

Reproductive bias, linkage learning and diversity preservation in bi-objective evolutionary optimization

Jean P. Martins*, Alexandre C. B. Delbem^a

Federal University of Technology - Paraná, Academic Department of Informatics, Pato Branco-PR, Brazil

^aUniversity of São Paulo, Institute of Mathematical Sciences and Computation, São Carlos-SP, Brazil

Abstract

Diversity preservation is a crucial component for any multiobjective evolutionary algorithm and its effectiveness defines how well an algorithm is able to find solutions to cover the whole extension of the Pareto-optimal front. In this paper, we show that traditional reproduction operators such as p -uniform and n -point crossover may sabotage the functioning of diversity preservation mechanisms by producing more solutions in certain regions of the objective space than in others, i.e they are biased. We argue that such reproductive bias is due to their high degree of disruptiveness which favor the generation of *average quality* offspring. From this perspective, we demonstrate that linkage learning helps to decrease disruptiveness and, consequently, decreases the reproductive bias, a result that better explains the benefits of estimation of distribution algorithms in bi-objective optimization. The experiments were performed on instances of the ρMNK -model, from which we concluded that unbiased reproduction operators for bi-objective optimization work in synergy with diversity preservation mechanisms and should be favored during design of multiobjective evolutionary algorithms.

Keywords: Reproductive bias, linkage learning, evolutionary optimization, diversity preservation.

1. Introduction

Most successful diversity preservation mechanisms provide ways to select, from a multiset of solutions, a subset that is representative of the whole Pareto front. Usually this is achieved by attributing to each solution a measurement of how isolated its objective vector is. Crowding-distance and hypervolume contribution are well-known measures used in this context, from such measures it is possible to assigning higher fitness to non-dominated solutions whose objective vectors are in less populated regions of the objective space, favoring reproduction to generate offspring in those regions. Such a separation between diversity preservation and reproduction was pervasive on early multiobjective evolutionary algorithms and still has a considerable influence on more recent approaches.

In this paper we analyze the influence of reproduction operators on the effectiveness of diversity preservation. Through the analysis of p -uniform and n -point crossover we evaluate a previously known hypothesis that suggests these operators contribute to the decrease of diversity in objective space (they are biased). Such bias would be an impediment for better diversity preservation, therefore, its study enables a better understanding of the requirements for multiobjective reproduction operators.

The reproductive bias was experimentally investigated on bi-objective instances of the ρMNK -model and the influence of linkage learning on such bias was also assessed. The results indicated that highly disruptive crossover operators, such as p -uniform and n -point (in their standard settings), are not able to preserve good genetic material. Such an inability produces an *averaging effect* on the quality of the offspring in relation to the parents, that bias the new objective vectors to occupy the central region of the objective space.

The use of a linkage-guided crossover, on the other hand, amended such bias by allowing a better preservation of good features from the parents. When reproduction is not too disruptive, the offspring produced tend to be of similar quality to at least one of the parents. As a result, regions in objective space covered by solutions will likely continue being covered by offspring in the next generations, an important characteristic for diversity preservation.

Section 2 reviews many mechanisms proposed to improve diversity preservation from the perspective of reproduction. Section 3 analyzes the bias in decision space of p -uniform and n -point crossover. Section 3 demonstrates how such bias might impact the positioning of the objective vectors. Section 5 defines our idea of unbiased reproduction for bi-objective evolutionary optimization. Section 6 analyzes experimentally the reproductive bias of different operators, highlighting the influence of linkage learning on such bias.

*Corresponding author

Email address: jeanmartins@utfpr.edu.br (Jean P. Martins)

URL: acbd@icmc.usp.br (Alexandre C. B. Delbem)

2. Background

Although the use of effective reproduction operators is recognized as essential for the success of any evolutionary algorithm, many contemporary algorithms still rely on standard reproduction operators such as n -point and uniform crossover [42, 3, 53, 52]. The assumption that such operators could be directly employed to multiobjective optimization problems, shaped, to some extent, a developmental methodology for multiobjective evolutionary algorithms, which established reproduction, fitness assignment and selection (diversity preservation) as procedures with independent goals. Many researchers, however, have noticed that such separation imposes certain limitations to the effectiveness of the evolutionary process, as a consequence, methods have been proposed to circumvent such limitations. This section reviews many proposals that implicitly explore the relationship between reproduction operators and diversity preservation in the multiobjective context. By making explicit such relations it becomes possible to identify the general goals of reproduction operators for multiobjective optimization.

2.1. Crossover-based Reproduction

Traditional operators, such as n -point and uniform crossover, are still the reproduction methods most commonly employed in multiobjective optimization. However, the amount of methods proposed to improve their effectiveness indicates a certain dissatisfaction with their performance. One of the main complains is their negative effect on the diversity (in objective space) of the solutions produced.

Most studies regarding the effectiveness of traditional crossover operators for multiobjective optimization have emerged from the perspective of mating restrictions [7]. The main idea underlying mating restrictions is the assumption that reproduction should occur between parents that share some similarities, rather than parents randomly chosen. Such hypothesis is usually justified experimentally by the ability of mating restrictions to improve the diversity of non-dominated solutions along the Pareto front [4, 53].

Watanabe et al. [47], for example, coined the term *Neighborhood cultivation* to indicate that individuals would not be chosen randomly for reproduction, but based on their distance in objective space, in an attempt to generate offspring close to the parents. As far as we know, this was one of the first references to the idea of generating offspring close to the parents but other authors have also investigated the same problem. Jaszekiewicz [17], for example, stated the following:

“If the objectives are not positively correlated, their optima will be, in general, completely different. This suggests that recombination of approximately efficient solutions distant in the objective space is very unlikely to yield good offspring.”

Based on this consideration, Jaszekiewicz and Kominek [18] proposed a distance preserving recombination operator for the vehicle routing problem, in which some common substructures of the parents were identified and preserved during reproduction. The authors argued that, in such conditions, the offspring produced would tend to be closer to their parents in objective space. From a general perspective, the success of their approach depends on how easily representative substructures can be identified.¹

Ishibuchi and Shibata [13], generalized the notion of distance preserving recombination, by assuming that close parents in objective space should share more substructures than randomly selected parents. Therefore, they proposed a simple scheme for implementing mating restrictions on top of tournament selection. First, one parent is chosen by tournament selection, next, a pool of potential other parents are selected by tournament selection. The individual in the pool that is the closest to the first parent (in decision or objective space) is then chosen as the second parent.

Many other studies has followed from mating restrictions. Ishibuchi and Shibata [12, 14], for example, showed that was possible, to some extent, to balance convergence and diversity of the Pareto fronts found by using their similarity-based mating method. Ishibuchi and Narukawa [10], evaluated the effectiveness of similarity-based crossover through experiments with knapsack and flowshop problem instances; whereas Ishibuchi et al. [11] expanded and summarized the results. Sato et al. [39], also determined the benefits of performing recombination based on the locality of solutions in objective space, their results showed that such strategies become more important as the number of objectives grows. As an attempt to provide an explanation for the effectiveness of mating restrictions, Ishibuchi et al. [16, 9] argued that reproduction operators tend to decrease the spread of solutions due to their geometric properties, therefore, the authors evaluated the use of non-geometric operators as an alternative to circumvent the drawbacks. In their own words:

“It should be noted that crossover does not decrease the diversity of binary strings whereas it decreases the diversity of objective vectors”

By that time, Ishibuchi’s studies assumed that controlled similarity-based recombination of parents improved the diversity of non-dominated solutions in the objective space. Sato et al. [38] investigated such conclusions from a different perspective, they argued that by recombining similar parents the chance of generating offspring close to parents in objective space is increased due to the genetic material involved in such recombination. Therefore, the most important feature of such operators would be their ability to generate offspring close to their parents and not,

¹Please notice the similarity between this idea and the notion of building-blocks in the context of linkage-learning and estimation of distribution algorithms.

necessarily, the parents similarity. To evaluate their hypothesis, Sato et al. [38] proposed a crossover operator that exchanges only a limited number of bits, guaranteeing a certain proximity between parents and offspring in objective space. Ishibuchi et al. [15] compared experimentally both assumptions, and obtained similar conclusions. Recent studies continue to investigate these properties, but most of them fits on some of these two ideas previous described [33, 34, 8].

2.2. Model-based Reproduction

A retrospective analysis of the model-based reproduction operators employed by most estimation of distribution algorithms makes evident a situation very similar to what is seen in more conventional multiobjective evolutionary algorithms. Since many of the initial efforts to transpose model-based reproduction operators from single- to multi-objective problems did not completely meet the expectations, some of the contingency measures taken to circumvent their ineffectiveness were assumed to be essential. Although many interesting developments were found, nevertheless, such a direction may have prevented additional research to focus on multiobjective reproduction operators. This section briefly reviews some studies regarding model-based reproduction operators for multiobjective optimization, focusing on their ability to enforce diversity preservation.

One of the first model-based operators for multiobjective optimization was proposed by Thierens and Bosman [44], inspired by its single-objective version. Regarding diversity of the solutions in objective space, the authors stated the following:

“Clustering is necessary for the covering to take place: without it the algorithm finds only a small part of the Pareto front.”

Therefore, their algorithm cluster the population into many subpopulations and estimates a probabilistic model for each of them. By sampling each of these models, the whole Pareto front is expected to be covered by offspring. Such an approach can be seen as the model-based counterpart for restricted mating, but its adoption it is not a complete consensus among researchers.

The multiobjective versions of the Bayesian Optimization Algorithm (BOA), for example, have been implemented using single- and multiple- probabilistic models. Laumanns and Ocenasek [21], whom proposed one of the first implementations, used only one probabilistic model for the whole population (mutation was employed to improve diversity). Pelikan et al. [36, 35], on the other hand, followed the idea of using multiple models to achieve a better diversity preservation. Sastry et al. [37] endorse the necessity for multiple models by demonstrating that, in some cases, massive multimodality could prevent model-building of producing accurate models. Although similar conclusions to Sastry et al. [37]’s have been recently found for single-objective

problems [28, 29, 30, 31, 32], the validity of his results are still controversial, given the interesting results posteriorly found for benchmark functions and for some NP-hard multiobjective problems by Martins et al. [27], Shah and Reed [40]. Later on we demonstrate how to conciliate these conclusions.

Zhou et al. [51] also opt for the use of multiple models in their algorithm, but differently from the previous proposals, model-sampling and crossover-mutation were used alternatively. In subsequent studies, Zhou et al. [49] have proposed a method to identify the proper moment to use crossover or model-sampling, after that Zhang et al. [48] gave up the use of crossover, but not without mentioning their surprise on the lack of studies regarding reproduction operators for multiobjective optimization.

“Surprisingly, not much work has been done on how to generate new solutions in MOEAs. The implementation of most current MOEAs directly adopt traditional genetic recombination operators such as crossover and mutation.”

Zhou et al. [50] continued their work by proposing a method to approximate the Pareto-set and the Pareto front simultaneously along the evolutionary process.

As far as we known, Martí et al. [26] was one of the first authors to investigate specific guidelines for developing model-based reproduction operators in the multiobjective context, their main criticism can be summarized by their statement:

“In particular, we have identified three of such issues: the incorrect treatment of outliers, the loss of population diversity, and the excess computational effort spend on finding an optimal model of the fittest population elements.”

Martí et al. [24, 23] argued that outliers represent solutions in underexplored, or recently discovered, regions of the objective space, therefore, their proper consideration is essential to obtain representative probabilistic models of the Pareto-front. Afterwards, Martí et al. [25] described their final results and algorithms.

Many innovative ways to deal with diversity preservation from the perspective of model-building are still being developed. Karshenas et al. [19, 20], for example, investigated the modeling of objectives values and decision variables simultaneously. Since the objective values are considered part of the model, outliers tend to be better represented, which is useful for diversity preservation. From other perspective, Gao et al. [6] and Shim et al. [41], for example, followed a decomposition-based framework to define the subpopulations in which model-building would be performed, which is also an effective way to deal with the lost of diversity faced in more traditional scenarios of evolutionary optimization.

3. Reproductive bias in decision space

The studies briefly described in the previous section illustrate the ineffectiveness of traditional reproduction operators for multiobjective problems. The most common argument being the tendency of traditional reproduction operators to generate solutions in the middle region of the objective space, a harmful characteristic in terms of diversity preservation. This section provides the foundation for our hypothesis and analyzes the n -point and p -uniform crossover operators in decision space.

3.1. p -uniform crossover

Given two parents $\mathbf{x}, \mathbf{y} \in \{0, 1\}^N$, the uniform crossover, with probability p , generates $\mathbf{v} \in \{0, 1\}^N$ such that

$$\begin{aligned} P(v_i = x_i) &= p, & \forall i = 1, \dots, N \\ P(v_i = y_i) &= 1 - p \end{aligned}$$

Therefore, the number of bits of \mathbf{v} that comes from \mathbf{x} is a random variable \mathcal{B}_p with expected value

$$\mathbb{E}[\mathcal{B}_p] = Np. \quad (1)$$

In the traditional parameterization, with $p = 1/2$, the offspring \mathbf{v} will differ, on average, by $N/2$ bits from its parent \mathbf{x} . The increase (decrease) of p , will increase (decrease) the similarity between \mathbf{x} and \mathbf{v} .

3.2. n -point crossover

Given two parents $\mathbf{x}, \mathbf{y} \in \{0, 1\}^N$, the n -point crossover cuts each parent into $n + 1$ segments; the offspring \mathbf{v} is then composed of interchangeable segments from \mathbf{x} and \mathbf{y} .

If n is an odd number, n -point crossover produces an even number of segments, therefore, \mathbf{v} will be composed of an equal number of segments from each parent. The size of each segment varies but the total expected number of bits transferred from each parent is, on average, $N/2$.

For example, assume $n = 1$. There are $N - 1$ possible crossover points, as illustrated below. Each of these crossover points occur with equal probability $1/(N - 1)$ and transmits s bits from \mathbf{x} to the offspring.

$$(N-1) \left\{ \begin{array}{ll} \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=1)} \\ \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=2)} \\ \dots \\ \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=N-1)} \end{array} \right.$$

Therefore, the expected number of bits transferred from \mathbf{x} to \mathbf{v} must account for all these possibilities. For $n = 1$ we have,

$$\mathbb{E}[\mathcal{B}_{n=1}] = \sum_{s=1}^{N-1} s \cdot (N-1)^{-1} = \frac{N}{2}. \quad (2)$$

If n is an even number, the n -point crossover cuts the parents into an odd number of segments.

For example, assume $n = 2$. There are $\binom{N-1}{2}$ different ways to choose the crossover points, each of them with equal probability of occurrence. Besides, there are $(N - s - 1)$ different ways to transmit s bits from one parent to the offspring. As illustrated below for $s = 1$:

$$(N-s-1) \left\{ \begin{array}{ll} \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=1)} \\ \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=1)} \\ \dots \\ \overbrace{x_1, x_2, x_3, x_4, x_5, \dots, x_{N-1}, x_N}^{(s=1)} \end{array} \right.$$

Therefore, the expected number of bits transferred from \mathbf{x} to \mathbf{v} , when $n = 2$ is given by

$$\mathbb{E}[\mathcal{B}_{n=2}] = \sum_{s=1}^{N-1} s(N-s-1) \cdot \binom{N-1}{2}^{-1} = \frac{N}{3}. \quad (3)$$

In words, the offspring \mathbf{v} will be composed (on average) by $N/3$ bits from \mathbf{x} and $2N/3$ bits from \mathbf{y} . However, as n assumes larger even values, the size of the segments will shrink and $\mathbb{E}[\mathcal{B}_n] \approx N/2$.

3.3. Empirical verification

To verify and extend the previous analysis, p -uniform and n -point crossover were performed multiple times on a $N = 100$ bit string, and the average number of bits transferred computed. Figure 1 shows the average results for $0 \leq p \leq 1$ and $1 \leq n < N$.

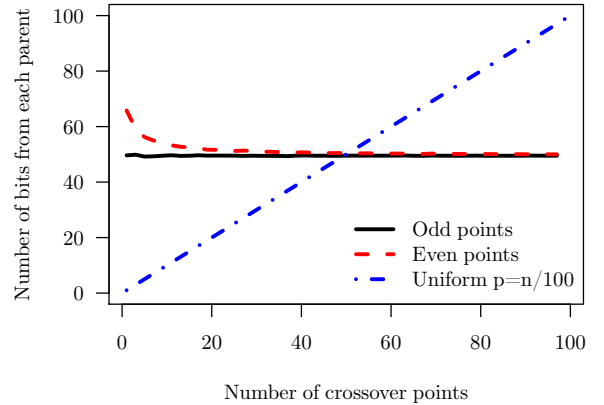


Figure 1: For n odd, half the bits are expected to come from each parent. For n even, the value decreases and also tends to $N/2$. For uniform crossover, \mathcal{B}_p varies linearly with p .

In summary, the n -point crossover transfer $\mathbb{E}(\mathcal{B}_n) \approx N/2$ bits to the offspring for all $n > 1$. Whereas the p -uniform crossover allows parent-offspring similarity to be controlled by increasing/decreasing p , i.e. $\mathbb{E}(\mathcal{B}_p) = Np$.

4. Reproductive bias in objective space

Many multiobjective evolutionary algorithms have been proposed relying on traditional crossover operators and their standard settings, i.e. $p = 0.5$ and $n = 1, 2$. In the previous section, we have shown that these settings produce similar results regarding the number of bits transferred from one parent to the offspring

$$\mathbb{E}(\mathcal{B}_p) \approx \mathbb{E}(\mathcal{B}_n) \approx N/2.$$

In this section we show how such property may induce a bias in objective space, hampering diversity preservation. Additionally, based on our hypothesis, we provide an explanation for the usefulness of restricted mating and linkage-learning in diversity preservation.

Disruptiveness, is the term usually employed to describe the inability of crossover operators to preserve the good features of the parents. In this sense, a too disruptive crossover operator would be likely to produce offspring whose quality is worst than the parents. For randomized operators such as p -uniform and n -point, it is reasonable to assume the expected number of transferred bits as a measure of disruptiveness.

In unimodal single-objective optimization problems, high-quality solutions always share the same good features, therefore, the effects of disruptiveness can be softened by choosing high-quality parents. In multimodal optimization, however, this is not guaranteed, since the good features of high-quality parents can also be conflicting.

In multiobjective optimization, where the problems are inherently multimodal, highly disruptive crossover operators are likely to destroy good features of the parents. We hypothesize that, as a result, the objective vectors of the offspring can be expected to inhabit the *inter-parents region* in objective space (see Figure 2(a)), due to an *averaging effect* on the quality of the parents.

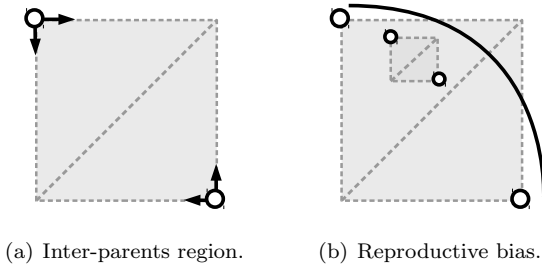


Figure 2: Highly disruptive operators tend to produce offspring whose objective vectors are more likely to inhabit the inter-parents region.

Such a characteristic requires attention mainly for the case of negatively correlated objective functions. If parents are chosen independently of their distance in objective space, the inter-parents region will frequently coincide with the central region of the objective space, producing a bias towards that region, see Figure 2(b).

According to that hypothesis, highly disruptive crossover, such as p -uniform (for $p = 0.5$) and n -point crossover, would produce more offspring in the central region of the objective space, hampering diversity preservation.

Although not describing their purposes in terms of reproductive bias, most of the methods described in Section 2 attempt to decrease crossover disruptiveness. As a result, those methods decrease the reproductive bias, which allows a better diversity preservation in objective space. Among those methods, we identified three main approaches.

1. Restricted mating: only similar parents reproduce (close in objective or decision space) [11],
2. Model-based mating: explicitly model the good features of the parents and try to preserve them [40],
3. Controlled exchange: decreases the expected number of bits transferred from parents to offspring [38, 15].

Although superficially different, we argue that these three approaches agree in their purpose: decrease disruption to generate offspring of similar quality to the parents, i.e. close in objective space.

Let us take as example, restricted mating, where only similar parents (close in decision or objective space) are allowed to reproduce.² The main idea here is that similar parents would be more likely to share the same good features. Therefore, by increasing the parents similarity, we would decrease the effects of crossover disruption and inhibit the generation of offspring too far from the parents (see Figure 3).

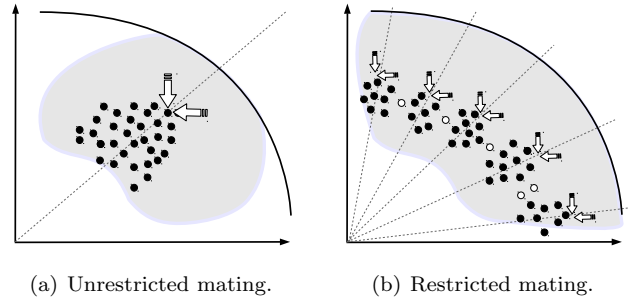


Figure 3: Illustrative example of the impact of restricted mating on the disruptiveness of crossover and the reproductive bias.

In summary, in a restricted mating scenario, the intrinsic bias of p -uniform and n -point crossover operators would not impact diversity preservation as intensely as in the unrestricted scenario. The positive effects of restricted mating, or any method that controls disruptiveness, are expected to be intensified when considering negatively correlated objective functions. That might be the reason why some researchers did not find significant improvements when applying restricted mating.

²It is reasonable to assume as forms of restricted mating: decomposition-based and some probabilistic-based methods [22, 36].

5. Unbiased multiobjective reproduction

So far we have argued that highly disruptive reproduction operators might bias the positioning of the offspring in objective space. As a consequence, traditional crossover operators, in their standard settings, would hamper the efficacy of diversity preservation. We also described how, many of the improvements regarding multiobjective reproduction operators can be interpreted as an attempt to decrease disruptiveness, and consequently, decrease bias.

From this point of view, it is natural to think on reproduction operators that would inherently avoid bias in objective space and contribute to diversity preservation. This section defines our concept of *unbiased multiobjective reproduction*.

Given a set of weight vectors \mathbf{W} , a reproduction operator is said *unbiased* in relation to \mathbf{W} if the offspring it produces are uniformly distributed according to \mathbf{W} , i.e., on average, the same amount of offspring is generated in each of the regions covered by each aggregated single-objective function. Figure 4 illustrates this idea.

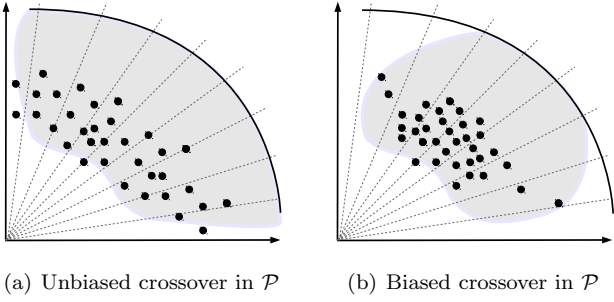


Figure 4: Points representing the offspring generated by unbiased (a) and biased (b) crossovers. The unbiased version exploits the regions of all possible aggregations uniformly and enables the exploration of a more well-spread set of objective vectors (shaded region).

In order to measure the reproductive bias of an operator, the amount of solutions it generates in each subregion of the objective space must be accounted. Let us consider the example illustrated by Figure 5.

In this example, there are only three weight-vectors, consequently, three aggregated single-objective functions. Each offspring produced is represented by its objective vector, which must be closer to one of the weight-vectors than to the others. Such interpretation, provides us a mechanism to count the number of solutions in each subregion. Therefore, we argue that an unbiased reproductive operator must preserve the distribution of objective vector in relation to \mathbf{W} , which is equivalent to say that every region, defined by an weight-vector, must be covered by approximately the same amount of objective vectors.

This idea can be easily measured experimentally in order to verify if the traditional operators are biased or not. At every generation t , a nadir vector $\mathbf{z}^{(nadir,t)}$ is estimated from the current population $P(t)$ and used as reference

point.

$$\mathbf{z}^{(nadir,t)} = (\min_{\mathbf{z} \in P(t)} z_1, \dots, \min_{\mathbf{z} \in P(t)} z_M), \quad (4)$$

where z_i is the objective value associated with a solution $\mathbf{x} \in P(t)$ for the i^{th} objective function.

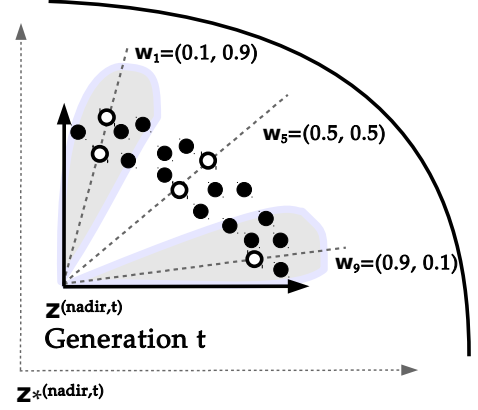


Figure 5: Illustrative example of weight vectors and the subregions they define. The white points are objective vectors whose ideal weight vectors \mathbf{w}^* coincide with some $\mathbf{w} \in \mathbf{W}$.

To every objective vector \mathbf{z} there is an ideal weight vector \mathbf{w}^* associated, which produces an aggregated single-objective function in the same positioning of \mathbf{z} (white points in Figure 5 represent \mathbf{z} vectors whose ideal weight vector \mathbf{w}^* coincides with some $\mathbf{w} \in \mathbf{W}$).

$$w_i^* = \frac{z_i - z_i^{(nadir,t)}}{\sum_{i=1}^M (z_i - z_i^{(nadir,t)})}, \quad \forall i = 1, \dots, M. \quad (5)$$

Such ideal weight vectors are then compared with the pre-defined \mathbf{W} . For our purposes we define ten well-distributed weight vectors $\mathbf{W} = \{\mathbf{w}_i : \mathbf{w}_i = (i/10, 1 - i/10), \forall i = 0, \dots, 10\}$. In this case, with $M = 2$, each $\mathbf{w}_i \in \mathbf{W}$ is uniquely identifiable from its first component, therefore, the closest \mathbf{w}_i to any \mathbf{w}^* can be computed as $i = \lceil \|\mathbf{w}_1 \times 10\| + 1 \rceil$, where $\lceil \cdot \rceil$ is the nearest integer value. In other words, an objective vector \mathbf{z} , with ideal weight vector \mathbf{w}^* is within the region defined by the weight vector $\mathbf{w}_{\lceil \|\mathbf{w}_1 \times 10\| + 1 \rceil}$. The number of objective vectors generated in each subregion characterize the bias of any specific crossover operator. Ideally, in unbiased operators, such a number should be uniformly distributed among the $\forall \mathbf{w} \in \mathbf{W}$.

6. Experiments

In this section we evaluate experimentally the hypothesis raised so far. In summary we are interested in answering three questions:

1. Are n -point and p -uniform crossover biased?
2. Does a decrease of disruptiveness decreases the bias?

3. Does linkage-learning decreases the bias?

In order to compare our hypotheses with experimental results, all alternative reproduction operators were implemented on top of the same algorithm, the Non-dominated Sorting Genetic Algorithm II (NSGA-II).

First, Section 6.1 describes the ρMNK -landscapes model used to generate problem instances with different degrees of multimodality and correlation. Then, Section 6.2 describes the settings used for traditional operators such as n -point and p -uniform crossover. Finally, Section 6.3 specifies the linkage-guided crossover that was used to assess the influence of linkage-learning in the reproductive bias.

6.1. The ρMNK -landscapes Problem

Multi-objective MNK -landscapes were proposed by Aguirre and Tanaka [1], whom studied the performance of one-bit climbers regarding: problem size N , degree of non-linearity K and number of objectives M [2]. As a continuation of these studies, Verel et al. [45] proposed and analyzed the fitness landscapes of MNK -landscapes with different degrees of correlation, ρ , among objective functions (ρMNK -landscapes).

$$\max f_{\rho MNK}^j(\mathbf{x}), \forall j \in [M] \quad (6)$$

Verel et al. [46] also provided an instance generator for the ρMNK model, which is used in our experiments³.

The experiments were performed on bi-objective instances of size $N = 500$ and epistasis degrees $K = 2, 4, 8, 12$. As discussed by Verel et al. [46], instances with negative correlations produce wider Pareto-fronts appropriate for the evaluation of diversity preservation mechanisms, whereas instances with zero or higher correlation produce Pareto-fronts of shorter extension and more distant from the initial population. Since diversity preservation is our main concern, results are shown for instances with $\rho = -0.9, 0$ correlation. For each instance, 30 runs were performed, with a limit of 500,000 fitness evaluations each.⁴

6.2. Traditional Crossover bias

As previously shown, p -uniform crossover is more flexible than n -point, therefore, we restrict our attention to it. Figures 6-7 show the EAFs for instances with $\rho = -0.9$ and $\rho = 0$, respectively, where we set $p = 0, 0.1, 0.5$.

Figure 6 shows that when no crossover is applied ($p = 0$) the NSGA-II faced difficulties in exploring the middle region of the Pareto-front but showed a much better performance in extreme regions. With slightly increase to $p = 0.1$, the Pareto-fronts found were mostly equivalent to the results of $p = 0.5$ in the middle region, but with advantages in the extreme regions. This result corroborates the hypothesis that small disruptiveness is beneficial to the exploration of extreme regions.

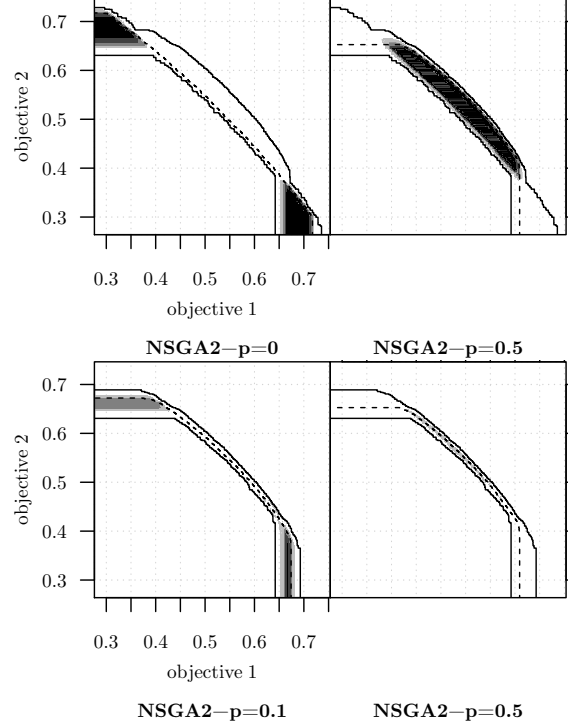


Figure 6: Comparison of p -uniform on instances with $\rho = -0.9$.

Figure 7 shows the results for instances with $\rho = 0$. These instances have a less extensive Pareto-optimal front and the importance of diversity preservation is not evident. In fact, for such instances, the use of crossover seems unattractive as shown for the case of $p = 0$.

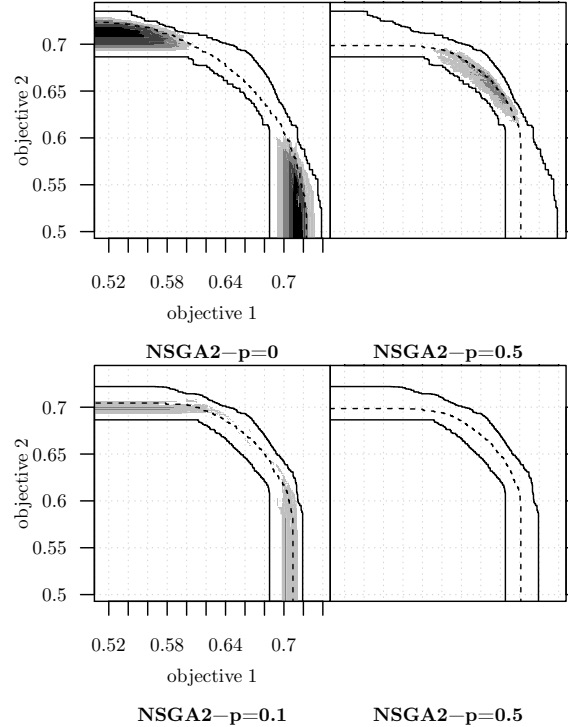


Figure 7: Comparison of p -uniform on instances with $\rho = 0$.

³<http://mocobench.sourceforge.net/>

⁴The results were compared by the Empirical Attainment Function (EAF) [5] computed with the R package provided by Manuel López-Ibáñez, <http://iridia.ulb.ac.be/~manuel/eaftools>.

6.3. Linkage-guided Crossover bias

Any linkage learning algorithm can be used to implement a linkage-guided crossover, i.e. a crossover operator that exchanges only bits identified as statistically dependent. In this context, it is of interest that we could predetermine the number of “linked” groups to be identified by linkage learning. Such an algorithm was implemented using as reference the linkage-tree learning algorithm [43]. Therefore, by defining a parameter $\Psi \geq 2$, we can easily obtain Ψ groups of “linked” variables to be used during crossover.

For example, for $\Psi = 2$ a linkage model consisting of $\mathcal{T} = \{\tau_1, \tau_2\}$ would be produced, where τ_1 and τ_2 are disjoint sets of variables, such that $|\tau_1| + |\tau_2| = N$. Since linkage learning is expected to identify good features of the solutions, by exchanging only the subgroups of variables indicated by \mathcal{T} , the offspring produced by linkage-guided crossover should be closer to its parents, in objective space, than those produced by p -uniform crossover. Obviously, this argument is only valid if \mathcal{T} represents meaningful information.

Let us first show the results for the linkage-guided crossover and only then to explain why they were reached. Figure 8 shows the EAFs generated by the linkage-guided crossover using linkage-models with $\Psi = 2, 9, 17$ groups of variables and uniform crossover with $p = 0.5$ in negatively correlated instances with $\rho = -0.9$.

Notice that with $\Psi = 2$, the Pareto-fronts found by linkage-guided crossover are much wider than those found by p -uniform crossover ($p = 0.5$). In one hand, this result corroborates the assumption that linkage-learning helps to preserve the good features of parents, enabling a better exploration of extreme regions. On the other hand, it evidences that proper exploration of the middle region of the Pareto-front seems to be related to higher degree of disruptiveness during reproduction.

As Ψ was increased to 9, then 17, the benefits of linkage-guided crossover regarding the extreme regions deteriorated but at the same time, a better exploration of the middle region was enabled. By relating Ψ to the degree of disruptiveness of linkage-guided crossover, such a result is easy to explain. As Ψ increases, the linkage groups decrease in average size, therefore, the expected number of bits transferred from each parent approaches $\mathbb{E}(\mathcal{B}_\Psi) \approx N/2$ and in the limit linkage-guided crossover will act similarly to uniform crossover with $p = 0.5$.

Analogous results were found for instances with

$$\rho = -0.5, 0, 0.5, 0.9.$$

However, as ρ increases to positive values, the optimal Pareto-front becomes narrower and the benefits of linkage learning are insignificant, since in such cases diversity preservation is not a real concern.

6.4. Bi-objective linkages

The previous results indicate that somehow linkage learning helps to decrease the reproductive bias and im-

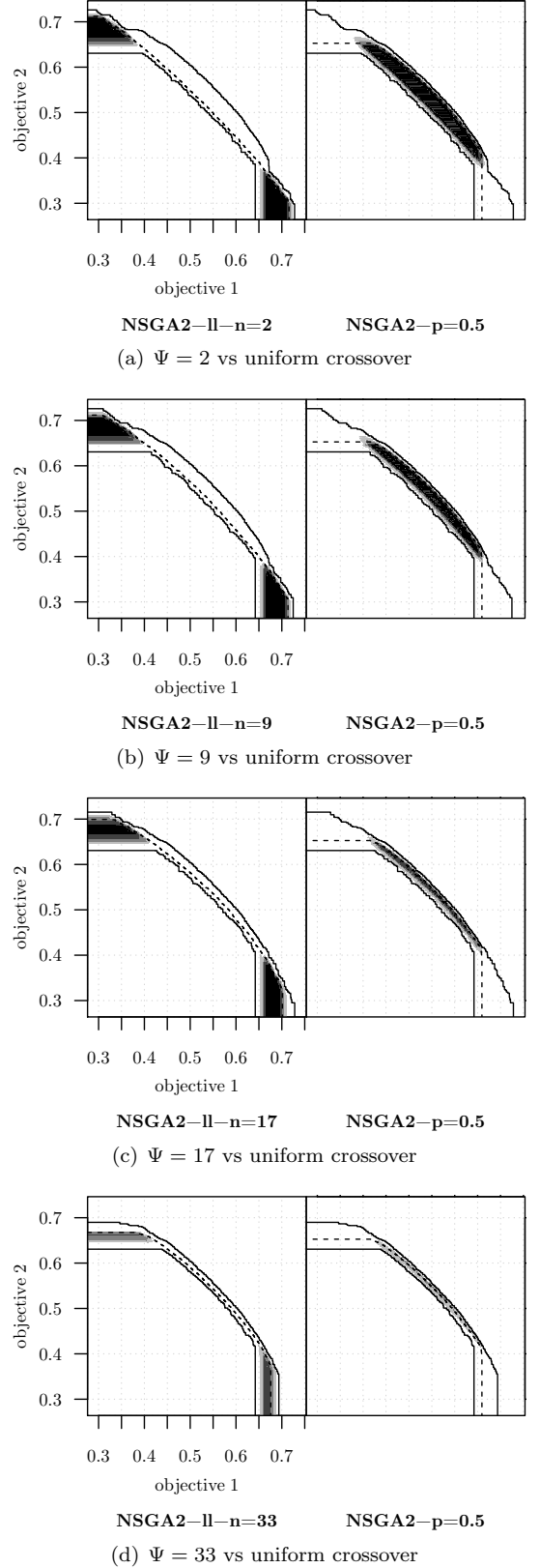


Figure 8: Comparison between the linkage-guided crossover (NSGA2-ll-n= Ψ) and uniform crossover with $p = 0.5$ (NSGA2-p=0.5) for instances with $\rho = -0.9$. Darker regions represent the better performance of each algorithm.

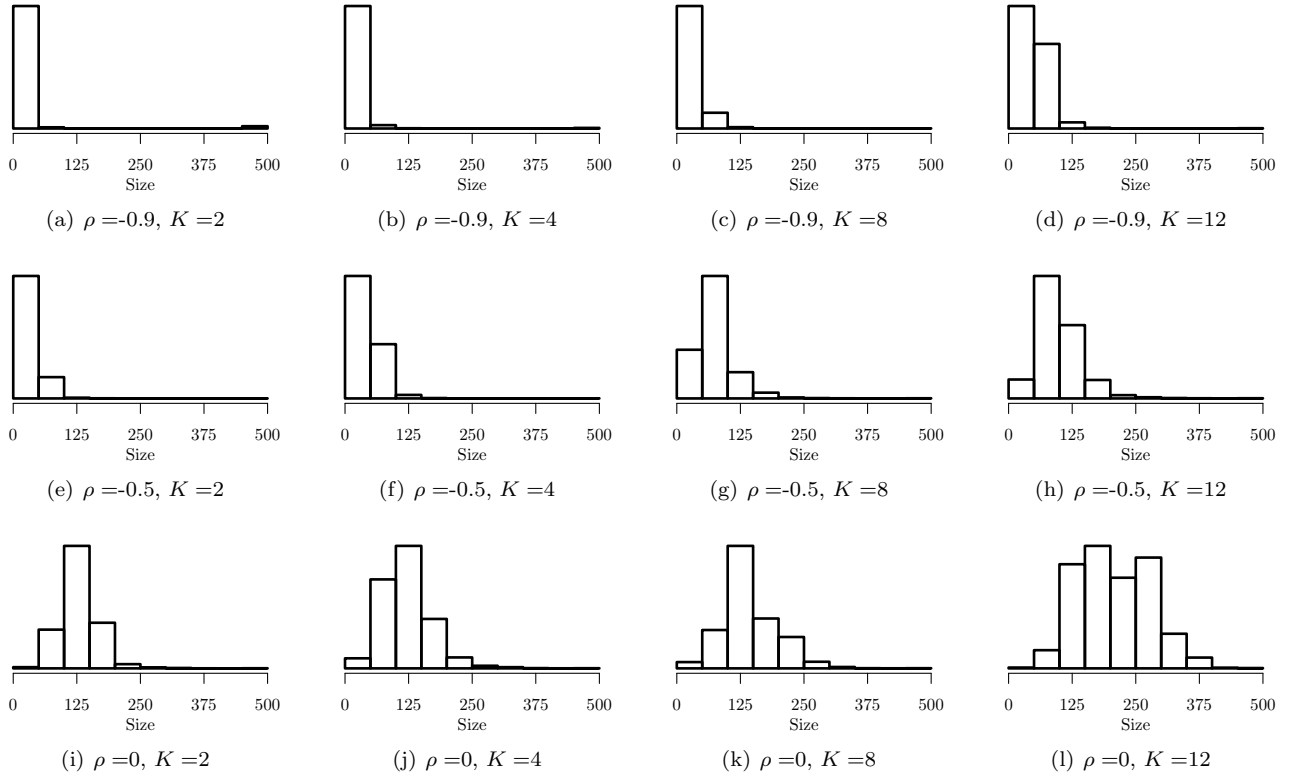


Figure 9: The distribution of $|\tau_1|$ along the generations in instances with $\rho = -0.9, -0.5, 0$ and $K = 2, 4, 8, 12$.

prove diversity preservation. From now on we investigate how linkages influence crossover. The first step in such a direction consists of characterizing bi-objective linkages.

Linkage learning in single-objective optimization has been shown to face limitations in highly multimodal search spaces [28, 29, 30, 31, 32], the same being valid for some artificial multiobjective problems [37]. Conversely, interesting results have been found for some NP-complete optimization problems [40]. This section shows that these conclusions are not contradictory, by noting that in each case linkage learning models different kinds of information.

Consider the case of $\Psi = 2$. At every generation a linkage model $\mathcal{T} = \{\tau_1, \tau_2\}$ is built. Since the linkages are built according to statistical information from the population, they must reflect some property of the objective functions. Observe Figure 9, which shows how the sizes of the linkage set τ_1 was distributed during the experiments for instances with $K = 2, 4, 8, 12$ and correlation $\rho = -0.9, -0.5, 0$.

In the case of $\rho = -0.9$, for example, $|\tau_1| \approx 50$ in many situations, which says something about what is being modeled. As K grows to 12, the size of τ_1 has a large variation, which indicates the greater difficulty of “learning” in the context of rugged search spaces. For $\rho = -0.5, 0$ the situation repeats. In summary, the size of τ_1 seems to be correlated with the number of correlated variables in the objective functions, given by ρ . Therefore, we conclude that is reasonable to assume that bi-objective linkages represent correlated decision variables between objective functions.

Such characterization helps to understand why linkage-guided crossover seems to better preserve the parents good features and decrease the reproductive bias. The next sections provide more evidences for this conclusion.

6.4.1. Linkages and reproductive bias

Assume the hypothesis that the linkages identified by linkage learning represent decision variables that are correlated in both objective functions. If this is true, linkage-guided crossover is expected to be less disruptive than uniform crossover for $\Psi = 2$, as a consequence, the offspring produced would tend to be closer of the parents in objective space. This is a reasonable assumption due to the fact that correlated decision variables have considerable influence on the objective values. This section evaluates experimentally if this assumption holds for instances of the ρMNK -model.

Using the same NSGA-II employed in the previous experiments, we extract additional information from the runs by the following procedure. Whenever a solution \mathbf{y} , with objective vector $\mathbf{z}_{\mathbf{y}}$, is produced from parents \mathbf{x}_1 and \mathbf{x}_2 , the parental objective vector which is the closest to $\mathbf{z}_{\mathbf{y}}$ is identified, let us assume it is $\mathbf{z}_{\mathbf{x}_1}$. By assuming $\mathbf{z}_{\mathbf{x}_1}$ as the origin of a coordinate system, the position of $\mathbf{z}_{\mathbf{y}}$ relative to $\mathbf{z}_{\mathbf{x}_1}$ provides a vector linking a parent to the child in the objective space, a *reproduction vector*. Therefore, for every reproduction we have: (1) the distance between the parents, (2) the positioning of the offspring in relation to

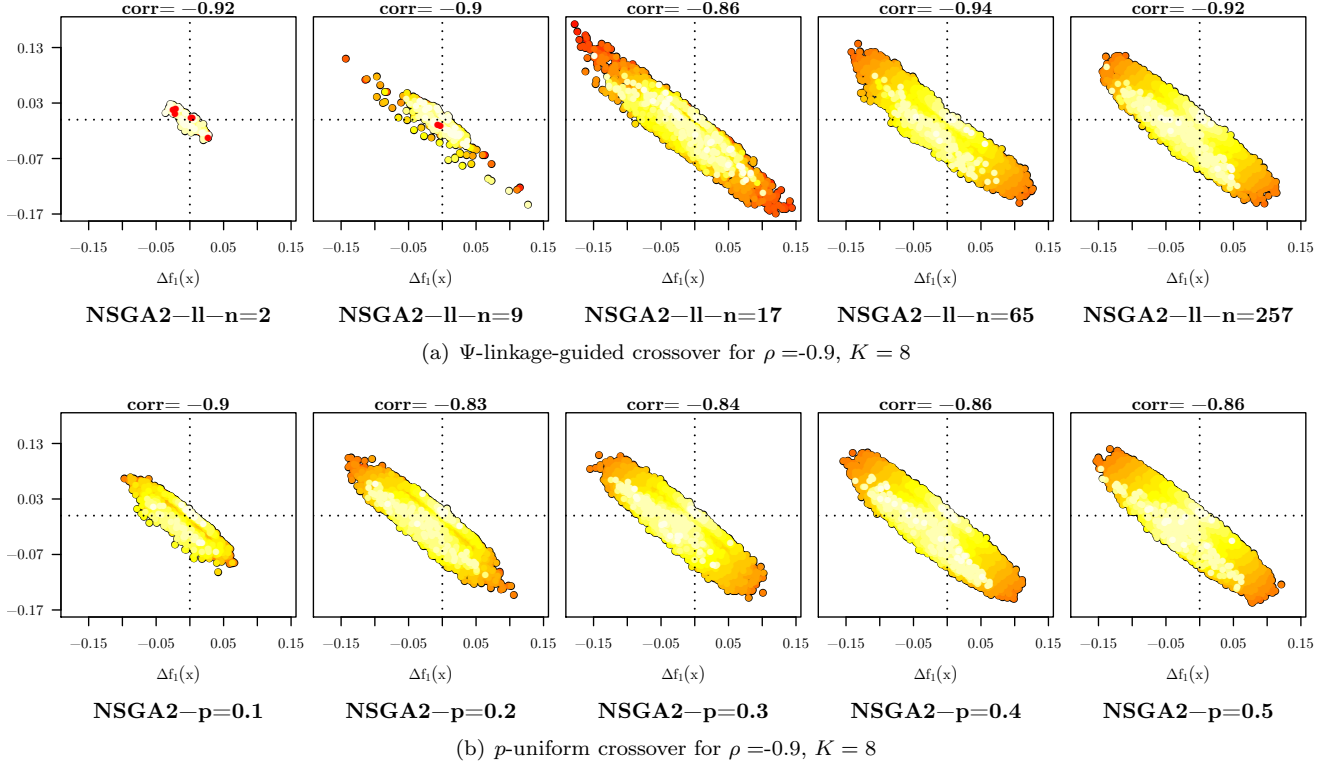


Figure 10: Reproduction vectors generated by linkage-guided and uniform crossovers, for instances with $\rho = -0.9$. The color of a point represents its distance to the closest parent, darker red implies more distant parents.

its closest parent; both measured in the objective space.

The reproduction vectors provides a practical way to verify if the solutions produced by linkage-guided crossover are really closer to the parents than those produced by uniform-crossover. Figure 10 illustrates the reproduction vectors generated for instances with $\rho = -0.9$ and $K = 8$.

The reproduction vectors are represented as points, whereas the distance between the parents is represented by their color (close = white, distant = red). Linkage-guided crossover with $\Psi = 2$ always produced offspring close to one of the parents. However, as Ψ was increased, disruptiveness also increased and the parent-to-offspring distance got bigger. Uniform-crossover followed the same pattern, with $p = 0.1$ it produced offspring reasonably close to the parents, which was not true for $p = 0.5$.

This is an interesting result since it shows how these reproductive operators behave when in conjunction with diversity preservation mechanisms. We believe that an operator that always generate offspring close to the parents, in objective space, tend to preserve the distribution of objective vectors, avoiding a bias towards certain regions. In the next section we use the method described in Section 5 to effectively measure the bias of different reproduction operators and verify if this conclusion holds.

6.4.2. Measuring the reproductive bias

We have argued that reproductive bias is a consequence of highly disruptive reproduction operators. It was shown

that such bias can be reduced by decreasing disruptiveness, in both linkage-guided and uniform crossover operators (though more effectively in linkage-guided crossover). We hypothesized that in situations of low disruptiveness, offspring are more likely to be generated close to the parents, which would help to preserve the current distribution of objective vectors and better explore the whole extension of the Pareto-front. This section evaluates such hypothesis by effectively measuring the amount of offspring generated in each region of the objective space, for this we use the method described in Section 5.

Every subregion of the objective space is associated with a weight-vector, whenever an offspring is produced in some region, this occurrence is accounted. At the end, this procedure results in an histogram of the number occurrences during the whole run. In an unbiased reproduction operator, all the subregions will be covered by approximately the same amount of solutions, whereas in a biased one, we expect the region defined by the weight-vector $(0.5, 0.5)$ to be covered by more solutions.

Figure 11 shows the reproductive bias averaged from 30 runs. Each run comprised 2,500 generations performed on a population of size $\mu = 100$. In a bi-objective space, every weight-vector can be uniquely identified by its first component, represented by the horizontal axis. The reproductive bias of n -point, p -uniform and linkage-guided crossovers endorses all the hypothesis and analysis made so far.

The peaks in this graphics indicates the middle region

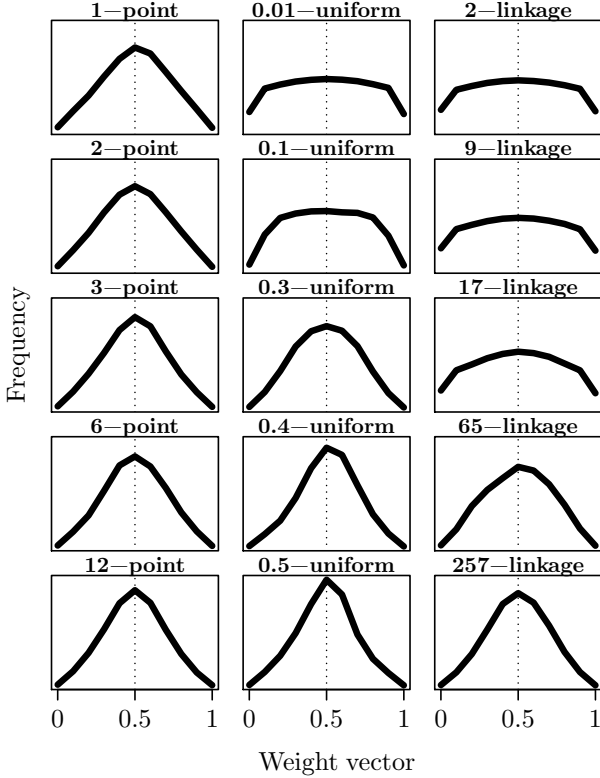


Figure 11: Reproductive bias of n -point p -uniform and Ψ -linkage guide crossover. The absence of a peak indicates the absence of bias.

of the Pareto-front. In this sense, the n -point crossover is biased towards the center, and no value for n changes this behavior. The p -uniform crossover, on the other hand, allow us control over disruptiveness and consequently its bias can be diminished by using a small probability p . Overall, linkage-guided crossover proved to be the best alternative as far as reproductive bias is concerned, it produced well distributed offspring for a wide range of Ψ values.

From these results we can conclude our chain of reasoning. Reproduction operators that are able to generate offspring close to one of the parents are also able to preserve the distribution of the current Pareto-front. Therefore, given that representative parents are chosen, all regions of interest will be covered by the new offspring.

7. Conclusions

The diversity preservation problem has always been a concern in multiobjective evolutionary optimization. However, most of the early proposals to solve this problem have considered it as an independent component within evolutionary algorithms. As a result, well-known algorithms such as NSGA-II and its contemporaries, worked by first, generating an offspring, and then assigning an adapted fitness to indicate how interesting a solution is in terms of diversity preservation. In this paper, we showed that such strategies might have their efficacy hampered by commonly used reproduction operators, since in their standard

settings they produce offspring that are biased to certain regions of the objective space. In this sense unbiased reproduction operators can considerably improve diversity preservation and should be a required characteristic for any general purpose multiobjective reproduction operator.

At first, we analyzed traditional reproduction operators, such as p -uniform and n -point crossover, and demonstrated that, due to their high disruptiveness, the offspring they produce (for $p = 0.5$ and any n) were biased towards the center of the objective space. We argued that such bias was the result of an averaging effect on the quality of the parents. Therefore, we hypothesized that by decreasing disruptiveness it would be possible to decrease the reproductive bias and produce an more well distributed offspring (a similar conclusion was found by Sato et al. [38]).

But how to decrease disruptiveness of reproduction operators? We evaluated two alternatives: (1) the use of a small p for uniform crossover (2) the use of a linkage-guided crossover. An uniform crossover with $p < 0.5$ exchanges less genetic material between the parents, therefore, it is less disruptive. A linkage-guided crossover, on the other hand, will only exchange probably good substructures, therefore, it is less disruptive. The experiments were performed on instances of the ρMNK -model from which results were found that corroborate the hypothesis, i.e. the reproductive bias decreased as the disruptiveness of the operators decreased.

Additionally, we observed that linkage learning when applied to the ρMNK instances, did not necessarily find substructures of the objective functions, instead we showed evidence that the linkage groups identified might in fact represent decision variables that are correlated in terms of the multiple objectives. This interpretation helps to conciliate two previous results that seemed contradictory in the literature of estimation of distribution algorithms. First, Sastry et al. [37] affirmed that due to the high multimodality of multiobjective problems, multiple linkage models would be required to properly model the whole Pareto-front. Shah and Reed [40], on the other hand, demonstrated that a multiobjective Bayesian optimization algorithm employing only one linkage model could outperform many other evolutionary algorithms. Our results propose that in multiobjective problems the correlation among the objective functions is another dimension for “learning” that can conciliate these interpretations and should be further explored.

This paper analyzed the influence of reproduction on diversity preservation from the perspective of well known crossover operators and linkage learning. We believe the results here discussed improves our understanding of the fundamentals for multiobjective reproduction operators and how estimation of distribution algorithms can be useful in this context.

Acknowledgments

The authors are indebted to the FAPESP for the financial support provided for this research (2011/07792-4).

References

- [1] Aguirre, H.E., Tanaka, K., 2004. Effects of elitism and population climbing on multiobjective MNK-landscapes, in: *Evolutionary Computation*, 2004. CEC2004. Congress on, pp. 449–456 Vol.1. doi:[10.1109/CEC.2004.1330891](https://doi.org/10.1109/CEC.2004.1330891).
- [2] Aguirre, H.E., Tanaka, K., 2007. Working principles, behavior, and performance of MOEAs on MNK-landscapes. *European Journal of Operational Research* 181, 1670–1690. doi:[10.1016/j.ejor.2006.08.004](https://doi.org/10.1016/j.ejor.2006.08.004).
- [3] Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE transactions on evolutionary computation* 6, 182–197.
- [4] Fonseca, C.M., Fleming, P.J., 1995. An Overview of Evolutionary Algorithms in Multiobjective Optimization. *Evolutionary Computation* 3, 1–16.
- [5] Fonseca, C.M., Guerreiro, A.P., López-Ibáñez, M., Paquete, L., 2011. On the Computation of the Empirical Attainment Function. Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 106–120. doi:[10.1007/978-3-642-19893-9_8](https://doi.org/10.1007/978-3-642-19893-9_8).
- [6] Gao, F., Zhou, A., Zhang, G., 2012. An estimation of distribution algorithm based on decomposition for the multiobjective tsp, in: *Natural Computation (ICNC)*, 2012 Eighth International Conference on, pp. 817–821. doi:[10.1109/ICNC.2012.6234618](https://doi.org/10.1109/ICNC.2012.6234618).
- [7] Goldberg, D.E., 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional.
- [8] He, Z., Yen, G.G., 2016. Many-objective evolutionary algorithm: Objective space reduction and diversity improvement. *IEEE Transactions on Evolutionary Computation* 20, 145–160. doi:[10.1109/TEVC.2015.2433266](https://doi.org/10.1109/TEVC.2015.2433266).
- [9] Ishibuchi, H., Akedo, N., Nojima, Y., 2012. Recombination of Similar Parents in SMS-EMOA on Many-Objective 0/1 Knapsack Problems, in: *Parallel Problem Solving from Nature - PPSN XII*. Springer. volume 7492 of *Lecture Notes in Computer Science*, pp. 132–142. doi:[10.1007/978-3-642-32964-7_14](https://doi.org/10.1007/978-3-642-32964-7_14).
- [10] Ishibuchi, H., Narukawa, K., 2005. Recombination of Similar Parents in EMO Algorithms, in: *Evolutionary Multi-Criterion Optimization*. Springer Berlin Heidelberg. volume 3410 of *Lecture Notes in Computer Science*, pp. 265–279. doi:[10.1007/978-3-540-31880-4_19](https://doi.org/10.1007/978-3-540-31880-4_19).
- [11] Ishibuchi, H., Narukawa, K., Tsukamoto, N., Nojima, Y., 2008. An empirical study on similarity-based mating for evolutionary multiobjective combinatorial optimization. *European Journal of Operational Research* 188, 57–75. doi:[10.1016/j.ejor.2007.04.007](https://doi.org/10.1016/j.ejor.2007.04.007).
- [12] Ishibuchi, H., Shibata, Y., 2003a. A Similarity-Based Mating Scheme for Evolutionary Multiobjective Optimization, in: *Genetic and Evolutionary Computation — GECCO 2003*. Springer Berlin Heidelberg. volume 2723 of *Lecture Notes in Computer Science*, pp. 1065–1076. doi:[10.1007/3-540-45105-6_116](https://doi.org/10.1007/3-540-45105-6_116).
- [13] Ishibuchi, H., Shibata, Y., 2003b. An Empirical Study on the Effect of Mating Restriction on the Search Ability of EMO Algorithms, in: *Evolutionary Multi-Criterion Optimization*. Springer Berlin Heidelberg. volume 2632 of *Lecture Notes in Computer Science*, pp. 433–447. doi:[10.1007/3-540-36970-8_31](https://doi.org/10.1007/3-540-36970-8_31).
- [14] Ishibuchi, H., Shibata, Y., 2004. Mating scheme for controlling the diversity-convergence balance for multiobjective optimization, in: *Genetic and Evolutionary Computation — GECCO 2004*. Springer, pp. 1259–1271. doi:[10.1007/978-3-540-24854-5_121](https://doi.org/10.1007/978-3-540-24854-5_121).
- [15] Ishibuchi, H., Tanigaki, Y., Masuda, H., Nojima, Y., 2014. Distance-based analysis of crossover operators for many-objective knapsack problems, in: *Parallel Problem Solving from Nature – PPSN XIII*, Springer International Publishing, Cham. pp. 600–610. doi:[10.1007/978-3-319-10762-2_59](https://doi.org/10.1007/978-3-319-10762-2_59).
- [16] Ishibuchi, H., Tsukamoto, N., Nojima, Y., 2010. Diversity Improvement by Non-Geometric Binary Crossover in Evolutionary Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation* 14, 985–998. doi:[10.1109/TEVC.2010.2043365](https://doi.org/10.1109/TEVC.2010.2043365).
- [17] Jaszkiewicz, A., 2002. On the performance of multiple-objective genetic local search on the 0/1 knapsack problem - a comparative experiment. *Evolutionary Computation*, IEEE Transactions on 6, 402–412. doi:[10.1109/TEVC.2002.802873](https://doi.org/10.1109/TEVC.2002.802873).
- [18] Jaszkiewicz, A., Kominek, P., 2003. Genetic local search with distance preserving recombination operator for a vehicle routing problem. *European Journal of Operational Research* 151, 352–364. doi:[10.1016/S0377-2217\(02\)00830-5](https://doi.org/10.1016/S0377-2217(02)00830-5). meta-heuristics in combinatorial optimization.
- [19] Karshenas, H., Santana, R., Bielza, C., Larrañaga, P., 2011. Multi-objective optimization with joint probabilistic modeling of objectives and variables, in: *Evolutionary Multi-Criterion Optimization: 6th International Conference*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 298–312. doi:[10.1007/978-3-642-19893-9_21](https://doi.org/10.1007/978-3-642-19893-9_21).
- [20] Karshenas, H., Santana, R., Bielza, C., Larraaga, P., 2014. Multiobjective estimation of distribution algorithm based on joint modeling of objectives and variables. *IEEE Transactions on Evolutionary Computation* 18, 519–542. doi:[10.1109/TEVC.2013.2281524](https://doi.org/10.1109/TEVC.2013.2281524).
- [21] Laumanns, M., Ocenasek, J., 2002. Bayesian optimization algorithms for multi-objective optimization. *Parallel Problem Solving from Nature – PPSN VII*, 298–307.
- [22] Li, H., Zhang, Q., 2009. Multiobjective optimization problems with complicated pareto sets, moea/d and nsga-ii. *IEEE Transactions on Evolutionary Computation* 13, 284–302. doi:[10.1109/TEVC.2008.925798](https://doi.org/10.1109/TEVC.2008.925798).
- [23] Martí, L., García, J., Antonio, B., Coello, C., Molina, J., 2011a. Scalable multi-objective optimization. Ph.D. thesis. Departamento de Informática. URL: <http://e-archivo.uc3m.es/handle/10016/11896>.
- [24] Martí, L., García, J., Berlanga, A., Coello, C.A.C., Molina, J.M., 2011b. Mb-gng: Addressing drawbacks in multi-objective optimization estimation of distribution algorithms. *Operations Research Letters* 39, 150–154. doi:[10.1016/j.orl.2011.01.002](https://doi.org/10.1016/j.orl.2011.01.002).
- [25] Martí, L., García, J., Berlanga, A., Molina, J.M., 2016. MONEDA: scalable multi-objective optimization with a neural network-based estimation of distribution algorithm. *Journal of Global Optimization*, 1–40doi:[10.1007/s10898-016-0415-7](https://doi.org/10.1007/s10898-016-0415-7).
- [26] Martí, L., Garca, J., Berlanga, A., Molina, J., 2008. Introducing MONEDA: scalable multiobjective optimization with a neural estimation of distribution algorithm, in: *Proceedings of the 10th annual conference on Genetic and evolutionary computation*, ACM. pp. 689–696. URL: <http://dl.acm.org/citation.cfm?id=1389230>.
- [27] Martins, J., Soares, A., Vargas, D., Delbem, A., 2011. Multi-objective Phylogenetic Algorithm: Solving Multi-objective Decomposable Deceptive Problems, in: *Evolutionary Multi-Criterion Optimization*. Springer Berlin / Heidelberg. volume 6576 of *Lecture Notes in Computer Science*, pp. 285–297. doi:[10.1007/978-3-642-19893-9_20](https://doi.org/10.1007/978-3-642-19893-9_20).
- [28] Martins, J.P., 2015. Analysis of linkage learning in evolutionary optimization. Ph.D. thesis. University of São Paulo. Brazil. doi:[10.11606/T.55.2015.tde-14092015-140718](https://doi.org/10.11606/T.55.2015.tde-14092015-140718).
- [29] Martins, J.P., Bríngel Neto, C., Crocorno, M.K., Vittori, K., Delbem, A.C.B., 2013. A Comparison of Linkage-learning-based Genetic Algorithms in Multidimensional knapsack Problems, in: *Proceedings of the Congress on Evolutionary Computation*, IEEE. pp. 502–509. doi:[10.1109/CEC.2013.6557610](https://doi.org/10.1109/CEC.2013.6557610).
- [30] Martins, J.P., Delbem, A.C., 2014. Multimodality and the Linkage-learning Difficulty of Additively Separable Functions, in: *Proceedings of the Genetic and Evolutionary Computation Conference*, ACM. pp. 365–372. doi:[10.1145/2576768.2598281](https://doi.org/10.1145/2576768.2598281).
- [31] Martins, J.P., Delbem, A.C., 2016. Pairwise independence and its impact on estimation of distribution algorithms. *Swarm and Evolutionary Computation* 27, 80–96. doi:[10.1016/j.swevo.2015.10.001](https://doi.org/10.1016/j.swevo.2015.10.001).
- [32] Martins, J.P., Fonseca, C.M., Delbem, A.C., 2014. On the performance of linkage-tree genetic algorithms for the multi-dimensional knapsack problem. *Neurocomputing* 146, 17–29. doi:[10.1016/j.neucom.2014.04.069](https://doi.org/10.1016/j.neucom.2014.04.069).
- [33] Miyakawa, M., Takadama, K., Sato, H., 2015. Control of crossed genes ratio for directed mating in evolutionary constrained multi-objective optimization, in: *Proceedings of the Companion Publication of the 2015 Annual Conference on Genetic and Evolution-*

- ary Computation, ACM, New York, NY, USA. pp. 1201–1204. doi:[10.1145/2739482.2768480](https://doi.org/10.1145/2739482.2768480).
- [34] Miyakawa, M., Takadama, K., Sato, H., 2016. Controlling selection areas of useful infeasible solutions for directed mating in evolutionary constrained multi-objective optimization. *Annals of Mathematics and Artificial Intelligence* 76, 25–46. doi:[10.1007/s10472-015-9455-9](https://doi.org/10.1007/s10472-015-9455-9).
- [35] Pelikan, M., Sastry, K., Goldberg, D., 2006. Multiobjective Estimation of Distribution Algorithms. volume 33 of *Studies in Computational Intelligence*. Springer Berlin / Heidelberg. doi:[10.1007/978-3-540-34954-9_10](https://doi.org/10.1007/978-3-540-34954-9_10).
- [36] Pelikan, M., Sastry, K., Goldberg, D.E., 2005. Multiobjective hboa, clustering, and scalability, in: *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ACM, New York, NY, USA. pp. 663–670. doi:[10.1145/1068009.1068122](https://doi.org/10.1145/1068009.1068122).
- [37] Sastry, K., Goldberg, D.E., Pelikan, M., 2005. Limits of scalability of multiobjective estimation of distribution algorithms, in: *2005 IEEE Congress on Evolutionary Computation*, pp. 2217–2224 Vol. 3. doi:[10.1109/CEC.2005.1554970](https://doi.org/10.1109/CEC.2005.1554970).
- [38] Sato, H., Aguirre, H., Tanaka, K., 2013. Variable space diversity, crossover and mutation in moea solving many-objective knapsack problems. *Annals of Mathematics and Artificial Intelligence* 68, 197–224. doi:[10.1007/s10472-012-9293-y](https://doi.org/10.1007/s10472-012-9293-y).
- [39] Sato, H., Aguirre, H.E., Tanaka, K., 2007. Local dominance and local recombination in MOEAs on 0/1 multiobjective knapsack problems. *European Journal of Operational Research* 181, 1708–1723. doi:[10.1016/j.ejor.2006.08.006](https://doi.org/10.1016/j.ejor.2006.08.006).
- [40] Shah, R., Reed, P., 2011. Comparative analysis of multiobjective evolutionary algorithms for random and correlated instances of multiobjective d-dimensional knapsack problems. *European Journal of Operational Research* 211, 466–479. doi:[10.1016/j.ejor.2011.01.030](https://doi.org/10.1016/j.ejor.2011.01.030).
- [41] Shim, V.A., Tan, K.C., Cheong, C.Y., 2012. A hybrid estimation of distribution algorithm with decomposition for solving the multiobjective multiple traveling salesman problem. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42, 682–691. doi:[10.1109/TSMCC.2012.2188285](https://doi.org/10.1109/TSMCC.2012.2188285).
- [42] Srinivas, N., Deb, K., 1994. Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation* 2, 221–248. doi:[10.1162/evco.1994.2.3.221](https://doi.org/10.1162/evco.1994.2.3.221).
- [43] Thierens, D., 2010. The Linkage Tree Genetic Algorithm, in: *Parallel Problem Solving from Nature, PPSN XI*. Springer Berlin Heidelberg. volume 6238 of *Lecture Notes in Computer Science*, pp. 264–273. doi:[10.1007/978-3-642-15844-5_27](https://doi.org/10.1007/978-3-642-15844-5_27).
- [44] Thierens, D., Bosman, P.A.N., 2001. Multi-objective mixture-based iterated density estimation evolutionary algorithms, in: *Proceedings of the 3rd Annual Conference on Genetic and Evolutionary Computation*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. pp. 663–670. URL: <http://dl.acm.org/citation.cfm?id=2955239.2955338>.
- [45] Verel, S., Liefvooghe, A., Jourdan, L., Dhaenens, C., 2011. Analyzing the Effect of Objective Correlation on the Efficient Set of MNK-Landscapes, in: Coello, C. (Ed.), *Learning and Intelligent Optimization*. Springer Berlin Heidelberg. volume 6683 of *Lecture Notes in Computer Science*, pp. 116–130. doi:[10.1007/978-3-642-25566-3_9](https://doi.org/10.1007/978-3-642-25566-3_9).
- [46] Verel, S., Liefvooghe, A., Jourdan, L., Dhaenens, C., 2013. On the structure of multiobjective combinatorial search space: MNK-landscapes with correlated objectives. *European Journal of Operational Research* 227, 331–342. doi:[10.1016/j.ejor.2012.12.019](https://doi.org/10.1016/j.ejor.2012.12.019).
- [47] Watanabe, S., Hiroyasu, T., Miki, M., 2002. Ncga: Neighborhood cultivation genetic algorithm for multi-objective optimization problems., in: *GECCO late breaking papers*, pp. 458–465.
- [48] Zhang, Q., Zhou, A., Jin, Y., 2008. RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation* 12, 41–63. doi:[10.1109/TEVC.2007.894202](https://doi.org/10.1109/TEVC.2007.894202).
- [49] Zhou, A., Jin, Y., Zhang, Q., Sendhoff, B., Tsang, E., 2006. Combining model-based and genetics-based offspring generation for multi-objective optimization using a convergence criterion, in: *2006 IEEE International Conference on Evolutionary Computation*, pp. 892–899. doi:[10.1109/CEC.2006.1688406](https://doi.org/10.1109/CEC.2006.1688406).
- [50] Zhou, A., Zhang, Q., Jin, Y., 2009. Approximating the set of pareto-optimal solutions in both the decision and objective spaces by an estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation* 13, 1167–1189. doi:[10.1109/TEVC.2009.2021467](https://doi.org/10.1109/TEVC.2009.2021467).
- [51] Zhou, A., Zhang, Q., Jin, Y., Tsang, E., Okabe, T., 2005. A model-based evolutionary algorithm for bi-objective optimization, in: *2005 IEEE Congress on Evolutionary Computation*, pp. 2568–2575 Vol. 3. doi:[10.1109/CEC.2005.1555016](https://doi.org/10.1109/CEC.2005.1555016).
- [52] Zitzler, E., Laumanns, M., Thiele, L., 2002. Spea2: Improving the strength pareto evolutionary algorithm for multiobjective optimization.
- [53] Zitzler, E., Thiele, L., 1999. Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. *Evolutionary Computation*, IEEE transactions on 3, 257–271.