMUSE employs the same crossover and mutation mechanisms as the AWHs, its selection operator favors both fitness and diversity contribution (rather than only fitness). As a result, ACROMUSE outperforms the AWHs from a diversity perspective. Fig. 9 illustrates that this diversity consistently exceeds the initial random diversity of the population. This is achieved by individuals congregating at peaks phenotypically distant from each other.

From the SPD trends (Fig. 9), the CHCs signature “converge-restart” cycle is the most distinctive. This large SPD oscillation is caused by a large mutation rate (0.35) being applied to copies of the best-individual when the population has fully converged. This essentially corresponds to a random re-initialization of the population in the general neighborhood of the best individual. The flat (low diversity) areas between spikes reflect, not a completely converged population, but a population where continued crossover is allowed by very low-hamming distances between individuals. As 30 bits are employed to represent each individual’s phenotype (in the range $[0, 1]$), a difference of several least significant bits between individuals is insignificant when observing the phenotypic diversity of the entire population. More frequent restarts could be initiated by increasing (from zero) the minimum hamming crossover threshold required to cause a soft-restart.

The TGA employs static selection pressure with a low-mutation rate and consequently converges very quickly, as illustrated by its low-SPD scores. This trend can result in premature convergence at a sub-optimal solution. As the M1 benchmark optima all share identical fitness, this is not an issue. BRS demonstrates slower convergence than the TGA, as initially weak selection pressure steadily rises during evolution. DC steadily loses diversity but at a much slower rate than BRS, TGA, and CHC. Loss of DC diversity is caused by the migration of individuals toward dominating peaks. This slow loss of diversity is discussed in detail in Section VII-A5 regarding the distribution of each algorithm’s population across M1s five peaks (*Note*: In this paper, the word peak and niche are used interchangeably).

4) *M1 HPD*: Fig. 10 illustrates the HPD scores achieved by each algorithm. ACROMUSE demonstrates the best healthy diversity performance (HPD) of all the algorithms tested. HPD is increased from the beginning and remains high throughout. This is the result of rewarding both fitness and diversity contributions in the selection process. Individuals tend to move toward niches that are phenotypically distant from each other, thus providing a broad diversity of potential solutions. HPD

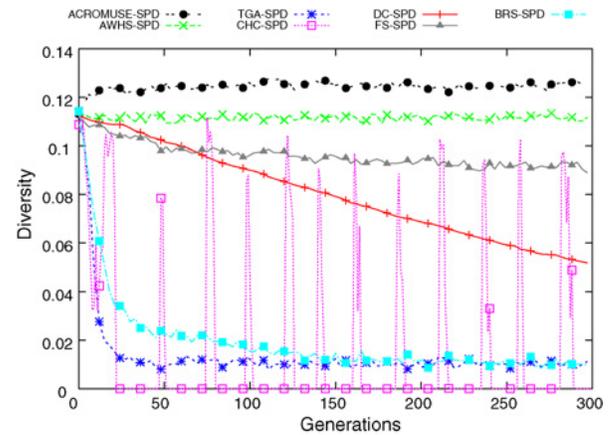


Fig. 9. M1 SPD.

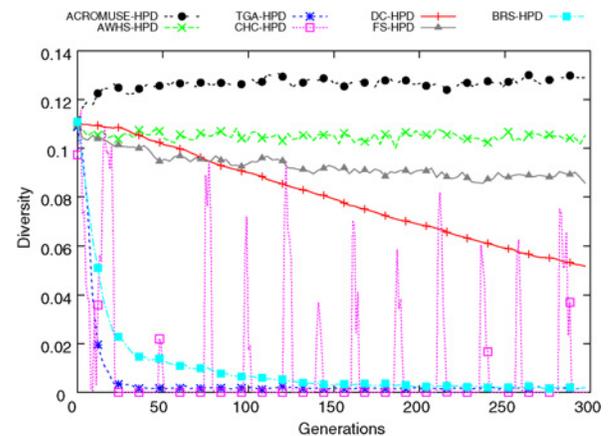


Fig. 10. M1 HPD.

trends for FS and AWHs are similar to those obtained from the SPD results (Fig. 9).

The CHCs lower HPD spikes (compared to its SPD spikes) reflect that standard (low-fitness) diversity, rather than healthy diversity is being introduced to the population through the restart process. These random mutations generally tend to have a negative effect on fitness, thereby leading to a smaller degree of healthy diversity (HPD).

The TGA and BRS algorithms again exhibit similar HPD behavior to that observed in the SPD graphs (Fig. 9). Fast TGA and slower BRS convergence demonstrate the properties of these algorithms. A notable difference, however, is the level of HPD maintained compared to SPD when the TGA and BRS populations are converged (stabilized post-generation 150). When the population is converged, SPD values for TGA and BRS hover around 0.01. This is evidence of the mutation rate ($P_m = 0.01$) continually introducing diversity. In contrast, the HPD trend rarely strays from zero, reflecting the generally fitness-destructive nature of this mutation. DC, similar to its SPD trend, also slowly loses HPD as evolution progresses. The reasons for this are outlined in Section VII-A5 when discussing DC’s population distribution across peaks.

5) *M1 Niches Maintained*: Fig. 14 displays the average number of peaks on which each algorithm maintains a presence during the course of evolution. The M1–M4 functions

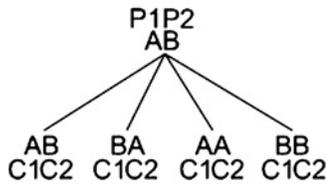


Fig. 11. Potential outcomes of a cross between an individual from class A and an individual from class B. If combination AA and BB were to occur, there is a strong possibility of diversity loss.

	P1	P2	C1	C2
Phenotype (x)	0.685	0.905	0.903	0.937
Fitness	0.845	0.982	0.993	0.341

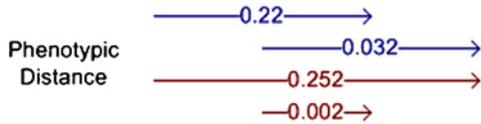


Fig. 12. DC selection and crossover: C1 will replace P1, thereby eliminating the individual near the 0.7 peak. Parent and children phenotypes and fitness scores are displayed alongside their inter-phenotypic distances.

comprise five peaks. The aim of niching GAs is to maintain a continued presence on as many peaks as possible. Fig. 14 displays the performance of each GA on this task. As in [11], a peak is considered maintained if at least one population element exists in the basin of attraction of that peak, and that individual has a fitness of at least 80% of that peak’s height.

The TGA is the quickest to converge to a single peak through strong selection pressure. CHC also converges quickly. However, its soft-restart introduces novel random diversity periodically. The number of peaks maintained by CHC, even during soft-restart rarely exceeds two.

DC begins by promoting its initial randomly distributed population to each of the five peaks in the landscape. However, this initial presence on each peak is slowly eroded by migration caused by a combination of the crossover operator and the cross-generational elitist selection mechanism. In [11], an analysis of DC makes ideal assumptions regarding the crossover operator. Mahfoud states that given three classes of individuals {A, B, C}, crossing A with B or C always yields one element from each class involved in the cross. These crosses will be true of an ideal crossover operator. However, a stochastic crossover operator between two parents P1 (A) and P2 (B) can produce with equal probability, any of the four combinations of children C1 and C2 as indicated in Fig. 11.

The rightmost permutations of AA and BB will automatically result in a loss of diversity if both children outperform their parents. This is illustrated with an example taken from DC experiments; Fig. 12 illustrates a microcosm of DC’s selection and crossover process. Parent and children phenotypes and fitness scores are displayed alongside their inter-phenotypic distances. As $\text{Distance}(P1, C1) + \text{Distance}(P2, C2) \leq \text{Distance}(P1, C2) + \text{Distance}(P2, C1)$, individual C1 will replace P1, because C1 is fitter.

Fig. 12 demonstrates DC genetic drift since a parent around the peak at 0.7 is lost in favor of a fitter offspring at the 0.9 peak. Extrapolating from this, it is evident how a group

	P1	P2	C1	C2
Phenotype (x)	0.919	0.118	0.978	0.31
Fitness	0.762	0.787	0.002	0.931

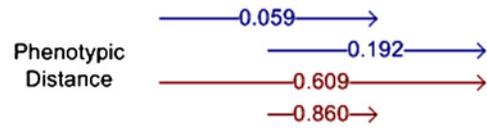


Fig. 13. DC selection and crossover: C1 will replace P1, resulting in the genetic drift of an individual from the peak at 0.1 to the peak at 0.3.

of individuals at one peak, which are slightly better than individuals at another peak, will slowly dominate the selection procedure.

In addition, aside from crossover creating two children similar to a single parent, children resulting from crossover may be from a different class than either parent. In Fig. 13, an example is illustrated of how crossover creates an individual (near the 0.3 peak) from neither of the parent classes (near the 0.9 and 0.1 peaks). As C2 is fitter than P2, this results in the loss of an individual from the 0.1 peak in favor of the peak at 0.3.

These particular examples demonstrate the possibilities for DC genetic drift and explain DC’s SPD, HPD, and niches maintained results. These examples also serve to outline the tremendous influence that crossover has on the DC algorithm. Regarding these crossover possibilities, Mahfoud [11] acknowledges that for functions with multiple global optima, the migration of individuals discussed above will occur freely, inevitably resulting in genetic drift. In addition, Mahfoud states that (as demonstrated by these results) DC will drift much more slowly than the TGA.

The drift demonstrated in M1 could be stalled somewhat through a mechanism warranting a child’s fitness to be better than its parent’s fitness by an amount greater than a threshold, in order to progress to the next generation. This results in the elimination of drift caused by small differences in fitness between parent and child. These small fitness differences are very common given that 30 bits are used to represent a number in the range [0, 1]. Mahfoud [11] deals with this problem by incorporating a slight design change that requires offspring be strictly fitter than their parents in order to progress to the next generation. By increasing the degree by which children must outperform their parents, further reduction in drift occurs at the expense of “survival of the fittest.” This mechanism has been implemented and it does indeed reduce drift in the M1 benchmark, where all optima are at equal heights. As expected, however, this fitness threshold does nothing to mitigate drift when peaks are of varying heights (i.e., in the M4 function).

FS and ACROMUSE perform similarly in terms of presence across peaks with roughly four peaks maintained consistently across generations (Fig. 14). This FS result differs from other work [11], [48] which demonstrates that FS can maintain all peaks throughout evolution. A reason why FS does not perform as well as [11] and [48] is that the population size used in this paper is significantly smaller. With larger population sizes, FS and ACROMUSE can maintain coverage over all five peaks. However, the goal of this research is to develop

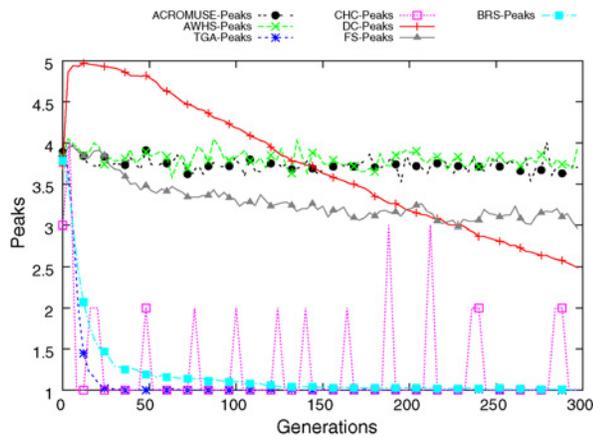


Fig. 14. M1 number of niches maintained. A niche is considered maintained if at least one population element exists in the basin of attraction of that peak, and that individual has a fitness of at least 80% of that peak's height.

and compare GAs that can search effectively using smaller population sizes; this is required because fitness evaluations of real world (e.g., robotics) controllers are extremely costly time-wise. ACROMUSE has been designed to effectively exploit and explore a population size of 40.

It is interesting to note that although DC initially maintains a higher number of peaks than ACROMUSE, its HPD value is lower. In addition, FS, ACROMUSE, and AWHs all demonstrate similar performance in terms of peaks maintained but ACROMUSE maintains much higher SPD and HPD than AWHs and FS. This is because of ACROMUSE's ability to cluster more individuals toward high-fitness peaks at the outskirts of the phenotypic fitness landscape. This may be of interest for real-world application (e.g., gaming or robotics predator-prey strategies) where maximal phenotypic diversity is of benefit.

B. M7 and M9 Functions

In order to perform challenging tests on the GAs, the M7 and M9 benchmark functions are employed. M7 is a massively multimodal deceptive function [11], [43] comprising of five 6-bit bimodal subfunctions. Each subfunction is a function of unitation with global optima located at $x = 000\,000$ and $x = 111\,111$. The total number of optima in M7 is 5 153 632 of which 32 are global. The maximum fitness score attainable is 5. Fig. 15 illustrates the fitness landscape of a single subfunction.

M9 is a highly challenging minimum distance function [11], [49] containing 2197 optima of which 27 are global. Overall fitness is the sum of fitnesses over three 8-bit deceptive subfunctions. The maximum fitness score attainable is 30. The M9 benchmark problem is described in more detail in [11].

1) *M7 and M9 Best Fitness*: The highly multimodal M7 and M9 test problems ideally serve to highlight the limitations of TGAs converging prematurely at sub-optimal solutions. This can be observed from the leveling-out of best fitness scores in both M7 (Fig. 16) and M9 (Fig. 17) best fitness results. These results further underpin the need for diversity-maintaining and diversity-introducing GAs. With the M7 test

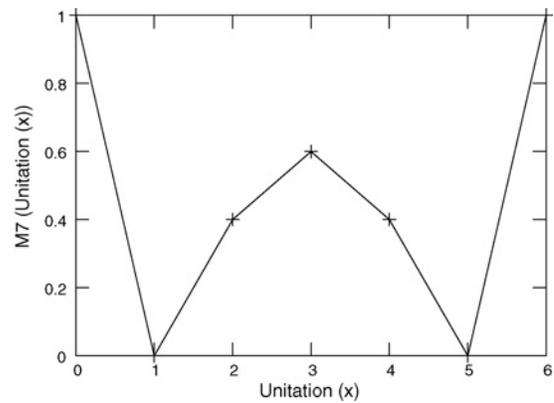


Fig. 15. M7 is comprised of 5 of these deceptive unitation subfunctions.

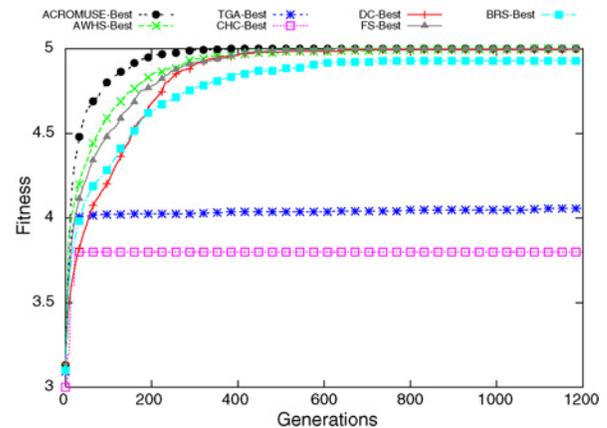


Fig. 16. M7 best fitness (max possible fitness = 5).

problem, the TGA and CHC algorithms become prematurely converged while all other test algorithms succeed in locating the best solution. ACROMUSE performs marginally best of these in locating a global optimum first.

With the more difficult M9 problem, ACROMUSE's diversity and fitness-rewarding selection scheme performs significantly better than the other algorithms in achieving the highest best fitness scores (locating a global optimum on all occasions). ACROMUSE owes its success to its novel employment of separate population subsections to simultaneously exploit and explore. In contrast, the massively multimodal and deceptive nature of this problem causes difficulties for DC and FS in escaping sub-optimal solutions. This is because FS and DC have a policy of maintaining a continued presence at these optima, rather than sufficiently introducing the novel diversity required to locate a deceptive solution.

Boxplots for each algorithm (Fig. 18) observed at generation 950 of the M9 problem reveal that FS and BRS attain similar best fitness scores with values ranging from 18 to 30 (max fitness). Both FS and BRS are negatively skewed with both medians found at 22. The mean best fitness scores of each algorithm are indicated by the crosshairs in Fig. 18. On average, DC achieves higher best fitness scores than FS and BRS for M9. While FS, BRS, CHC, ACROMUSE and AWHs all successfully locate a global optimum on at least one occasion, DC and TGA fail on each of the 100 GA runs

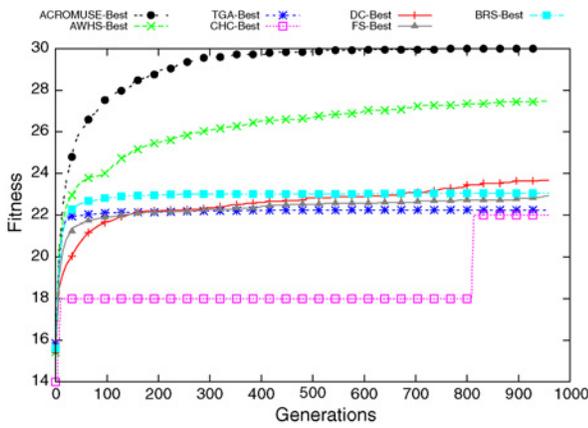


Fig. 17. M9 best fitness (max possible fitness = 30).

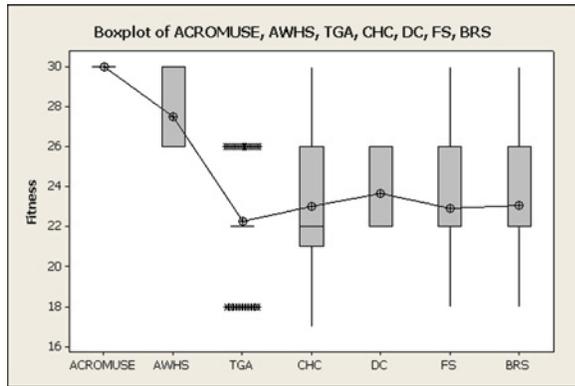


Fig. 18. M9 best fitness boxplots observed from 100 samples taken at generation 950. The mean of each group is illustrated by the crosshairs.

allowed (Fig. 18). ACROMUSE and AWHs perform best, with ACROMUSE locating a global optimum on every occasion. CHC exhibits the largest inter-quartile range, reflecting convergence to multiple different solutions across the 100 runs.

2) *M7 and M9 Average Fitness*: M7 and M9 average fitness results (Figs. 19 and 20) resemble in some respects the trends from the easier M1 problem. Firstly, ACROMUSE and AWHs fitness scores are quite low as a result of the negative effects of mutation in the exploration segment of their populations. Similarly FS’s selection of individuals lying outside more heavily populated niches reduces average fitness scores. DC and BRS exhibit steadily rising average fitness scores through different mechanisms. BRS’s steadily rising selection pressure ensures that evolution’s search has a better chance of escaping local optima that might prematurely trap the TGA’s population. DC’s cross-generational elitist selection ensures that average fitness always monotonically increases. The TGA average fitness trend plateaus quickly as its population prematurely converges to a single optimum. This convergence may be confirmed by studying the low SPD (Figs. 22 and 23) and HPD (Figs. 25 and 26) values and identical best (Figs. 16 and 17) and average (Figs. 19 and 20) fitness scores. CHC also exhibits a flattening out of average fitness as the population converges about a single solution. This is illustrated by almost identical average and best fitness

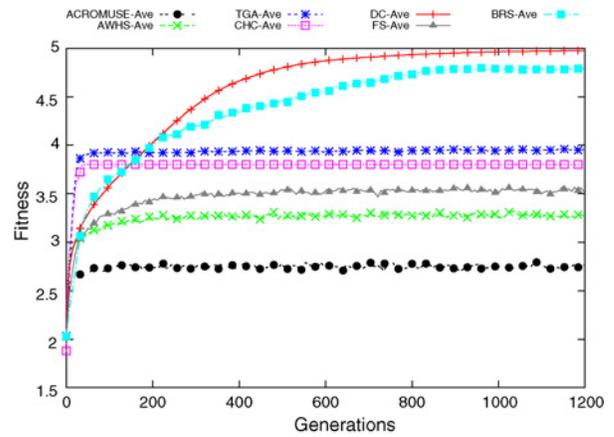


Fig. 19. M7 average population fitness.

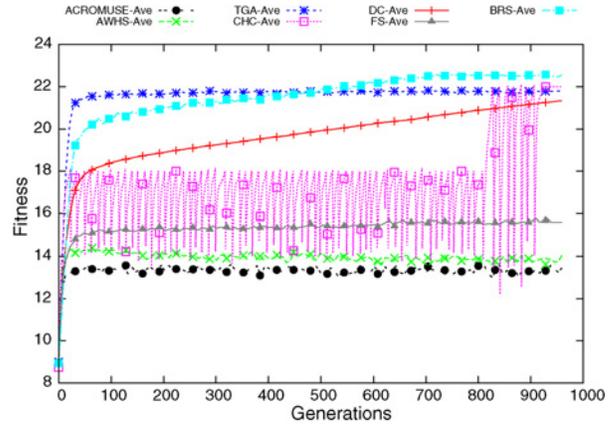


Fig. 20. M9 average population fitness.

scores. The soft-restart oscillation, although observable for the M9 (Fig. 20) problem, is notably absent in the M7 result (Fig. 19). This absence of the restart is caused by the highly multimodal nature of the fitness landscape (see Section VII-B3 for discussion).

Average fitness score boxplots (Fig. 21) observed at generation 950 of the M9 function reveal that TGA, CHC, and BRS have the largest ranges. This indicates that these algorithms tend to converge to different optima on different runs. This converging tendency is reminiscent of Fig. 1. FS exhibits lower variation in values, indicating more consistent behavior across runs while DC, AWHs, and ACROMUSE demonstrate the smallest ranges, indicating a narrower variability and more consistency in the algorithms’ average fitness behaviors.

3) *M7 and M9 SPD*: M7 (Fig. 22) and M9 (Fig. 23) SPD trends include many similarities to the M1 function (Fig. 9) SPD results. ACROMUSE demonstrates the best diversity performance. FS and AWHs exhibit high-SPD scores while BRS and TGA demonstrate low SPD/HPD, reflecting what would be expected from their differing selection strategies (i.e., high-TGA convergence with slower convergence from the BRS algorithm). DC’s SPD results for M7 mirror its M1 experiment behavior. However, the DC SPD trend for the M9 differs somewhat from other experiments as it quickly falls

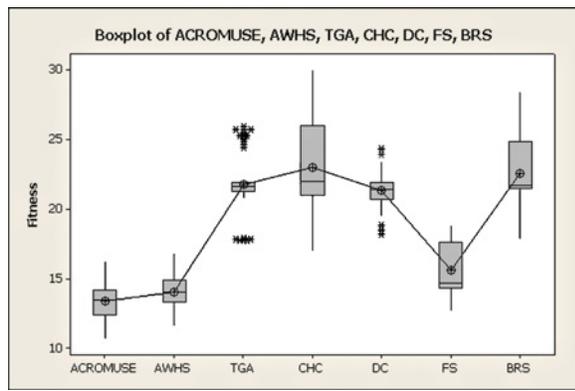


Fig. 21. M9 average fitness boxplots observed from 100 samples taken at generation 950.

before rising slowly. This is due to the deceptive nature of the problem, where the basins of attraction of the various local optima lead solutions further away from each other thereby causing a slow increase in SPD. The progressively rising average fitness scores confirm this migration. This average-fitness trend also explains the corresponding slow increase in HPD (Fig. 26) values (higher average fitness leads to higher fitness-weighted healthy diversity).

It is interesting to note from the M9 average fitness graph (Fig. 20) that both BRS and DC exhibit steadily rising average fitness scores. However, from observing the different SPD/HPD trends in Fig. 23 it is clear that while DC's population is moving toward multiple optima, BRS is converging toward a single solution. This serves to highlight that fitness-behavior alone is not enough to classify an algorithm's overall behavior.

The CHC algorithm's diversity behavior is of interest because of the lack of random restarts in the M7 benchmark. This is due to the highly-multimodal nature of the M7 problem, where many solutions share the same fitness score but possess radically different genotypes. The CHCs cross-generational truncation selection does not exhibit a strong enough selection pressure to focus search on any one of these solutions. As all individuals share the same fitness (as indicated by identical best and average fitness scores), no selection bias is afforded toward any single individual. In addition, as these individuals have different genotypes, crossover is always possible, meaning that no soft-restart is triggered.

All GAs, except DC, applied to the M9 problem reveal narrow SPD ranges (Fig. 24), reflecting consistent diversity behavior from each algorithm. ACROMUSE consistently maintains the highest diversity with AWHs in second place. TGA, BRS, and CHC consistently maintain low levels of diversity. CHC outliers reflect occasional random restarts.

DC exhibits the highest SPD variability of the algorithms investigated. Although DC generally maintains high-SPD levels, on occasion its population converges dramatically. As observed from the average fitness boxplots (Fig. 21), DC maintains a narrow range of average fitness scores. Also notable from Fig. 18 is that DC has a narrow variability of best fitness scores. This convergence of average and best fitness is caused by strong cross-generational selection pressure. The

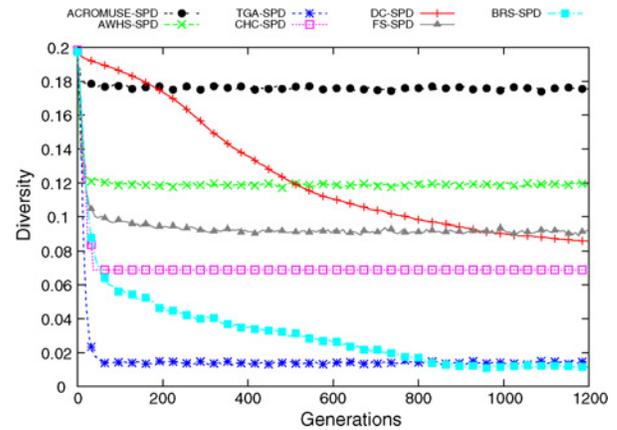


Fig. 22. M7 SPD.

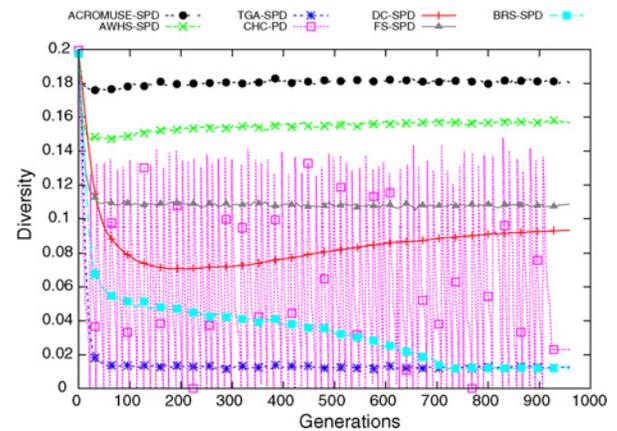


Fig. 23. M9 SPD.

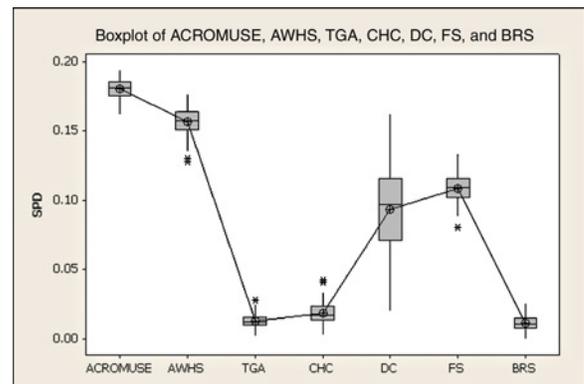


Fig. 24. M9 SPD boxplots observed from 100 samples at generation 950.

highly multimodal nature of the M9 function means that individuals can maintain identical fitness scores and be situated on different peaks. This explains the normally high measures of SPD. On occasion, however, many individuals converge toward a single peak, resulting in much lower SPD measures.

4) *M7 and M9 HPD*: The HPD (Figs. 25 and 26) trends of each GA's performance in solving the M7 and M9 functions are broadly similar to the SPD results (Figs. 22 and 23). The most notable difference is the narrowing of the gap between AWHs HPD and FS HPD, compared to that between AWHs

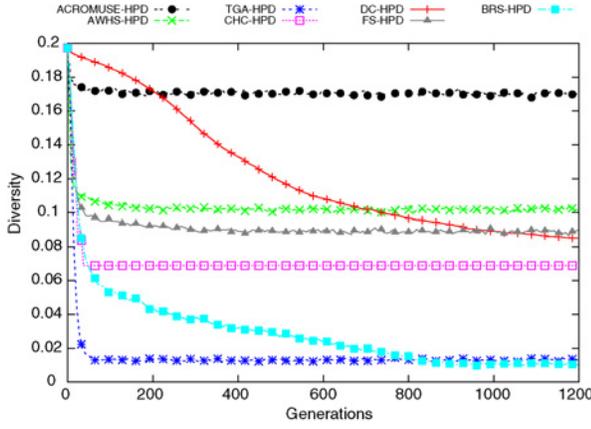


Fig. 25. M7 HPD.

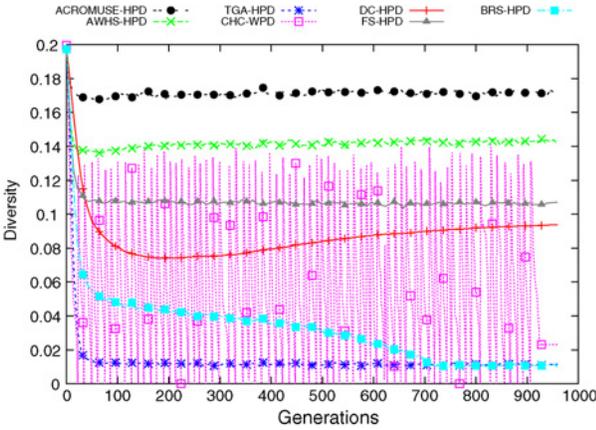


Fig. 26. M9 HPD.

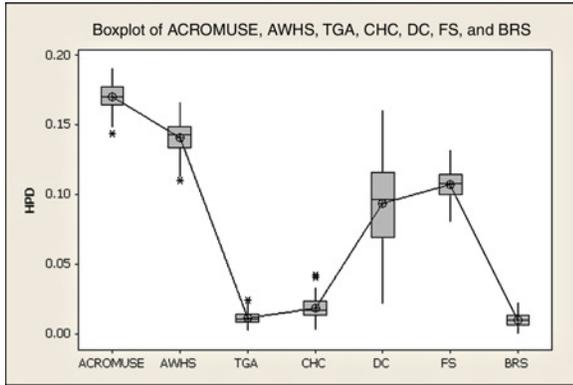


Fig. 27. M9 HPD boxplot observed from 100 samples at generation 950.

SPD and FS SPD. This may be attributed to the destructive nature of mutation in the AWHS’s exploration section of the population.

Narrow HPD ranges revealed in Fig. 27 reflect consistent healthy diversity behavior from each algorithm. ACROMUSE performs best with TGA and BRS maintaining the lowest healthy diversity. Similar to the SPD results in Fig. 24, DC exhibits the largest variability in HPD. This, as explained in the previous section, is caused by varying degrees of population convergence across GA runs.

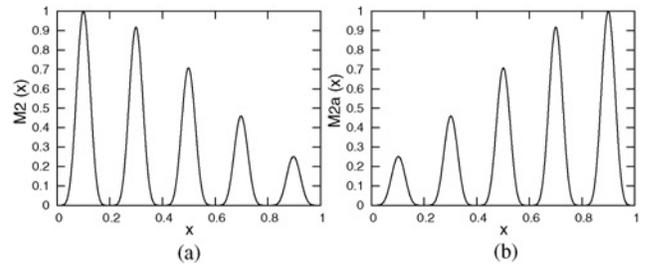


Fig. 28. (a) M2 and (b) M2a fitness landscapes.

VIII. FITNESS LANDSCAPE CHANGE RESULTS

A key motivation in designing the ACROMUSE GA is the incorporation of an ability to adapt quickly to fitness landscape change. In the real world, fitness landscape change is usually manifested as an environmental change, e.g., an evolved robot being transferred into a different environment. An example in the natural world is the difficult challenge facing the polar bear as the Arctic Sea ice melts. A population’s ability to react quickly to such environmental changes is crucial for successful survival of the species.

FS and DC both implement clever mechanisms for promoting and maintaining diversity. FS encourages exploration outside heavily populated niches while DC preserves already existing niches across generations. BRS delays convergence by slowly increasing selection pressure, while the TGA almost immediately eliminates any diversity present in the population. Aside from a very low-mutation rate, none of these algorithms include a facility for the rapid introduction of novel diversity.

The ACROMUSE and CHC algorithms differ in that they have an explicit mechanism for extensive exploration. ACROMUSE achieves this through an adjustably sized exploration section in the population. The CHC, akin to other saw-tooth GAs [35], achieves diversity introduction through a population re-initialization process.

The following results document the response of each algorithm to fitness landscape change. The first experiment deals with the M2 to M2a fitness landscape transition while the second scales up complexity by investigating an M6–M6a function transition. For the M2–M2a experiment, the time-frame of most interest is post-150 generations (at which point the fitness landscape change occurs). For the M6–M6a test, the fitness landscape change occurs at generation 250.

Since the M2 and M6 test functions possess multiple optima of differing heights, a challenge is posed to converging GAs in introducing the novel diversity required to relocate the optimal solution after a fitness landscape change.

A. M2–M2a Transition

The fitness landscapes (Fig. 28) and function definitions of the M2 and M2a benchmarks are introduced in this section. Equations (16) and (17) describe the M2 and M2a benchmark functions respectively. The real-valued variable x is restricted to the range $[0, 1]$ and is encoded using 30 bits

$$M2(x) = e^{-2(\ln 2)\left(\frac{x-0.1}{0.8}\right)^2} \sin^6(5\pi x) \quad (16)$$

$$M2a(x) = e^{-2(\ln 2)\left(\frac{0.9-x}{0.8}\right)^2} \sin^6(5\pi(1-x)). \quad (17)$$

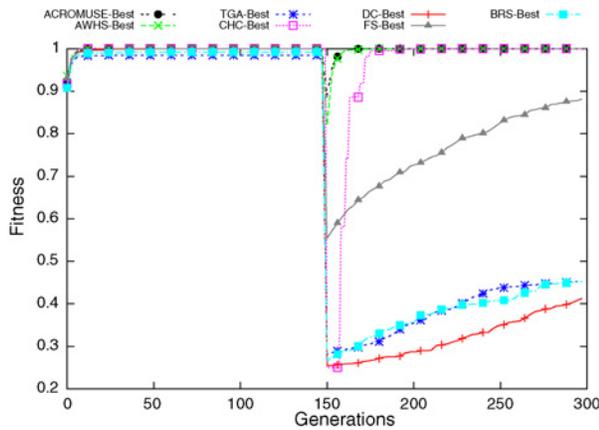


Fig. 29. M2-M2a best fitness.

1) *M2-M2a Best Fitness*: Fig. 29 illustrates the best fitness scores of each GA applied to the M2-M2a function transition. It is interesting to note that the ACROMUSE algorithm performs best in the event of fitness landscape change. This is due to the diversity of solutions already preserved in the population and the ability of the exploration section to search for new and better optima.

The TGA and BRS algorithms perform poorly after environmental change, failing to introduce the diversity required to locate better optima. DC does not fare any better. Due to the slow genetic drift inherent with the DC algorithm, particularly in landscapes where different peaks have differing heights, the converged DC population performs badly in introducing new diversity.

The converged CHC population initially has a dramatic fall in best fitness. However, a soft-restart triggered shortly after environmental change successfully introduces the diversity required for recovering the global optimum.

FS best fitness scores do not immediately suffer as much as those of the TGA, BRS, DC or CHC due to its higher diversity levels prior to the environmental change. However, FS's inability to introduce novel diversity seriously affects its capacity to relocate the global optimum.

2) *M2-M2a Average Fitness*: Fig. 30 illustrates the average fitness scores for each test algorithm applied to the M2-M2a benchmark. Rapid deterioration in average fitness occurs for all GAs immediately after fitness landscape change at generation 150. This is caused by the disruption of the high-fitness points previously occupied by each GA's population.

3) *M2-M2a SPD and HPD*: SPD and HPD scores for each test algorithm (Fig. 31) verify that the level of best fitness degradation is positively correlated with the level of diversity present in the population. DC, TGA, BRS, and CHC all suffer catastrophic immediate loss of fitness, through maintaining the lowest diversity of the GAs in question. This serves to highlight the vital importance of genetic diversity within a population.

As illustrated in Fig. 32, ACROMUSE again performs best in terms of HPD performance. This is due to the maintenance of healthy diversity through the adaptive HPD_i selection operator and through the introduction of fresh diversity by the exploration division of the population.

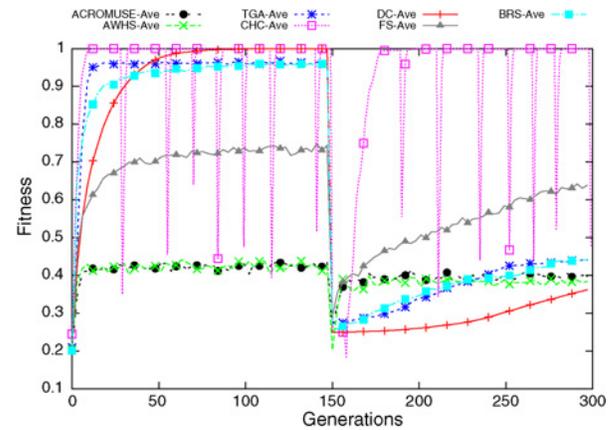


Fig. 30. M2-M2a average population fitness.

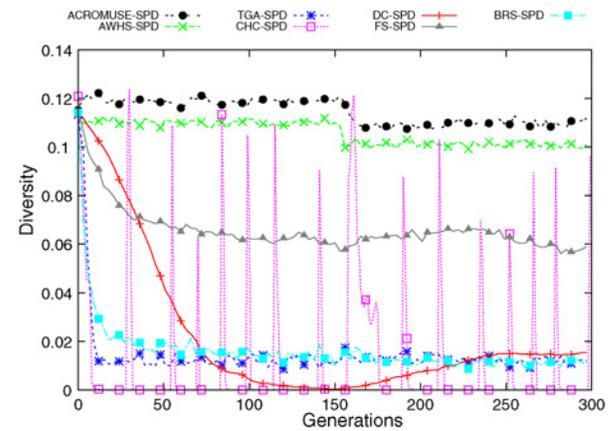


Fig. 31. M2-M2a SPD.

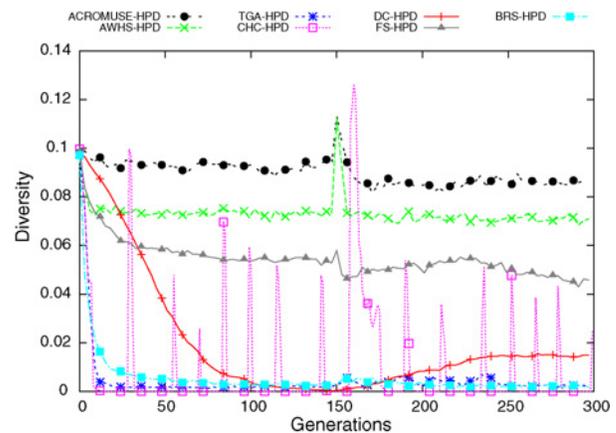


Fig. 32. M2-M2a HPD.

CHC HPD spikes are again lower than CHC SPD spikes, due to the (low-fitness) random nature of diversity being introduced during soft-restart.

4) *M2-M2a Niches Maintained*: ACROMUSE and AWHS consistently maintain the widest coverage of M2's peaks (Fig. 33). However, although ACROMUSE and AWHS maintain a similar number of peaks, their distribution of individuals is not equal. As illustrated in Fig. 32, ACROMUSE maintains higher HPD values than AWHS meaning that

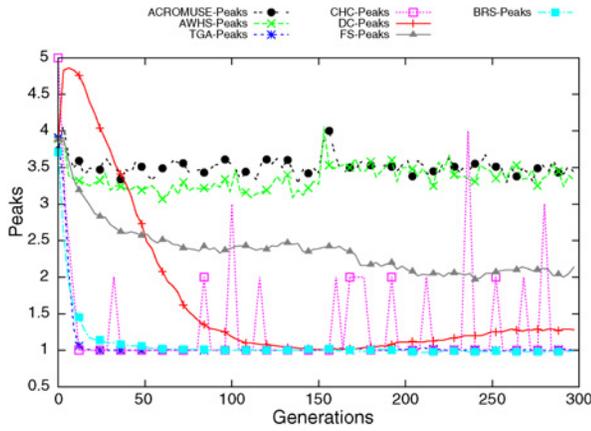


Fig. 33. M2-M2a number of niches maintained.

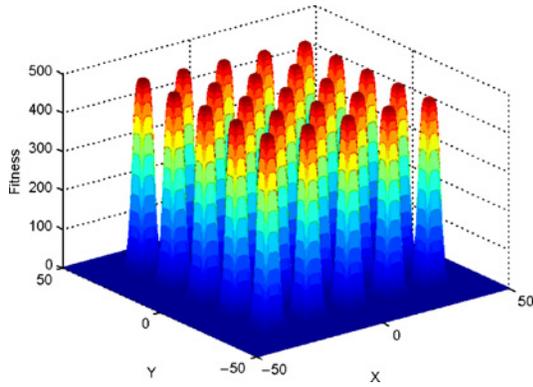


Fig. 34. M6 fitness landscape.

healthy individuals tend strongly toward more phenotypically distant peaks in the landscape.

For the M2 problem, a similar number of peaks is maintained (Fig. 33) by each algorithm compared to the M1 result (Fig. 14). A key difference is the more rapid convergence of the DC population, caused by genetic drift through crossover and selection toward the single global optimum.

FS also suffers relative to its performance on the M1 problem. FS exhibits drift toward the higher peaks. This drift of individuals is related to the fitness carrying capacity of the niches. As outlined earlier (Section VII-A5), full coverage of the landscape could be achieved by increasing population size. This research, however, is interested in developing an algorithm that performs well with small population sizes, due to the costly nature of fitness evaluation in the real world.

B. M6-M6a Transition

To investigate a more difficult example of fitness landscape change, the complexity of the task is scaled up to the M6-M6a function transition. The M6 peaks range from 476.191 to 499.002. Fig. 34 illustrates the fitness landscape for M6 while the definitions for M6 and M6a are detailed as

$$M6(x, y) = 500 - \frac{1}{0.002 + \sum_{i=0}^{24} 1/[1 + i + (x - a(i))^6 + (y - b(i))^6]} \quad (18)$$

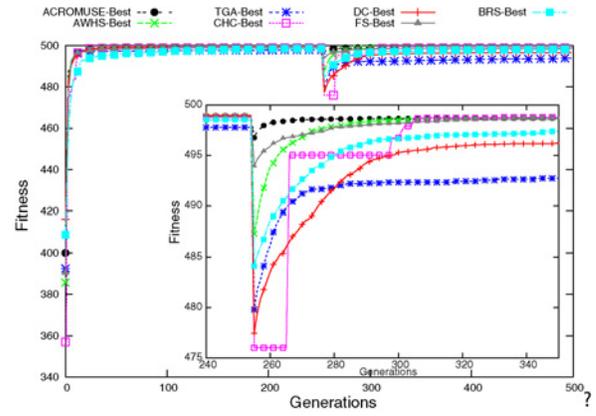


Fig. 35. M6-M6a best fitness.

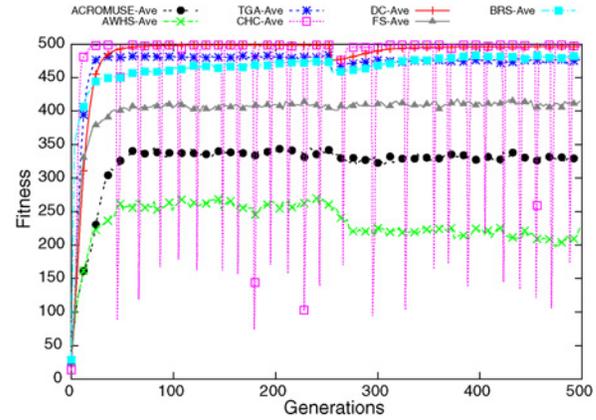


Fig. 36. M6-M6a average population fitness.

$M6a(x, y)$

$$= 500 - \frac{1}{0.002 + \sum_{i=0}^{24} 1/[1 + i + (-x - a(i))^6 + (-y - b(i))^6]} \quad (19)$$

where, $a(i) = 16[(i\%5) - 2]$ and $b(i) = 16(\lfloor i/5 \rfloor - 2)$.

1) *M6-M6a Best Fitness*: The best fitness graph (Fig. 35) includes an inset highlighting the environmental change. ACROMUSE deals best with fitness landscape change, as best-fitness is barely affected. An interesting point to note is how, though the AWHs suffers fitness degradation in the initial stage after environmental change, it quickly recovers to optimal fitness, due to the exploration section of the population. DC, BRS, and TGA all fail to rediscover the global optimum after the fitness landscape change.

2) *M6-M6a Average Fitness*: Average fitness trends (Fig. 36) reflect the same behavior exhibited earlier on simpler problems, with ACROMUSE and AWHs exhibiting the lowest average fitness scores.

3) *M6-M6a SPD and HPD*: SPD (Fig. 37) and HPD (Fig. 38) also reflect similar trends to those previously observed. DC behavior is interesting in that it experiences a sudden increase in diversity after the environmental change. This can be explained by the spreading out of individuals across the fitness landscape as they move toward the global

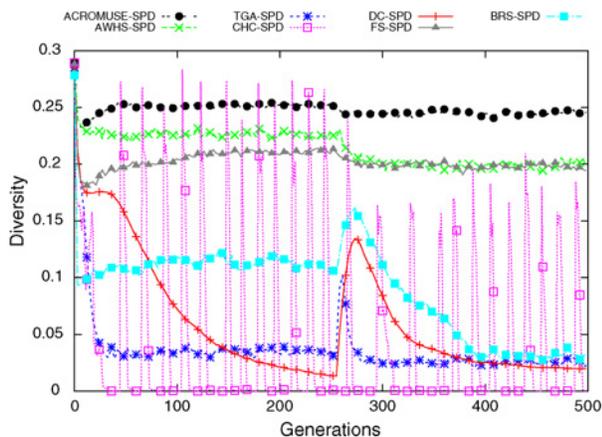


Fig. 37. M6–M6a SPD.

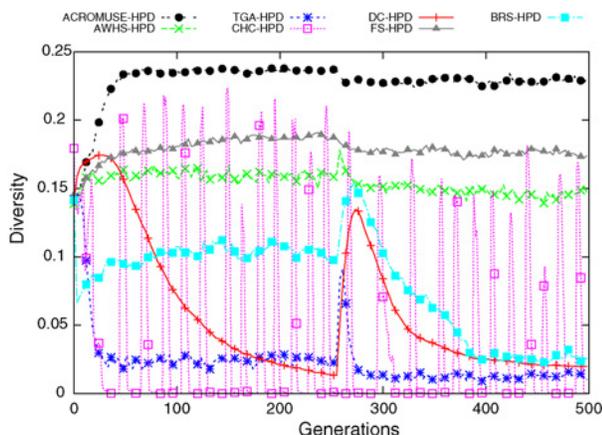


Fig. 38. M6–M6a HPD.

optimum. Once a new optimum has been located, diversity drops as individuals converge again around this new solution.

The most interesting point to note in the HPD graph (Fig. 38) is how FS maintains a higher HPD than AWHS's HPD. This can be contrasted with the SPD graph (Fig. 37) where AWHS exceeds FS SPD scores. This reversal of SPD/HPD scores between AWHS and FS can be explained by the negative affects of mutation, resulting from the exploration section of AWHS's population. This is confirmed by AWHS's low-average fitness scores.

This reversal is interesting when it is viewed in relation to best fitness degradation, directly after fitness landscape change (Fig. 35). The most interesting result from the above experiment is that the level of fitness degradation after an environmental change is not directly linked to the level of SPD in the population but instead is negatively correlated with the level of HPD. This result should assist the design of other algorithms for problems with changeable fitness landscapes.

IX. ACROMUSE OPERATOR ANALYSIS

The results presented indicate that ACROMUSE performs strongly when measured against a range of GAs on the chosen multimodal benchmark functions. However, it is not obvious

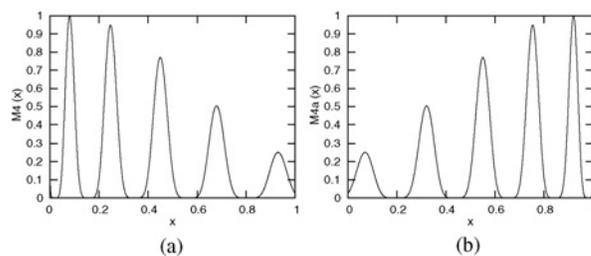


Fig. 39. (a) M4 and (b) M4a fitness landscapes.

from results whether the improvement comes from diversity introducing (adaptive mutation), or diversity maintaining (adaptive selection) operators. Section IV-C1 describes how, on implementing adaptive crossover and adaptive mutation (AC+AM), GA performance does not improve as much as expected. This is because strong nonadaptive selection keeps choosing from the same converged group of relatively high-fitness individuals in the exploitation section of the population.

Figs. 34–36 demonstrate that the key benefit of ACROMUSE is the synergy of the adaptive operators working together to achieve a balance between exploration and exploitation in offering the best results. Diversity maintenance when there is no diversity to maintain is futile, as is the introduction of diversity when such diverse individuals have no chance of survival.

The following experiment compares the various ACROMUSE adaptive operators. AC+AM are evaluated in isolation. The AC+AM algorithm employs TGA tournament selection ($Tsize = 2$). ACROMUSE and AWHS adaptive selection operators (ACROMUSE-AS and AWHS-AS) are also tested independently. ACROMUSE-AS and AWHS-AS employ traditional crossover ($P_c = 0.9$) and mutation ($P_m = 0.01$). These crossover and mutation rates are identical to those employed by the TGA in earlier results. The AC+AM, ACROMUSE-AS and AWHS-AS operator results are compared to those from the ACROMUSE and AWHS GAs.

The M4 and M4a test functions have been chosen to compare the various operators. Fig. 39 illustrates the M4 and M4a fitness landscapes while (20) and (21) define the functions

$$M4(x) = e^{-2(\ln 2)\left(\frac{x-0.08}{0.854}\right)^2} \sin^6(5\pi [x^{0.75} - 0.05]) \quad (20)$$

$$M4a(x) = e^{-2(\ln 2)\left(\frac{0.92-x}{0.854}\right)^2} \sin^6(5\pi [(1-x)^{0.75} - 0.05]). \quad (21)$$

Best fitness (Fig. 40) illustrates that ACROMUSE-AS and AWHS-AS struggle to recover diversity after environmental change. In contrast, the AC+AM, ACROMUSE and AWHS recover diversity very quickly through the exploration section of the population. The high levels of healthy diversity in the ACROMUSE and AWHS populations ensure that they perform best (smallest best fitness degradation) directly after fitness landscape change. Another interesting observation is that, although the AWHS-AS performance is not as good as that of ACROMUSE-AS directly after environmental change, AWHS-AS recovers best fitness faster than the ACROMUSE-AS. This is because the AWHS-AS selection operator selects

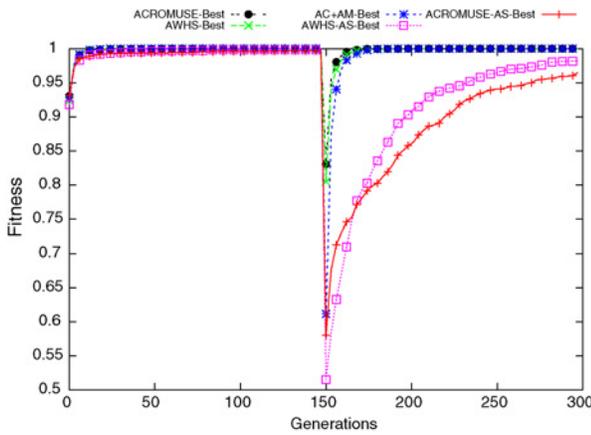


Fig. 40. Operator analysis: M4–M4a best fitness.

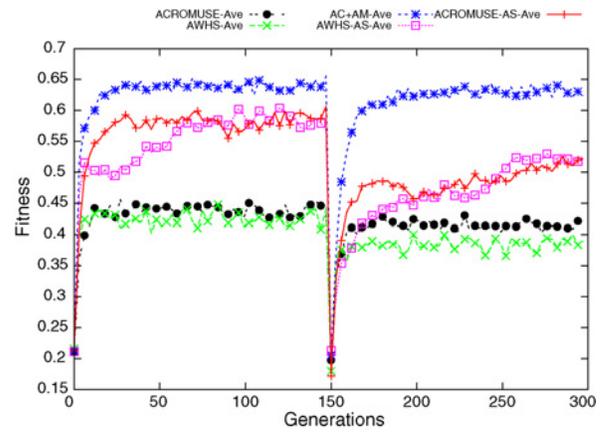


Fig. 41. Operator analysis: M4–M4a average population fitness.

according to fitness scores while ACROMUSE-AS selects according to both fitness and diversity contribution. This bias in favor of fitness assists the AWHs-AS in concentrating on potential high-fitness areas. This concentration is also reflected in the falling AWHs-AS SPD scores (Fig. 42; generation 150–300) as individuals migrate toward the higher fitness peaks.

The clear discrepancies between the behaviors of the operators applied individually and in unison demonstrate that the increased performance of the ACROMUSE algorithm comes from the balance of: 1) exploration and exploitation, and 2) diversity creation and diversity maintenance.

Average fitness results are presented in Fig. 41. AC+AM demonstrates high-average fitness as the constant strong selection pressure gives little chance of selection to low-lying or mutated individuals. This essentially negates the exploration section of the population. Much work is done by this exploration section to introduce new diversity. This diversity, however, is being immediately eradicated by strong selection pressure. The adaptive selection operators (ACROMUSE-AS and AWHs-AS) exhibit slower population convergence as low-fitness or diverse individuals have more chance of being selected. This is achieved by adaptively reducing tournament size according to HPD. The ACROMUSE and AWHs algorithms with adaptive mutation, crossover and selection combined exhibit the lowest average fitness scores as the adaptive selection pressure allows for selection of individuals from the exploration section of the population.

The SPD chart (Fig. 42) further reveals the dynamics of the algorithms. ACROMUSE and AWHs maintain the highest SPD scores, through the synergy of adaptive selection retaining diversity and adaptive mutation introducing new diversity.

The ACROMUSE-AS SPD trend demonstrates the slow convergence of the algorithm. Although selection is making an effort to maintain diversity, the inevitable genetic drift of selection, coupled with an inability to dramatically reintroduce diversity, results in a slowly converging population.

The AWHs-AS population converges a lot more quickly because its fitness-favoring selection mechanism results in quicker diversity loss. In addition, the inability to reintroduce diversity results in population convergence. An interesting observation is the rise in AWHs-AS SPD from generation 80

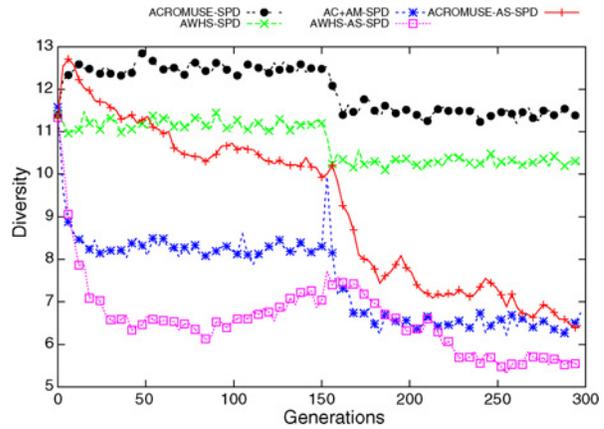


Fig. 42. Operator analysis: M4–M4a SPD.

to generation 150. This rise in diversity is a direct product of the adaptive tournament selection operator. As HPD (Fig. 43) shrinks and approaches 0.03, tournament size approaches 1. This effectively means random selection, giving individuals as a result of mutation ($P_m = 0.01$) equal chance of survival. The selection of these individuals gives rise to the increasing AWHs-AS SPD values between generations 80 and 150 (Fig. 42). This introduction of diversity is interestingly not reflected in the HPD graph (Fig. 43), meaning such mutated individuals are generally unfit.

In comparison to the SPD chart (Fig. 42), the HPD scores (Fig. 43) for the AC+AM algorithm are much lower. This is indicative of the destructive nature of mutation employed in the exploration section of the population. The ACROMUSE and AWHs algorithms maintain higher fitness-weighted healthy diversity by adapting selection pressure according to HPD. This adaptation of selection pressure allows for the protection of promising novelty in unexplored or remote areas of the landscape. This contrasts with AC+AM where novelty introduced is swiftly eliminated by the selection process. The AC+AM result echoes the statement (from Section IV-C2) that “a diverged parent population does not automatically result in diverged offspring.”

The results presented serve to demonstrate the individual and collective performance of the ACROMUSE adaptive mu-

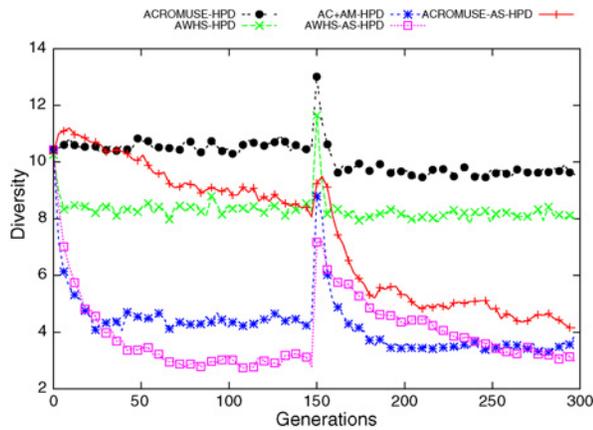


Fig. 43. Operator analysis: M4–M4a HPD.

tation, crossover and selection operators. The key contribution of ACROMUSE is the synergy of the adaptive operators working together to achieve a successful balance between: 1) exploration and exploitation, and 2) diversity creation and diversity maintenance in offering the best results.

X. CONCLUSION

This paper presents ACROMUSE; an adaptive GA which creates and maintains a diverse population of healthy individuals through adaptation of mutation, crossover and selection parameters.

The No Free Lunch theorem [58] states that “any two algorithms are equivalent when their performance is averaged across all possible problems.” This means that the goal of this research cannot be to implement a GA that will outperform other algorithms in all cases. The objective of the proposed ACROMUSE algorithm is to create and maintain a diverse population of healthy individuals, capable of adapting quickly to environmental change and well suited to the efficient navigation of highly multimodal fitness landscapes (i.e., those present in many real world problems). Results indicated that ACROMUSE achieves this objective.

For the multimodal functions employed, ACROMUSE outperforms all other investigated algorithms with regard to locating and recovering the best solution in the shortest time. With regard to diversity and healthy diversity performance, ACROMUSE achieves the highest levels of both SPD and HPD diversity types. This healthy diversity greatly assists recovery of good solutions quickly in the event of fitness landscape change. The principal contributions of this paper are summarized as follows.

- 1) *HPD measure*: By weighting each individual’s contribution to diversity according to its fitness, a measure of the population’s healthy diversity rather than standard diversity is determined. This measure provides a new method for quantifying diversity and may be used as an additional means of performance analysis for population-based heuristics. An interesting result is that the level of fitness degradation after an environmental change is not directly linked to the level of SPD in the

population but rather is negatively correlated with the level of HPD. This result should assist in the design of GAs for problems with changeable fitness landscapes.

- 2) *An adaptive tournament selection operator which regulates selection pressure, according to HPD*: This mechanism allows for the survival of outlying low-fitness individuals when the population is converged from a fitness perspective. This protection of innovation is very important for escaping local-optima. Conversely, with too much diversity, tournament size is increased to promote “survival of the fittest.” This new selection operator supports the goal of maintaining a healthy population of diverse individuals, scattered throughout the solution space.
- 3) *HPD_i selection*: By selecting individuals not according to fitness performance but to both fitness and diversity contribution, the selection mechanism can explicitly pursue both objectives of a healthy and diverse population.
- 4) *An adaptive crossover operator which divides the population into two sections, namely, an exploration section and an exploitation section*: The size of each is determined by the SPD measure. The exploration section grows as a result of low diversity, while the exploitation section expands with increased diversity. Mutation is employed as a local-search mechanism in the exploitation section of the population while it is applied with higher probability (as an explorer) in the exploration section.
- 5) *ACROMUSE*: A novel adaptive GA that combines adaptive crossover, mutation and selection operators to achieve a successful balance between exploitation and exploration. ACROMUSE’s objective is to balance these opposing forces, while targeting the creation and maintenance of a diverse population of healthy individuals. Such a diverse healthy population is capable of adapting quickly to environmental change and is well suited to the efficient navigation of highly multimodal fitness landscapes (typically present in real-world problems).
- 6) *An in depth performance analysis and comparison of several well-established GA implementations, namely: a TGA, a GA with Boltzmann selection, CHC, DC, and FS*. The SPD and HPD analysis described in this paper, coupled with average and best fitness performance results, provide insightful observations on the interplay between fitness and diversity for these algorithms.

The performance of ACROMUSE has been evaluated using multimodal function optimization benchmarks and compared to existing benchmark GA implementations. Statistically significant results demonstrate that for the multimodal benchmark functions employed, ACROMUSE outperforms other algorithms both from a fitness and diversity perspective.

Maintaining a diverse population is important for increasing population search coverage and for dealing with fitness landscape change. Results demonstrate that through maintaining high HPD, ACROMUSE copes well with environmental change, restoring better fitness scores faster than all other investigated algorithms.

Analyzing the reference GAs, results show that FS performs well from both a fitness and diversity perspective. However, FS fails to maintain the same levels of SPD or HPD as ACROMUSE. DC exhibits good fitness results, comparable to those obtained from FS, though DC diversity is lost much more steadily than with FS or ACROMUSE, due to genetic drift inherent within the algorithm. In addition to this drift, DC's inability to introduce novel diversity results in poor performance in the event of environmental change. CHC with its population re-initialization mechanism performs competitively (fitness and diversity-wise) with the other niching algorithms, particularly in introducing novel diversity after fitness landscape change. BRS demonstrates good performance in escaping local optima in the initial stages of search but can still become trapped at later generations, particularly after fitness landscape change. Indeed, BRS was not designed to maintain long-term diversity. TGA exhibits all the negative side effects associated with strong, nondynamic selection pressure and low-static mutation rates (i.e., loss of diversity/premature convergence at local optima).

Future work will investigate the application of ACROMUSE to real-world tasks, e.g., evolution of robotics controllers, where phenotypic rather than genotypic diversity will be employed to control the GA. The HPD measure introduced will be used in further experiments as a useful indicator of a population's healthy diversity.

REFERENCES

- [1] A. Eiben and J. Smith, "Multi-modal problems and spatial distribution," in *Introduction to Evolutionary Computing*. Berlin, Germany: Springer, 2003, pp. 153–172.
- [2] J. Holland, "Reproductive plans and genetic operators," in *Adaptation in Natural and Artificial Systems*, ch. 6. Cambridge, MA: MIT Press, 1992, pp. 89–118.
- [3] D. Whitley, "The GENITOR algorithm and selection pressure: Why rank-based allocation of reproductive trials is best," in *Proc. 3rd Int. Conf. Genetic Algorithms*, vol. 1. 1989, pp. 116–121.
- [4] A. Eiben, R. Hinterding, and Z. Michalewicz, "Parameter control in evolutionary algorithms," *IEEE Trans. Evol. Comput.*, vol. 3, no. 2, pp. 124–141, Jul. 1999.
- [5] K. Zhu, "A diversity-controlling adaptive genetic algorithm for the vehicle routing problem with time windows," in *Proc. 15th IEEE Int. Conf. Tools Artif. Intell.*, 2003, pp. 176–183.
- [6] R. Ursem, "Diversity-guided evolutionary algorithms," in *Proc. 7th PPSN*, vol. 2439. 2002, pp. 462–471.
- [7] H. Shimodaira, "A diversity-control-oriented genetic algorithm (DCGA): Performance in function optimization," in *Proc. Congr. Evol. Comput. (CEC)*, vol. 1. 2001, pp. 44–51.
- [8] D. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Proc. Foundations Genetic Algorithms*, vol. 1. 1991, pp. 69–93.
- [9] J. Hooker, "Testing heuristics: We have it all wrong," *J. Heuristics*, vol. 1, no. 1, pp. 33–42, 1995.
- [10] G. Greenwood, "So many algorithms. So little time," *ACM SIGSOFT Softw. Eng. Notes*, vol. 22, no. 2, pp. 92–93, 1997.
- [11] S. Mahfoud, "Niching methods for genetic algorithms," Ph.D. dissertation, Dept. General Eng., Univ. Illinois at Urbana-Champaign, Champaign, IL, 1996.
- [12] K. De Jong, "Analysis of the behavior of a class of genetic adaptive systems," Ph.D. dissertation, Dept. Comput. Commun. Sci., Univ. Michigan, Ann Arbor, 1975.
- [13] L. Chambers, "The Boltzman selection procedure," in *Practical Handbook of Genetic Algorithms: New Frontiers*, vol. 2, ch. 5. Boca Raton, FL: CRC Press, 1995, pp. 111–139.
- [14] L. Eshelman, "The CHC adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination," in *Proc. Foundations Genetic Algorithms*, vol. 1. 1991, pp. 265–283.
- [15] D. Goldberg and J. Richardson, "Genetic algorithms with sharing for multimodal function optimization," in *Proc. 2nd Int. Conf. Genetic Algorithms Their Applicat.*, 1987, pp. 41–49.
- [16] I. Rechenberg, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart, Germany: Frommann-Holzboog, 1973.
- [17] D. Liu and S. Feng, "A novel adaptive genetic algorithms," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 1. 2004, pp. 414–416.
- [18] P. Palmes, T. Hayasaka, and S. Usui, "Mutation-based genetic neural network," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 587–600, May 2005.
- [19] R. Krohling, Y. Zhou, and A. Tyrrell, "Evolving FPGA-based robot controllers using an evolutionary algorithm," in *Proc. 1st Int. Conf. Artif. Immune Syst.*, 2002, pp. 41–46.
- [20] D. Thierens, "Adaptive mutation rate control schemes in genetic algorithms," in *Proc. Congr. Evol. Comput.*, 2002, pp. 980–985.
- [21] H. Hagnas, A. Pounds-Cornish, M. Colley, V. Callaghan, and G. Clarke, "Evolving spiking neural network controllers for autonomous Robots," in *Proc. IEEE Int. Conf. Robotics Automat.*, vol. 5. 2004, pp. 4620–4626.
- [22] J. Zhang, H. Chung, and W. Lo, "Clustering-based adaptive crossover and mutation probabilities for genetic algorithms," *IEEE Trans. Evol. Comput.*, vol. 11, no. 3, pp. 326–335, Jun. 2007.
- [23] S. Yuen and C. Chow, "A genetic algorithm that adaptively mutates and never revisits," *IEEE Trans. Evol. Comput.*, vol. 13, no. 2, pp. 454–472, Apr. 2009.
- [24] M. Srinivas and L. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 4, pp. 656–667, Apr. 1994.
- [25] Z. Liu, J. Zhou, and S. Lai, "New adaptive genetic algorithm based on ranking," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 3. 2003, pp. 1841–1844.
- [26] T. Kuo and S. Hwang, "A genetic algorithm with disruptive selection," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 26, no. 2, pp. 299–307, Apr. 1996.
- [27] M. Affenzeller and S. Wagner, "Offspring selection: A new self-adaptive selection scheme for genetic algorithms," in *Proc. Int. Conf. Adapt. Natural Comput. Algorithms*, 2005, pp. 218–221.
- [28] A. Eiben, A. de Wilde, and M. Schut, "Boosting genetic algorithms with self-adaptive selection," in *Proc. IEEE Conf. Evol. Comput.*, 2006, pp. 477–482.
- [29] A. Eiben, M. Schut, and A. de Wilde, "Is self-adaptation of selection pressure and population size possible? A case study," in *Proc. 9th Parallel Problem Solving Nature (PPSN)*, vol. 4193. 2006, pp. 900–909.
- [30] P. Vajda, A. Eiben, and W. Hordijk, "Parameter control methods for selection operators in genetic algorithms," in *Proc. 10th Parallel Problem Solving Nature (PPSN)*, 2008, pp. 620–630.
- [31] A. Cioppa, C. De Stefano, and A. Marcelli, "Where are the niches? Dynamic fitness sharing," *IEEE Trans. Evol. Comput.*, vol. 11, no. 4, pp. 453–465, Aug. 2007.
- [32] S. Mahfoud, "Crowding and preselection revisited," in *Proc. 2nd Parallel Problem Solving Nature (PPSN)*, 1992, pp. 27–36.
- [33] C. Jassadaporn and P. Chongstitvatana, "Diversity control to improve convergence rate in genetic algorithms," in *Proc. 4th Int. Conf. Intell. Data Eng. Automat. Learn. (IDEAL)*, 2003, pp. 421–425.
- [34] M. Lozano, F. Herrera, and J. Cano, "Replacement strategies to maintain useful diversity in steady-state genetic algorithms," *Soft Comput.: Methodologies Applicat.*, vol. 32, pp. 85–96, May 2006.
- [35] V. Koumousis and C. Katsaras, "A saw-tooth genetic algorithm combining the effects of variable population size and reinitialization to enhance performance," *IEEE Trans. Evol. Comput.*, vol. 10, no. 1, pp. 19–28, Feb. 2006.
- [36] M. Rocha and J. Neves, "Preventing premature convergence to local optima in genetic algorithms via random offspring generation," in *Proc. 12th Int. Conf. Ind. Eng. Applicat. Artif. Intell. Expert Syst.*, vol. 1611. 1999, pp. 127–136.
- [37] M. Laumanns, L. Thiele, K. Deb, and E. Zitzler, "Combining convergence and diversity in evolutionary multiobjective optimization," *Evol. Comput.*, vol. 10, no. 3, pp. 263–282, 2002.
- [38] T. Back and F. Homeister, "Extended selection mechanisms in genetic algorithms," in *Proc. 4th Int. Conf. Genetic Algorithms*, 1991, pp. 92–99.
- [39] D. Curran, C. O'Riordan, and H. Sorensen, "The effects of lifetime learning on the diversity and fitness of populations," in *Proc. 9th Ann. Conf. Genetic Evol. Comput.*, 2007, p. 337.
- [40] J. Grefenstette, "Optimization of control parameters for genetic algorithms," *IEEE Trans. Syst., Man, Cybern.*, vol. 16, no. 1, pp. 122–128, Jan. 1986.

- [41] J. Schaffer, R. Caruana, L. Eshelman, and R. Das, "A study of control parameters affecting online performance of genetic algorithms for function optimization," in *Proc. 3rd Int. Conf. Genetic Algorithms*, 1989, pp. 51–60.
- [42] T. Blickle and L. Thiele, "A mathematical analysis of tournament selection," in *Proc. 6th Int. Conf. Genet. Algorithms*, 1995, pp. 9–16.
- [43] D. Goldberg, K. Deb, and J. Horn, "Massive multimodality, deception, and genetic algorithms," in *Proc. 2nd Parallel Problem Solving Nature (PPSN)*, vol. 51. 1992, p. 61 801.
- [44] X. Du, Y. Li, W. Chen, Y. Zhang, and D. Yao, "A Markov random field based hybrid algorithm with simulated annealing and genetic algorithm for image segmentation," in *Advances in Natural Computation*, LNCS 4221. Berlin, Germany: Springer-Verlag, 2006, pp. 706–715.
- [45] K. Srinivasa, K. Venugopal, and L. Patnaik, "A self-adaptive migration model genetic algorithm for data mining applications," *Inf. Sci.*, vol. 177, no. 20, pp. 4295–4313, 2007.
- [46] D. Zhang, Q. Chen, and L. Wei, "Building behavior scoring model using genetic algorithm and support vector machines," in *Proc. Int. Conf. Comput. Sci. (ICCS)*, LNCS 4488. 2007, pp. 482–485.
- [47] R. Poli, "Tournament selection, iterated coupon-collection problem, and backward-chaining evolutionary algorithms," in *Proc. Foundations Genetic Algorithms*, vol. 3469. 2005, pp. 132–155.
- [48] K. Deb and D. Goldberg, "An investigation of niche and species formation in genetic function optimization," in *Proc. 3rd Int. Conf. Genetic Algorithms*, 1989, pp. 42–50.
- [49] J. Horn and D. Goldberg, "Genetic algorithm difficulty and the modality of fitness landscapes," in *Proc. 3rd Foundations Genetic Algorithms*, 1994, pp. 243–269.
- [50] R. Brits, A. Engelbrecht, and F. Van den Bergh, "Locating multiple optima using particle swarm optimization," *Appl. Math. Comput.*, vol. 189, no. 2, pp. 1859–1883, 2007.
- [51] G. Singh, "Comparison of multimodal optimization algorithms based on evolutionary algorithms," in *Proc. 8th Ann. Conf. Genetic Evol. Comput.*, 2006, pp. 1305–1312.
- [52] S. Bird and X. Li, "Adaptively choosing niching parameters in a PSO," in *Proc. 8th Ann. Conf. Genetic Evol. Comput.*, 2006, pp. 3–10.
- [53] G. Dick, P. Whigham, and N. Dunedin, "A weighted local sharing technique for multimodal optimization," in *Proc. 7th Int. Conf. Simulated Evol. Learn.*, 2008, pp. 452–461.
- [54] D. Beasley, D. Bull, and R. Martin, "A sequential niche technique for multimodal function optimization," *Evol. Comput.*, vol. 1, no. 2, pp. 101–125, 1993.
- [55] D. Montgomery, "Experiments with a single factor: The analysis of variance," in *Design and Analysis of Experiments*, ch. 3. New York: Wiley, 2006, pp. 60–114.
- [56] W. Kruskal and W. Wallis, "Use of ranks in one-criterion variance analysis," *J. Am. Stat. Assoc.*, vol. 47, no. 260, pp. 583–621, 1952.
- [57] ACROMUSE-Statistical-Analysis-Results [Online]. Available: <http://sites.google.com/site/brianmcginley/>
- [58] D. Wolpert and W. Macready, "No free lunch theorems for search," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.



Brian Mc Ginley received the B.Eng. (with Hons.) degree in electronic and computer engineering, and the Ph.D. degree from National University of Ireland (NUI), Galway, Ireland, in 2003 and 2009, respectively. His Ph.D. research investigated spiking neural networks and genetic algorithm parameter adaptation.

Since 2008, he is working with NUI, as a Post-Doctoral Researcher with the efficient embedded digital signal processing (EEDSP) research cluster. His current research interests include heuristic algorithms, embedded systems, reconfigurable hardware, and digital signal processing applications for mobile digital healthcare.



John Maher received the B.Eng. (with Hons.) in electronic and electrical engineering from Dublin City University, Dublin, Ireland, in 1999, and the M.Eng. in electronic engineering from the National University of Ireland, Galway, Ireland, in 2002, where he is currently working toward the Ph.D. degree in the area of cellular development and reconfigurable hardware bio-inspired systems.



Colm O'Riordan is with the Computational Intelligence Research Group, National University of Ireland, Galway, Ireland, as a Lecturer. His main research interests are in the fields of artificial life, evolutionary computation and information retrieval. He has published over 100 conference and journal papers in these areas. His current research interests include cooperation in artificial life societies and multiagent systems.



Fearghal Morgan received the B.S. (with Hons.) degree in electrical and electronic engineering, and the Ph.D. degree from Queens University, Belfast, U.K., in 1981 and 1986, respectively. His Ph.D. research investigated Buried Channel MOSFET design, fabrication, and characterisation.

He is the Director of the Bio-Inspired and Reconfigurable Computing Research Group, Electrical and Electronic Engineering, National University of Ireland (NUI), Galway, Ireland. He was a Design Engineer of network and router products with Digital Equipment Corporation, Ireland, for seven years. He has been with NUI, since 1996, following four years with the Institute of Technology Tallaght, Dublin, Ireland.