

An Effective Real-Parameter Genetic Algorithm with Parent Centric Normal Crossover for Multimodal Optimisation

Pedro J. Ballester and Jonathan N. Carter

Imperial College London, Department of Earth Science and Engineering,
RSM Building, Exhibition Road, London SW7 2AZ, UK
{p.ballester,j.n.carter}@imperial.ac.uk

Abstract. Evolutionary Algorithms (EAs) are a useful tool to tackle real-world optimisation problems. Two important features that make these problems hard are multimodality and high dimensionality of the search landscape.

In this paper, we present a real-parameter Genetic Algorithm (GA) which is effective in optimising high dimensional, multimodal functions. We compare our algorithm with two previously published GAs which the authors claim gives good results for high dimensional, multimodal functions. For problems with only few local optima, our algorithm does not perform as well as one of the other algorithm. However, for problems with very many local optima, our algorithm performed significantly better. A wider comparison is made with previously published algorithms showing that our algorithm has the best performance for the hardest function tested.

1 Introduction

Many real-world optimisation problems, particularly in engineering design, have a number of key features in common: the parameters are real numbers; there are many of these parameters; and they interact in highly non-linear ways, which leads to many local optima in the objective function. Clearly it is useful to have optimisers that are effective at solving problems with these characteristics. It has been shown elsewhere [1] that Genetic Algorithms (GAs) are good at solving multimodal functions. In this work, we describe and demonstrate a GA that appears to be good at solving problems where the objective function is characterised as being: high dimensional; real variable; continuous and smooth; many local optima.

Deb et al. [2] have recently produced a comprehensive review of optimisation methods. They included in their study: real parameter GAs, Self-Adaptative ESs (Evolution Strategies), DE (Differential Evolution) and GMs (Gradient Methods). All of these methods were tested on a set of high dimensional analytical test functions. They concluded that a real-parameter GA known as G3-PCX had the best overall performance. It is worth noting that G3-PCX was shown

to have better convergence than GMs on some unimodal functions. G3-PCX also obtained the best results, in the published literature, for the Schwefel and Rosenbrock functions. An excellent result on the Rastrigin function was also reported.

Ballester and Carter [1] investigated the performance of a variety of GA formulations over a set of 2-variable multimodal functions. It was found that GAs which used random selection with crowding replacement strategies were robust optimisers. The same authors showed [3] that one of those GAs (named SPC-vSBX) was also effective in optimising high dimensional real-variable functions. Of particular importance were the results obtained with the Rastrigin and rotated Rastrigin functions, the hardest of all tested in terms of number of local minima. On these functions, SPC-vSBX achieved the best performance, starting with a skewed initialisation, reported in the literature. A version of SPC-vSBX has been also successful in a real-world optimisation problem [4]. The algorithm was applied to the direct inversion of a synthetic oil reservoir model with three free parameters. In experiments where the optimal model was a priori known, it was observed that the algorithm was able to find the global minimum

In this paper, a new family of crossovers is presented which, unlike vSBX, are not biased with respect to the coordinate directions. One of this crossovers will be studied in combination with the SPC model. A benchmark will be set up to test the behaviour of the new algorithm together with G3-PCX and SPC-vSBX. This benchmark consists of a set of analytical test functions that are known to be difficult to many optimisation algorithms. These functions have been widely used, which will allow a wider comparison with previously published studies.

We arrange the rest of the paper as follows. Section 2 discusses the approach to testing algorithms. Section 3 describes the structure of the proposed GA. In Sect. 4, the experimental setup is explained. Results are presented and a comparison with G3-PCX and SPC-vSBX made in Sect. 5. Section 6 reviews the presented results with respect to past studies. Lastly, we present our conclusions and discuss future work in Sect. 7.

2 Testing Algorithms on Analytical Functions

When tackling a real-world problem, the conventional approach is to test first the algorithm on a set of analytical functions. Most real-world applications involve objective functions considerably more expensive to evaluate than an analytical one. Consequently, it is usually unviable to test the effectiveness of an algorithm directly on the real problem. There is the assumption that the chosen set of functions share some characteristics with the target problem. Based on this assumption, one applies the algorithm that performed well on the benchmark in the expectation that it will do also well on the real-world problem.

This approach to evaluating algorithms is not without drawbacks. As pointed out by Whitley et al. [5], there is the potential danger that algorithms ‘become overfitted to work well on benchmarks and therefore that good performance on benchmarks does not generalize to real world problems’. An example of this is

algorithms that exploit benchmark symmetries unlikely to be present on the target problem. For instance, many algorithms are tested by initialising the population symmetrically around the global optimum. The algorithm might have an inherent tendency to create children near the centroid of the parents (eg. mean-centric recombination in GAs). Deb et al. [2] argued that this is unfair since: ‘a mean-centric recombination of two solutions at either side of $x_j = 0$ is likely to result in a children near $x_j = 0$. Moreover, in most real-world problems, the knowledge of the exact optimum is usually not available, and the performance of an Evolutionary Algorithm (EA) on a symmetric initialisation may not represent the EA’s true performance in solving the same problem with a different initialisation or other problems’. Consequently it is important that algorithms are tested with skewed initialisations, so as to give a better indication of their performance on real-world problems. There are a number of additional studies that have also pointed out the need of using a skewed initialisation [6, 7, 8, 9, 10, 3]. In this work, we have chosen a skewed initialisation that does not bracket the global minimum. This has been done to test the algorithms under the hardest situation. However, we feel that an initialisation bracketing the global minimum, but in a sufficiently asymmetrical way, is also a valid approach.

A different example of a bias that results in improved performance is given by Ballester and Carter [3]. In that study, the vSBX crossover was used. vSBX has a preference for searching along the coordinate directions. It was pointed out that this may give the GA an advantage on test functions with minima aligned with the axis. The latter is a property of the Rastrigin function. The algorithm success in solving a 50-variable Rastrigin could have benefitted from this characteristic. Consequently, the authors introduced a rotation in the function to neutralise the algorithm’s advantage. A significantly inferior performance on the rotated Rastrigin was reported.

3 GA Description

Our real parameter GA uses a steady state population model. In each generation, two parents are selected from the current population to produce λ children through crossover. Offspring and current populations are then combined so that the population remains at a constant size.

This GA combines the following features: parental selection is not fitness biased, a self-adaptative unbiased crossover operator, implicit elitism and locally scaled probabilistic replacement. We will refer to this GA as SPC-PNX (Scaled Probabilistic Crowding Genetic Algorithm with Parent Centric Normal crossover). Below we describe the details of SPC-PNX selection, replacement and crossover schemes:

3.1 Selection

We use uniform random selection, without replacement, to select two parents from the current population. Unusually for a GA, fitness is not taken into account during the selection process.

3.2 Scaled Probabilistic Crowding Replacement

We use a scaled probabilistic crowding scheme for our replacement policy. First, NREP individuals from the current population are selected at random. These individuals then compete with the offspring for a place in the population.

In the probabilistic crowding scheme [11], the closest preselected individual (\mathbf{x}^{cst}) enters a probabilistic tournament with the offspring (\mathbf{x}^{ofp}), with culling likelihoods (survival, if we were in a maximisation problem) given by

$$p(\mathbf{x}^{ofp}) = \frac{f(\mathbf{x}^{ofp})}{f(\mathbf{x}^{ofp}) + f(\mathbf{x}^{cst})}, \quad p(\mathbf{x}^{cst}) = \frac{f(\mathbf{x}^{cst})}{f(\mathbf{x}^{ofp}) + f(\mathbf{x}^{cst})} . \quad (1)$$

where $f(\mathbf{x})$ is the objective function value for an individual \mathbf{x} .

If the differences in function values across the population are small with respect to their absolute values, these likelihoods would be very similar in all cases. The scaled probabilistic crowding replacement is introduced to avoid this situation. It operates with culling likelihoods

$$p(\mathbf{x}^{ofp}) = \frac{f(\mathbf{x}^{ofp}) - f_{best}}{f(\mathbf{x}^{ofp}) + f(\mathbf{x}^{cst}) - 2f_{best}}, \quad p(\mathbf{x}^{cst}) = \frac{f(\mathbf{x}^{cst}) - f_{best}}{f(\mathbf{x}^{ofp}) + f(\mathbf{x}^{cst}) - 2f_{best}} . \quad (2)$$

where f_{best} is the function value of the best individual in the offspring and selected group of NREP individuals.

This replacement scheme has several beneficial features. The fittest individual does not always win, which helps to prevent premature convergence. Crowding schemes such as this promote the creation of subpopulations that explore different regions of the search space. This has been shown [1] [3] to be beneficial for creating multiple optimal solutions and to increase the effectiveness in finding the global minimum. It implements elitism in an implicit way. If the best individual in either offspring or current parent population enters this replacement competition will have probability zero of being culled.

3.3 Crossovers

In this work, we test two crossovers in combination with the SPC model: a version of the Simulated Binary Crossover (SBX) [12] [13] called vSBX [1] [3] and a new crossover called PNK. These crossovers are self-adaptative in the sense that the spread of the possible offspring solutions depends on the distance between the parents, which decreases as the population converge.

In SBX, children have zero probability of appearing in some regions of the parameter space, as shown in Fig 1. vSBX does not exclude any regions, while preserving the good SBX properties. This may allow a better exploration of the search space. It should be noted that SBX and vSBX preferentially search along the coordinate directions. This may give an advantage on test functions where minima are aligned along coordinate directions.

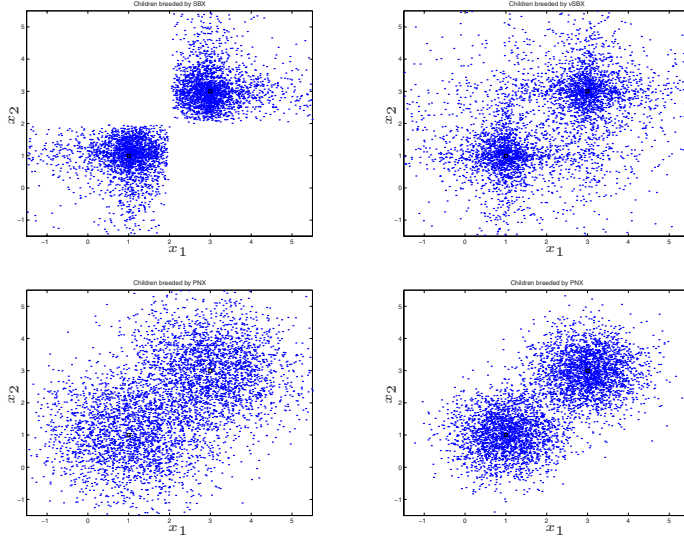


Fig. 1. Children bred from parents $\mathbf{x}^{(1)} = (1, 1)$ and $\mathbf{x}^{(2)} = (3, 3)$ for (clockwise starting from upper left plot) a) SBX ($\eta = 1$), b) vSBX ($\eta = 1$), c) PNX ($\eta = 3$) and d) PNX ($\eta = 2$)

Like vSBX, PNX does not exclude any regions, while creating offspring close to the parents. However, unlike SBX and vSBX, PNX does not preferentially search along the axis and hence it is not biased towards coordinate directions. In PNX, for each of the λ children, we proceed as follows to determine its j^{th} gene (y_j). First, we draw a single random number $w \in [0, 1]$, we use the form $y_j^{(1)}$ if $w < 0.5$ and $y_j^{(2)}$ if $w \geq 0.5$. Once this choice is made, the same selected form is used for every component j . The forms are

$$y_j^{(1)} = N(x_j^{(1)}, |x_j^{(2)} - x_j^{(1)}|/\eta), \quad y_j^{(2)} = N(x_j^{(2)}, |x_j^{(2)} - x_j^{(1)}|/\eta) . \quad (3)$$

where $N(\mu, \sigma)$ is a random number drawn from a gaussian distribution with mean μ and standard deviation σ , $x_j^{(i)}$ is the j^{th} component of the i^{th} parent and η is a tunable parameter. The larger is the value of η the more concentrated is the search around the parents.

4 Experimental Setup

We use the same experimental setup as in Deb et al. [2], allowing a direct comparison with their results. The stopping criteria are: either a maximum of 10^6 function evaluations or an objective value of 10^{-20} is obtained.

Our benchmark consists in six analytical 20-variable functions: ellipsoidal (f_{elp}), Schwefel (f_{sch}), Generalized Rosenbrock (f_{ros}), Ackley (f_{ackl}), Rastrigin

(f_{rtg}) and a rotated Rastrigin function (f_{rrtg}). Views of the two-dimensional versions of these functions are given in Figs. 2 to 7.

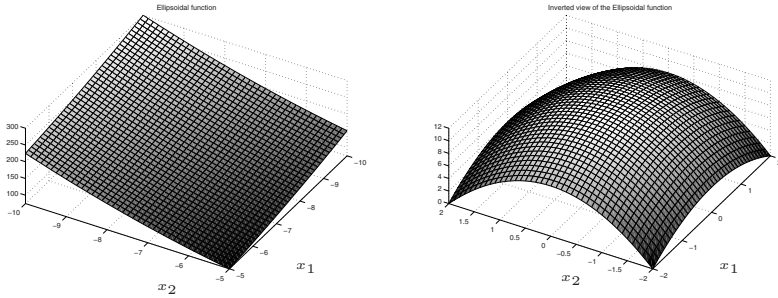


Fig. 2. Initialisation (left) and global minimum (right, inverted view) for the Ellipsoidal function. $f_{elp}(\mathbf{x}) = \sum_{j=1}^M jx_j^2$

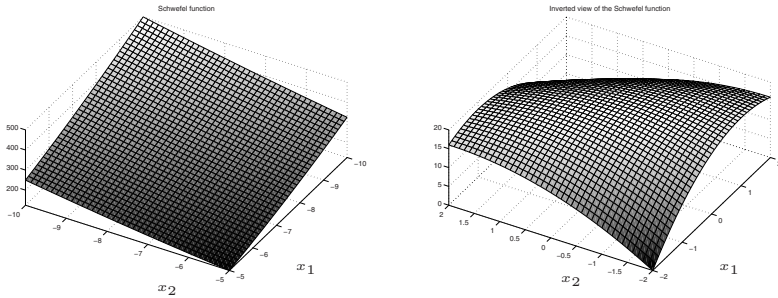


Fig. 3. Initialisation (left) and global minimum (right, inverted view) for the Schwefel function: $f_{sch}(\mathbf{x}) = \sum_{j=1}^M \left(\sum_{k=1}^j x_k \right)^2$

These functions were selected for several reasons. First, they have been widely used, which will allow an extensive comparison with previously published algorithms. Also, these functions have a number of features that are known to be hard for optimisation algorithms and believed to be present in many real-world problems. The Ellipsoidal is a unimodal function with different weights for each variable. This will serve to test the algorithms with a badly scaled objective function. The Schwefel function is also unimodal, but its variables are correlated. The Generalized Rosenbrock has been regarded as a unimodal function, but there is

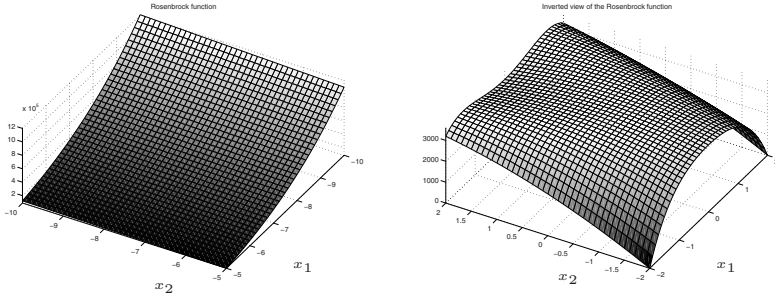


Fig. 4. Initialisation (left) and global minimum (right, inverted view) for the Rosenbrock function: $f_{ros}(\mathbf{x}) = \sum_{j=1}^{M-1} (100(x_j^2 - x_{j+1})^2 + (x_j - 1)^2)$

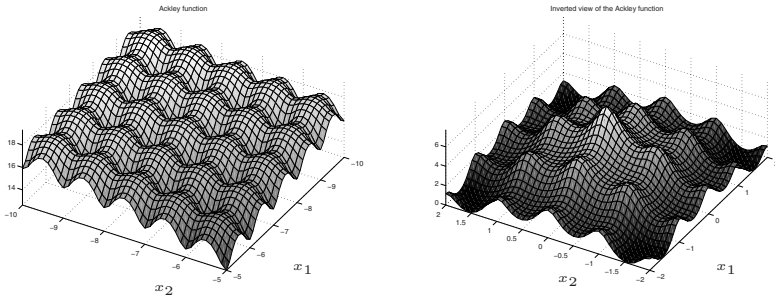


Fig. 5. Initialisation (left) and global minimum (right, inverted view) for the Ackley function: $f_{ackl}(\mathbf{x}) = 20 + e - 20 \exp(-0.2 \sqrt{\frac{1}{M} \sum_{j=1}^M x_j^2}) - \exp(\frac{1}{M} \sum_{j=1}^M \cos(2\pi x_j))$

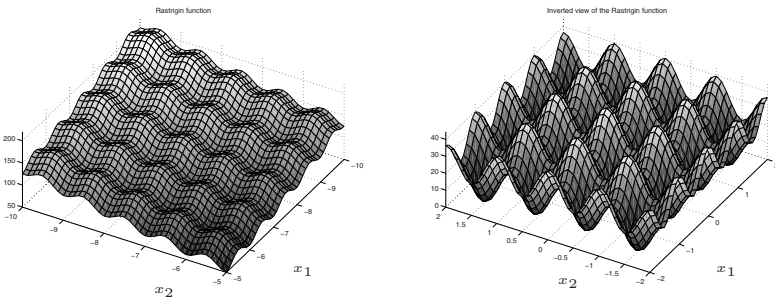


Fig. 6. Initialisation (left) and global minimum (right, inverted view) for the Rastrigin function: $f_{rtg}(\mathbf{x}) = 10M + \sum_{j=1}^M (x_j^2 - 10 \cos(2\pi x_j))$

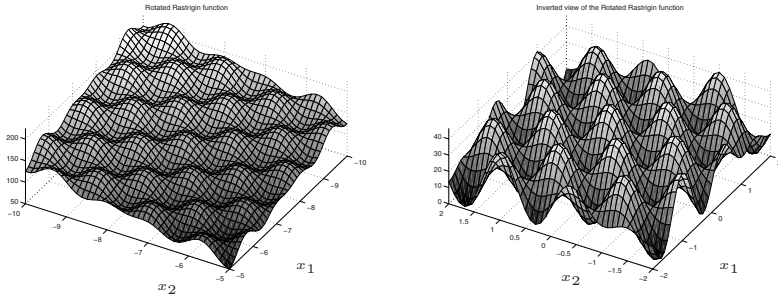


Fig. 7. Initialisation (left) and global minimum (right, inverted view) for the Rotated Rastrigin function: $f_{rrtg}(\mathbf{y}) = 10M + \sum_{j=1}^M (y_j^2 - 10 \cos(2\pi y_j))$, $\mathbf{y} = \mathbf{A}\mathbf{x}$ with $A_{j,j} = 4/5$, $A_{j,j+1} = 3/5$ (j odd), $A_{j,j-1} = -3/5$ (j even), $A_{j,k} = 0$ (the rest)

evidence [2] suggesting that it contains several minima in high dimensional instances. This will test the behaviour of the algorithm with objective functions having a couple of minima and an almost flat region near the global. The Ackley function is highly multimodal. The basin of these local minima increase in size as one moves away from the global minimum, as discussed by [8]. Thus, this function will be useful to study the behaviour of the algorithms when initialised in a highly multimodal region. In the Rastrigin function, the opposite is observed. Away from the global minimum, the landscape has a parabolic structure. As we move towards the global minimum, the size of the basins increase [8]. Therefore, an algorithm has to discard many local minima of similar quality before reaching the global minimum. This is known to be difficult for many optimisation algorithms, specially in high dimensional Rastrigin instances. Lastly, a rotation is carried out on the Rastrigin to make it non-separable, while still being highly multimodal. The resulting rotated function has no longer local minima arranged along the axis. The rotated Rastrigin function is expected to help to avoid overestimating the performance of algorithms using separable objective functions. All functions have a single global minimum with value zero. The global minimum is located at $x_j = 1$ (Rosenbrock) or $x_j = 0$ (the rest).

As an experiment has some dependence on the initial conditions, we repeat each experiment, each time with a different initial population. We do not initialise the population symmetrically around the global minimum, all variables are initialised at random within $[-10, -5]$. The purpose of this skewed initialisation is two fold. First, it ensures that the algorithms generally have to overcome a number of local minima before reaching the global minimum. Second, it neutralises the advantage enjoyed by algorithms that have an inherent tendency to create solutions near the centroid of the parents.

5 Discussion of the Results

SPC-vSBX and SPC-PNX contain four tunable parameters: N , λ , NREP and η . In this study, we fix $\eta = 0.01$ (for vSBX), $\eta = 2.0$ (for PNX) and NREP=2. G3-PCX's results for the Ellipsoidal, Schwefel, Generalized Rosenbrock and Rastrigin are extracted from the original study [2]. For the rest of functions, we use the G3-PCX code downloaded from the KanGAL website [14]. The procedure consists in doing some preliminary runs to determine the best N and λ for each function. Due to the limited computing precision, the accuracy for the Ackley function was set as 10^{-10} .

In Table 1, we compare G3-PCX, SPC-vSBX and SPC-PNX. G3-PCX reports the best results in the literature for the Schwefel and Generalised Rosenbrock. It also obtained an excellent result for the Rastrigin starting with a skewed initialisation. For the Ellipsoidal, only a Gradient Method (the BFGS quasi-Newton algorithm with a mixed quadratic-cubic polynomial line search approach achieved a solution in the order of 10^{-24} in 6,000 function evaluations [2]) was shown to outperform it. Over the unimodal functions SPC-vSBX and SPC-PNX are not competitive in terms of number of function evaluations, although they reached the required accuracy in all runs.

In the 20-variable Rosenbrock function, there are two known local minima [2] with function values of 3.986624 and 65.025362. G3-PCX found solutions better than 10^{-20} in 36 out of 50 runs, but in the other 14 got stuck in the best local minimum. SPC-vSBX only found a best solution of 10^{-4} in 50 runs. It found solutions below the best local minimum (ie. within the global basin) in 48 out of 50 runs. SPC-PNX found a best solution of 10^{-10} in 50 runs. It found solutions below the best local minimum in 38 out of 50 runs. By incrementing N , SPC-PNX reached the global basin in 47 out of 50 runs. Since the SPC model is not strongly fitness biased, we conjecture that the slow convergence observed is due to the function's flat regions.

In the 20-variable Ackley function, G3-PCX was not able to find the global basin in any of the ten runs, finding a best value of 3.959. In most of the runs, the algorithm could not escape the highly multimodal initialisation region. Whereas SPC-vSBX and SPC-PNX found the global minimum in all runs, with SPC-PNX outperforming SPC-vSBX in terms of required number of evaluations.

In the 20-variable Rastrigin function, SPC-vSBX could find a solution better than 10^{-20} in 6 out of 10 runs, whereas G3-PCX was reporting an overall best solution of 15.936 within the prescribed limits. In the other 4 runs, SPC-vSBX always found one of the best local minima with value 0.9949591. However, the authors of this algorithm warned that vSBX may benefit of an advantage when applied to the Rastrigin function because of its preferential search along the axis. To neutralise this advantage, SPC-vSBX was tested on the rotated Rastrigin function and a best value of 8.955 was found, whereas a best value of 309.429 was reported with G3-PCX. By contrast, SPC-PNX found best values of 4.975 and 3.980 for the Rastrigin and rotated Rastrigin, respectively.

It has been generally observed that by incrementing N , in both SPC-vSBX and SPC-PNX, better results are found at a cost of taking longer to converge.

Table 1. Performance comparison between G3-PCX, SPC-vSBX and SPC-PNX over the test functions. The best, median and worst columns refer to the number of function evaluations required to obtain a value of 10^{-20} . If the target is not reached then the best found function value within 10^6 evaluations is given. ‘Success’ refers to how many runs reach the target accuracy (unimodal) or end up within the global basin (multimodal). The latter is determined by checking if the best found solution is below the function’s best local minimum. ‘?’ accounts for information not specified in the original study [2].

Model	Crossover	(N, λ)	Function	Best	Median	Worst	Best Found	Success
G3	PCX-(0.1,0.1)	(100,2)	Elp	5,826	6,800	7,728	10^{-20}	10/10
SPC	vSBX-0.01	(6,1)	Elp	49,084	50,952	57,479	10^{-20}	10/10
SPC	PNX-2.0	(35,1)	Elp	36,360	39,360	40,905	10^{-20}	10/10
G3	PCX-(0.1,0.1)	(150,2)	Sch	13,988	15,602	17,188	10^{-20}	10/10
SPC	vSBX-0.01	(6,1)	Sch	260,442	294,231	334,743	10^{-20}	10/10
SPC	PNX-2.0	(35,1)	Sch	236,342	283,321	299,301	10^{-20}	10/10
G3	PCX-(0.1,0.1)	(150,4)	Ros	16,508	21,452	25,520	10^{-20}	36/50
SPC	vSBX-0.01	(12,1)	Ros	10^6	-	-	10^{-4}	48/50
SPC	PNX-2.0	(35,1)	Ros	10^6	-	-	10^{-10}	38/50
SPC	PNX-2.0	(80,1)	Ros	10^6	-	-	10^{-6}	47/50
G3	PCX-(0.1,0.1)	(150,2)	Ackl	10^6	-	-	3.959	0
SPC	vSBX-0.01	(8,1)	Ackl	57,463	63,899	65,902	10^{-10}	10/10
SPC	PNX-2.0	(50,1)	Ackl	45,736	48,095	49,392	10^{-10}	10/10
G3	PCX-(?,?)	(?,?)	Rtg	10^6	-	-	15.936	0
SPC	vSBX-0.01	(20,3)	Rtg	260,658	306,819	418,482	10^{-20}	6/10
SPC	vSBX-0.01	(40,3)	Rtg	639,102	721,401	800,754	10^{-20}	10/10
SPC	PNX-2.0	(400,4)	Rtg	10^6	-	-	4.975	0
G3	PCX-(0.1,0.1)	(300,3)	Rot. Rtg	10^6	-	-	309.429	0
SPC	vSBX-0.01	(75,3)	Rot. Rtg	10^6	-	-	8.955	0
SPC	PNX-2.0	(400,4)	Rot. Rtg	10^6	-	-	3.980	0

Also, a restricted search (through a higher value of η) seems to be beneficial in the highly multimodal functions. Based on these observations, we investigate the performance of SPC-PNX with different combinations of N, λ and η , allowing a higher number of function evaluations and using the same initialisation. As a result, we solved the 20-variable Rastrigin (N=2,000, $\lambda = 4$, PNX-3.0 and $2 \cdot 10^6$ evaluations, obtaining a function value of $3.634 \cdot 10^{-12}$) and the 20-variable rotated Rastrigin (N=2,500, $\lambda = 3$, PNX-3.0 and $2.25 \cdot 10^6$ evaluations, obtaining a function value of $2.438 \cdot 10^{-2}$).

6 Review of Results with Respect to Other Studies

In this section, other previous studies reporting results on the used test functions are reported. This will allow a wider comparison with previously published studies.

Eiben and Bäck [7] used an (μ, λ) -ES to optimise 30-variable Schwefel, Ackley and Rastrigin functions. On the Schwefel, the ES was initialised within [60,65] and a best solution greater than 1.0 was reported. The initialisation for the Ackley function was [15,30] and the best found values was greater than 10^{-13} . The Rastrigin was initialised within [4,5] and a solution better than 10.0 was reported. Storn and Price [15] used DE on a testbed including Ackley and Rastrigin functions with symmetric initialisations. The 30-variable Ackley, 100-variable Ackley, 20-variable Rastrigin and 100-variable Rastrigin functions were solved. However, as the authors admit for these multimodal functions: ‘As many symmetries are present, the main difficulty of these test functions lies in their dimensionality’. Chellapilla and Fogel [8] solved the 10-variable Rastrigin and Ackley functions starting from [8,12]. Compared to a symmetric initialisation, this study showed negative improvement in best function values with the skewed initialisation. Patton et al. [9] used also a skewed initialisation (but bracketting the global minimum) and solved the 10-variable instances of the Schwefel, Rosenbrock, Ackley and Rastrigin. Wakunda and Zell [16] apply a number of CMA-ESs (Covariance Matrix Adaptation Evolutionary Strategies) and solved the 20-variable Ellipsoidal, Schwefel, Rosenbrock and Ackley functions. As the initialisation was not stated, it is not possible to compare with these results. Kita [17] using a real-parameter GA (known as MGG-UNDX) and a ES solves the 20-variable Rosenbrock function. Also, the symmetric initialisation [-5.12,5.12] was used to solve a 5-variable rotated Rastrigin. Hansen and Ostermeier [18] applied a CMA-ES to the Ellipsoidal, Schwefel, Rosenbrock and Rastrigin. Starting from a unit away from the global minimum, the Ellipsoidal, Schwefel and Rosenbrock functions were solved with up to 320 variables. The algorithm was applied on the 20-variable Rastrigin. Starting with a solution initialised in [-5.12,5.12], function values within 30.0 and 100.0 were found.

7 Conclusions and Future Work

We have presented a GA (SPC-PNX) which has been shown to be effective in optimising high dimensional real-variable functions. This algorithm incorporates the new parent-centric crossover PNX (Parent-centric Normal Crossover).

SPC-PNX’s performance has been tested on a set of high dimensional real-variable functions. These functions have a number of features that are known to be hard for optimisation algorithms and believed to be present in many real-world problems. By using PNX instead of vSBX, a better convergence was obtained while maintaining practically the same average performance. This was observed for all test functions but the Rastrigin, where SPC-vSBX is known to enjoy an advantage. In comparison with G3-PCX, SPC-PNX does not perform as well as this algorithm for problems with few local minima. However, for problems with very many local optima, our algorithm performed significantly better. In the hardest test function (rotated Rastrigin) in terms of separability and multimodality, SPC-PNX widely overcomes the performance previously reported.

In future work, we will investigate the effect of varying the parameter NREP, which seems to affect the ability of maintaining several subpopulations during the GA run. Also, we plan to apply it to carry out the calibration of the model parameters corresponding to a real petroleum reservoir.

References

1. Ballester, P.J., Carter, J.N.: Real-parameter genetic algorithms for finding multiple optimal solutions in multi-modal optimization. In: Genetic and Evolutionary Computation Conference, Lecture Notes in Computer Science 2723. (2003)
2. Deb, K., Anand, A., Joshi, D.: A computationally efficient evolutionary algorithm for real-parameter optimization. *Evolutionary Computation* **10** (2002) 371–395
3. Ballester, P.J., Carter, J.N.: An effective real-parameter genetic algorithms for multimodal optimization. In Parmee, I.C., ed.: Proceedings of the Adaptive Computing in Design and Manufacture VI. (2004) In Press.
4. Carter, J.N., Ballester, P.J., Tavassoli, Z., King, P.R.: Our calibrated model has no predictive value: An example from the petroleum industry. In: Proceedings of the Sensitivity Analysis and Model Output Conference (SAMO-2004), Santa Fe, New Mexico, U.S.A. (2004) In Press.
5. Whitley, D., Watson, J., Howe, A., Barbulescu, L.: Testing, evaluation and performance of optimization and learning systems. In Parmee, I.C., ed.: Proceedings of the Adaptive Computing in Design and Manufacture V, Springer Verlag (2002) 27–39
6. Fogel, D.B., Beyer, H.G.: A note on the empirical evaluation of intermediate recombination. *Evolutionary Computation* **3** (1996) 491–495
7. Eiben, A.E., Bäck, T.: Empirical investigation of multiparent recombination operators in evolution strategies. *Evolutionary Computation* **5** (1998) 347–365
8. Chellapilla, K., Fogel, D.B.: Fitness distributions in evolutionary computation: Analysis of local extrema in the continuous domain. In Angeline, P.J., Michalewicz, Z., Schoenauer, M., Yao, X., Zalzal, A., eds.: Proceedings of the Congress on Evolutionary Computation. Volume 3., Mayflower Hotel, Washington D.C., USA, IEEE Press (1999) 1885–1892
9. Patton, A.L., Goodman, E.D., III, W.F.P.: Scheduling variance loss using population level annealing for evolutionary computation. In Angeline, P.J., Michalewicz, Z., Schoenauer, M., Yao, X., Zalzal, A., eds.: Proceedings of the Congress of Evolutionary Computation. Volume 1., Mayflower Hotel, Washington D.C., USA, IEEE Press (1999) 760–767
10. Deb, K., Beyer, H.G.: Self-adaptive genetic algorithms with simulated binary crossover. *Evolutionary Computation* **9** (2001) 197–221
11. Mengshoel, O., Goldberg, D.: Probabilistic crowding: Deterministic crowding with probabilistic replacement. In Banzhaf, W., Daida, J., Eiben, A., Garzon, M., V.Honavar, Jakiela, M., Smith, R., eds.: Proceedings of the Genetic and Evolutionary Computation Conference, Morgan Kaufmann (1999) 409–416
12. Deb, K., Agrawal, S.: Simulated binary crossover for continuous search space. *Complex Systems* **9** (1995) 115–148
13. Deb, K., Kumar, A.: Real-coded genetic algorithms with simulated binary crossover: Studies on multi-modal and multi-objective problems. *Complex Systems* **9** (1995) 431–454
14. KanGAL: (January 2004) <http://www.iitk.ac.in/kangal/soft.htm>.

15. Storn, R., Price, K.: Differential evolution a simple and efficient heuristic for global optimisation over continuous spaces. *Journal of Global Optimization* **11** (1997) 341–359
16. Wakunda, J., Zell, A.: Median-selection for parallel steady-state evolution strategies. In Schoenauer, M., Deb, K., Rudolph, G., Yao, X., Lutton, E., Merelo, J.J., Schwefel, H.P., eds.: *Proceedings of the 6th International Conference "Parallel Problem Solving from Nature - PPSN VI*. Volume 1917 of *Lecture Notes in Computer Science.*, Paris, France, Springer Verlag (2000) 405–414
17. Kita, H.: A comparison study of self-adaptation in evolution strategies and real-coded genetic algorithms. *Evolutionary Computation* **9** (2001) 223–241
18. Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation* **9** (2001) 159–195