Theoretical Aspects of Evolutionary Multiobjective Optimization—A Review

Dimo Brockhoff

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Theoretical Aspects of Evolutionary Multiobjective Optimization—A Review

Dimo Brockhoff

N° 7030

September 2009

ISSN 0249-6399 ISRN INRIA/RR--7030--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SACLAY - ÎLE-DE-FRANCE

Theoretical Aspects of Evolutionary Multiobjective Optimization—A Review

Dimo Brockhoff*

Thème : Mathématiques appliquées, calcul et simulation - Optimisation, apprentissage et méthodes statistiques
Équipes-Projets Thème Apprentissage et Optimisation (TAO)

Rapport de recherche n° 7030 — September 2009 — 38 pages

Abstract: Optimization problems in practice often involve the simultaneous optimization of 2 or more conflicting objectives. Evolutionary multiobjective optimization (EMO) techniques are well suited for tackling those multiobjective optimization problems because they are able to generate a set of solutions that represent the inherent trade-offs between the objectives. In the beginning, multiobjective evolutionary algorithms have been seen as single-objective algorithms where only the selection scheme needed to be tailored towards multiobjective optimization. In the meantime, EMO has become an independent research field with its specific research questions—and its own theoretical foundations. Several important theoretical studies on EMO have been conducted in recent years which opened up a better understanding of the underlying principles and resulted in the proposition of better algorithms in practice.

Besides a brief introduction about the basic principles of EMO, the main goal of this report is to give a general overview of theoretical studies published in the field of EMO and to present some of the theoretical results in more detail†. Due to space limitations, we only focus on three main aspects of previous and current research here: (i) performance assessment with quality indicators, (ii)

This work has been supported by the French national research agency (ANR) within the SYSCOMM project ANR-08-SYSC-017. In addition, the author would like to thank his former employer ETH Zurich for the support during the literature research and Anne Auger for her assistance in writing the mandatory French title and abstract.

* INRIA Saclay—Île-de-France, projet TAO, Bat 490, Université Paris-Sud, 91405 Orsay Cedex, France, dimo.brockhoff@lri.fr, http://www.lri.fr/∼brockho/

† The selection of the papers presented here is made as broad and objective as possible although such an overview can never be exhaustive. In particular, the author tried to collect all studies containing theoretical results on EMO published at the major conferences in the field, i.e., FOGA (1999–2009), EMO (2003–2009), GECCO, and CEC (2005–2009) as well as in all journal volumes of the IEEE Transactions on Evolutionary Computing and the Evolutionary Computation Journal. Furthermore, the EMOO web page http://www.lania.mx/∼ccoello/EMOO/ built the basis of the selection.

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université
4, rue Jacques Monod, 91893 ORSAY Cedex
Téléphone : +33 1 72 92 59 00

hypervolume-based search, and (iii) rigorous runtime analyses and convergence properties of multiobjective evolutionary algorithms.

Key-words: evolutionary multiobjective optimization, theory

Un état de l'art sur la théorie des algorithmes évolutionnaires multi-objectifs

Résumé : En pratique, les problèmes d'optimisation impliquent souvent l'optimisation simultanée de 2 ou plus objectifs contradictoires. Les techniques d'optimisation multi-objectif évolutionnaire (EMO) sont adaptées à de tels problémes parce qu'elles savent produire un ensemble de solutions qui représentent les compromis inhérents entre les objectifs. Les premiers algorithmes evolutionnaires multi-objectifs ont été vus comme des algorithmes mono-objectifs où seulement le sélection devait être adaptée pour l'optimisation multi-objectif. Depuis, l'EMO est devenu un domaine de recherche indépendant avec ses questions de recherche spécifiques—et ses propres fondations théoriques. Plusieurs études théoriques importantes sur l'EMO ont été accomplies au cours des dernières années. Ces derniéres ont rendu possible une meilleure compréhension des principes fondamentaux et cela a permis de proposer de meilleurs algorithmes en pratique.

Le but principal de ce rapport est de donner une vue d'ensemble des études théoriques publiées dans le domaine de l'EMO et présenter des résultats théoriques choisis en détail. Par manque de place, nous nous concentrons seulement sur trois aspects principaux : (i) l'évaluation des performances avec les indicateurs de qualité, (ii) les algorithmes fondés sur l'hypervolume et (iii) les analyses rigoureuses de complexité du temps d'atteinte d'un optimum au sens de Pareto.

Mots-clés : optimisation multi-objectifs évolutionnaires, théorie

1 Overview

Before we present the main theoretical studies in the field of evolutionary multiobjective optimization (EMO), we first introduce basic principals of multiobjective optimization in Sec. 2 and give a very brief overview of EMO methods in Sec. 3. Then, specific theoretical results on performance assessment (Sec. 4), hypervolume-based search (Sec. 5), and convergence and runtime analyses (Sec. 6) are presented and corresponding future research directions are pointed out. Section 8 concludes the report.

2 Multiobjective Optimization

In the following, we assume without loss of generality that k objective functions $f_i : X \to \mathbb{R}$, $1 \leq i \leq k$, mapping a solution x in decision space X to its objective vector $f(x) = (f_1(x), \ldots, f_k(x))$ in the objective space $\mathbb{R}^k$, have to be minimized simultaneously. This yields a major difference to single-objective optimization tasks: in most problems with more than one objective function, no solution exists that minimizes all objective functions simultaneously. Instead of finding or approximating the best objective function value, we consider in the following to find or approximate the set of so-called *Pareto-optimal solutions* representing the best trade-offs between the objectives. To this end, we define the Pareto dominance relation as follows. A solution $x \in X$ is said to *dominate* another solution $y \in X$ iff $\forall 1 \leq i \leq k : f_i(x) \leq f_i(y)$ and $\exists 1 \leq i \leq k : f_i(x) < f_i(y)$. We also write $x \prec y$. A solution $x^* \in X$ is then called *Pareto-optimal* iff there is no other solution in X that dominates x^*. In the same way, the *weak dominance relation* $\preceq$ can be defined. A solution $x \in X$ weakly dominates a solution $y \in X$ ($x \preceq y$) iff $\forall 1 \leq i \leq k : f_i(x) \leq f_i(y)$. Two solutions that are mutually weakly dominating each other are called *indifferent* whereas they are called *non-dominated* iff none is weakly dominating the other. Both dominance relations can be generalized to relations between sets of solutions. For example, a solution set $A \subseteq X$ weakly dominates a solution set $B \subseteq X$ iff $\forall b \in B \, \exists a \in A : a \preceq b$. Specific sets of solutions are the so-called *Pareto set approximations*, which are solution sets of pairwisely non-dominated solutions. More general definitions of the Pareto dominance concepts, e.g., via cones, exist but due to space limitations we refer the interesting reader to text books like [Ehrgott, 2005].

When comparing the differences between single-objective and multiobjective optimization, one viewpoint is the order-relation based view. In single-objective optimization, every solution is mapped to a real value and solutions can always be pairwisely compared via the less or equal relation $\leq$ on $\mathbb{R}$. In other words, the total order $\leq \subseteq \mathbb{R} \times \mathbb{R}$ induces via f an order on the search space X that is a total preorder, i.e., a reflexive, transitive, and total relation. In a multiobjective scenario, the $\leq$ relation is generalized to objective vectors, i.e., $\leq$ is a subset of $\mathbb{R}^k \times \mathbb{R}^k$. Here, the totality is not given due to vectors $a, b \in \mathbb{R}^k$ where $f_1(a) < f_1(b)$ but $f_2(a) > f_2(b)$—the relation $\leq$ on the set of objective vectors is only a partial order, i.e., reflexive, antisymmetric, and transitive. Moreover, the induced Pareto dominance relation and the weak Pareto dominance relation in the decision space X are not total like in the single-objective case. In the terms of order relations, the search for all Pareto-optimal solutions can be seen

as the search for all minimal elements of the order relation $\prec$. This set of minimal elements, also denoted as the *Pareto set*, is well-defined for a finite search space but also in the case of continuous optimization, the existence of a non-empty Pareto set can be proven if some assumptions on the problem are given, see for example [Henig, 1982] or [Miettinen, 1999, p. 35]. The image of the Pareto set under f is called *Pareto front*.

In practice, finding or approximating the Pareto set is not enough—usually a *decision maker* or a group of decision makers decides which of the non-dominated solutions found by an optimization algorithm is implemented in the end. Depending on when the decision maker is involved in the optimization process, three main approaches can be identified: *a priori*, *a posteriori* and *interactive* methods [Miettinen, 1999]. Until recently, research in evolutionary multiobjective optimization focused on a posteriori methods—assuming that the decision making process is postponed after the optimization. During the optimization, a good approximation of the Pareto set that both maps to a region close to the Pareto front and that is diverse in objective space is computed before a decision maker is involved, cf. [Deb, 2001, p. 24]. Due to the absence of decision making in most of the EMO research in general, also theoretical EMO studies assumed the a posteriori scenario as we will do in the remainder of this chapter.

A recently proposed view on multiobjective optimization should also be mentioned here. As the goal in an a posteriori scenario is to find a *set* of solutions, multiobjective problems can be seen as set problems [Zitzler et al., 2008]: the search space Ψ is then the set of all Pareto set approximations, i.e., a set of sets, and a set preference relation on Ψ is leading an optimization algorithm towards the minimal elements of this set preference relation. The advantage of this viewpoint is the simple integration of user preferences into the search, see [Zitzler et al., 2008] for details.

3 Evolutionary Multiobjective Optimization—A Very Brief Overview

Since the first usage of evolutionary algorithms for multiobjective optimization by Schaffer [1985], researchers have proposed many evolutionary algorithms that are tailored towards the simultaneous optimization of several objectives. Among the well-established ones, NSGA-II [Deb et al., 2002] and SPEA2 [Zitzler et al., 2002] have to be mentioned here. Both use the Pareto dominance concept as the main selection criterion in an elitist manner where non-dominated solutions are favored over dominated ones. In addition, a second selection criterion establishes diversity among the solutions. However, experimental studies have shown that both algorithms do not scale well if the number of objectives increases and that a cyclic behavior can be observed. This means that—although non-dominated solutions are preferred over dominated ones—over time, previously dominated solutions enter the population again, resulting in an oscillating distance to the Pareto front. During the last years, several attempts have been made to circumvent this behavior of which the indicator-based algorithms, especially those based on the ε-indicator and the hypervolume indicator, have been shown to produce better results in practice [Zitzler and Künzli, 2004, Beume et al., 2007b]. Besides the various approaches to improve the selection criterion

of evolutionary multiobjective algorithms, only minor attention has been paid to the variation operators so far where the multiobjective version of the CMA-ES is the most noticeable approach [Igel et al., 2007].

Nowadays, also several new research areas that are specific to EMO can be identified, of which interactive EMO [Jaszkiewicz and Branke, 2008], the combination of EMO and classical multicriteria decision making (MCDM) approaches, and the set-based view of EMO [Zitzler et al., 2008] are the probably most fruitful ones. For a more comprehensive view of the field of evolutionary multiobjective optimization, we refer the interested reader to the text books by Deb [2001] and Coello Coello et al. [2007].

4 Performance Assessment, the Attainment Function and Quality Indicators

With the huge amount of different evolutionary multiobjective optimizers, it is required to be able to compare the performance of them with respect to both certain test functions and on real-world applications. Unlike in single-objective optimization, where the outcome of an algorithm run is usually the real-valued best function value found so far, standard statistical methods are not applicable in the case of multiobjective optimization where the outcome of an evolutionary multiobjective optimizer is not directly describable by a single real-valued random variable but by a random *set* of real-valued *vectors*. This results in several difficulties in comparison to single-objective optimization: (i) not all resulting sets of objective vectors are comparable; (ii) the size of the sets can vary between different algorithms and even between different runs; (iii) standard statistical approaches cannot be applied directly, e.g., the mean of the sets of objective vectors generated by different runs of an algorithm might lie beyond the Pareto front if the front is concave; (iv) moreover, the comparison of evolutionary multiobjective optimizers' performance always needs to take into account the preferences of a decision maker. To tackle the mentioned difficulties, the attainment function approach and the idea of quality indicators have been proposed which we present from a theoretical point of view here.

4.1 The Attainment Function

The attainment function approach can be seen as the generalization of the cumulative distribution function $F_{\mathcal{X}}(z) = P(\mathcal{X} \leq z)$ of a real-valued random variable $\mathcal{X}$ with $z \in \mathbb{R}$ to the multiobjective domain. This generalization is necessary since the outcome of an evolutionary multiobjective optimizer cannot be modeled as a single real-valued random variable $\mathcal{X}$ for which the above distribution function allows to define common statistical measures such as the mean or the variance. Instead, the outcome of an evolutionary multiobjective algorithm needs to be modeled by a random *set* $\mathcal{A} = \{A_i \,|\, A_i \in \mathbb{R}^k, 1 \leq i \leq M\}$ of non-dominated objective vectors the size M of which is also a random variable. The assumption that $\mathcal{A}$ does not contain any dominated solution, i.e., that $A_i \not\leq A_j$ for any $i \neq j$ is not crucial here but simplifies the notations. The attainment function of such a set of non-dominated objective vectors $\mathcal{A}$, first proposed by Fonseca and Fleming [1996] and later on further investigated and generalized by Grunert da Fonseca et al. [2001] and Fonseca et al. [2005], is defined by the

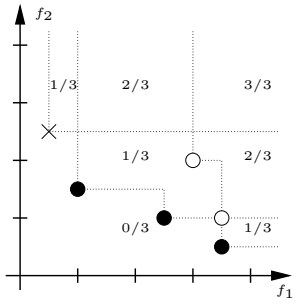

Figure 1: Illustration of the empirical attainment function for three different sets of objective vectors. The numbers indicate the value of the empirical attainment function $\alpha_3(z)$, i.e., the number of runs that produce objective vectors dominating the corresponding objective vector z.

function $\alpha_{\mathcal{A}} : \mathbb{R}^k \to [0, 1]$ with

$$\alpha_{\mathcal{A}}(z) = P(A_1 \leq z \vee A_2 \leq z \vee \ldots \vee A_M \leq z) =: P(\mathcal{A} \leq z) \ ,$$

where $\leq \subseteq \mathbb{R}^k \times \mathbb{R}^k$ is the less or equal relation on the set of all objective vectors, $z \in \mathbb{R}^k$, and $P(E)$ is the probability of an event E.

For an arbitrary objective vector z, the attainment function corresponds to the probability that at least one of the objective vectors produced by an algorithm run dominates z. It is a generalization of the multivariate cumulative distribution function $F_A(z) = P(A \leq z)$ of a vector $A \in \mathbb{R}^k$ with $z \in \mathbb{R}^k$ to which it reduces if only $M = 1$ objective vector per run is produced by the algorithm. In case of only one objective it reduces further to the standard cumulative distribution function $F_{\mathcal{X}}(z) = P(\mathcal{X} \leq z)$ of a real-valued random variable $\mathcal{X}$ where $z \in \mathbb{R}$. In practice, the attainment function can be estimated by the *empirical attainment function*

$$\alpha_n(z) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{I}\{\mathcal{A}_i \leq z\} \ ,$$

where $\mathbf{I}\{\cdot\}$ is the indicator function, giving 1 iff the argument is true and where the random sets $\mathcal{A}_1, \ldots, \mathcal{A}_n$ correspond to the outcomes of n independent runs of the optimizer, see Fig. 1. Statistical tests can then be run on this empirical attainment function for different algorithms to reject a null hypothesis like "algorithms Alg_1 and Alg_2 are performing equally", see for example [Fonseca and Fleming, 1996] for details.

Besides the definition and interpretation of the attainment function, Grunert da Fonseca et al. [2001] also pointed out an interesting relation to random closed set theory: Instead of the set $\mathcal{A} = \{A_i \,|\, A_i \in \mathbb{R}^k, 1 \leq i \leq M\}$, the outcome of an algorithm can also be described by the region $\mathcal{Y}$ in objective space that is weakly dominated by the solutions associated with $\mathcal{A}$ or in other words that is *attained* by $\mathcal{A}$:

$$\mathcal{Y} = \{y \in \mathbb{R}^k \,|\, A_1 \leq y \vee A_2 \leq y \vee \ldots \vee A_M \leq y\} \ .$$

With this alternative representation, the attainment function can be written as $\alpha_{\mathcal{A}}(z) = P(z \in \mathcal{Y})$ for any $z \in \mathbb{R}^k$ and represents the expected value of

the binary random field $\mathbf{I}\{\mathcal{Y} \cap \{z\} \neq \emptyset\} = \mathbf{I}\{\mathcal{A} \leq z\}$, see [Grunert da Fonseca et al., 2001] for details. In other words, the attainment function represents the first order moment of the *location* of the objective vectors produced by an algorithm outcome $\mathcal{A}$. The generalization of the attainment function to higher order moments has been investigated as well [Fonseca et al., 2005] but its applicability in practice failed up to now—mainly due to the high dimensionality of the approach[1]. Nevertheless, according to its general properties and its links to other fields of mathematics, the attainment function approach is one of the main results in the field of the theory of evolutionary multiobjective optimization.

4.2 Quality Indicators

Quality indicators are functions that map one or several Pareto set approximations to a real value. With this, the performance assessment of multiobjective optimizers can be done in the real-valued domain. Since with quality indicators only real numbers have to be compared, two Pareto set approximations can be made always comparable by investigating the quality indicator values of them—even if they are incomparable with respect to the Pareto dominance relation. Moreover, a quality indicator also allows for statements of *how much better* a Pareto set approximation is compared to another one.

Definition 1 (Quality indicator)**.** *Let $\Psi \subseteq 2^X$ be the set of all possible Pareto set approximations. An m-ary quality indicator is a function $I : \Psi^m \to \mathbb{R}$, assigning each vector $(A_1, \ldots, A_m)$ of m Pareto set approximations a real value $I(A_1, \ldots, A_m)$.*

An example of a (unary) quality indicator is the general distance measure [Veldhuizen and Lamont, 2000] that assigns a Pareto set approximation A the average Euclidean distance in objective space between all objective vectors $f(a)$ with $a \in A$ and the Pareto front. Another unary quality indicator is the hypervolume indicator [Zitzler and Thiele, 1998b, see also Sec. 5] that assigns A the hypervolume of the objective space dominated by A but not by a specified reference set R. The ε-indicator of Zitzler et al. [2003] is one example of a binary quality indicator. For two Pareto set approximations A, B, the ε-indicator value

$$I_\varepsilon(A, B) = \inf_{\varepsilon \in \mathbb{R}} \{\forall b \in B \exists a \in A \forall 1 \leq i \leq k : f(a) \leq \varepsilon \cdot f(b)\}$$

can be interpreted as the smallest value by which all objective vectors of solutions in A have to be divided such that B is weakly dominated—if A weakly dominates B without any rescaling of the objective vectors, the ε-indicator value gives a value of ≤ 1 whereas $I_\varepsilon(A, B) > 1$ indicates that $A \not\preceq B$.

With the help of quality indicators like the described ones, the performance assessment of multiobjective stochastic optimizers usually follows the same theoretical framework:

Definition 2 (Comparison Method [Zitzler et al., 2003])**.** *Let $A, B \in \Psi$ be two Pareto set approximations, $\mathbf{I} = \{I_1, \ldots, I_l\}$ a set of quality indicators and $E : \mathbb{R}^k \times \mathbb{R}^k \to \{false, true\}$ an interpretation function that maps two real vectors*

[1] The attainment function itself can be plotted in k-dimensional space, but the second order moment of it, comparing two approximation sets, needs already $2k$ dimensions which is problematic for visualization even if only $k = 2$ objectives are considered.

to a Boolean value. In case, $\mathbf{I}$ *contains only unary quality indicators, we define the* comparison method $C_{\mathbf{I},E}$ *as*

$$C_{\mathbf{I},E}(A,B) = E(\mathbf{I}(A),\mathbf{I}(B))$$

and as

$$C_{\mathbf{I},E}(A,B) = E(\mathbf{I}(A,B),\mathbf{I}(B,A))$$

in case $\mathbf{I}$ *contains only binary indicators where* $\mathbf{I}(A') = (I_1(A'),\ldots,I_l(A'))$ *and* $\mathbf{I}(A',B') = (I_1(A',B'),\ldots,I_l(A',B'))$ *for arbitrary* $A',B' \in \Psi$*.*

Typical examples of the usage of quality indicators fit to this formalization of a comparison method. For a single unary indicator I for example, $\mathbf{I}(A') = I(A')$ and $E(I(A),I(B))$ is usually defined as $I(A) > I(B)$ if we assume maximization of the indicator function. Thus, we can compare all types of Pareto set approximations by comparing their quality indicator value and interpreting a "true" of the interpretation function E as "A is better than B"—even if the solution sets are incomparable with respect to the (weak) Pareto dominance relation. In this case, the performance assessment of multiobjective optimizers can be done similar to the single-objective case and standard statistical approaches can be used to investigate the differences in indicator values between different algorithms.

However, there is an important fact that one should keep in mind when using quality indicators. Although we are always able to come to a statement about whether a Pareto set approximation is better than another one with the help of this approach, one would like to use quality indicators that do not contradict decisions that can be made by the weak Pareto dominance relation itself as it represents the most general form of "outperformance": whenever a Pareto set approximation A is better than another set B with respect to the weak Pareto dominance relation, i.e., when $A \preceq B \wedge B \not\preceq A$, which we denote by $A \lhd B$, we do not want to get a different statement with the comparison method $C_{\mathbf{I},E}(A,B)$. In other words, we would like $C_{\mathbf{I},E}(A,B)$ to be a sufficient criterion for $A \lhd B$ such that $C_{\mathbf{I},E}(A,B)$ can state *that* A is better than B. We say, the comparison method is compatible with the dominance relation $\lhd$ [2]. If $C_{\mathbf{I},E}(A,B)$ is in addition also a necessary condition for $A \lhd B$, i.e., if $C_{\mathbf{I},E}(A,B) \Leftrightarrow A \lhd B$, the comparison method can indicate *whether* A is better than B which is formalized in the following definition.

Definition 3 (Compatibility and Completeness [Zitzler et al., 2003])**.** *Let* $\blacktriangleleft$ *be an arbitrary binary relation on the set* Ψ *of Pareto set approximations. The comparison method* $C_{\mathbf{I},E}$ *is denoted as* $\blacktriangleleft$-compatible *if either*

$$\forall A,B \in \Psi : C_{\mathbf{I},E}(A,B) \Rightarrow A \blacktriangleleft B \quad \text{or} \quad \forall A,B \in \Psi : C_{\mathbf{I},E}(A,B) \Rightarrow B \blacktriangleleft A$$

and as $\blacktriangleleft$-complete *if either*

$$\forall A,B \in \Psi : A \blacktriangleleft B \Rightarrow C_{\mathbf{I},E}(A,B) \quad \text{or} \quad \forall A,B \in \Psi : B \blacktriangleleft A \Rightarrow C_{\mathbf{I},E}(A,B) \ .$$

Zitzler et al. [2003] also theoretically investigate the restrictions of using a set of only unary indicators, the main result of which we state here while referring to the paper for the proof.

[2] Although our notation follows [Zitzler et al., 2003] here, other studies define the compatibility in the same manner [Hansen and Jaszkiewicz, 1998, Knowles and Corne, 2002, Knowles, 2002, Farhang-Mehr and Azarm, 2003] or relate it to so-called refinements of the weak Pareto dominance relation in set-based EMO [Zitzler et al., 2008].

Theorem 1 ([Zitzler et al., 2003]). *For the case of an optimization problem with $k \geq 2$ objectives and objective space $\mathbb{R}^k$, there exists no comparison method $C_{\mathbf{I},E}$ that is based on a finite combination $\mathbf{I}$ of unary quality indicators that is $\lhd$-compatible and $\lhd$-complete at the same time.*

Although the previous theorem tells us that there is no comparison method based on unary indicators that yields the equivalence $C_{\mathbf{I},E}(A,B) \Leftrightarrow A \lhd B$, the *unary hypervolume indicator* I_H, further investigated in the next section, is both $\not\lhd$-compatible and $\lhd$-complete such that whenever $I_H(A) > I_H(B)$ we know that B is not better than A ($C_{\mathbf{I},E}(A,B) \Rightarrow B \not\lhd A$) and on the other hand, the comparison method also detects that A is better than B whenever this is the case ($A \lhd B \Rightarrow C_{\mathbf{I},E}(A,B)$). In this case, we also say the hypervolume indicator is *compliant* with the weak Pareto dominance relation. Note that many other quality indicators proposed in the literature, e.g., the mentioned generational distance measure, do not respect the Pareto dominance relation, i.e., are neither $\blacktriangleleft$-compatible nor $\blacktriangleleft$-complete for $\blacktriangleleft \in \{\lhd, \preceq, \prec\}$, cf. [Zitzler et al., 2003]. Also note that the restriction of the previous theorem does not hold for binary indicators such that with several binary indicators, e.g., the above mentioned binary ε-indicator, comparison methods that are at the same time $\lhd$-compatible and $\lhd$-complete can be constructed, see [Zitzler et al., 2003] for details.

4.3 Future Research Directions

With the help of quality indicators, it is possible to use techniques known from single-objective optimization also for multiobjective performance assessment. However, far the most studies nowadays fix the number of evaluations and consider the achieved quality indicator values. Instead, in single-objective optimization, the recent trend is to fix a target value (of the objective function, which would be the quality indicator function in the multiobjective case) and report the number of evaluations to reach it. Like that, the reported numbers have an absolute meaning: if an algorithm needs twice as many function evaluations to reach a certain target than another one it is twice as fast, whereas reaching an objective function or quality indicator value of $2x$ instead of x, one can only give a relative statement of the first being better than the latter. Here, a rethinking needs to take place in the EMO field and theoretical investigations of this so-called *horizontal view*[3] of performance assessment might help here.

Another theoretical question is whether there exists an indicator that is both $\blacktriangleleft_1$-compatible and $\blacktriangleleft_2$-complete for a certain combination of set preference relations $\blacktriangleleft_1$ and $\blacktriangleleft_2$. For some combinations, Zitzler et al. [2003] gave already the answer, but for some combinations, the question is still open, compare Table III in [Zitzler et al., 2003].

Last and in anticipation of the following section, the question which Pareto set approximations yield a maximal value of a given unary quality indicator needs to be investigated. With respect to the mentioned horizontal view of performance assessment, appropriate target values for a fixed test function, a fixed quality indicator, and an upper bound on the Pareto set approximation size

[3] If one plots the achieved objective function or quality indicator values for several algorithms over time, fixing a target value can be seen as comparing the runtimes of the algorithms at a horizontal line located at the fixed target value [Hansen et al., 2009].

can only be given if the maximum value of the quality indicator among all Pareto set approximations with a fixed number of solutions is known. Furthermore, the knowledge about where these Pareto set approximations, that maximize a quality indicator, are located in objective space is necessary to investigate the bias of the indicator when the (weak) Pareto dominance relation is replaced by the indicator as selection criterion. The following section investigates this question further for the special case of the hypervolume indicator.

5 Hypervolume-Based Search

One of the latest areas of evolutionary multiobjective optimization where theoretical investigations have been made is the area of hypervolume-based search. The hypervolume indicator, initially proposed for performance assessment, has nowadays been used to guide the search in several multiobjective evolutionary algorithms such as the SMS-EMOA of Beume et al. [2007b], the multiobjective version of CMA-ES [Igel et al., 2007], or HypE [Bader and Zitzler, 2008, 2009].

As mentioned above, the hypervolume indicator is a unary quality measure mapping a set of solutions to the hypervolume in objective space that is dominated by the corresponding objective vectors of the solutions but not by a predefined set of so-called *reference points*. More formally, for a given set $A \in X$ of solutions and a set $R \in \mathbb{R}^k$ of reference points in objective space, we define the set of objective vectors that are dominated by A but not by R according to Zitzler et al. [2008] as

$$H(A,R) = \{h \mid \exists a \in A \exists r \in R : f(a) \leq h \leq r\}$$

and the corresponding hypervolume indicator $I_H(A)$ as the Lebesgue measure of this set

$$I_H(A) = \lambda(H(A,R)) = \int_{\mathbb{R}^n} \mathbf{1}_{H(A,R)}(z)dz$$

where $\mathbf{1}_{H(A,R)}(z)$ is the indicator function which equals 1 iff $z \in H(A,R)$ and 0 otherwise. Zitzler et al. [2007] generalized the hypervolume indicator to a weighted version, where a user-specified weight distribution on the objective space can be used to guide the search towards certain regions and or points of high interest, so-called preference points[4]. With respect to a weight function $w : \mathbb{R}^k \to \mathbb{R}$, the weighted hypervolume indicator $I_H^w(A)$ for a solution set $A \subseteq X$ is defined by

$$I_H^w(A) = \int_{\mathbb{R}^n} w(z)\mathbf{1}_{H(A,R)}(z)dz \ .$$

In the remainder of this chapter, we will assume, for simplicity, one reference point only, i.e., $R = \{r\}$ where $r \in \mathbb{R}^k$ is the hypervolume's reference point in objective space.

The main idea behind all hypervolume-based algorithms is the same, be it with respect to the original or the weighted version: applying the (weighted) hypervolume indicator to a set of solutions reformulates a multiobjective problem

[4]Instead of the term *reference point* frequently used in the area of Multicriteria Decision Making, see [Miettinen, 1999], we use the term preference point here to reduce the confusion with the hypervolume indicator's reference point.

as a single-objective one and the aim is to find a solution set, the hypervolume indicator of which is as high as possible. In case of population-based algorithms where the maximal number of solutions is finite and corresponds often to the population size μ, the goal is to find or approximate a set of μ solutions that maximizes the hypervolume indicator [Auger et al., 2009a].

In the light of hypervolume-based search, two main lines of theoretical research can be identified. On the one hand, fast and effective algorithms to compute the hypervolume indicator are needed and theoretical research in this area focused on the computational complexity of the hypervolume computation and the search for fast algorithms. On the other hand, a theoretical characterization of how the μ solutions, optimizing the hypervolume indicator, are distributed on the Pareto front builds the basis for hypervolume-based evolutionary algorithms in general and the comparison of algorithms with respect to the hypervolume indicator as performance criterion in particular. The characterization of such so called *optimal μ-distributions* is one of the more recent branches in theoretical evolutionary multiobjective optimization and is going to be presented together with the results on the computational complexity of the hypervolume computation in more detail.

5.1 Computational Complexity of Computing the Hypervolume Indicator

When introducing the hypervolume indicator for performance assessment, Zitzler and Thiele [1998a] did not publish an explicit algorithm for computing it but independently, Knowles [2002] proposed a similar algorithm and proved the runtime of his algorithm to be $\mathcal{O}(n^k+1)$ for n solutions and k objectives. Later on, While et al. [2006] studied algorithms for exactly computing the hypervolume indicator in more detail. Besides showing that a previously proposed algorithm for the hypervolume calculation has exponential runtime instead of the claimed polynomial one[5], While et al. [2006] proposed the *hypervolume by slicing objectives* (HSO) algorithm and thoroughly proved the number of generated hypercubes to be $\binom{k+n-2}{n-1}$ by solving appropriate recurrence equations. Fonseca et al. [2006] improved the exact hypervolume calculation further by proposing an improved dimension-sweep algorithm with a runtime of $\mathcal{O}(n^{k-2}\log n)$. Furthermore, Beume and Rudolph [2006] proposed the current best exact algorithm with a runtime of $\mathcal{O}(n\log n+n^{k/2})$ by using a relation between the hypervolume indicator calculation and Klee's measure problem.

With Klee's measure problem, introduced by Klee [1977], we associate the calculation of the size of the union of a set of n real-valued intervals. Generalized to an arbitrary dimension k, we ask for the computation of the union of a set of n axis-parallel hypercuboids. The relation to the hypervolume indicator calculation is straightforward, cf. [Beume and Rudolph, 2006]: given a set $A \in X$ of n solutions and the hypervolume's reference point $r \in \mathbb{R}^k$, one can build an instance for Klee's measure problem by using the objective vectors as lower bounds and the reference point as upper bound for n k-dimensional intervals or hypercuboids respectively. Therefore, known algorithms for Klee's measure problem can be directly used for the hypervolume indicator computation. Beume and Rudolph [2006] use a slightly modified version of the best

[5] This proof was originally published in [While, 2005].

known algorithm for Klee's measure problem proposed by Overmars and Yap [1991] which has a runtime of $\mathcal{O}(n \log n + n^{k/2} \log n)$. This algorithm is tailored towards the hypervolume indicator computation using the fact that the hypercuboids all have one corner, i.e., the hypervolume's reference point in common. Note here that Beume and Rudolph [2006] argue that this property of the hypercuboids reduces the runtime of the algorithm to $\mathcal{O}(n \log n + n^{k/2})$ but Bringmann and Friedrich [2009a] state that there is still some ongoing discussion whether there are counter examples.

After the problem of the worst-case complexity of the hypervolume indicator computation remained open for years—only Beume et al. [2007a] proved a non-trivial lower bound of $\Omega(n \log n)$ for a fixed number of k objectives—Bringmann and Friedrich [2008] showed that the problem of computing the hypervolume indicator value for n solutions and k objectives is $\#\mathcal{P}$-hard. This result indicates that we cannot hope for an exact algorithm for computing the hypervolume indicator that is polynomial in the number of objectives unless $\mathcal{P} = \mathcal{NP}$[6].

The proof of the $\#\mathcal{P}$-hardness of the hypervolume indicator computation problem, denoted by `HYP` in the following, is a simple reduction from the $\#\mathcal{P}$-hard problem `#MON-CNF` that is defined as follows.

Definition 4. *Given a monotone Boolean formula in conjunctive normal form, i.e., of the form* $\texttt{CNF} = \bigwedge_{j=1}^{n} \bigvee_{i \in C_j} x_i$ *where* $C_j \subset \{1, \ldots, k\}$ *are the clauses and the* k *variables* x_i *are only used in their non-negated forms. Then the problem* `#MON-CNF` *(*`SATISFIABILITY PROBLEM FOR MONOTONE BOOLEAN FORMULAS IN CONJUNCTIVE NORMAL FORM`*) asks for the number of satisfying variable assignments for CNF.*

Theorem 2 (Bringmann and Friedrich [2008])**.** `HYP` *is* $\#\mathcal{P}$*-hard.*

Proof. We show the $\#\mathcal{P}$-hardness of `HYP` by a polynomial reduction from the $\#\mathcal{P}$-hard problem `#MON-CNF`. To this end, let $\texttt{CNF} = \bigwedge_{j=1}^{n} \bigvee_{i \in C_j} x_i$ be the `#MON-CNF` instance where $C_j \subset \{1, \ldots, k\}$ are the clauses. We compute the number of satisfying variable assignments—instead for `CNF`—for the negated formula $\overline{\texttt{CNF}} = \bigvee_{j=1}^{n} \bigwedge_{i \in C_j} \neg x_i$ and return the number 2^k of all possible variable assignments minus the computed number of variable assignments for $\overline{\texttt{CNF}}$. For constructing the `HYP` instance, we introduce for each clause C_j a solution a^j with objective vector $(a_1^j, a_2^j, \ldots, a_k^j) \in \mathbb{R}^k$ where $a_i^j = 1$ if $i \in C_j$ and $a_i^j = 0$ otherwise for each $i \in \{1, \ldots, k\}$ and set the hypervolume's reference point to $(2, \ldots, 2) \in \mathbb{R}^k$. We observe that the hypervolume indicator value of the set of all solutions a^j, $1 \leq j \leq n$, can be written as the hypervolume of a union of boxes of the form

$$B_{(x_1,\ldots,x_k)} = [1 - x_1, 2 - x_1] \times [1 - x_2, 2 - x_2] \times \cdots \times [1 - x_k, 2 - x_k]$$

with $x_i \in \{0, 1\}$ for all $1 \leq i \leq k$. Moreover, $B_{(x_1,\ldots,x_k)}$ is a subset of the hypervolume dominated by all solutions a^j iff it is a subset of the hypervolume contribution $[a_1^j, 2] \times \cdots \times [a_k^j, 2]$ of at least one solution a^j iff we have $a_i^j \leq 1 - x_i$ for all $i \in \{1, \ldots, k\}$ iff $a_i^j = 0$ for all $i \in \{1, \ldots, k\}$ with $x_i \in 1$ iff $i \notin C_j$ for all

[6] A problem A is called $\#\mathcal{P}$-hard if all problems in $\#\mathcal{P}$ are Turing reducible to A where $\#\mathcal{P}$ is the set of all counting problems that can be solved by a *counting Turing machine*—a non-deterministic Turing machine that outputs its number of accepting paths in polynomial time, see [Valiant, 1979] for details.

i with $x_i = 1$ iff $(x_1, \ldots, x_k)$ satisfies $\bigwedge_{i \in C_j} \neg x_i$ for at least one $j \in \{1, \ldots, n\}$ iff $(x_1, \ldots, x_k)$ satisfies $\overline{\texttt{CNF}}$.

Since the hypervolume of a box $B_{x_1,\ldots,x_k}$ equals 1, we have therefore

$$I_H(\bigcup_{1 \leq j \leq n} a^j) = |\{(x_1, \ldots, x_k) \in \{0,1\}^k \,|\, (x_1, \ldots, x_k) \text{ satisfies } \overline{CNF}\}| \ .$$

Thus, a polynomial time algorithm for HYP would result in a polynomial time algorithm for #MON-CNF which proves the theorem. □

Along with this first complexity result, Bringmann and Friedrich [2009b] also proved the computational complexity for related problems occurring in hypervolume-based multiobjective optimization, such as finding the solution of a given set with the smallest hypervolume contribution or computing the hypervolume of this smallest contribution. Also those two problems have been shown to be difficult to solve exactly; depending on the problem type, Bringmann and Friedrich [2009a] showed the $\#\mathcal{P}$-hardness or the $\mathcal{NP}$-hardness of these problems.

Furthermore, Bringmann and Friedrich also proposed efficient approximation algorithms for HYP and the related problems and proved bounds on their approximation quality and runtime [Bringmann and Friedrich, 2008, 2009b]. Also an exact algorithm for computing the set of μ solutions out of $n \geq \mu$ that maximizes the hypervolume indicator has been proposed—reducing the runtime of all previously known algorithms by a factor of $n^{\min\{n-\mu, d/2\}}$ to $\mathcal{O}(n^{k/2} \log n + n^{n-\mu})$ for $k > 3$ by exploiting the algorithm by Overmars and Yap [Bringmann and Friedrich, 2009a]. We refrain from presenting more details here due to space limitations and refer the interested reader to the corresponding publications.

5.2 Optimal μ-Distributions

Due to the Pareto dominance compliance property of the hypervolume indicator (see Sec. 4.2), we know that maximizing the hypervolume indicator will result in a set of solutions the objective vectors of which (i) lie on the Pareto front and (ii) even cover the entire Pareto front if the number of solutions is larger or equal than the number of Pareto-optimal solutions as was shown by Fleischer [2003].

However, until recently, it was not clear how the objective vectors of μ solutions that maximize the hypervolume indicator are distributed on the Pareto front. Furthermore, the question arises whether evolutionary algorithms that aim at finding such a set of μ solutions maximizing the hypervolume indicator are really converging to the optimal hypervolume value.

In the light of this discussion, we denote a set of μ points that have the largest hypervolume indicator value among all sets of μ points as *optimal μ-distribution*.

Definition 5. *A set $X^\mu \subseteq X$ of μ solutions that maximizes the hypervolume indicator, i.e., with*

$$X^\mu = \operatorname*{argmax}_{\substack{A \subseteq X \\ |A| = \mu}} \{I_H(A)\} \ ,$$

is called optimal μ-distribution.

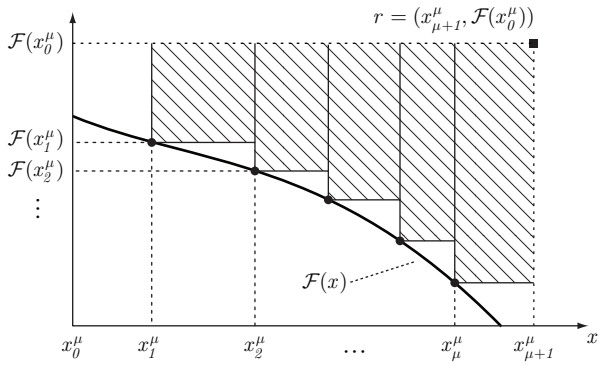

Figure 2: Hypervolume (hatched area) for a set of objective vectors (filled circles) on a biobjective continuous Pareto front. The reference point is depicted by a filled square.

In case of a finite Pareto front, the existence of an optimal μ-distribution is trivial and for the case of a continuous Pareto front, Auger et al. [2009a] gave a simple existence proof based on the Mean Value Theorem. Auger et al. [2009a] have also been the first to rigorously study where the points in an optimal μ-distribution are located on the front. This helps to understand the bias, the hypervolume (and especially the weighted hypervolume) is introducing as well as to investigate whether hypervolume-based algorithms converge towards the optimum of the formulated optimization problem, i.e., to an optimal μ-distribution.

5.2.1 Properties of Optimal μ-Distributions

To state the main results about optimal μ-distributions, we introduce the basic notations from [Auger et al., 2009a]. All published results in this area only deal with biobjective problems, i.e., $k = 2$ in the remainder of this section. We furthermore assume that the Pareto front can be described by a continuous function $x \in [x_{\min}, x_{\max}] \mapsto \mathcal{F}(x)$ where x lies in the image of the decision space under the first objective f_1 and $\mathcal{F}(x)$ lies in the image of the decision space under f_2, see Fig. 2. For simplicity, we neglect the decision space completely, i.e., identifying a solution with its objective vector. A solution on the Pareto front can then be—due to the restriction to biobjective problems—unambiguously identified with its first coordinate $x \in [x_{\min}, x_{\max}]$. The hypervolume indicator value of a set of μ solutions on the Pareto front can then be written as

$$I^\mu_H((x^\mu_1, \ldots, x^\mu_\mu)) := \sum_{i=1}^{\mu} (x^\mu_{i+1} - x^\mu_i)(\mathcal{F}(x^\mu_0) - \mathcal{F}(x^\mu_i))$$

where we define $x^\mu_{\mu+1} := r_1$ and $\mathcal{F}(x^\mu_0) := r_2$ and $r = (r_1, r_w)$ is the reference point of the hypervolume indicator, cf. Fig. 2.

The first property of optimal μ-distributions, stated in [Auger et al., 2009a], is its existence: if the function $\mathcal{F}(x)$ is continuous there is at least one optimal μ-distribution. How an optimal μ-distribution can be characterized and whether

there is a unique optimal μ-distribution are questions that have been further investigated in [Auger et al., 2009a]. Two of the general results on optimal μ-distributions can be stated as follows.

Proposition 1. (Necessary condition for optimal μ-distributions) *If $\mathcal{F}$ is continuous, differentiable and $(x_1^\mu, \ldots, x_\mu^\mu)$ denote the x-coordinates of a set of μ points maximizing the hypervolume indicator, then for all $x_{min} < x_i^\mu < x_{max}$*

$$\mathcal{F}'(x_i^\mu)\left(x_{i+1}^\mu - x_i^\mu\right) = \mathcal{F}(x_i^\mu) - \mathcal{F}(x_{i-1}^\mu),\ i = 1 \ldots \mu \tag{1}$$

where $\mathcal{F}'$ denotes the derivative of $\mathcal{F}$, $\mathcal{F}(x_0^\mu) = r_2$ and $x_{\mu+1}^\mu = r_1$.

Proof. First of all, we observe that the position of a point $(x_i^\mu, \mathcal{F}(x_i^\mu))$ between its neighbors $(x_{i-1}^\mu, \mathcal{F}(x_{i-1}^\mu))$ and $(x_{i+1}^\mu, \mathcal{F}(x_{i+1}^\mu))$ only influences the hypervolume with respect to the objective space that is solely dominated by $(x_i^\mu, \mathcal{F}(x_i^\mu))$. The volume of this hypervolume contribution of the point $(x_i^\mu, \mathcal{F}(x_i^\mu))$ equals

$$H_i = (x_{i+1}^\mu - x_i^\mu)(\mathcal{F}(x_{i-1}^\mu) - \mathcal{F}(x_i^\mu)) \ .$$

For an optimal μ-distribution, each of the hypervolume contributions H_i has to be maximal with respect to x_i^μ since otherwise one would be able to increase I_H^μ by moving x_i^μ. Since the hypervolume contribution becomes zero for x_i^μ at the boundary of $[x_{i-1}^\mu, x_{i+1}^\mu]$, the maximum must lie in the interior of the domain. Therefore, the necessary condition holds that the derivative of H_i with respect to x_i^μ is zero or x_i^μ is an endpoint of $\mathcal{F}$, i.e., either $x_i^\mu = x_{min}$ or $x_i^\mu = x_{max}$. The derivative of H_i with respect to x_i^μ equals

$$H_i^{'}(x_i^\mu) = -x_{i+1}^\mu \mathcal{F}'(x_i^\mu) - \mathcal{F}(x_{i-1}^\mu) + \mathcal{F}(x_i^\mu) + x_i^\mu \mathcal{F}'(x_i^\mu) \ .$$

Reorganizing the terms and setting $H_i^{'}(x_i^\mu)$ to zero, we obtain Equation 1. □

Corollary 1. *If $x_i^\mu, i = 2 \ldots \mu - 1$ is a point of a set of μ points maximizing the hypervolume indicator and x_i^μ is not an endpoint of the Pareto front, then $\mathcal{F}'(x_i^\mu) \neq 0$.*

Here, as well as for the following results, we refrain from showing the proof due to space restrictions and refer to the original publication instead. Besides these general results, Auger et al. [2009a] argue that the characterization of optimal μ-distributions is not an easy task and therefore pursue their study on the one hand by proving results for special front shapes, i.e., for linear fronts, and on the other hand by investigating the case where μ goes to infinity and the optimal distribution of μ points on the Pareto front converges to a density. For the case of linear fronts, Auger et al. [2009a] proved the uniqueness of the optimal μ-distribution and gave a formula describing the location of the corresponding μ points exactly.

Theorem 3 ([Auger et al., 2009a])**.** *If the Pareto front is a (connected) line, optimal μ-distributions are such that the distance is the same between all neighbored solutions.*

This result covers earlier published results on even more special cases by Emmerich et al. [2007] and Beume et al. [2007a] where the slope of the linear

front equals 1 and where the extreme points of the front are always assumed to be included in the optimal μ-distribution.

As to the results for μ going to infinity, i.e., results on the density $\delta(x)$ of points on the Pareto front at a point $(x, \mathcal{F}(x))$, we only state the most general result for the weighted hypervolume indicator informally due to space limitations [Auger et al., 2009c]. For a fixed integer μ, Auger et al. [2009c] consider a sequence of μ ordered points on the Pareto front, for which their x-coordinates are $x_1^\mu, \ldots, x_\mu^\mu$. Then, the authors assume that this sequence converges—when μ goes to ∞—to a density $\delta(x)$ where the density in a point x is formally defined as the limit of the number of points contained in a small interval $[x, x+h[$ normalized by the total number of points μ when both μ goes to ∞ and h to 0, i.e., $\delta(x) = \lim_{\substack{\mu\to\infty \\ h\to 0}} \left(\frac{1}{\mu h} \sum_{i=1}^{\mu} \mathbf{1}_{[x,x+h[}(x_i^\mu) \right)$.

Then, one can prove that the density δ corresponding to an optimal μ-distribution equals

$$\delta(x) = \frac{\sqrt{-\mathcal{F}'(x) w(x, \mathcal{F}(x))}}{\int_0^{x_{max}} \sqrt{-\mathcal{F}'(x) w(x, \mathcal{F}(x))} dx} \ .$$

When looking carefully at this formula for the density, we see that for a fixed weight w, the density of points only depends on the slope of the front and not on whether the front is convex or concave as in previous belief. The analyses in [Auger et al., 2009a,b] further show that the hypervolume indicator prefers regions of the front the slope of which is -1 and supplementary experimental studies show that even for a small number μ of points, the density is a good approximation for an optimal μ-distribution. Numerical results on optimal μ-distributions, as obtained in [Auger et al., 2009a] for some test functions, can help in performance assessment of multiobjective evolutionary algorithms by comparing the approximation of the optimal μ-distribution and the distribution of points, found by the algorithms if the hypervolume indicator is the underlying quality indicator. However, if the hypervolume indicator is the assumed underlying quality indicator and therefore, an optimal μ-distribution is sought by an evolutionary algorithm, the additional question arises how the convergence to an optimal μ-distribution of a multiobjective evolutionary algorithm can be guaranteed which immediately brings us to the topic of hypervolume-based selection.

5.2.2 Hypervolume-Based Selection and the Convergence to Optimal μ-Distributions

Optimally, having generated λ offspring from μ parents in a hypervolume-based evolutionary algorithm, one would like to choose the next population as the set of μ solutions out of all $\mu + \lambda$ solutions that maximizes the hypervolume indicator. Like this, one can easily guarantee the convergence to an optimal μ-distribution if $\lambda \geq \mu$ and every point in the search space can be sampled with a positive probability[7]. This strategy, however, is computationally expensive as, in principle, $\binom{\mu+\lambda}{\mu}$ many solution sets have to be considered[8].

[7] Under these circumstances, the probability to sample μ new solutions with a better hypervolume indicator value than the current μ parents is always positive when no optimal μ-distribution is found.

[8] Bringmann and Friedrich [2009a] proposed an algorithm that computes the solution set with the highest hypervolume indicator among all sets with μ solutions in time $\mathcal{O}(n^{k/2} \log n +$

Algorithm 1 Simple Indicator-Based Evolutionary Algorithm (SIBEA)

Given: population size μ; number of generations N
Step 1 (Initialization): Generate an initial set of decision vectors P of size μ; set the generation counter $m := 0$
Step 2 (Environmental Selection): Iterate the following three steps until the size of the population does no longer exceed μ:

1. Rank the population using dominance rank (number of dominating solutions) and determine the set of solutions $P' \subseteq P$ with the worst rank
2. For each solution $x \in P'$ determine the loss of hypervolume $d(x) = I_H(P') - I_H(P' \setminus \{x\})$ if it is removed from P'
3. Remove the solution with the smallest loss $d(x)$ from the population P (ties are broken randomly)

Step 3 (Termination): If $m \geq N$ then output P and stop; otherwise set $m := m+1$.
Step 4 (Mating): Randomly select elements from P to form a temporary mating pool Q of size λ. Apply variation operators such as recombination and mutation to the mating pool Q which yields Q'. Set $P := P + Q'$ (multi-set union) and continue with Step 2.

To circumvent this costly strategy of taking the best set of μ solutions, most known hypervolume-based evolutionary algorithms use a greedy strategy that is not always optimal: When a population P of $\mu + \lambda$ solutions has to be reduced to μ solutions, the solution s with the smallest hypervolume loss $d(s) := I_H(P) - I_H(P \setminus \{s\})$ is deleted iteratively until the desired size is reached; the hypervolume loss is recalculated every time a solution is deleted. Algorithm 1 shows a general framework of such a hypervolume-based algorithm with greedy strategy in terms of the Simple Indicator Based Evolutionary Algorithm (SIBEA) [Zitzler et al., 2007]. To be even more efficient, a Pareto dominance based ranking is often used before the reduction, such that only pairwisely incomparable solutions are taken into account within P[9]. Note that in general, many other strategies are possible to optimize the hypervolume indicator in an evolutionary algorithm, e.g., the k-greedy strategy of [Zitzler et al., 2008] or the one in HypE [Bader and Zitzler, 2008]. However, practically relevant algorithms such as SMS-EMOA and MO-CMA-ES are using the described greedy strategy which explains the interest of theoretical studies into the greedy approach.

One of the most basic questions for such hypervolume-based algorithms with greedy environmental selection, for which often $\lambda = 1$ is chosen to reduce the runtime further, is whether they are able to find an optimal μ-distribution.

Theorem 4 ([Zitzler et al., 2008])**.** *Hypervolume-based evolutionary algorithms with a greedy environmental selection step as in Algorithm 1 do not guarantee to find an optimal μ-distribution in general.*

Proof. We consider a simple biobjective problem with 4 solutions $x_1, \ldots, x_4$, where $f(x_1) = (1, 6)$, $f(x_2) = (6, 2)$, $f(x_3) = (5, 3)$, $f(x_4) = (7, 1)$ and the reference point is $r = (10, 7)$. By simply computing the hypervolume indi-

n^λ). However, this is still exponential in the number of objectives and might be inapplicable in practice.

[9] Although in SIBEA the dominance rank is used, also non-dominated sorting or other ranking techniques can be used and do not change the theoretical results shown below.

cator values of all possible solution sets with 2 solutions, we prove that a hypervolume-based algorithm as in Algorithm 1 with population size $\mu = 2$ that produces $\lambda = 1$ offspring per generation cannot find the optimal solution set $\{x_3, x_4\}$ when initialized with $\{x_1, x_2\}$: starting from $I_H(\{x_1, x_2\}) = 25$, only solution sets with hypervolume $I_H(\{x_1, x_3\}) = 24$, $I_H(\{x_1, x_4\}) = 24$, $I_H(\{x_2, x_3\}) = 24$, and $I_H(\{x_2, x_4\}) = 24$ are reachable, but not the optimal set with $I_H(\{x_3, x_4\}) = 26$. □

This result can be seen as a major drawback of current hypervolume-based evolutionary algorithms. However, algorithms that use the greedy strategy such as SMS-EMOA and MO-CMA-ES are known to work well for applications in practice which might indicate that the above example is quite artificial and rare. To further investigate how rare such a non-convergence result is, Beume et al. [2009] generalized the example to a continuous one which shows that the non-convergence of the greedy strategy is not only observed on a pathological example of a discrete Pareto front with 4 solutions. Beume et al. [2009] further showed that if the hypervolume indicator function for μ solutions mapping to the Pareto front is concave, the greedy strategy converges. For the case of linear Pareto fronts the hypervolume indicator function itself is proven to be concave such that the greedy strategy always finds an optimal μ-distribution on this type of fronts.

As a last result of hypervolume-based search, we mention the fact that the greedy strategy is not only unable to converge to an optimal μ-distribution in general but that the set of λ deleted solutions can even have a hypervolume indicator value that is arbitrarily far from the best achievable value. For a proof, we refer to the corresponding paper.

Theorem 5 ([Bringmann and Friedrich, 2009a])**.** *For all $\kappa \geq 1$, $k > 3$, $n > k$, and $2 \leq \lambda \leq k$, there is a set of n solutions for which the greedy reduction to $n - \lambda$ solutions deletes a set of λ solutions with a hypervolume contribution that is κ times larger than the contribution of the optimal set.*

5.3 Future Research Directions

Although the recent complexity results for the hypervolume indicator calculation reduced the chances of a fast exact algorithm, further improvements of the current algorithms are desirable in terms of practical applications—using Monte Carlo sampling is a first step towards this goal [Everson et al., 2002, Bader and Zitzler, 2008, Bringmann and Friedrich, 2008, 2009b]. Also the use of specialized data structures to store the hypervolume boxes might decrease the actual runtime of exact algorithms further as it was done in [Bringmann and Friedrich, 2009a].

With respect to optimal μ-distributions, the generalization of the biobjective results to problems with more than 2 objectives might give insights on the differences between problems with a few and many objectives. Although Friedrich et al. [2009] already investigated the connection between optimal μ-distributions and the approximation quality with respect to the multiplicative ε-indicator, the high number of practically relevant quality indicators leaves enough room for further studies. The transfer of the optimal μ-distribution concept to other quality indicators will also help to better understand the influence of quality indicators on the performance assessment. Furthermore, research about convergence

properties and runtime analyses of indicator based evolutionary algorithms in general are needed to argue towards their application in practice.

Last, we would like to mention to investigate the influence of different selection strategies in hypervolume-based algorithms. It is known that theoretically, the often used greedy selection can yield arbitrary bad approximation sets [Bringmann and Friedrich, 2009a] and also the non-convergence results of Zitzler et al. [2008] and Beume et al. [2009] indicate that hypervolume-based algorithms might fail in general. However, further theoretical studies about *when* this happens and whether other strategies, such as HypE [Bader and Zitzler, 2008], can circumvent such a behavior need to be carried out.

6 Runtime Analyses and Convergence Properties

For the sake of completeness, we also want to state the main results achieved in the most active branch of theoretical evolutionary multiobjective optimization: convergence and runtime analyses. Due to the limited space, we refrain from stating too many details but instead give an extensive overview of what has been achieved in the last years.

6.1 Convergence Results

The first theoretical papers in evolutionary multiobjective optimization dealt with the convergence properties of evolutionary algorithms. Convergence of an algorithm A in the multiobjective case is given if the population of A contains the Pareto set with probability 1 if time goes to infinity. In other words, the distance between the population and the minimal elements of the Pareto dominance relation goes to zero if time goes to infinity. However, depending on whether the search space is discrete and finite or continuous, the used distance measures and therefore the mathematical formulation of convergence differs. We refer to the papers mentioned below for details.

Many convergence results of multiobjective evolutionary algorithms are due to Rudolph. He and his co-authors investigated algorithms for which at least one solution is Pareto-optimal after a finite number of steps in finite and discrete search spaces [Rudolph, 1998b, Rudolph and Agapie, 2000] as well as algorithms with a fixed population size on finite search spaces where all solutions in the population converge to the Pareto set [Rudolph, 2001b]. Rudolph also provided convergence results of multiobjective optimizers in continuous domain [Rudolph, 1998a] and investigated noisy and interval-based multiobjective evolutionary algorithms and their convergence [Rudolph, 2001a]. Early results on the convergence of multiobjective evolutionary algorithms are also provided by Hanne [1999]. Here, the focus lies on different selection schemes, the possibilities of temporary fitness deterioration, and problems with unreachable solutions.

Later on, Yan et al. [2003] proved the convergence of a multiobjective evolutionary algorithm which has been specifically developed for solving the traveling salesperson problem. Also differential evolution approaches have been investigated [Xue et al., 2005a,b]. Villalobos-Arias et al. [2005] proved the convergence of several meta-heuristics for multiobjective optimization problems including a general evolutionary algorithm and recently, even the convergence properties

of quantum-inspired algorithms have been investigated [Li et al., 2007]. The latest result is the convergence proof for an algorithm called PDMOEA that approaches the Pareto set both from the feasible and the infeasible region [Hanne, 2007].

All mentioned theoretical studies on the convergence properties of multiobjective algorithms have in common that they use standard probability theory and/or the theory of Markov chains. As the first main result, we can state that one necessary condition for convergence in finite search spaces is—similar to single-objective optimization—the fact that the variation operators can produce any solution with positive probability [Rudolph, 1998b]. A second interesting result states that a simple elitist strategy does not converge to the Pareto set if the population size is fixed to μ and the N non-dominated solutions among the μ parents and the λ offspring are used to fill the next population by randomly deleting non-dominated solutions (if $N > \mu$) or filling the population with randomly selected dominated individuals (if $N < \mu$) [Rudolph, 2001b]. It is worth to mention here that two of the most used algorithms, namely NSGA-II and SPEA2, use a similar type of elitism and are known to diverge in practice if the number of objectives is large. Last, we mention that for continuous search spaces, single-objective results, e.g., regarding the choice of the step-size rule, cannot be fully transferred to the multiobjective case [Rudolph, 1998a].

In contrast to the convergence results mentioned above, Teytaud [2007] investigated the *convergence rate* of multiobjective evolutionary algorithms, i.e., the speed of convergence or in other words the (asymptotic) time until the distance between the population and the Pareto set reaches a certain precision[10]. More precisely, Teytaud showed a general lower bound for the computation time of comparison-based algorithms until a certain precision of the Haussdorff distance is found and an upper bound for a simple random search. The results imply that no comparison-based multiobjective evolutionary algorithm performs better than random search if the number of objectives is large.

6.2 The First Runtime Analyses, Common Algorithms, and Special Proof Techniques

In terms of runtime analyses of multiobjective evolutionary algorithms in discrete domain such as $X = \{0,1\}^n$, the goal is to prove upper and lower bounds on the *expected runtime*, i.e., the expected number of objective function evaluations until the population of an evolutionary algorithm contains the entire Pareto set or an approximation thereof, especially if the Pareto set is too large.

The main difference between runtime analyses of single- and multiobjective evolutionary algorithms is that in single-objective optimization only one search point with the smallest objective function value has to be found whereas in the multiobjective case almost always a set of solutions is sought. The investigated algorithms for evolutionary multiobjective optimization are therefore population-based. The simplest evolutionary algorithms, comprising their single-objective counterparts randomized local search (RLS) and (1+1)EA, are the local and the global version of the Simple Evolutionary Multiobjective Optimizer. The local version is, according to the literature, denoted by SEMO

[10] In general, any distance between the population and the Pareto set or the Pareto front can be considered. In his work, Teytaud actually used the Haussdorff distance in objective space, i.e., a distance measure with respect to the Pareto front.

Algorithm 2 Simple Evolutionary Multiobjective Optimizer (SEMO, [Laumanns et al., 2002b])

```
Choose an initial individual x uniformly at random from X = {0,1}^n
P ← {x}
loop
  Select one element x uniformly at random from P
  Create offspring x' by flipping a randomly chosen bit of x
  P ← P \ {z ∈ P | x' ≺ z}
  if ∄z ∈ P such that (z ≺ x' ∨ f(z) = f(x')) then
    P ← P ∪ {x'}
  end if
end loop
```

and is detailed for minimization in Algorithm 2 whereas Global SEMO refers to the global version. After sampling a first solution uniformly at random in the search space X, in each iteration of the algorithms, a parent x is drawn uniformly at random from the population P and an offspring x' is generated from x by either flipping one randomly chosen bit (SEMO) or each bit with probability $1/n$ (Global SEMO). If x' is not weakly dominated by any other solution in P, the new solution x' is added to P and all solutions dominated by x' are deleted from P. This guarantees that the algorithms converge with probability 1 to a set covering the entire Pareto front which, at the same time, implies that the population size is not bounded by a constant as in most of the algorithms used in practice.

The algorithm SEMO was introduced by Laumanns et al. [2002b] to perform the first runtime analysis of a multiobjective evolutionary algorithm on a simple test problem called LeadingOnesTrailingZeros (LOTZ)[11]. Here, as well as in several other studies, the analyses of the algorithm is divided into two parts—the first phase covers the time until the first Pareto-optimal point is found and the second phase stops if the entire Pareto set is found.

Theorem 6 ([Laumanns et al., 2002b])**.** *Let the biobjective maximization problem* $\text{LOTZ} : \{0,1\}^n \to \mathbb{N}^2$ *be defined as*

$$\text{LOTZ}(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} \prod_{j=1}^{i} x_j, \sum_{i=1}^{n} \prod_{j=i}^{n} (1 - x_j) \right) \quad .$$

The expected runtime until SEMO finds the entire Pareto set for LOTZ *is bounded by* $\Theta(n^3)$*.*

Proof. First of all, we investigate the problem itself in more detail; see also Fig. 3. The decision space can be partitioned into $n+1$ sets $S_0 \ldots, S_n$ with $S_i = \{x \in X \mid f_1(x) + f_2(x) = i\}$, i.e, S_i contains all decision vectors of the form $1^a 0 *^{(n-i-2)} 10^b$ with $a + b = i$ for $i < n$ and $1^a 0^b$ with $a + b = n$ for the set S_n which coincides with the Pareto set of size $n+1$. The set S_{n-1} is empty. All solutions within a set S_i are incomparable whereas a solution in set S_i dominates a solution in S_j iff $i > j$.

[11] Although in the remainder of the chapter we consider minimization problems only, the following result is stated in its original version of maximization.

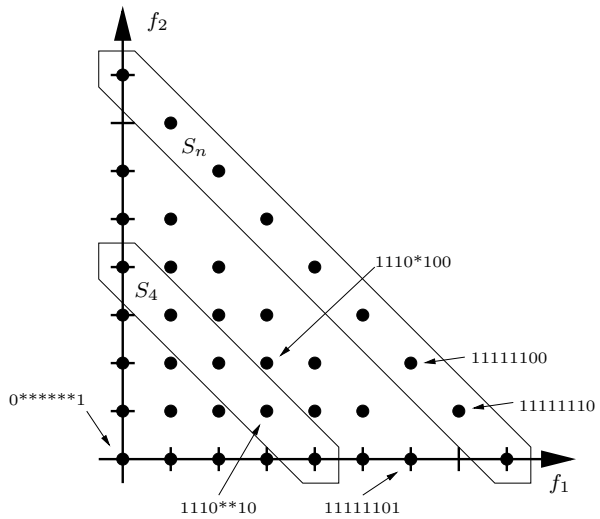

Figure 3: Objective space and corresponding decision vectors for the problem LOTZ with $n = 8$ bits; a '$*$' can be either a one or a zero. Exemplary, the sets S_4 and S_n are shown.

For the runtime analyses, we first consider the time until the first Pareto-optimal point is found. During this first phase, the population will consist of one individual only since the mutation of a single bit either produces a dominating or a dominated solution in another set S_i. Therefore, a new solution x' that is added to the population of SEMO will replace its parent x iff $x \in S_i$ and $x' \in S_j$ with $i < j$. As there is always a one-bit mutation with probability $1/n$ from the current solution in S_i to one in a higher set S_j, the waiting time to leave S_i is $\mathcal{O}(n)$. The overall time until S_n is reached is therefore $\mathcal{O}(n^2)$ since there are at most $\mathcal{O}(n)$ such steps necessary.

Once a Pareto-optimal solution is found, we can derive a lower and an upper bound for the time until the entire Pareto set is found. We observe that the set of already found Pareto-optimal solutions is always contiguous (on the Pareto front) and that we can generate a new Pareto-optimal point by sampling a solution the objective vector of which is one of the outmost of the population and flipping one single bit. The probability to sample an outmost point is at least $1/i$ and at most $2/i$ if i Pareto-optimal solutions are found and the probability to sample a new Pareto-optimal point from there is at least $1/n$ and at most $2/n$. By summing up the expected runtimes $E(T_i)$ for finding the $(i+1)$th Pareto-optimal solution which is at least $i/2 \cdot n/2 = in/4$ and at most in, we bound the expected waiting time $E(T)$ until the entire Pareto set is found by $1/8n^3 - 1/8n^2 \leq \sum_{i=1}^{n-1} ni/4 \leq E(T) = \sum_{i=1}^{n-1} E(T_i) \leq \sum_{i=1}^{n-1} ni \leq 1/2n^3 - 1/2n^2$. □

The proof techniques in this very first proof on the runtime of multiobjective evolutionary algorithms are seemingly the same than in runtime analyses of single-objective algorithms. The same holds for the first analyses of Global SEMO on LOTZ by Giel [2003] and the runtime analysis of SEMO on another simple biobjective test problem called multiobjective counting ones

Algorithm 3 Fair Evolutionary Multiobjective Optimizer (FEMO, [Laumanns et al., 2002b])

```
Choose an initial individual x uniformly at random in X = {0,1}^n
w(x) ← 0                                          {Initialize offspring count}
P ← {x}
loop
  Select parent x uniformly at random in {y ∈ P | ∀z ∈ P : w(y) ≤ w(z)}
  w(x) ← w(x) + 1                                 {Increment offspring count}
  Create offspring x' by flipping a randomly chosen bit of x
  P ← P \ {z ∈ P | x' ≺ z}
  if ∄z ∈ P such that (z ≺ x' ∨ f(z) = f(x')) then
    P ← P ∪ {x'}
    w(x') ← 0                                     {Initialize offspring count}
  end if
end loop
```

(MOCO, [Thierens, 2003]). Also in the first study comparing single-objective approaches and SEMO by Giel and Lehre [2006], common proof techniques of single-objective runtime analyses have been used to show that a real multi-objective approach like SEMO can improve the runtime drastically compared to single-objective evolutionary algorithms. However, there are a few studies that provide new proof techniques which are tailored towards multiobjective problems and which we will describe briefly in the following.

Two proof techniques have been proposed by Laumanns et al. [2004a] when SEMO and two other simple algorithms called Fair Evolutionary Multiobjective Optimizer (FEMO, [Laumanns et al., 2002b]) and Greedy Evolutionary Multiobjective Optimizer (GEMO) have been analyzed on the test problems LOTZ, Counting Ones Counting Zeros (COCZ), and their variants mLOTZ and mCOCZ with more than two objectives. FEMO, detailed in Algorithm 3, comprises the same operations as in SEMO with an additional mechanism that allows only solutions to be selected for mating that have been used most infrequently as parents. To this end, the algorithm carries a count along with every solution indicating how often it was selected as parent. Newly generated solutions that are not contained in the current population get a count of 0. GEMO is a specialized version of FEMO aiming at faster progress towards the Pareto front and is not discussed further here.

The first proof technique, mentioned in Laumanns et al. [2004a] can be seen as a generalization of the fitness-based partitioning approach used frequently in single-objective runtime analyses, see [Wegener, 2003]. The idea is to partition the decision space into l disjoint sets $S_1, \ldots, S_l$ where S_l coincides with the Pareto set. If for arbitrary two sets S_i and S_j with $i, j \neq l$ all solutions in S_j dominate all solutions in S_i (written $S_j \unlhd S_i$), a lower bound on the probability to leave S_i gives an upper bound on the expected number of times mutations are performed on non-Pareto-optimal solutions.

Lemma 1 (Decision Space Partitioning [Laumanns et al., 2004a])**.** *Let $S_1, \ldots, S_l \subset X$ be a partitioning of the decision space, i.e., $\bigcup_{i=1}^{l} S_i = X$, where S_l is the Pareto set. Let $\unlhd$ be defined on the partition as $S_i \unlhd S_j$ iff $\forall (x, y) \in S_i \times S_j : x \prec y$. Let then $d(S_i) = \{S_j \mid S_j \unlhd S_i\}$ contain all sets that dominate S_i with respect*

to $\trianglelefteq$. Moreover, we assume an algorithm that iteratively modifies a population P by mutation and selection operations with the following properties.

(1) The probability that the mutation operator creates a solution $x' \in d(S_i)$ for any $x \in S_i$ can be bounded below for all $1 \leq i < l$ by a constant $p(S_i) > 0$, i.e., $0 < p(X_i) \leq \min_{x \in S_i}\{Prob\{x' \in d(S_i) | x \in S_i\}\}$.

(2) A newly generated solution will enter the population P iff it is not dominated by another solution in P.

(3) A solution is deleted from P iff a dominating solution is entering P.

Then, the expected number of times, the mutation operator of the above described algorithm is applied to non-Pareto-optimal solutions is bounded above by $\sum_{i=1}^{l-1} p(X_l)^{-1}$.

The decision space partitioning approach can mainly be used to bound the time until the first Pareto-optimal solution is found as for example in [Laumanns et al., 2004a,b]. To prove bounds on the time until all Pareto-optimal solutions are found if one of them is known already, the following technique of general graph search [Laumanns et al., 2004b] is helpful[12].

Given a fully connected weighted graph $G = (V, E)$ with nodes V, edges $E = V \times V$, and weights $w : E \to \mathbb{R}$ for the edges, we identify the nodes with the Pareto-optimal solutions and the edges with mutations between them. The weights of the edges correspond to the mutation probabilities of the associated mutations. Instead of analyzing the full optimization process on the entire search space, the general graph search method restricts the analysis to the mutations between the Pareto-optimal solutions by analyzing a simplified graph search algorithm that starts by visiting a randomly chosen node, marks all visited nodes, and iteratively jumps from a—according to the evolutionary algorithm's mating selection operator—selected marked node to a neighbor that is chosen uniformly at random until all nodes are visited.

Instead of investigating *all* mutation probabilities between Pareto-optimal solutions to do the runtime analysis, the following General Graph Search Lemma allows to bound the runtime until all Pareto-optimal solutions are found even if the graph's weights, i.e., the mutation probabilities, are known only on a *spanning tree* on G.

Lemma 2 (General Graph Search [Laumanns et al., 2004b])**.** *If the edge weights of a spanning tree on $G = (V, E)$ can be lower bounded by p then the above described graph search algorithm has found all nodes and edges of G after $(c+1)\frac{|V|}{p} \ln |V|$ jumps with probability at least $1 - |V|^{-c}$ and the expected number of jumps is bounded by $\mathcal{O}(\frac{|V|}{p} \log |V|)$.*

The general graph search approach is nicely applied to the runtime analyses of FEMO with a global mutation operator on a 0-1 knapsack problem in [Laumanns et al., 2004b]. There, the considered spanning tree on the Pareto set of size $\Theta(n^2)$ is constructed by connecting search points with Hamming distance 2 which bounds the mutation probability by $\Omega(1/n^2)$ and results in a runtime

[12] A specialized version of the method has been proposed before by Laumanns et al. [2004a] which is only applicable in case of local mutations and, thus, not presented here.

bound of $\mathcal{O}(n^4 \log n)$. Note, that the way a marked node is selected by the graph search algorithm depends on the investigated evolutionary algorithm.

A further proof technique that is tailored towards the analysis of the diversity maintaining evolutionary multiobjective optimizer (DEMO) has been proposed by Horoba [2009]. Similar to the fitness-based partitions method, Horoba [2009] proved a theorem that upper bounds the optimization time of DEMO until it finds an approximation of the Pareto set if three assumptions on the population and the mutation operator are fulfilled. Due to space limitations, we refer the interested reader to the corresponding paper [Horoba, 2009] for details.

6.3 Other Runtime Analysis Results

While the first runtime analyses of multiobjective evolutionary algorithms only investigated simple test functions, several practically relevant combinatorial optimization problems have been investigated in terms of runtime analyses or the algorithms' ability to find good approximations of the Pareto front. Also other aspects of evolutionary multiobjective optimization, such as multiobjectivization and hypervolume-based search, have been investigated by means of rigorous runtime analyses. Due to space limitations, we only mention the studies and their results briefly here.

Combinatorial Optimization Problems Similar to the development in single-objective optimization, most of the recent studies on the runtime analyses of evolutionary multiobjective optimization investigate combinatorial optimization problems. Table 1 presents an overview of the main results, the used algorithms and the used proof techniques for all publications on multiobjective runtime analyses that can be found in the literature[13].

Multiobjectivization Multiobjectivization, i.e., the addition of objectives to a problem or the decomposition of a single objective into two or more objectives, has been originally proposed from a practical point of view to speed-up the performance of evolutionary multiobjective algorithms by providing additional information that might guide the search towards the Pareto set. However, several studies questioned this approach and argued that increasing the number of objectives will in general increase the difficulty of the problem for evolutionary multiobjective algorithms due to the increased number of incomparable solution pairs. The recent runtime analyses by Brockhoff et al. [2007, 2009] for the addition of objectives and by Handl et al. [2008] for the decomposition of objectives show that multiobjectivization can provable reduce the runtime of Global SEMO—depending on the objective function that is added or how the decomposition is performed. Furthermore, Brockhoff et al. [2007, 2009] showed that two equally difficult single-objective problems can be solved faster simultaneously when combined to a biobjective problem.

Hypervolume-Based Search Although many algorithms aiming at optimizing the hypervolume indicator are known in practice, only one study analyzed

[13] For a description of the proof techniques that are not described in this chapter, we refer to [Wegener, 2003] and [Neumann and Wegener, 2007]; a description of the algorithms as well as the problem definitions can always be found in the corresponding papers.

Table 1: Overview of runtime analyses for combinatorial multiobjective optimization problems in the literature. The abbreviations for the proof techniques are coupon collector theorem (CCT), variants of decision space partitioning (DSP), expected multiplicative weight decrease (EMWD), general graph search (GGS), potential functions (PF), and typical runs (TR).

problem	publications	main result	proof techniques
Knapsack problems (KP)	[Laumanns et al., 2004b]	runtime of Global SEMO and Global FEMO on a biobjective KP (max. two specific profit functions)	DSP,TR, GGS
	[Kumar and Banerjee, 2005, 2006]	runtime of REMO on LOTZ and a quadratic function; runtime of REMO until $(1+\varepsilon)$-approximation for 0-1-KP (max. profit, min. weight) is found	DSP,TR
Minimum spanning trees (MST)	[Neumann, 2004, 2007]	runtime of Global SEMO on two multiobjective formulations of MST	DSP
	[Neumann and Wegener, 2006]	multiobjective formulation results in lower runtime for SEMO and Global SEMO than for RLS/(1+1)EA	DSP,PF
	[Neumann and Wegener, 2008]	book chapter with results of [Neumann and Wegener, 2006] and [Scharnow et al., 2004]	—
Cutting problems	[Neumann et al., 2008]	runtime of RLS and (1+1)EA on minimum cut problem; runtime of Global SEMO and DEMO on a biobjective version (min. costs and flow)	TR,EMWD
	[Neumann and Reichel, 2008]	runtime until Global SEMO and DEMO find an approximation of the front for a multiobjective formulation of a minimum multicut problem	DSP,PF, EMWD
Shortest path problems (SP)	[Scharnow et al., 2002, 2004]	runtime of a specialized (1+1)EA on multiobjectivized version of single source SP with $n-1$ objectives	DSP
	[Horoba, 2009]	DEMO is FPRAS for single source multiobjective SP with k weights per edge; new fitness based partition approach for ≥ 2 objectives	DSP
covering problems and plateau functions	[Friedrich et al., 2007a]	runtime and approximation results for RLS and (1+1)EA with lexicographic ordering and SEMO and Global SEMO on a biobjective vertex cover problem (min. number of uncovered vertices and ones in bitstring) and on a biobjective set cover problem (min. cost and number of uncovered elements)	CCT,PF, TR,EMWD
	[Friedrich et al., 2007b]	runtime of (1+1)EA with lexicographic order is polynomial and for Global SEMO exponential on a plateau function and on a biobjective SETCOVER instance (min. number u of uncovered elements and sum of cost and u)	TR,DSP, EMWD
	[Kratsch and Neumann, 2009]	parameterized complexity results of Global SEMO on two biobjective formulations of vertex cover	DSP,PF
Graph bisectioning problem	[Greiner, 2009]	runtime analyses on two instances of the graph bisectioning problem where either (1+1)EA or Global SEMO (with the number of free nodes as second objective) is faster	DSP,PF, TR

the runtime of SIBEA on two simple test problems so far [Brockhoff et al., 2008]. The results show[14] that optimizing the hypervolume indicator by always generating $\lambda = 1$ new offspring per generation might be sufficient for some problems although the general convergence of such algorithms cannot be guaranteed, cf. Sec. 5.2.2.

EMO Operators and Mechanisms The impact of single operators or of a specific concept of EMO algorithms on the runtime has been the focus of other recent studies where different algorithms that only differ in a single operator or mechanism have been analyzed theoretically. Horoba and Neumann [2008, 2009], for example, compared the diversity mechanisms in Global SEMO, in an algorithm called Global DEMO$_\varepsilon$ (Diversity Evolutionary Multi-objective Optimizer) and in a simplified version of SPEA2. The rigorous analyses of the time until an ε-approximation of the Pareto front is found indicate that there is at least one problem for which every investigated diversity mechanism fails to find a good approximation in polynomial time whereas the other two algorithms are fast. In another study, Friedrich et al. [2008] investigated the concept of fairness in the same manner: the analyses of two different versions of FEMO on two test functions show that each of the algorithms can outperform the other with a runtime gap that is exponential in the bitstring length.

6.4 Future Research Directions

As we have pointed out, various convergence studies and runtime analyses of multiobjective evolutionary algorithms have been performed in recent years. However, the theoretical understanding of why and when multiobjective evolutionary algorithms are efficient and how general parameters should be chosen lies way behind the knowledge in current research on single-objective algorithms. The first approximation results on hard problems and the comparison of several operators like in [Horoba and Neumann, 2009] are first steps towards a better understanding of the algorithms' principles and work in this direction should be continued, e.g., by deriving approximation results for other combinatorial problems or by comparing algorithms with different selection schemes. Hypervolume-based search is another area where theoretical studies are needed. Further studies about when hypervolume-based algorithms are beneficial and when obstructive are needed as well as detailed analyses about how basic parameters such as the number of offsprings should be chosen or how other selection schemes, such as the one in HypE [Bader and Zitzler, 2008], are working. Moreover, other indicator-based algorithms could be analyzed in the near future which might result in a better understanding of the underlying principles of quality indicators in general.

7 Other Areas of Interest

Although the title of this report gives the impression that it covers all theoretical aspects of evolutionary multiobjective optimization, the limited space does not allow for an exhaustive discussion of all theoretical studies in this area.

[14] Note that there is some ongoing discussion among the authors whether the proof for the large Pareto front is fully correct.

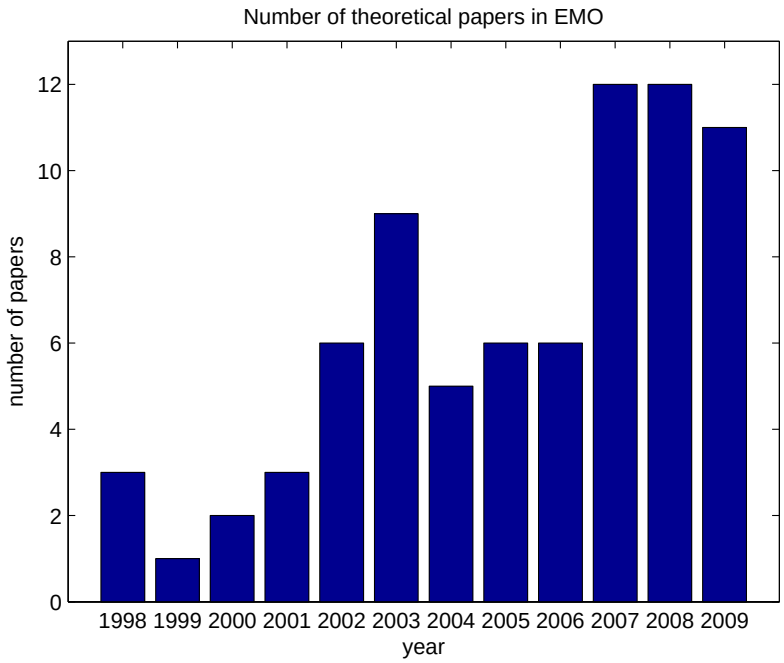

Figure 4: Histogram of theoretical papers on EMO cited in this review.

Besides the mentioned topics, theoretical research has also focused on archiving [Laumanns et al., 2002a, Knowles and Corne, 2003, 2004, Laumanns, 2007, Schütze et al., 2007a,b, 2008a,b], resulted in runtime improvements for some operators [Jensen, 2003], transferred no free lunch results to the multiobjective case [Corne and Knowles, 2003], or investigated the connections to relational algebra [Diedrich et al., 2008].

8 Summary

The number of theoretical studies about evolutionary multiobjective optimization has been growing quickly in the recent years. Where in the exhaustive overview of publications about evolutionary multiobjective optimization by Coello Coello et al. [2007], only 20 publications on theoretical aspects of evolutionary multiobjective algorithms are cited, both in 2008 and in 2009, more than 10 papers with theoretical investigations in the field of evolutionary multiobjective optimization can be reported, see Fig. 4. This chapter presents an extensive overview over these developments in the theory of evolutionary multiobjective optimization where the focus lies on quality indicators, hypervolume-based search, and runtime analyses. The detailed list of references and the identification of interesting open research questions makes it a good starting point to further advance our fundamental understanding of the underlying principles of evolutionary multiobjective optimization.

References

A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Theory of the Hypervolume Indicator: Optimal μ-Distributions and the Choice of the Reference Point. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 87–102. ACM, 2009a.

A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Articulating User Preferences in Many-Objective Problems by Sampling the Weighted Hypervolume. In G. Raidl et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 555–562. ACM, 2009b.

A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Investigating and Exploiting the Bias of the Weighted Hypervolume to Articulate User Preferences. In G. Raidl et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 563–570. ACM, 2009c.

J. Bader and E. Zitzler. HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. TIK Report 286, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, November 2008.

J. Bader and E. Zitzler. A Hypervolume-Based Multiobjective Optimizer for High-Dimensional Objective Spaces. In *Conference on Multiple Objective and Goal Programming (MOPGP 2008)*, Lecture Notes in Economics and Mathematical Systems. Springer, 2009. to appear.

N. Beume and G. Rudolph. Faster S-Metric Calculation by Considering Dominated Hypervolume as Klee's Measure Problem. Technical Report CI-216/06, Sonderforschungsbereich 531 Computational Intelligence, Universität Dortmund, 2006. a shorter version has been published at IASTED International Conference on Computational Intelligence (CI 2006).

N. Beume, C. M. Fonseca, M. Lopez-Ibanez, L. Paquete, and J. Vahrenhold. On the Complexity of Computing the Hypervolume Indicator. Technical Report CI-235/07, University of Dortmund, December 2007a.

N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective Selection Based on Dominated Hypervolume. *European Journal on Operational Research*, 181:1653–1669, 2007b.

N. Beume, B. Naujoks, M. Preuss, G. Rudolph, and T. Wagner. Effects of 1-Greedy $\mathcal{S}$-Metric-Selection on Innumerably Large Pareto Fronts. In M. Ehrgott et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2009)*, volume 5467 of *LNCS*, pages 21–35. Springer, 2009.

K. Bringmann and T. Friedrich. Approximating the Volume of Unions and Intersections of High-Dimensional Geometric Objects. In S. H. Hong, H. Nagamochi, and T. Fukunaga, editors, *International Symposium on Algorithms and Computation (ISAAC 2008)*, volume 5369 of *LNCS*, pages 436–447. Springer, 2008.

K. Bringmann and T. Friedrich. Don't Be Greedy When Calculating Hypervolume Contributions. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 103–112. ACM, 2009a.

K. Bringmann and T. Friedrich. Approximating the Least Hypervolume Contributor: NP-hard in General, But Fast in Practice. In M. Ehrgott et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2009)*, volume 5467 of *LNCS*, pages 6–20. Springer, 2009b.

D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler. Do Additional Objectives Make a Problem Harder? In D. Thierens et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2007)*, pages 765–772. ACM, 2007.

D. Brockhoff, T. Friedrich, and F. Neumann. Analyzing Hypervolume Indicator Based Algorithms. In G. Rudolph et al., editors, *Conference on Parallel Problem Solving From Nature (PPSN X)*, volume 5199 of *LNCS*, pages 651–660. Springer, 2008.

D. Brockhoff, T. Friedrich, N. Hebbinghaus, C. Klein, F. Neumann, and E. Zitzler. On the Effects of Adding Objectives to Plateau Functions. *IEEE Transactions on Evolutionary Computation*, 13(3):591–603, 2009.

C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen. *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer, 2007.

D. Corne and J. Knowles. No Free Lunch and Free Leftovers Theorems for Multiobjective Optimisation Problems. In C. M. Fonseca et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 327–341. Springer, 2003.

K. Deb. *Multi-Objective Optimization Using Evolutionary Algorithms*. Wiley, Chichester, UK, 2001.

K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

F. Diedrich, B. Kehden, and F. Neumann. Multi-objective Problems in Terms of Relational Algebra. In *Conference on Relational Methods in Computer Science*, volume 4988 of *LNCS*, pages 84–98. Springer, 2008.

M. Ehrgott. *Multicriteria Optimization*. Springer, 2nd edition, 2005.

M. Emmerich, A. Deutz, and N. Beume. Gradient-Based/Evolutionary Relay Hybrid for Computing Pareto Front Approximations Maximizing the S-Metric. In *Hybrid Metaheuristics*, pages 140–156. Springer, 2007.

R. Everson, J. Fieldsend, and S. Singh. Full Elite-Sets for Multiobjective Optimisation. In I.C. Parmee, editor, *Conference on adaptive computing in design and manufacture (ADCM 2002)*, pages 343–354. Springer, 2002.

A. Farhang-Mehr and S. Azarm. Minimal Sets of Quality Metrics. In C. M. Fonseca et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 405–417. Springer, 2003.

M. Fleischer. The measure of Pareto optima. Applications to multi-objective metaheuristics. In C. M. Fonseca et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 519–533. Springer, 2003.

C. M. Fonseca, V. Grunert da Fonseca, and L. Paquete. Exploring the Performance of Stochastic Multiobjective Optimisers with the Second-Order Attainment Function. In C. A. Coello Coello et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2005)*, pages 250–264. Springer, 2005.

C. M. Fonseca, L. Paquete, and M. López-Ibáñez. An Improved Dimension-Sweep Algorithm for the Hypervolume Indicator. In *Congress on Evolutionary Computation (CEC 2006)*, pages 1157–1163. IEEE Press, 2006.

Carlos M. Fonseca and Peter J. Fleming. On the Performance Assessment and Comparison of Stochastic Multiobjective Optimizers. In *Parallel Problem Solving from Nature (PPSN IV)*, pages 584–593. Springer, 1996.

T. Friedrich, J. He, N. Hebbinghaus, F. Neumann, and C. Witt. Approximating Covering Problems by Randomized Search Heuristics Using Multi-Objective Models. In *Genetic and Evolutionary Computation Conference (GECCO 2007)*, pages 797–804. ACM, 2007a.

T. Friedrich, N. Hebbinghaus, and F. Neumann. Plateaus Can Be Harder in Multi-objective Optimization. In *Congress on Evolutionary Computation (CEC 2007)*, pages 2622–2629. IEEE Press, 2007b.

T. Friedrich, C. Horoba, and F. Neumann. Runtime Analyses for Using Fairness in Evolutionary Multi-Objective Optimization. In G. Rudolph et al., editors, *Conference on Parallel Problem Solving from Nature (PPSN X)*, volume 5199 of *LNCS*, pages 671–680. Springer, 2008.

T. Friedrich, C. Horoba, and F. Neumann. Multiplicative Approximations and the Hypervolume Indicator. In G. Raidl et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 571–578. ACM, 2009.

O. Giel. Expected Runtimes of a Simple Multi-objective Evolutionary Algorithm. In *Congress on Evolutionary Computation (CEC 2003)*, IEEE Press, pages 1918–1925, 2003.

O. Giel and P. K. Lehre. On The Effect of Populations in Evolutionary Multi-objective Optimization. In *Genetic and Evolutionary Computation Conference (GECCO 2006)*, pages 651–658. ACM, 2006.

G. Greiner. Single- and Multi-Objective Evolutionary Algorithms for Graph Bisectioning. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 29–38. ACM, 2009.

V. Grunert da Fonseca, C. M. Fonseca, and A. O. Hall. Inferential Performance Assessment of Stochastic Optimisers and the Attainment Function. In E. Zitzler et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2001)*, volume 1993 of *LNCS*, pages 213–225. Springer, 2001.

J. Handl, S. C. Lovell, and J. Knowles. Multiobjectivization by Decomposition of Scalar Cost Functions. In G. Rudolph et al., editors, *Conference on Parallel Problem Solving From Nature (PPSN X)*, volume 5199 of *LNCS*, pages 31–40. Springer, 2008.

T. Hanne. On the convergence of multiobjective evolutionary algorithms. *European Journal Of Operational Research*, 117(3):553–564, 1999.

T. Hanne. A Primal-Dual Multiobjective Evolutionary Algorithm for Approximating the Efficient Set. In *Congress on Evolutionary Computation (CEC 2007)*, pages 3127–3134. IEEE Press, 2007.

M. P. Hansen and A. Jaszkiewicz. Evaluating the quality of approximations of the non-dominated set. IMM Technical Report IMM-REP-1998-7, Institute of Mathematical Modeling, Technical Universit of Denmark, 1998.

N. Hansen, A. Auger, S. Finck, and R. Ros. Real-Parameter Black-Box Optimization Benchmarking 2009: Experimental Setup. INRIA Research Report RR-6828, INRIA Saclay—Ile-de-France, May 2009.

M. I. Henig. Existence and Characterization of Efficient Decision With Respect to Cones. *Mathematical Programming*, 23:111–116, 1982.

C. Horoba. Analysis of a Simple Evolutionary Algorithm for the Multiobjective Shortest Path Problem. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 113–120. ACM, 2009.

C. Horoba and F. Neumann. Benefits and Drawbacks for the Use of Epsilon-dominance in Evolutionary Multi-objective Optimization. In *Conference on Genetic and Evolutionary Computation (GECCO 2008)*, pages 641–648. ACM, 2008.

C. Horoba and F. Neumann. Additive Approximations of Pareto-Optimal Sets by Evolutionary Multi-Objective Algorithms. In *Foundations of Genetic Algorithms (FOGA 2009)*, pages 79–86. ACM, 2009.

C. Igel, N. Hansen, and S. Roth. Covariance Matrix Adaptation for Multi-objective Optimization. *Evolutionary Computation*, 15(1):1–28, 2007.

A. Jaszkiewicz and J. Branke. Interactive Multiobjective Evolutionary Algorithms. In J. Branke, K. Deb, K. Miettinen, and R. Slowinski, editors, *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 179–193. Springer, 2008.

M. T. Jensen. Reducing the Run-Time Complexity of Multiobjective EAs: The NSGA-II and Other Algorithms. *IEEE Transactions on Evolutionary Computation*, 7(5):503–515, 2003.

V. Klee. Can the Measure of $\bigcup_i^n \{a_i, b_i\}$ be Computed in Less Than $O(n \log n)$ Steps? *The American Mathematical Monthly*, 84(4):284–285, 1977.

J. Knowles and D. Corne. On Metrics for Comparing Non-Dominated Sets. In *Congress on Evolutionary Computation (CEC 2002)*, pages 711–716. IEEE Press, 2002.

J. Knowles and D. Corne. Properties of an Adaptive Archiving Algorithm for Storing Nondominated Vectors. *IEEE Transactions on Evolutionary Computation*, 7(2):100–116, 2003.

J. Knowles and D. Corne. Bounded Pareto Archiving: Theory and Practice. In X. Gandibleux et al., editors, *Metaheuristics for Multiobjective Optimisation*, volume 535 of *Lecture Notes in Economics and Mathematical Systems*, pages 39–64. Springer, 2004.

J. D. Knowles. *Local-Search and Hybrid Evolutionary Algorithms for Pareto Optimization*. PhD thesis, University of Reading, 2002.

S. Kratsch and F. Neumann. Fixed-Parameter Evolutionary Algorithms and the Vertex Cover Problem. In G. Raidl et al., editors, *Genetic and Evolutionary Computation Conference (GECCO 2009)*, pages 293–300. ACM, 2009.

R. Kumar and N. Banerjee. Running Time Analysis of a Multiobjective Evolutionary Algorithm on Simple and Hard Problems. In A. H. Wright, M. D. Vose, K. A. De Jong, and L. M. Schmitt, editors, *Foundations of Genetic Algorithms (FOGA 2005)*, volume 3469 of *LNCS*, pages 112–131. Springer, 2005.

R. Kumar and N. Banerjee. Analysis of a Multiobjective Evolutionary Algorithm on the 0-1 Knapsack Problem. *Theoretical Computer Science*, 358(1):104–120, 2006.

M. Laumanns. Stochastic Convergence of Random Search to Fixed Size Pareto Set Approximations. arXiv.org:0711.2949, 2007.

M. Laumanns, L. Thiele, K. Deb, and E. Zitzler. Combining Convergence and Diversity in Evolutionary Multiobjective Optimization. *Evolutionary Computation*, 10(3):263–282, 2002a.

M. Laumanns, L. Thiele, E. Zitzler, E. Welzl, and K. Deb. Running Time Analysis of Multi-Objective Evolutionary Algorithms on a Simple Discrete Optimization Problem. In *Conference on Parallel Problem Solving From Nature (PPSN VII)*, volume 2439 of *LNCS*, pages 44–53. Springer, 2002b.

M. Laumanns, L. Thiele, and E. Zitzler. Running Time Analysis of Multiobjective Evolutionary Algorithms on Pseudo-Boolean Functions. *IEEE Transactions on Evolutionary Computation*, 8(2):170–182, 2004a.

M. Laumanns, L. Thiele, and E. Zitzler. Running Time Analysis of Evolutionary Algorithms on a Simplified Multiobjective Knapsack Problem. *Natural Computing*, 3(1):37–51, 2004b.

Z. Li, Z. Li, and G. Rudolph. On the Convergence Properties of Quantum-Inspired Multi-Objective Evolutionary Algorithms. In D.-S. Huang, L. Heutte, and M. Loog, editors, *Conference on Intelligent Computing (ICIC 2007)*, pages 245–255. Springer, 2007.

K. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer, 1999.

F. Neumann. Expected Runtimes of a Simple Evolutionary Algorithm for the Multi-Objective Minimum Spanning Tree Problem. In X. Yao et al., editors, *Conference on Parallel Problem Solving from Nature (PPSN VIII)*, volume 3242 of *LNCS*, pages 80–89. Springer, 2004.

F. Neumann. Expected runtimes of a simple evolutionary algorithm for the multi-objective minimum spanning tree problem. *European Journal of Operational Research*, 181(3):1620–1629, 2007.

F. Neumann and J. Reichel. Approximating Minimum Multicuts by Evolutionary Multi-objective Algorithms. In G. Rudolph et al., editors, *Conference on Parallel Problem Solving From Nature (PPSN X)*, volume 5199 of *LNCS*, pages 72–81. Springer, 2008.

F. Neumann and I. Wegener. Minimum Spanning Trees Made Easier Via Multi-Objective Optimization. *Natural Computing*, 5(3):305–319, 2006.

F. Neumann and I. Wegener. Randomized local search, evolutionary algorithms, and the minimum spanning tree problem. *Theoretical Computer Science*, 378 (1):32–40, 2007.

F. Neumann and I. Wegener. Can Single-Objective Optimization Profit from Multiobjective Optimization? In J. Knowles, D. Corne, and K. Deb, editors, *Multiobjective Problem Solving from Nature: From Concepts to Applications*, pages 115–130. Springer, 2008.

F. Neumann, J. Reichel, and M. Skutella. Computing Minimum Cuts by Randomized Search Heuristics. In *Genetic and Evolutionary Computation Conference (GECCO 2008)*, pages 779–786. ACM, 2008.

M. H. Overmars and C. K. Yap. New Upper Bounds in Klee's Measure Problem. *SIAM Journal on Computing*, 20(6):1034–1045, 1991.

G. Rudolph. On A Multi-objective Evolutionary Algorithm and Its Convergence To The Pareto Set. In *Proceedings of the IEEE International Conference on Evolutionary Computation*, pages 511–516. IEEE Press, 1998a.

G. Rudolph. Evolutionary Search for Minimal Elements in Partially Ordered Finite Sets. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, editors, *Conference on Evolutionary Programming VII*, volume 1447 of *LNCS*, pages 345–353. Springer, 1998b.

G. Rudolph. Some Theoretical Properties of Evolutionary Algorithms under Partially Ordered Fitness Values. In *Evolutionary Algorithms Workshop (EAW-2001)*, pages 9–22, 2001a.

G. Rudolph. Evolutionary Search under Partially Ordered Fitness Sets. In M. F. Sebaaly et al., editors, *International NAISO Congress on Information Science Innovations (ISI 2001)*, pages 818–822. ICSC Academic Press, 2001b.

G. Rudolph and A. Agapie. Convergence Properties of Some Multi-Objective Evolutionary Algorithms. In A. Zalzala and R. Eberhart, editors, *Congress on Evolutionary Computation (CEC 2000)*, volume 2, pages 1010–1016. IEEE Press, 2000.

J. D. Schaffer. Multiple Objective Optimization with Vector Evaluated Genetic Algorithms. In John J. Grefenstette, editor, *Conference on Genetic Algorithms and Their Applications*, pages 93–100, 1985.

J. Scharnow, K. Tinnefeld, and I. Wegener. Fitness Landscapes Based on Sorting and Shortest Path Problems. In *Conference on Parallel Problem Solving from Nature (PPSN VII)*, volume 2439 of *LNCS*, pages 54–63. Springer, 2002.

J. Scharnow, K. Tinnefeld, and I. Wegener. The Analysis of Evolutionary Algorithms on Sorting and Shortest Paths Problems. *Journal of Mathematical Modelling and Algorithms*, 3(4):349–366, 2004.

O. Schütze, C. A. Coello Coello, and E.-G. Talbi. Approximating the ε-Efficient Set of an MOP with Stochastic Search Algorithms. In *Mexican International Converence on Artificial Intelligence (MICAI 2007)*, volume 4827 of *LNCS*, pages 118–128. Springer, 2007a.

O. Schütze, M. Laumanns, E. Tantar, C. A. Coello Coello, and E.-G. Talbi. Convergence of Stochastic Search Algorithms to Gap-Free Pareto Front Approximations. In *Genetic and Evolutionary Computation Conference (GECCO 2007)*, pages 892–899. ACM, 2007b.

O. Schütze, C. A. Coello Coello, E. Tantar, and E. G. Talbi. Computing Finite Size Representations of the Set of Approximate Solutions of an MOP with Stochastic Search Algorithms. In C. Ryan et al., editor, *Genetic and Evolutionary Computation Conference (GECCO 2008)*, pages 713–720. ACM, 2008a.

O. Schütze, M. Laumanns, C. A. Coello Coello, M. Dellnitz, and E.-G. Talbi. Convergence of Stochastic Search Algorithms to Finite Size Pareto Set Approximations. *Journal of Global Optimization*, 41:559–577, 2008b.

O. Teytaud. On the Hardness of Offline Multi-objective Optimization. *Evolutionary Computation*, 15(4):475–491, 2007.

D. Thierens. Convergence Time Analysis for the Multi-objective Counting Ones Problem. In C. M. Fonseca, P. J. Fleming, E. Zitzler, K. Deb, and L. Thiele, editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 355–364. Springer, 2003.

L. G. Valiant. The Complexity of Computing the Permanent. *Theoretical Computer Science*, 8:189–201, 1979.

D. A. Van Veldhuizen and G. B. Lamont. On Measuring Multiobjective Evolutionary Algorithm Performance. In *Congress on Evolutionary Computation (CEC 2000)*, volume 1, pages 204–211. IEEE Service Center, 2000.

M. Villalobos-Arias, C. A. Coello Coello, and O. Hernández-Lerma. Asymptotic Convergence of Some Metaheuristics Used for Multiobjective Optimization. In A. H. Wright, M. D. Vose, K. A. De Jong, and L. M. Schmitt, editors, *Foundations of Genetic Algorithms (FOGA 2005)*, volume 3469 of *LNCS*, page 95111. Springer, 2005.

I. Wegener. Methods for the Analysis of Evolutionary Algorithms on Pseudo-Boolean Functions. In *Evolutionary Optimization*, International Series in Operations Research & Management Science, pages 349–369. Springer, 2003.

L. While. A New Analysis of the LebMeasure Algorithm for Calculating Hypervolume. In *Conference on Evolutionary Multi-Criterion Optimization (EMO 2005)*, volume 3410 of *LNCS*, pages 326–340, Guanajuato, México, 2005. Springer.

L. While, P. Hingston, L. Barone, and S. Huband. A Faster Algorithm for Calculating Hypervolume. *IEEE Transactions on Evolutionary Computation*, 10(1):29–38, 2006.

F. Xue, A. C. Sanderson, and R. J. Graves. Modeling and Convergence Analysis of a Continuous Multi-Objective Differential Evolution Algorithm. In *Congress on Evolutionary Computation (CEC 2005)*, pages 228–235. IEEE Press, 2005a.

F. Xue, A. C. Sanderson, and R. J. Graves. Multi-Objective Differential Evolution—Algorithm, Convergence Analysis, and Applications. In *Congress on Evolutionary Computation (CEC 2005)*, pages 743–750. IEEE Press, 2005b.

Z. Yan, L. Zhang, L. Kang, and G. Lin. A New MOEA for Multi-objective TSP and Its Convergence Property Analysis. In C. M. Fonseca et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2003)*, volume 2632 of *LNCS*, pages 342–354. Springer, 2003.

E. Zitzler and S. Künzli. Indicator-Based Selection in Multiobjective Search. In X. Yao et al., editors, *Conference on Parallel Problem Solving from Nature (PPSN VIII)*, volume 3242 of *LNCS*, pages 832–842. Springer, 2004.

E. Zitzler and L. Thiele. An Evolutionary Approach for Multiobjective Optimization: The Strength Pareto Approach. TIK Report 43, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, May 1998a.

E. Zitzler and L. Thiele. Multiobjective Optimization Using Evolutionary Algorithms - A Comparative Case Study. In *Conference on Parallel Problem Solving from Nature (PPSN V)*, pages 292–301, 1998b.

E. Zitzler, M. Laumanns, and L. Thiele. SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multiobjective Optimization. In K.C. Giannakoglou et al., editors, *Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001)*, pages 95–100, 2002.

E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. Grunert da Fonseca. Performance Assessment of Multiobjective Optimizers: An Analysis and Review. *IEEE Transactions on Evolutionary Computation*, 7(2):117–132, 2003.

E. Zitzler, D. Brockhoff, and L. Thiele. The Hypervolume Indicator Revisited: On the Design of Pareto-compliant Indicators Via Weighted Integration. In S. Obayashi et al., editors, *Conference on Evolutionary Multi-Criterion Optimization (EMO 2007)*, volume 4403 of *LNCS*, pages 862–876. Springer, 2007.

E. Zitzler, L. Thiele, and J. Bader. On Set-Based Multiobjective Optimization (Revised Version). TIK Report 300, Computer Engineering and Networks Laboratory (TIK), ETH Zurich, December 2008.

Centre de recherche INRIA Saclay – Île-de-France
Parc Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 Orsay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399