

# Evolutionary Multiobjective Optimization Algorithm as a Markov System

Ewa Gajda<sup>1</sup>, Robert Schaefer<sup>1</sup>, and Maciej Smolka<sup>2</sup>

<sup>1</sup> AGH University of Science and Technology, Department of Computer Science  
Al. Mickiewicza 30, 30-059 Kraków, Poland  
{gajda,schaefer}@agh.edu.pl  
<http://www.iisg.agh.edu.pl/>

<sup>2</sup> Jagiellonian University, Institute of Computer Science  
ul. Łojasiewicza 6, 30-348 Kraków, Poland  
smolka@ii.uj.edu.pl

**Abstract.** In the paper we consider the ranking given by the Pareto dominance relation as a basis to create a selection operator for the Evolutionary Multiobjective Optimization Algorithm (EMOA). Assuming that sampling to the next epoch is performed according to the generalized Bernoulli schema with regard to a selected type of the rank selection, a heuristic operator for EMOA is introduced. Having defined the heuristic operator, the transition probability matrix of the uniform Markov chain modeling EMOA can be explicitly obtained as in the Vose's theory of the Simple Genetic Algorithm (SGA). This chain is ergodic if the mixing operator following the EMOA selection operator in each epoch is strictly positive. Moreover, we show that the measure on the space of populations imposed by the EMOA infinite population concentrates on the set of fixed points of the heuristic operator after infinite number of epochs, assuming that the heuristic operator is focusing.

**Keywords:** evolutionary algorithm, multi-objective optimization, Markov system.

## 1 Introduction

Evolutionary Multiobjective Optimization Algorithms (EMOAs) have been studied by several groups of researchers. Different types of selection were introduced i.a. by Goldberg in [5], Fonseca and Fleming in [4] and Zitzler and Thiele in [16]. Nondominated sorting was also used by Srinivas and Deb (see. e.g. [12]). Theoretical properties of EMOAs applied to discrete problems were studied i.a. by Rudolph in [9], [10], Hanne in [6] and Laumanns in [7]. Authors of these papers base on the Markov description of populations processing and use an archive in which an approximation of the Pareto front is stored. Convergence with regard to  $\varepsilon$ -Pareto dominance relation was analyzed by Laumanns in [7].

We build a Markov model of EMOA basing on the introduced heuristic operator, similar to the Vose's genetic operator for the Simple Genetic Algorithm

(SGA) (see e.g. [13]). We assume the set of genes to be finite but general (individuals are not necessarily strings over any alphabet). The proposed heuristic is created with regard to some special types of rank selection. We will analyze asymptotic features of the evolved population according to such EMOA selection rules and mixing operations (crossover, mutation, etc.) that return a strictly positive sampling probability.

Because the asymptotic results obtained for EMOA Markov model are similar as those obtained for SGA by Vose, they can be used for verifying two-phase strategies in the same manner as for the single criteria ones (see e.g. [11]). Such strategies consist of finding the approximation of the connected components of the Pareto set (using EMOA combined with the proper population clustering) in the first phase, and the parallel, detailed search in each of them.

## 2 Evolutionary Approach to the Multiobjective Optimization

### 2.1 Pareto Dominance

In the multiobjective optimization, we are given  $k \geq 2$  objective functions

$$f_i : U \rightarrow [0, M] \subset \mathbb{R}, \quad M < +\infty, \quad i \in \{1, \dots, k\} \quad (1)$$

defined over some search space  $U$ , which might be implicitly defined by constraints. We assume the search space  $U$  to be finite  $\#U = r < +\infty$  and that all objectives shall be maximized. Therefore we are interested in solving

$$\max \{f(p) = (f_1(p), \dots, f_k(p))^T \mid p \in U\}. \quad (2)$$

**Definition 1.** (*Pareto dominance*) For any pair  $(p, q) \in U \times U$ ,  $p$  is said to dominate  $q$ , denoted as  $p \succ q$ , if and only if

$$f(p) \geq f(q) \text{ and } \exists_{i=1, \dots, k} f_i(p) \neq f_i(q). \quad (3)$$

*Remark 1.* The definition can be easily adapted to the minimization problem, when in formula (3) inequality changes form  $\geq$  to  $\leq$ . It can be also adapted to mixed min-max problems by changing inequalities for certain coordinates representing different objective functions.

### 2.2 Evolutionary Multiobjective Optimization (EMOA)

One of the possible ways of solving (2) is finding the *Pareto set*  $\mathcal{P}$  being the set of non-dominated elements from  $U$  and its image  $f(\mathcal{P}) \subset [0, M]^k$  called the *Pareto front*.

A popular class of stochastic algorithms designed for finding Pareto set is called Evolutionary Multiobjective Optimization (EMOA) (see e.g. [14]). Their simplest instances operate on the single population being the multiset  $P = (U, \eta)$  of the search space members called *individuals*, while  $U$  is called now *genetic*

*universum*. The occurrence function  $\eta : U \rightarrow \mathbb{Z}_+ \cup \{0\}$  returns  $\eta(i)$  being the number individuals with the genotype  $i \in U$  and  $\mu = \sum_{i \in U} \eta(i) < +\infty$  stands for the population cardinality.

As other genetic algorithms, EMOA consists in producing the sequence of populations  $\{P^t\}$  in the consecutive *genetic epochs*  $t = 1, 2, \dots$  starting from the population  $P^0$  uniformly sampled from  $U$ .

Later, we will consider only the scheme  $(\mu, \lambda)$  where  $\lambda = \mu$ , such that the transformation between  $P^t$  and its successor  $P^{t+1}$  is obtained by the composition of two groups of random operations: selection and mixing. Hereafter  $(\mu, \lambda)$  stands for the Schwefel's symbol, where  $\lambda$  is the offspring cardinality and  $\mu$  is the population cardinality.

While the mixing operations utilized in EMOA do not differ significantly from those applied in other groups of evolutionary algorithms, the EMOA selection is performed with regard to the Pareto dominance (see e.g. [2]). The algorithm terminates after a predefined number of epochs or when another stopping criterion is satisfied (see e.g. [14]).

### 2.3 Selection Schemes

Several important EMOAs with different selection schemes will be briefly described in the following section. A comparison of these methods can be found e.g. in [14].

The idea of calculating an individual's fitness according to Pareto-dominance was first suggested by Goldberg in [5]. The procedure of NSGA (*Nondominated Sorting Genetic Algorithm*) is based on ranking individuals in an iterative way: firstly nondominated solutions are assigned rank one and temporarily removed from the population. After that, next nondominated solutions are given rank two and so forth. The rank of an individual determines its fitness value. Goldberg's concept was implemented e.g. by Srinivas and Deb [12].

Fonseca and Fleming in [4] proposed a Pareto-based ranking procedure (*FFGA*), where an individual's rank equals the number of solutions by which it is dominated. After sorting population according to the rank, new fitness values are assigned to individuals by interpolating from the best (with the lowest rank) to the worst (with the highest rank) according to some function. Fitness of individuals with the same rank should be equal, so that all of them will be sampled at the same rate. We used this type of selection as a basis for creating the selection operator.

Later on we will refer mainly to NSGA and FFGA selection schemes in preparing the EMOA Markov model. In the following paragraphs we will mention two important strategies which seem to be difficult or impossible to model in the way presented in next sections.

One of these methods is aimed to construct an algorithm in which the *hypervolume measure* (see e.g. [3]) governs the selection operator of an EMOA in order to find a set of solutions well distributed on the Pareto front. Hypervolume measure or *S-metric* corresponds to the size of dominated space [16]. Individuals are rated according to their contribution to the dominated hypervolume of the

current population, therefore ranks are not based on relations between pairs of individuals but on relation between an individual and the whole population.

*Strength Pareto Evolutionary Algorithm* (SPEA, see [16]) uses a regular population and an external set (archive) into which all nondominated solutions are copied in each iteration. If the size of the archive exceeds a predefined limit, further archive members are deleted by a clustering strategy which preserves the characteristics of the nondominated front. Ranks of solutions are calculated basing on strength values of individuals stored externally. SPEA was later improved and introduced as SPEA2 in [15]. The selection in SPEA cannot be described by our selection operator because of the existence of the archive. In order to model this selection scheme one should consider a different space of states.

### 3 EMOA Markov Model

#### 3.1 Evolutionary Algorithms with Heuristic

Each finite population represented as the multiset  $P = (U, \eta)$  may be identified with its frequency vector  $x = \{\frac{1}{\mu} \eta(p)\}, p \in U$  and all such vectors belong to the finite subset  $X_\mu$  of the well-known Vose simplex

$$\Lambda^r = \left\{ x = \{x_p\}; 0 \leq x_p \leq 1, p \in U, \sum_{p \in U} x_p = 1 \right\}. \quad (4)$$

Such construction has several advantages:

1. Although the frequency vector represents unambiguously only the finite populations ( $\mu < +\infty$ ), it is possible to represent also the infinite size populations. We will identify the population with its frequency vector if it does not lead to the ambiguity.
2. Each  $x \in \Lambda^r$  being the population frequency vector (possibly infinite one) belongs to  $\mathcal{M}(U)$  being the set of probabilistic measures on the set  $U$ .
3. The set containing representations of all populations is compact in  $\mathbb{R}^r$ .

The above settings allow to define the class of evolutionary algorithms which are characterized by the same genetic universum  $U$  and fitness as well as the same set of genetic operations that do not depend on the genetic epoch number. They can differ only by the population size  $\mu$ . This class of EA's may be also characterized as "stationary" because the evolutionary rule does not change during computations.

**Definition 2.** *The mapping  $\mathcal{H} \in C(\Lambda^r \rightarrow \Lambda^r)$  will be called the heuristic of the particular class of genetic algorithms if:*

1. *Each coordinate  $(\mathcal{H}(x))_p$  is equal to the sampling probability of the individual with the genotype  $p \in U$  in the epoch that immediately follows the epoch in which the population  $x \in \Lambda^r$  appears.*

2. The value  $\mathcal{H}(x)$  is the expected population in the epoch that immediately follows the epoch in which the population  $x \in \Lambda^r$  appeared, for all algorithms from the considered class.
3. It stands for the law of evolution of the abstract, deterministic, infinite population algorithm (we assume that it exists in the considered class). In other words, the infinite population algorithm is the dynamic system that starts from a particular initial population  $x^0 \in \Lambda^r$  and then passes consecutively by  $\mathcal{H}(x^0), \mathcal{H}^2(x^0), \mathcal{H}^3(x^0), \dots$ .

If a particular class of genetic algorithms admits a heuristic operator, we will call those algorithms *the genetic algorithms with heuristic*. The heuristic operator was introduced by Vose and his collaborators for the class of Simple Genetic Algorithms (SGA) (see e.g. [13]). This operator was equivalently called *genetic operator* in this case. Some further comments are contained in [11].

Furthermore, we restrict ourselves to the evolutionary algorithms in which the next epoch population  $x^{t+1} \in \Lambda^r$  is obtained by the  $\mu$ -times sampling with return according to the polynomial scheme (generalized Bernoulli scheme, see e.g. Billingsley [1]), assuming the probability distribution on the set of genotypes. Of course, in case of the GA class with heuristic, the value of  $\mathcal{H}(x^t)$  stands for such probability distribution.

**Observation 1.** *If the next population  $x^{t+1}$  is obtained using generalized Bernoulli scheme, then the condition 1 of the Definition 2 implies the condition 2.*

The Observation 1 can be motivated as follows. Let us denote by  $P^{t+1} = (U, \eta_{t+1})$  the random variable being the population in the  $t + 1$  epoch. Because  $P^{t+1}$  is obtained using the generalized Bernoulli scheme associated with the probability distribution  $\mathcal{H}(x^t) \in \mathcal{M}(U)$ , we have that  $EP^{t+1} = (U, \bar{\eta}_{t+1})$  with  $\bar{\eta}_{t+1}(p) = \mu \mathcal{H}(x^t)_p$ , where  $E$  is the proper expected value operator. Therefore the expected coordinate of the frequency vector  $x^{t+1}$  satisfies  $(Ex^{t+1})_p = \frac{1}{\mu} \bar{\eta}_{t+1}(p) = \mathcal{H}(x^t)_p$  for all  $p \in U$ .

**Observation 2.** *If the next population  $x^{t+1}$  is obtained using generalized Bernoulli scheme, then the condition 1 of the Definition 2 implies also the condition 3.*

The motivation of the Observation 2 is not so trivial as the previous one. It may be drawn from the following theorem.

**Theorem 1.**  $\forall k > 0, \varepsilon > 0, \nu < 1 \exists N$  independent upon  $x^0 \in \Lambda^r$  such that

$$\mu > N \Rightarrow \Pr\{\|x^t - \mathcal{H}^t(x^0)\| < \varepsilon\} > \nu \quad \forall t \in [0, k] \cap \mathbb{N}.$$

This theorem is a generalization of the well known Nix and Vose result (see Theorem 2 in [8]) proved originally for the SGA heuristic only. This theorem states that the finite population algorithm spends arbitrarily large number of epochs arbitrarily close to the heuristic trajectory with the probability arbitrarily close to 1 if the population size is large enough, so the heuristic trajectory might be understood as the trajectory of infinite population algorithm in this sense.

**Observation 3.** *If the particular EA has the heuristic  $\mathcal{H}$  and the next epoch population  $P^{t+1}$  is obtained using generalized Bernoulli scheme associated with the probability distribution  $\mathcal{H}(x^t) \in \mathcal{M}(U)$ , then it can be modeled as the stationary Markov chain with the finite space of states  $X_\mu$  and with the transition probability matrix  $\mathbf{Q}$  given by the formula similar to the formula introduced by Vose for SGA Markov model (see Theorem 1 in [8])*

$$(\mathbf{Q})_{x,y} = \mu! \prod_{p \in U} \frac{(\mathcal{H}(x)_p)^{\mu y_p}}{(\mu y_p)!} \quad \forall x, y \in X_\mu. \tag{5}$$

The above observation is a simple issue of the polynomial sampling distribution.

### 3.2 The EMOA Selection Operator

Let us start with the definition of the *binary Pareto dominance matrix*

$$\Xi \in \{0, 1\}^r \times \{0, 1\}^r; \quad \Xi_{p,q} = \begin{cases} 1 & \text{if } q \succ p \\ 0 & \text{otherwise.} \end{cases}, \quad \forall p, q \in U. \tag{6}$$

which completely characterizes the Pareto dominance relation among the genotypes from  $U$  for the particular multiobjective optimization (2). The above definition is appropriate also for different cases of problems, not only maximization (see Remark 1). NSGA selection scheme can be represented in a similar way but with a different  $\Xi$  matrix.

It is easy to observe that the  $p$ -th entry of the vector  $(\Xi \eta)$  represents the number of individuals which dominate the individual with the genotype  $p$  belonging to the population  $P = (U, \eta)$  (i.e.  $\eta(p) > 0$ ).

Next, we introduce function  $\xi : A^r \rightarrow [0, 1]^r$  of the form

$$\xi(x) = \Xi x, \tag{7}$$

so that  $\xi(x)_p$  defines the rank of all individuals with the genotype  $p \in U$  contained in the population  $P$  represented by its frequency vector  $x$ .

This function is well defined for both finite and infinite populations. In case of finite population of the cardinality  $\mu < +\infty$  the entry  $\xi(x)_p$  may be interpreted as the relative number of individuals that dominate the individual with the genotype  $p$  because  $\xi(x) = \frac{1}{\mu} (\mu \Xi x) = \frac{1}{\mu} (\Xi \mu x) = \frac{1}{\mu} (\Xi \eta)$ .

**Observation 4.** *It may be easily checked that*

$$\forall x \in A^r \exists p \in U : \xi(x)_p = 0, \quad x_p > 0.$$

*It follows from the fact that there is at least one non-dominated individual in each population.*

As usual, it is necessary to introduce the *validating function* in order to obtain the probability distribution of the rank selection

$$g \in C([0, 1] \rightarrow [0, 1]); \quad \forall \zeta, \gamma \in [0, 1], \quad \zeta > \gamma \Rightarrow g(\zeta) < g(\gamma). \tag{8}$$

As a simple example of a function correlated with the rank-based fitness assignment method [4] we can take

$$g(\zeta) = 1 - \zeta. \quad (9)$$

To obtain more control on the selection pressure the function

$$g(\zeta) = e^{-\alpha\zeta}, \quad \alpha \in \mathbb{R}_+ \quad (10)$$

may be chosen.

For technical purposes we introduce a next function  $G : [0, 1]^r \rightarrow [0, 1]^r$  such that  $G(x)_p = g(x_p)$ ,  $p \in U$ .

The probability of selecting the individual  $p \in U$  from the current EMOA population  $P$  represented by the vector  $x \in \Lambda^r$  equals to

$$\Pr(p) = \frac{1}{x^T G(\xi(x))} g((\xi(x))_p) x_p. \quad (11)$$

We are now ready to define the selection operator  $F : \Lambda^r \rightarrow \Lambda^r$  for the EMOA rank selection

$$F(x) = \frac{1}{x^T G(\Xi x)} \text{diag}(x) G(\Xi x), \quad (12)$$

where  $\text{diag}(x)$  denotes the  $r \times r$  diagonal matrix with the diagonal  $x$ .

**Observation 5.** *Taking into account the features of the function  $\xi$  (see Observation 4) and the features of the function  $g$  (see formula (8)) the EMOA rank selection operator (12) as well as the formula (11) are well defined, because the denominator  $x^T G(\Xi x)$  is strictly positive for all  $x \in \Lambda^r$ .*

**Observation 6.** *Because  $g$  is continuous in its domain the EMOA rank selection operator (12) is continuous on the whole  $\Lambda^r$ . If we additionally assume, that  $g$  is continuously differentiable, as in case of both examples (9), (10), then the EMOA rank selection operator is also continuously differentiable on the whole  $\Lambda^r$ .*

**Observation 7.** *If the validating function  $g$  is strictly positive, then the EMOA rank selection is always "soft", which means that each individual from the current population can be selected (with the positive probability) as a parental one. If we relax conditions of  $g$  contained in (8), assuming only that  $g$  is weakly decreasing ( $g(x) \leq g(y)$  for  $x > y$ ) and exists  $\gamma \in (0, 1)$  so that  $g(x) > 0$  only for  $x \in [0, \gamma]$ , then the EMOA rank selection may become partially hard, elitist one. Such a relaxation does not contradict the well-posedness of the formula (12) because the denominator  $x^T G(\Xi x)$  remains strictly positive for all  $x \in \Lambda^r$ .*

### 3.3 The EMOA Heuristic

The selection is followed by the genetic operations (e.g. mutation, crossover) in each EMOA epochs. They can be represented by the *mixing operator*  $M \in C^1(\Lambda^r \rightarrow \Lambda^r)$ . Currently, we do not impose any specific restrictions for this mapping.

The well known example of mixing operator was introduced by Vose and collaborators [13]. It expresses the binary mutation and positional crossover utilized in SGA.

$$M(x)_p = (\sigma_p x)^T \mathbf{M} \sigma_p x, \quad \forall x \in \Lambda^r, p \in U \quad (13)$$

where  $\sigma_i$  stands for the  $r \times r$  dimension permutation matrix with the entries  $(\sigma_p)_{q,k} = [q \oplus k = p]$ ,  $p, q, k \in U$ . The entries  $\mathbf{M}_{p,q}$  of the symmetric  $r \times r$  matrix  $\mathbf{M}$  express the probability of obtaining the genotype  $0 \in U$  (the genotype being the string of zeros) from the parents  $p, q \in U$  by the crossover and mutation.

Similarly, like in case of SGA, the composition

$$\mathcal{H} = M \circ F \quad (14)$$

may be considered as the candidate for the heuristic of the particular class of EMOA considered in this paper.

**Observation 8.** *As an immediate consequence of its construction,  $\mathcal{H}$  is well defined and continuous on the whole  $\Lambda^r$ . Assuming additionally that the function  $g$  (see formula (8)) is continuously differentiable on  $[0, 1]$  we have also that  $\mathcal{H}$  is continuously differentiable on the whole  $\Lambda^r$ .*

**Observation 9.** *Again, as the result of the construction,  $\mathcal{H}$  described by (14) satisfies the condition 1 of the Definition 2, because each coordinate of its value  $(\mathcal{H}(x))_p$  stands for the probability of sampling the genotype  $p \in U$  to the population following the population associated with the frequency vector  $x$ . If we moreover assume, that the sampling to the next epoch population is performed according to the generalized Bernoulli model (according to the polynomial probability distribution) then also the conditions 2 and 3 of the Definition 2 are satisfied (see Observations 1 and 2).*

**Observation 10.** *Immediately from the Observations 9 and 3 it follows that the considered class of EMOA transforming the finite populations of the cardinality  $\mu < +\infty$  can be modeled as the stationary Markov chain with the finite space of states  $X_\mu$  and with the transition probability matrix  $\mathbf{Q}$  given by the formula (5).*

### 3.4 Asymptotic Features

**Observation 11.** *If the mixing operator (13) is strictly positive, e.g.  $M(x)_p > 0$ ,  $\forall x \in \Lambda^r$ ,  $\forall p \in U$ , then the Markov chain describing EMOA is ergodic. The algorithm possesses the asymptotic guarantee of success, e.g. it will reach the population (state) which contain all points lying in the Pareto set after an infinite number of epochs.*

Let us denote by  $\pi_\mu^t \in \mathcal{M}(X_\mu)$  the probability distribution of the random variable representing EMOA population of the size  $\mu$  after  $t$  epochs. Assuming that EMOA is modeled by the ergodic Markov chain, each sequence  $\pi_\mu^0, \pi_\mu^1, \dots$  has a strictly positive limit  $\pi_\mu$  that does not depend on the initial distribution  $\pi_\mu^0$ , which is the simple issue of the ergodic theorem (see e.g. [1]). Measures  $\pi_\mu$  originally defined over  $X_\mu$  can be easily extended to  $\Lambda^r$ .

Next, according to the Prokhorov theorem (see Theorem 29.3 in [1]), the sequence  $\{\pi_\mu\}$  has a weak limit  $\pi^*$  in  $A^r$ , because  $A^r$  is compact.

Let us assume now, that the EMOA heuristic is *focusing* i.e. if for all  $x \in A^r$  the sequence  $\{\mathcal{H}^t(x)\}$  converges in  $A^r$  for  $t \rightarrow +\infty$ . Let  $w \in A^r$  be the limit of such sequence for some starting point  $x \in A^r$ . The continuity of  $\mathcal{H}$  guarantees, that  $\mathcal{H}(w) = \mathcal{H}(\lim_{t \rightarrow +\infty} \mathcal{H}^t(x)) = \lim_{t \rightarrow +\infty} \mathcal{H}^{t+1}(x) = w$ , so  $w$  is the fixed point of  $\mathcal{H}$ . Let us denote the set of fixed points of  $\mathcal{H}$  by  $\mathcal{K} \subset A^r$ . If  $\mathcal{H}$  is focusing, then obviously  $\mathcal{K} \neq \emptyset$ .

**Theorem 2.** *Assuming that the EMOA heuristic is focusing and the Markov chains associated with family of EMOA with various population sizes are ergodic, we obtain  $\pi^*(\mathcal{K}) = 1$ .*

The above theorem is a formal extension of the well known Vose and Nix result (see Theorem 3 in [8]) and it can be proved in an analogous way.

## 4 Conclusions and Further Research

- EMOA can be modeled as the ergodic Markov chain given some reasonable assumptions upon the type of selection and the presence of mutation in the mixing step of each epoch.
- Alternative selection types might be considered and formalized. The main step to adapt the current model to other EMOA selection schemes will consist of redefining the Pareto dominance matrix (6).
- In the proposed model a particular form of genotypes has not been assumed: the most common form of strings over an alphabet is appropriate but genotypes may be graphs as well.
- It was also proved that EMOA has the heuristic which fixed points are the only ones visited by the infinite population algorithm (see Theorem 2). It is possible to prove the theorem of a fixed point approximation and the theorem of the convergence of sampling measures, similar to those proved for SGA (see Theorems 4.54 and 4.66 in [11]).
- Furthermore, these results might be used for verifying two-phase strategies in the same manner as for the single criteria ones (see e.g. [11]). Such strategies consist in finding the approximation of the connected components of the Pareto set (using EMOA combined with the proper population clustering) in the first phase and the parallel, detailed search in each of them. It seems that the obtained results might also be useful in the analysis of the  $\varepsilon$ -Pareto dominance problem (see [7]).

## References

1. Billingsley, P.: Probability and Measure. John Wiley and Sons, Chichester (1979)
2. Coello Coello, C.A., Lamont, G.B.: Applications of Multi-objective Evolutionary Algorithms. World Scientific, Singapore (2004)

3. Emmerich, M., Beume, N., Naujoks, B.: An EMO algorithm using the hypervolume measure as selection criterion. In: Coello Coello, C.A., Hernández Aguirre, A., Zitzler, E. (eds.) EMO 2005. LNCS, vol. 3410, pp. 62–76. Springer, Heidelberg (2005)
4. Fonseca, C.M., Fleming, P.J.: Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In: Genetic Algorithms: Proceedings of the Fifth International Conference, pp. 416–423 (1993)
5. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learnings. Addison-Wesley, Reading (1989)
6. Hanne, T.: On the convergence of multiobjective evolutionary algorithms. European Journal of Operational Research 117(3), 553–564 (1999)
7. Laumanns, M.: Stochastic convergence of random search to fixed size Pareto set approximations (2007)
8. Nix, A.E., Vose, M.D.: Modelling genetic algorithms with Markov chains. Annals of Mathematics and Artificial Intelligence 5, 79–88 (1992)
9. Rudolph, G.: On a multi-objective evolutionary algorithm and its convergence to the Pareto set. In: Proceedings of the 5th IEEE Conference on Evolutionary Computation, pp. 511–516. IEEE Press, Los Alamitos (1998)
10. Rudolph, G., Agapie, A.: Convergence properties of some multi-objective evolutionary algorithms. In: Congress on Evolutionary Computation (CEC 2000), pp. 1010–1016. IEEE Press, Los Alamitos (2000)
11. Schaefer, R., Telega, H.: Foundation of Global Genetic Optimization. Springer, Heidelberg (2007)
12. Srinivas, N., Deb, K.: Multiobjective optimization using nondominated sorting in genetic algorithms. Evolutionary Computation (1994)
13. Vose, M.D.: The Simple Genetic Algorithm. MIT Press, Cambridge (1999)
14. Zitzler, E.: Evolutionary Algorithms for Multiobjective Optimization: Methods and Applications. PhD thesis, ETH Zurich, Switzerland (1999)
15. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In: Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001), pp. 95–100. International Center for Numerical Methods in Engineering, CIMNE (2002)
16. Zitzler, E., Thiele, L.: Multiobjective evolutionary algorithms: A comparative case study and the strength pareto evolutionary algorithm. IEEE Transactions on Evolutionary Computation 3(4), 257–271 (1999)