Multi-Parent Recombination Operators in Continuous Search Spaces*

A.E. Eiben†
Dept. of Maths. and Comp. Sci.
Leiden University
Niels Bohrweg 1
NL-2333 CA Leiden

T. Bäck‡
Informatik Centrum Dortmund
Joseph-von-Fraunhofer-Str. 20
D-44227 Dortmund
and
Dept. of Maths. and Comp. Sci.
Leiden University
Niels Bohrweg 1
NL-2333 CA Leiden

Abstract

An extension of evolution strategies to multi-parent recombination involving a variable number ϱ of parents to create an offspring individual is proposed. The extension is experimentally evaluated on a number of test functions, including unimodal and multimodal functions of high dimensionality. Multi-parent diagonal crossover, uniform scanning crossover, and a multi-parent version of intermediary recombination are considered in the experiments.

Algorithm performance is observed to strongly depend on the particular combination of recombination operator and objective function. In some cases, a significant increase of performance is observed even for multimodal objective functions as the number of parents increases, but there might also be no significant impact of recombination at all. Furthermore, the algorithm might also exhibit a divergent behavior in case of a unimodal optimization problem when the recombination operator is chosen inappropriately.

*This report is available at http://www.wi.leidenuniv.nl/~gusz/esmultip.ps.gz
†gusz@wi.leidenuniv.nl
‡baeck@ls11.informatik.uni-dortmund.de

1 Introduction

In natural evolution reproduction mechanisms are either asexual, or sexual. In the case of asexual reproduction one parent creates one (or more) offspring, whereas sexual reproduction requires two parents. It should be noted, however, that allelic recombination occurs during meiosis, i.e., the formation of a gamete within the parental organisms. In other words, rather than the genetic information of two different parents, it is the genetic information of diploid (in case of the human genome) homologous chromosomes within each parent that is subject to a rearrangement or recombination [12]. Although it is biologically incorrect, in the field of evolutionary computation the term *crossing-over* or *crossover* is often used synonymously with recombination, and this convention is adopted in this paper as well.

In simulated evolution, that is in evolutionary algorithms, many technical features are inspired by natural mechanisms. In particular, abstract variants of sexual and asexual reproduction are implemented as search operators. Some evolutionary techniques, e.g. evolutionary programming, work exclusively with mutation (i.e., they implement a simplification of asexual reproduction), while others, e.g. genetic algorithms and evolution strategies, use recombination (i.e., they implement a simplification of sexual reproduction) and mutation. There are several papers investigating the advantages and disadvantages of mutation with respect to crossover [8, 10, 11, 13, 18, 21]. At the moment the question whether mutation or crossover is preferable (or rather, which one is preferable under certain circumstances) is still an open research issue.

Technically, the question concerns the arity of the reproduction operators. Mutation and crossover have arity one and two, respectively and the question is whether unary or binary operators are preferable for typical instances of practically relevant optimization problems. From a purely technical point of view there is no need to restrict the arity of reproduction operators to one or two. In general, a reproduction operator can have an arity from one up to the population size (or even more, if we allow repetition among the parents). Hereby the analogy with natural evolution breaks down, to our knowledge there are no species on Earth that would apply multi-parent reproduction mechanisms where genetic material of more than two parents is mixed in *one* reproductive action. Simulating ϱ-ary reproduction operators, however, is no problem.

In evolution strategies, recombination has a local and a global form [2, 19]. In global recombination the i-th parameter of the child is determined choosing one of the parents randomly anew from the parent population for each value of i. Thus, global recombination is a multi-parent operator, although its arity is undefined (or it has a random arity). Recently, Schwefel and Rudolph [20] proposed an extension of recombination in evolution strategies to allow a variable number ϱ (with $2 \leq \varrho \leq \mu$) of parents to be involved in recombination. This generalization of recombination in evolution strategies has strong similarities to the multi-parent crossover operators introduced earlier by Eiben *et al.* [6]. The latter operators form the basis for our multi-parent recombination operators that will be discussed in detail in sections 2.1 and 2.2. Beyer's $(\mu/\varrho_I, \lambda)$- and $(\mu/\varrho_D, \lambda)$-strategies also introduce a variable number of parents [4], but it should be noted that this variant of intermediary recombination (select ϱ parents and yield the arithmetic average of these as the result) is different from the original intermediary recombination [2], which selects the two parents anew (at random) for each object variable from the set of ϱ available (i.e., preselected from the population) parents. Only for $\varrho = 2$ are both recombination schemes identical, while for $\varrho = \mu$ Beyer's intermediary recombination always creates the same offspring when applied to the parent population, thus drastically reducing the diversity in the population after recombination. This kind of recombination operator was introduced just for the purpose of making possible a theoretical analysis of the algorithm.

Recently, Voigt and Mühlenbein introduced so-called Gene Pool Recombination (GPR) in genetic algorithms [17, 22]. The basic mechanism of GPR is identical to global recombination in ES, thus the arity of the reproduction operator is again undefined. Note that in ES and in GAs with GPR, sexuality is a Boolean feature: recombination is either on or off, but its arity cannot be tuned.

Eiben *et al.* [6] generalized the traditional (binary) 1-point crossover and uniform crossover to ϱ parents. According to these definitions (see section 2.2 for details) the reproduction operator has an arity that can be set by the user. Hereby sexuality becomes a graded feature: by tuning the arity of crossover the 'amount of sex' can also be varied in an evolutionary algorithm.

In previous papers, the effect of using more parents within diagonal and uniform scanning crossover was investigated. In [5], numerical optimization problems were bit-coded and solved by a GA using these operators. In [7], pure bit-problems (NK-landscapes [14]) were investigated. The current paper investigates numerical optimization problems with floating point representation and using an evolution strategy with generalized arity of the recombination operators.

There are a number of hypotheses and questions to be investigated. The main working hypothesis is the following.

- **H1** Increasing the number of parents leads to increased EA performance in terms of achieved accuracy, i.e., distance from the global optimum (measured as quality of the result) at termination.

Clearly, further refinements of this hypothesis are needed. Below are a number of questions concerning particular refinements.

- **Q1** On what (type of) functions does H1 hold? What characteristics of a given objective function facilitate the increase of performance in the case of using more parents?

- **Q2** For which multi-parent operators does H1 hold?

To this end, note that three different recombination mechanisms are investigated and there is no reason to expect that they behave the same way, i.e. show the same response to increasing the number of parents.

Particular attention will be paid to the transition from one to two in the number of parents. Namely, this step from one to two amounts to introducing sexual reproduction in the system. Further increases in the number of parents only intensify the already present sexual character of reproduction. From this viewpoint the main working hypothesis can be broken into two components.

- **H1a** Increasing the number of parents from one to two leads to increased EA performance, i.e. 'sex is good'.

- **H1b** Increasing the number of parents from two to larger numbers leads to increased EA performance, i.e. 'more sex is better'.

Investigations on NK-landscapes showed that neither H1a nor H1b hold for very rugged landscapes with many randomly distributed local optima [7]. The experiments, however, suggested that on landscapes where H1a holds, H1b holds as well, thus H1a implies H1b; in common parlance 'if sex is good, then more sex is even better'.

2 Evolutionary Algorithm

This section presents the evolutionary algorithm used for the experiments, as well as the investigated multi-parent recombination operators.

2.1 Evolution Strategies

In the following, only a brief overview of the basic principles of evolution strategies is presented. The interested reader is referred to more thorough introductions to evolution strategies such as [1, 20].

As fundamental characteristics of evolution strategies, their emphasis on strategy parameter self-adaptation (i.e., on-line adaptation of mutation variances and covariances by evolutionary principles), a mutation operator working with normally distributed variations of real-valued vectors $\vec{x} \in \Re^n$, a deterministic selection operator that selects μ individuals from a surplus of $\lambda > \mu$ offspring, and on the utilization of recombination both for object variables x_i and strategy parameters are to be identified. The presentation in this paper is restricted to the application of evolution strategies where either one or n variances of the normally distributed variation of object variables are self-adapted (i.e., so-called *correlated mutations* are not taken into account here). Individuals $\vec{a} = (\vec{x}, \vec{\sigma})$ then consist of the object variable vector $\vec{x}$ and $n_\sigma \in \{1, n\}$ standard deviations $\vec{\sigma} = (\sigma_1, \ldots, \sigma_{n_\sigma})$, and mutation proceeds by modifying standard deviations and object variables according to

$$\begin{aligned} \sigma_i' &= \sigma_i \cdot \exp(\tau' N(0,1) + \tau N_i(0,1)) & (1) \\ x_i' &= x_i + \sigma_i' N_i(0,1) \ , & (2) \end{aligned}$$

if $n_\sigma > 1$, while $\sigma' = \sigma \cdot \exp(\tau_0 N(0,1))$ if $n_\sigma = 1$. For the so-called learning rates τ, τ', and τ_0, the settings $\tau' \approx (2n)^{-1/2}$, $\tau \approx (2\sqrt{n})^{-1/2}$, and $\tau_0 \approx n^{-1/2}$ are robust and effective recommended values (see e.g. [1]). It should be noted, however, that for particular problems a fine-tuning of these parameters might yield a considerable improvement in performance. The notation $N(0,1)$ denotes a realization of a normally distributed one-dimensional random variable with expectation zero and standard deviation one; $N_i(0,1)$ indicates that it is sampled anew for each value of i.

Concerning recombination, evolution strategies have typically been restricted to involving either two or potentially all μ parents (in case of so-called global recombination types) in the creation of new individuals either by randomly deciding the individual from which an object variable is copied to the offspring (discrete recombination, analogous with uniform crossover in genetic algorithms) or by arithmetic averaging of pairs x_{i_j}, x_{i_k} of corresponding object variables that come from parents j, k randomly selected from the set of either two or μ parents (intermediary recombination). Notice that intermediary recombination of two parents reduces to simply averaging each of their corresponding pairs of object variables.

The generalization to multi-parent recombination involving ϱ with $2 \leq \varrho \leq \mu$ parent individuals has recently been proposed by Schwefel and Rudolph [20], independently of the work of Eiben *et al.* [6], who were the first who formulated and tested multi-parent recombination operators of arbitrary arity. For evolution strategies, experimental investigations have not been performed yet with multi-parent operators. The generalized operators proceed by first picking ϱ parents *uniformly at random*, without repetition, and then mixing characters from these ϱ parents to form *one offspring*. The precise working mechanism of the multi-parent operators investigated here will be discussed in more detail in section 2.2.

It should be noted that, in addition to the object variables, also the strategy parameters typically undergo recombination in an evolution strategy. The recombination type, however, might (and typically will) differ between object variables and strategy parameters, and in the experiments discussed here we restrict ourselves to investigating the impact of multi-parent recombination on the object variables. In this paper, strategy parameters are always recombined using global intermediary recombination.

Finally, the selection operator of contemporary (μ,λ)-evolution strategies deterministically picks the μ best out of $\lambda > \mu$ offspring individuals to form the parent population of the next generation. The λ offspring individuals are created by λ-fold application of recombination, followed by mutation, to the parent population (i.e., recombination is always applied in evolution strategies, not just with a certain probability p_c such as in genetic algorithms — although this could also be introduced, as proposed in [20]).

At present, even the relative benefits of global recombination compared to two-parent recombination are neither theoretically understood nor experimentally investigated. Certain idealistic variants of global intermediary and global discrete recombination (i.e., using μ parents), however, have been analysed and shown to yield a μ-fold speedup on convex objective functions when compared to no recombination [4]. It is not clear, however, how these results relate to the real implementations of discrete and intermediary recombinations and how they are affected by the parameter ϱ. This study aims at giving some experimental hints about the impact of a varying number of parents involved in recombination on the accuracy of the evolution strategy.

2.2 Operators

The multi-parent operators investigated in this paper are intermediary recombination, scanning crossover, and diagonal crossover, and they are all designed or modified to produce one offspring individual as is common in evolution strategies.

The intermediary crossover creates one child from ρ parents. For each variable ($i = 1, \ldots, n$) two 'donors' are chosen uniformly from among the ρ parents and their genetic material is mixed by averaging them, i.e., the value $(x_i^{donor_1} + x_i^{donor_2})/2$ is passed to the child. Scanning crossover generalizes uniform crossover, although creating only one child. The idea behind it is to take ϱ parents and to create one child by scanning the parents' variable vectors from left to right and deciding at each position which parent can deliver its value to the child. The choice of the parent delivering its value can be random, based on a uniform distribution (uniform scanning), or biased by the fitness of the parents (fitness-based scanning). It can also be deterministic, choosing the most frequently occurring parent allele (occurrence based scanning).[1] In this investigation uniform scanning crossover is used. Note that uniform scanning crossover corresponds directly to the generalization of discrete recombination (according to [20]). For $\varrho = 2$ and $\varrho = \mu$, scanning crossover is equivalent to the well-known variants of local discrete and global discrete recombination. Likewise, the generalized intermediary recombination operator introduced here is equivalent to local intermediary recombination for $\varrho = 2$ and global intermediary recombination for $\varrho = \mu$. For both operators, a gradual variation between the two extremes $\varrho = 2$ and $\varrho = \mu$ is facilitated by the generalization to arbitrary values of ϱ, while keeping the resulting evolution strategy as closely related to the standard variant as possible.

Diagonal crossover generalizes 1-point crossover and to some extent n-point crossover. On ϱ parents it works by selecting $(\varrho - 1)$ crossover points (identical for each parent) and composing ϱ children by taking the resulting ϱ chromosome segments from the parents 'along the diagonals'. Note that using ϱ parents, the number of children is 1 for intermediate recombination and scanning crossover, while it equals ϱ for diagonal crossover. This means that for creating μ offspring μ recombination operations are needed for intermediate recombination and scanning crossover, implying that all together information from $\mu \cdot \varrho$ parents is utilized. Since the number of children equals the number of parents for diagonal crossover, only μ parents are needed to be utilized for creating μ offspring This might cause unintended effects that disturb fair comparisons between the operators. Therefore a slightly modified version of diagonal crossover that creates one child instead of ϱ is used. Figure 1 illustrates this idea for $\varrho = 3$.

To summarize, the most important properties of the multi-parent recombination operators introduced here into evolution strategies are the following:

- The selection of ϱ potential parents for recombination is performed uniformly at random.
- The value selection in scanning crossover, the parent selection in intermediary recombination, and the crossover point selection in diagonal crossover are performed uniformly at random.

[1] In case of order-based representation a more general scanning mechanism is needed, which will not be used here. For the definition see [6].

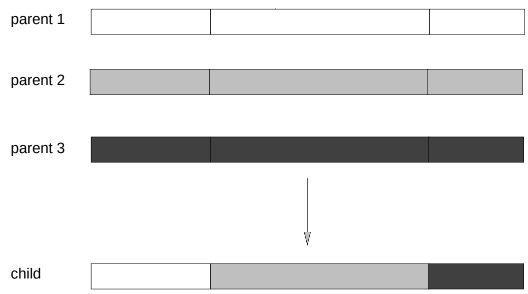

Figure 1: Diagonal crossover with three parents and one child. For three parents, two crossover points are chosen randomly, but identical on all parents. The offspring individual is then produced by concatenating the first segment of the first parent, the second segment of the second parent, and the third segment of the third parent (in general, the ith segment of the ith parent; $i = 1, \ldots, \varrho$).

- Only one offspring individual is created per application of the recombination operator.

These properties are emphasized because they assure that selective biases are excluded from the recombination operator and that the multi-parent version of recombination stays as close as possible to the original recombination operator used in evolution strategies.

3 Test Functions and Experimental Setups

The test functions used for a first experimental assessment of the characteristics of multi-parent recombination in evolution strategies are selected to reflect a certain basic diversity of topological characteristics, including unimodal and multimodal objective functions of scalable dimensionality. Experience led to the selection of the following six functions:

- Sphere model:

$$f_1(\vec{x}) = \sum_{i=1}^{n} x_i^2 \ ,$$

  where $n = 30$, $n_\sigma = 1$, and the object variables x_i are initialized in the range $-5.12 \leq x_i^0 \leq 5.12$.

- Schwefel's double sum:

$$f_2(\vec{x}) = \sum_{i=1}^{n} \left(\sum_{j=1}^{i} x_j \right)^2 ,$$

where $n = 30$, $n_\sigma = 30$, and $-65.536 \le x_i^0 \le 65.536$.

- Generalized Ackley's function:

$$f_3(\vec{x}) = 20 + e - 20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \right) - \exp\left(\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi x_i) \right) ,$$

where $n = 30$, $n_\sigma = 30$, $e = \exp(1)$, and $-20 \le x_i^0 \le 30$.

- Generalized Rastrigin's function:

$$f_4(\vec{x}) = 10n + \sum_{i=1}^{n} x_i^2 - 10\cos(2\pi x_i) ,$$

where $n = 30$, $n_\sigma = 30$, and $-5.12 \le x_i^0 \le 5.12$.

- Generalized Griewangk function:

$$f_5(\vec{x}) = 1 + \sum_{i=1}^{n} \frac{x_i^2}{400n} - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}} \right) ,$$

where $n = 30$, $n_\sigma = 30$, and $-600 \le x_i^0 \le 600$.

- Fletcher-Powell function:

$$\begin{aligned} f_6(\vec{x}) &= \sum_{i=1}^{n} (A_i - B_i)^2 \\ A_i &= \sum_{j=1}^{n} (a_{ij} \sin\alpha_j + b_{ij} \cos\alpha_j) \\ B_i &= \sum_{j=1}^{n} (a_{ij} \sin x_j + b_{ij} \cos x_j) , \end{aligned}$$

where $n = 30$, $n_\sigma = 30$, and $-\pi \le x_i^0 \le \pi$. The $a_{ij}, b_{ij} \in \{-100, \ldots, 100\}$ are random integers, and $\alpha_j \in [-\pi, \pi]$ is the randomly chosen global optimum position. Reference values for matrices $\mathbf{A}, \mathbf{B}$ as well as the vector $\vec{\alpha}$ are published in [1] (pp. 265–267).

For all experiments reported in section 4, a typical (16,100)-evolution strategy was used with $1 \leq \varrho \leq 16$ ($\varrho = 1$ means no recombination), performing 100 independent runs for each setting of ϱ and each of the three recombination operators. For intermediary recombination, an additional set of 100 independent runs was also performed for all objective functions, using a skewed initialization where the initial population is located in a subset of the search space far apart from the global optimum. This skewed initalization takes into account a recent result by Fogel and Beyer [9] stating that if the global optimum is located in the center of the search region covered by initializing the population uniformly at random, intermediary recombination generates offspring individuals which are essentially unbiased estimates of the global optimum with lower variance than offspring generated solely by Gaussian mutation. In other words, the uniform initialization technique is assumed to introduce a bias that favors a successful identification of the global optimum, such that the success of a strategy employing intermediary recombination might be just an artefact of a useful combination of initialization and global optimum location. To check whether this assumption holds we attempt to mislead the EA by using the following initialization intervals for an additional set of runs:

Function	Interval
f_1	$[4.0, 5.0]$
f_2	$[60.0, 65.0]$
f_3	$[4.0, 5.0]$
f_4	$[15.0, 30.0]$
f_5	$[580.0, 600.0]$
f_6	$[2.0, 3.0]$

The experiments were run for 2000 generations on f_4 and f_6, for 1000 generations on f_2, for 500 generations on f_3, for 300 generations on f_5, and for 200 generations on f_1, and the average final best objective function values over the 100 runs are reported as the accuracy measure of the algorithm.

4 Experimental Results

The experimental results obtained on the six test functions according to the experimental setup described in section 3 are summarized in figure 2 for f_1 and f_2, in figure 3 for f_3 and f_4, and in figure 5 for f_5 and f_6 by plotting the average final best objective function value as a function of the number of recombinants ϱ involved in recombination, for each of the three different recombination types and for intermediary recombination also with the skewed initialization.

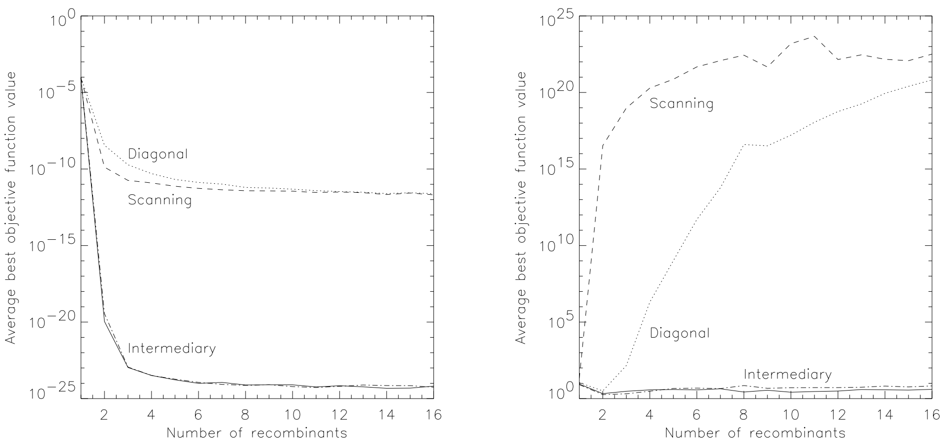

Figure 2: Average final best objective function value depending on the number of recombinants for the sphere model (left) and the double sum (right). The unlabeled dashed curves belong to intermediary recombination and skewed initialization.

As to the first question, Q1, concerning the relationship between objective function and the effect of more parents the following can be observed. The effect of increasing the number of parents can be clearly different on different objective functions. The effect of using high arity operators can be positive (e.g. Rastrigin's function), negative (scanning and diagonal crossover on the double sum), or there might even be no clear relationship between the number of parents and performance (e.g., the Fletcher-Powell function). Interesting, however, is that on five of the six test functions (the double sum being the exception) the three operators show the same kind of response to increasing ϱ.

As figure 2 clearly demonstrates we might obtain different behaviors on different unimodal objective functions. The hypothesis H1 is supported on f_1 for all recombination operators, while this is not the case on f_2. On this function H1 holds for intermediary recombination, neither H1a nor H1b holds for scanning crossover, while for diagonal crossover H1a holds and H1b does not. The unlabeled curves showing the behavior of intermediary recombination after skewed initialization are practically identical to those with normal initialization.

On the multimodal objective functions f_3 and f_4, as shown in figure 3, both H1a and H1b hold, i.e., an increase of the number of parents involved generally improves the solution accuracy. Moreover, the difference between

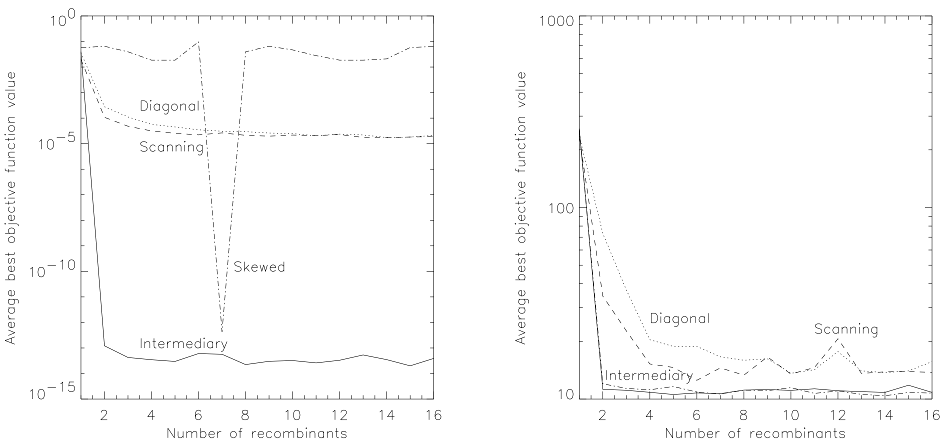

Figure 3: Average final best objective function value depending on the number of recombinants for Ackley's function (left) and Rastrigin's function (right). The unlabeled dashed curve on the left plot belongs to intermediary recombination and skewed initialization.

algorithms based on mutation only (i.e., with $\varrho = 1$) and those using recombination ($\varrho > 1$) is striking. A further increase of the number of parents for diagonal and scanning crossovers gives a substantial advantage on f_3, while only a small improvement on f_4. For intermediary recombination the advantage of more than two parents is visible, but small on both functions. One might argue from these findings that the multimodal landscapes f_3 and f_4 have a structure that facilitates the exploitation of recombination, e.g. for reasons of their regular arrangement of local optima and a global structure that is similar to a unimodal landscape. Looking at the curves belonging to skewed initialization we see no difference in behavior for Rastrigin's function (the unlabeled curve in figure 3, left). On the Ackley function, however, we see a steep valley in a plateau. From a statistical point of view this suggests the presence of outliers in the data. The disturbing effect of the outliers can be filtered out by disregarding the best and/or worst results and plotting the curves again. In figure 4 we provide trimmed mean curves. The left side figure is obtained by omitting the best and worst 5% of the data, the right side figure is created by disregarding the worst 10%.

The trimmed curves show that the outliers are the bad runs causing the plateau. The best-worst 5% trimmed curve for skewed initialization and intermediary recombination still shows data distortion, but worst 10% trimming, keeping 90% of the data yields a regular curve. This shows that the be-

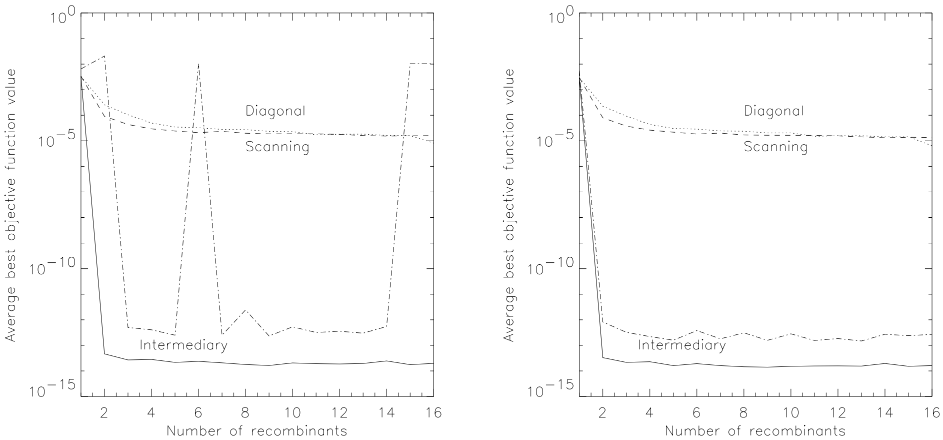

Figure 4: Best-worst 5% trimmed mean (left) and worst 10% trimmed mean curves for Ackley's function. The unlabeled dashed curves belong to intermediary recombination and skewed initialization.

havior of intermediary recombination is not changed significantly by skewed initialization on Ackley's function either.

The performance of the evolution strategy on Griewangk's function as shown in the left part of figure 5 shows no regularities at all except an indication that H1a holds, and the variances of the measured data points are extremely large. This plot suggets that further increasing the number of parents beyond two has no measurable impact on the performance.

Given the results from figure 3, this is counterintuitive, because Griewangk's function also has a global structure that is supposed to make this function quite easy, especially for high dimensionality. In fact, the noisy character of the results for f_5 is caused by the fact that most runs found the global optimum, but usually a few runs stagnated in local optima of bad quality, such that the averaged result shows no statistical significance in favor of a particular parameter setting. To see whether this explanation is valid the data was trimmed again. In figure 6 the best-worst 5%, respectively worst 10% trimmed mean curves are provided. These curves show a more regular character than the ones without trimming. Best-worst 5% trimming for scanning and diagonal crossover leads to more regular curves showing the H1 effect on Griewangk's function. Using worst 10% trimming this becomes clearly visible. Also for intermediary recombination trimming results in decreasing curves. For $\varrho = 4$, $\varrho = 13$, and $\varrho = 16$ the results after worst 10% trimming are not shown in the curves because they are as small as 10^{-13},

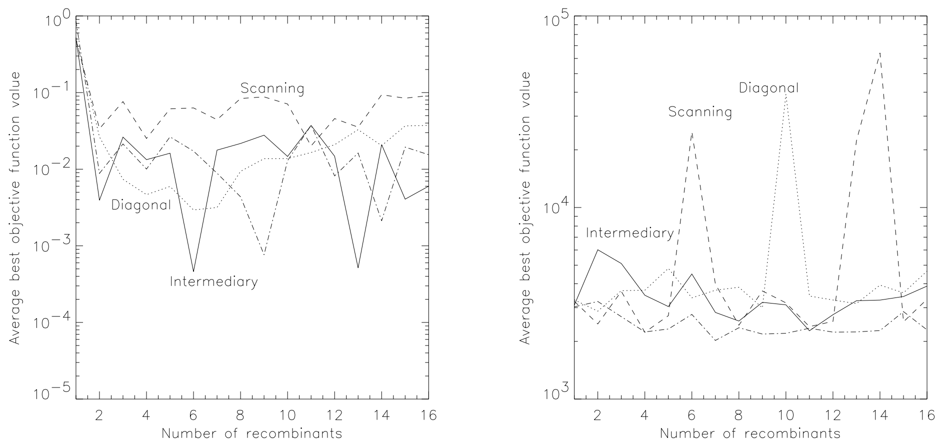

Figure 5: Average final best objective function value depending on the number of recombinants for Griewangk's function (left) and the Fletcher-Powell function (right). The unlabeled dashed curves belong to intermediary recombination and skewed initialization.

indicating that a reasonably small percentage of outliers is responsible for the deterioration of results in case of this recombination operator.

For f_6 the local optima are randomly distributed. Thus, this function has a structure different from the other ones. The results shown in the right part of figure 5 have no statistically significant structure at all. The trimmed curves given in figure 7 exhibit the same lack of structure, so the argument about the outliers cannot be applied here. The only plausible conclusion is that on this function the number of parents has no effect on the performance, i.e., H1 does not hold.

One might argue that it is the very irregular character of the fitness landscape that does not allow recombination to exploit any implicit regularities in the sense of combining good 'building blocks' from different parents. The correlation coefficient analysis of Manderick *et al.* [15] showed that the performance of (two-parent) crossover is inversely related to the correlation between the fitness of parent and offspring chromosomes. For an irregular landscape as the Fletcher-Powell function this correlation is low and our results show that Manderick's argument holds for multi-parent crossovers too. Furthermore, (binary) NK-landscapes, where K is relatively high with respect to N have a similar chaotic structure and the performance curves of multi-parent operators on rugged landscapes in [7] are very similar to those on the left of figure 5 and in figure 7 above. Nevertheless, it would be nec-

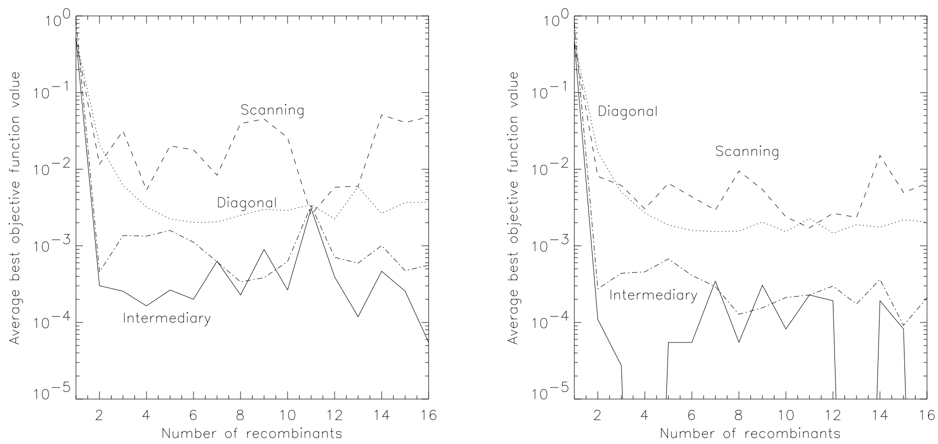

Figure 6: Best-worst 5% trimmed mean (left) and worst 10% trimmed mean for Griewangk's function. The unlabeled dashed curves belong to intermediary recombination and skewed initialization.

essary to perform more experiments with continuous landscapes where the optima are arranged in an irregular way in order to gain more insight into the general working principles of recombination operators.

Looking at the results with our second question, Q2, in mind leads to the following observations. On those functions where the H1 effect occurs at all, it occurs for every crossover, with one exception. On the double sum f_2, the operators respond differently. A divergence of the $(16/\varrho,100)$-evolution strategy is observed for scanning crossover with $\varrho > 1$, i.e. both H1a and H1b are falsified in a strong sense, i.e., increasing the arity of the operator decreasing performance is obtained. For diagonal crossover deterioration of performance occurs with $\varrho > 2$, meaning that H1a holds, while H1b does not. This result is of special importance, because it demonstrates that the optimum number of parents might be neither one (i.e., mutation only) nor μ (the maximum), but something in between. This can be explained because for the double sum function f_2 there is a strong correspondance between consecutive objective variables. Diagonal crossover becomes increasingly disruptive when increasing ϱ, and with $\varrho = 2$, corresponding to an ordinary one-point crossover known from genetic algorithms, this operator is likely to transfer a large number of already good consecutive objective variables of one parent to the offspring in such a way, that it is combined with a helpful 'building block' of the other parent. The results for intermediary recombination on f_2 confirm the working hypothesis that the increase of the number of parents

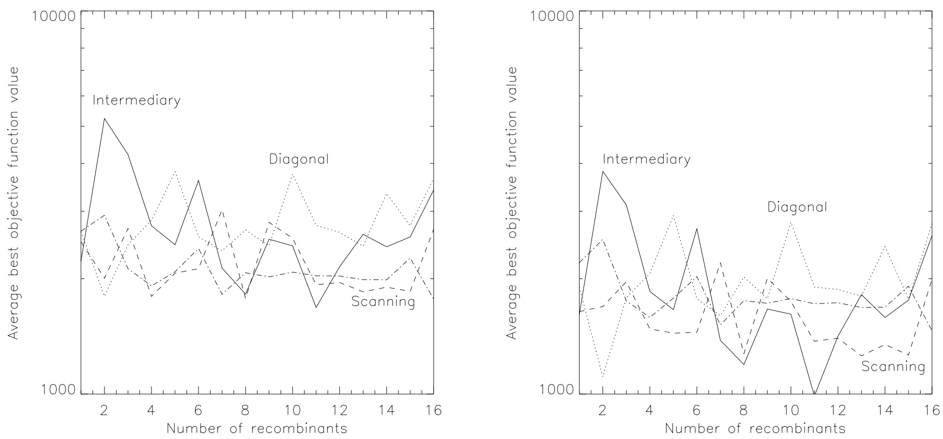

Figure 7: Best-worst 5% trimmed mean (left) and worst 10% trimmed mean for the Fletcher-Powell function. The unlabeled dashed curves belong to intermediary recombination and skewed initialization.

also increases the solution accuracy for this problem. Nevertheless, the difference between different numbers of parents is smaller on f_2 than on other functions. Summarizing, based on the experimental results available at the moment the conjecture is obtained that using multi-parent recombination does not lead to increased performance

1. if the objective function is irregular, having many randomly distributed local optima and

2. if consecutive parameters are strongly correlated and the increase in the number of parents implies an increase of disruptiveness of the operator.

An extension of this conjecture is based on reformulating the working hypothesis H1 as having a positive correlation between the number of parents and EA performance. This formulation allows a distinction between falsifying the working hypothesis in two possible ways:

i. there is no (significant) correlation between the arity of the recombination operator and EA performance;

ii. there is a negative correlation between operator arity and EA performance, i.e., the performance is inversely proportional to the arity.

The experiments described in this paper suggest that on objective functions of type 1 the EA shows type (i) behavior, whereas on type 2 functions the behavior is of type (ii).

5 Conclusions

In this paper we described experiments on a typical 'evolutionary' test suite with an evolution strategy using a typical parameter setting and new recombination operators with tunable arity. The following table summarizes the observed behavior of the three multi-parent recombination operators with respect to the working hypotheses H1a and H1b by indicating whether a hypothesis holds (marked by a "+"-sign), does not hold ("$\circ$"), or does not hold in the strong sense, i.e. type (ii) behavior ("-").

	f_1	f_2	f_3	f_4	f_5	f_6
Intermediary:						
H1a	+	+	+	+	+	$\circ$
H1b	+	$\circ$	$\circ$	+	$\circ$	$\circ$
Scanning:						
H1a	+	-	+	+	+	$\circ$
H1b	+	-	+	+	$\circ$	$\circ$
Diagonal:						
H1a	+	+	+	+	+	$\circ$
H1b	+	-	+	+	+	$\circ$

As the summary table again clarifies, a diversity of possible outcomes is observed, and whether the working hypotheses hold or not depends strongly on the particular combination of objective function and recombination operator. The results demonstrate that hypothesis H1a holds in more than 70% of the cases studied, but a further increase of the number of parents beyond two does not necessarily have a significant impact on the accuracy achieved. Even worse, in case of f_2 with diagonal crossover, H1a holds but a further increase of the number of parents results in a catastrophic deterioration of solution accuracy.

Although the results are encouraging concerning the usefulness of recombination, the objective function f_6 — which is the only one with a completely irregular, random arrangement of the locations of local optima — suggests that recombination operators might be advantageous only in case of objective functions with regularly arranged local optima and a superimposed unimodal topology (as in case of f_3, f_4, and f_5). More experimental work with such irregular multimodal objective functions needs to be performed in order to obtain a clearer picture concerning the validity of this new hypothesis.

Unfortunately, as f_2 demonstrates, there even exist objective functions where a certain choice of a recombination operator (scanning crossover, in this case) is generally harmful and causes a non-elitist algorithm such as the

(μ,λ)-evolution strategy to diverge. And, as diagonal crossover clarifies for this function, even if H1a holds, a further increase of the number of parents might still invert the behavior and cause divergence again.

Concerning the combination of a skewed initialization of the population with intermediary recombination, no significant impact on the final solution accuracy is observed for all but one objective function, namely f_3, i.e., Ackley's function. In case of this function, the skewed initialization causes a certain percentage — up to less than 10%, as figure 4 reveals — of all runs to stagnate in local optima, such that the averaged results are strongly deteriorated by these outliers. The majority of the runs, however, still find the global optimum of this function, where starting in the almost totally flat region $x_i \in [15, 30]$ implies a serious handicap for any kind of optimization algorithm.

Acknowledgments

The authors are grateful for David Fogel for his valuable comments and criticism on earlier versions of this paper. The second author gratefully acknowledges support by the German BMBF (project EVOALG, a cooperation of Informatik Centrum Dortmund, Siemens AG München, and Humboldt-Universität zu Berlin), grant 01 IB 403 A.

References

[1] Th. Bäck. *Evolutionary Algorithms in Theory and Practice.* Oxford University Press, New York, 1996.

[2] Th. Bäck and H.-P. Schwefel. An overview of evolutionary algorithms for parameter optimization. *Evolutionary Computation*, 1(1):1–23, 1993.

[3] R.K. Belew and L.B. Booker, editors. *Proceedings of the 4th International Conference on Genetic Algorithms.* Morgan Kaufmann, 1991.

[4] H.-G. Beyer. Toward a theory of evolution strategies: On the benefits of sex — the ($\mu/\mu, \lambda$) theory. *Evolutionary Computation*, 3(1):81–111, 1995.

[5] A.E. Eiben, C.H.M. van Kemenade, and J.N. Kok. Orgy in the computer: Multi-parent reproduction in genetic algorithms. In Morán et al. [16], pages 934–945.

[6] A.E. Eiben, P-E. Raué, and Zs. Ruttkay. Genetic algorithms with multi-parent recombination. In H.-P. Schwefel Y. Davidor and R. Männer, editors, *Proceedings of the 3rd Conference on Parallel Problem Solving from Nature*, number 866 in Lecture Notes in Computer Science, pages 78–87. Springer-Verlag, 1994.

[7] A.E. Eiben and C.A. Schippers. Multi-parent's niche: n-ary crossovers on NK-landscapes. In H.-M. Voigt, W. Ebeling, I. Rechenberg, and H.-P. Schwefel, editors, *Proceedings of the 4th Conference on Parallel Problem Solving from Nature*, number 1141 in Lecture Notes in Computer Science, pages 319–328. Springer, Berlin, 1996.

[8] L.J. Eshelman and J.D. Schaffer. Crossover's niche. In S. Forrest, editor, *Proceedings of the 5th International Conference on Genetic Algorithms*, pages 9–14. Morgan Kaufmann, 1993.

[9] D. B. Fogel and H.-G. Beyer. A note on the empirical evaluation of intermediate recombination. *Evolutionary Computation*, 3(4):491–495, 1995.

[10] D.B. Fogel and J.W. Atmar. Comparing genetic operators with gaussian mutations in simulated evolutionary processes using linear systems. *Biological Cybernetics*, 63:111–114, 1990.

[11] D.B. Fogel and L.C. Stayton. On the effectiveness of crossover in simulated evolutionary optimization. *Biosystems*, 32:171–182, 1994.

[12] W. Gottschalk. *Allgemeine Genetik.* Thieme, Stuttgart, 1994.

[13] W. Hordijk and B. Manderick. The usefulness of recombination. In Morán et al. [16], pages 908–919.

[14] S. A. Kauffman. *The Origins of Order. Self-Organization and Selection in Evolution.* Oxford University Press, New York, NY, 1993.

[15] B. Manderick, M. de Weger, and P. Spiessens. The genetic algorithm and the structure of the fitness landscape. In Belew and Booker [3], pages 143–150.

[16] F. Morán, A. Moreno, J. J. Merelo, and P. Chacón, editors. *Advances in Artificial Life. Third International Conference on Artificial Life*, volume 929 of *Lecture Notes in Artificial Intelligence.* Springer, Berlin, 1995.

[17] H. Mühlenbein and H.-M. Voigt. Gene pool recombination in genetic algorithms. In *Proc. of the Metaheuristics Conference*. Kluwer Academic Publishers, 1995.

[18] J.D. Schaffer and L.J. Eshelman. On crossover as an evolutionary viable strategy. In Belew and Booker [3], pages 61–68.

[19] H.-P. Schwefel. *Evolution and Optimum Seeking*. Wiley, New York, 1995.

[20] H.-P. Schwefel and G. Rudolph. Contemporary evolution strategies. In Morán et al. [16], pages 893–907.

[21] W.M. Spears. Crossover or mutation? In G. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 221–238. Morgan Kaufmann, 1993.

[22] H.-M. Voigt and H. Mühlenbein. Gene pool recombination and utilization of covariances for the Breeder Genetic Algorithm. In *Proceedings of the 2nd IEEE Conference on Evolutionary Computation*, pages 172–177. IEEE Press, 1995.