

Optimized Approximation Sets for Low-dimensional Benchmark Pareto Fronts

Tobias Glasmachers

Institut für Neuroinformatik, Ruhr-Universität Bochum, Germany
tobias.glasachers@ini.rub.de

Abstract. The problem of finding sets of bounded cardinality maximizing dominated hypervolume is considered for explicitly parameterized Pareto fronts of multi-objective optimization problems. A parameterization of the Pareto front is often known (by construction) for synthetic benchmark functions. For the widely used ZDT and DTLZ families of benchmarks close-to-optimal sets have been obtained only for two objectives, although the three-objective variants of the DTLZ problems are frequently applied. Knowledge of the dominated hypervolume theoretically achievable with an approximation set of fixed cardinality facilitates judgment of (differences in) optimization results and the choice of stopping criteria, two important design decisions of empirical studies. The present paper aims to close this gap. An efficient optimization strategy is presented for two and three objectives. Optimized sets are provided for standard benchmarks.

1 Introduction

Empirical benchmark studies play a major role for performance comparisons of nature inspired (optimization) algorithms. For many benchmark problems in widespread use the optimum is known analytically. This allows to compare algorithms not only relative to each other but also in relation to the actual optimum. This is a prerequisite, e.g., for the empirical investigation of convergence rates. It is practically useful for the design of meaningful stopping criteria in benchmark studies comparing the runtime of different (black-box) optimization algorithms for reaching a predefined solution accuracy.

The situation in multi-objective optimization differs in various respects from the single-objective case. The optimum is a set—the Pareto front—which can be of uncountably infinite cardinality. In practice optimal subsets of a priori bounded cardinality are of primary interest. There are multiple performance indicators in common use, and the optimal set of course depends on the indicator. In recent years the hypervolume indicator has advanced to the most widely applied performance measure at least for up to three or four objectives. Hence this study is focused on maximization of dominated hypervolume.

For the commonly used ZDT and DTLZ benchmark suites [13,7] sets with close-to-optimal hypervolume coverage are known only for the simplest case of

two objectives [1]. However, the DTLZ benchmarks are scalable to any number of objectives, and they are most often applied with three objective functions.

The present study aims to close this gap. For this purpose an efficient gradient-based hypervolume maximization algorithm based on an explicit parameterization of the Pareto front is proposed. With this algorithm we compute optimized Pareto front approximations for bi-objective and tri-objective benchmark problems.

2 Multi-objective Optimization

Consider a search space X and a set of m scalar objective functions $f_1, \dots, f_m : X \rightarrow \mathbb{R}$, each of which (w.l.o.g.) is to be minimized. Instead of aggregating the various goals encoded by the different functions into a single objective (e.g. by means of a weighted combination), the goal of multi-objective optimization (MOO) is to obtain the set of Pareto optimal (or non-dominated) points, which is the set of Pareto optimal compromises. This set is often huge or even infinite and we aim for a representative approximation set of a-priori bounded cardinality.

2.1 Dominance Order and Dominated Hypervolume

The objectives are collected in the vector-valued objective function $f : X \rightarrow \mathbb{R}^m$, $f(x) = (f_1(x), \dots, f_m(x))$. Let $Y = \{f(x) \mid x \in X\} = f(X) \subset \mathbb{R}^m$ denote the image of the objective function (also called the attainable objective space). For values $y, y' \in \mathbb{R}^m$ we define the Pareto dominance relation

$$\begin{aligned} y \preceq y' &\Leftrightarrow y_k \leq y'_k \text{ for all } k \in \{1, \dots, m\} , \\ y \prec y' &\Leftrightarrow y \preceq y' \text{ and } y \neq y' . \end{aligned}$$

This relation defines a partial order on Y , incomparable values y, y' fulfilling $y \not\preceq y'$ and $y' \not\preceq y$ remain. The relation is pulled back to the search space X by the definition $x \preceq x'$ iff $f(x) \preceq f(x')$.

The Pareto front is defined as the set of values that are optimal w.r.t. Pareto dominance, i.e., the set of non-dominated values

$$Y^* = \left\{ y \in Y \mid \nexists y' \in Y : y' \prec y \right\}$$

and the Pareto set is $X^* = f^{-1}(Y^*)$.

For generic objectives f_k without simultaneous critical points the Pareto front is a manifold of dimension $m - 1$. As such its cardinality is uncountably infinite. We are often interested in picking a “representative subset” or approximation set of fixed cardinality n . The approximation quality of a set $S = \{y_1, \dots, y_n\}$ can be judged with different set quality indicators of which the dominated hypervolume

$$H_r(S) = \Lambda \left(\left\{ y' \in \mathbb{R}^m \mid \exists y \in S \text{ s.t. } y \preceq y' \preceq r \right\} \right)$$

(where A denotes the Lebesgue measure on \mathbb{R}^m) is distinguished (up to weighting of the Lebesgue measure) for its property of being compliant with Pareto dominance [14,12]. The hypervolume indicator depends on a reference point $r \in \mathbb{R}^m$ that needs to be set by the user. It defines an objective-wise cut-off for the quality assessment.

One standard formalization of the goal of MOO is to produce a set $\{x_1, \dots, x_n\}$ of n points so that the corresponding values $y_k = f(x_k)$ maximize the hypervolume indicator for given f , r , and n . It is easy to see that the elements of the set S^* maximizing the hypervolume $H_r(S)$ are Pareto optimal, i.e., $S^* = \{y_1, \dots, y_n\} \subset Y^*$.

The optimal set S^* as well as the dominated hypervolume $H_r(S^*)$ depend on X and f only via Y^* . Since we are interested only in S^* and $H_r(S^*)$ we simplify the problem statement in the following by assuming that the set Y^* is known. This means that we will ignore a large part of the complexity of the underlying MOO problem related to the black-box setting and the potentially involved form of f . For practical purposes we may assume that a (surjective, sometimes bijective) parameterization $\varphi : U \rightarrow Y^*$, $U \subset \mathbb{R}^{m-1}$, is available. This parameterization then replaces the (often much harder to optimize) objective function.

2.2 Benchmark Problems and Existing Results

In this study we focus on well-established benchmark suits of problems with continuous variables. The so-called ZDT functions [13] ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6 (ZDT5 is a discrete problem) are scalable to any search space dimension $X = \mathbb{R}^d$ but come with only two objectives. The conceptually similar DTLZ functions [7] DTLZ1, DTLZ2, DTLZ3, and DTLZ4 (the other members of the family are rarely used) improve on this situation by being scalable to many objectives, with the “recommended default” of three objectives.

For the ZDT and DTLZ families of problems the sets Y^* are known analytically. For $m = 2$ near optimal sets S of cardinalities 2, 3, 4, 5, 10, 20, 50, 100, 1000 for the reference point (11, 11) have been obtained, see [1]. The one-dimensional fronts are visualized in figure 1 (a) to (f).

For $m = 3$ the DTLZ fronts become two-dimensional. They are depicted in figure 1 (g) and (h). Fixed size sets with maximal hypervolume coverage are not known. We provide such near optimal sets in section 5 and in the supplementary material.

The problem of optimizing $H_r(S)$ has been investigated theoretically, but analysis is mostly restricted to two objectives [3,4,11]. Results include conditions for the inclusion of extremal points in the optimal solution set, monotonicity of $H_r(S^*)$ in $n = |S^*|$, optimal sets on linear fronts, and asymptotically optimal distributions on smooth fronts in the limit $n \rightarrow \infty$. Basic results have been obtained for $m = 3$ objectives [2]; but even asymptotically optimal distributions based on local shape (derivative) features are unknown.

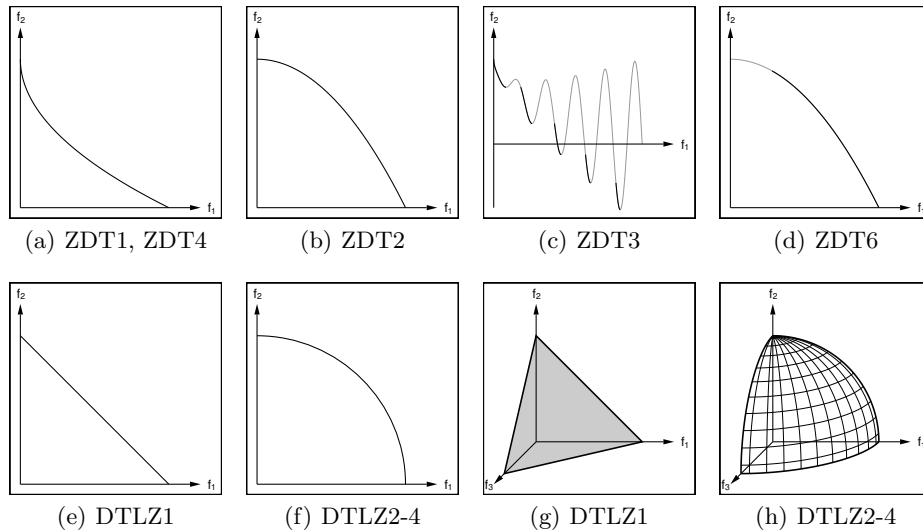


Fig. 1. Pareto fronts of the ZDT and DTLZ problems for $m = 2$ and $m = 3$.

Finally, the most relevant precursors for obtaining optimal sets of fixed cardinality n in practice are algorithms for the fast computation of the hypervolume and its derivatives [10,9,8].

3 Hypervolume Calculation

The computation of the dominated hypervolume $H_r(S)$ of n values $S = \{y_1, \dots, y_n\} \subset Y^*$ is known to be NP-hard, and the time complexity of the best known algorithms for this problem is exponential in m [5,6]. However, for few objectives ($m \leq 3$) it can be carried out in $\mathcal{O}(n \log(n))$ operations [10,9,8].

In the present paper we consider only subsets $S \subset Y^*$. Thus any pair of different points $y_i, y_j \in S$ is strictly incomparable (i.e., it holds $y_i \not\leq y_j$ and $y_j \not\leq y_i$).

3.1 Gradient of Dominated Hypervolume

It is easy to see that the dominated hypervolume $H_r(S)$ for $S = \{y_1, \dots, y_n\}$ considered as a function of y_k is nearly everywhere differentiable. Efficient algorithms for the computation of the hypervolume and its derivatives for $m \leq 4$ objectives have been proposed in [8]. It is sufficient for the purpose of practical optimization to consider the vector of derivatives $\frac{\partial H_r(S)}{\partial y_k}$, $k \in \{1, \dots, n\}$, and to ignore issues of the argument S being an *unordered* set of values (refer to [8] for a more careful derivation). With $y_k = \varphi(u_k)$ the chain rule gives rise to

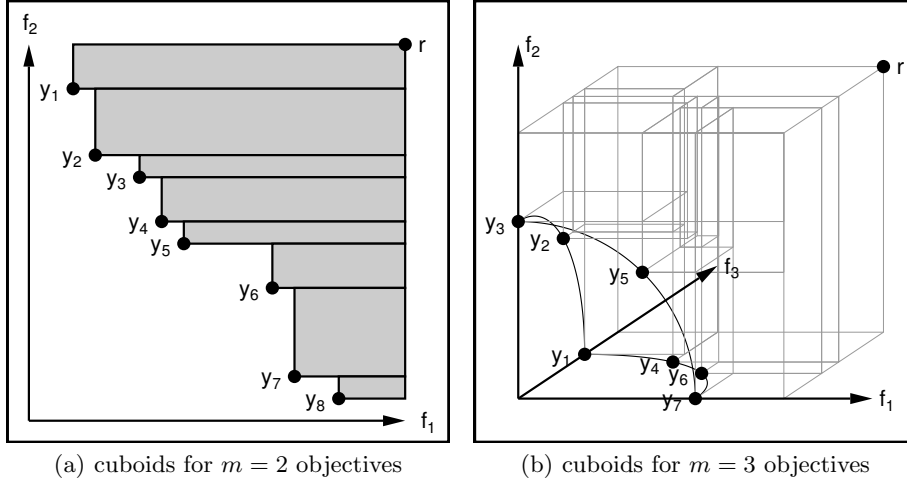


Fig. 2. Decomposition of the dominated hypervolume into disjoint cuboids.

$\frac{\partial H_r(S)}{\partial u_k} = \frac{\partial H_r(S)}{\partial y_k} \frac{\partial y_k}{\partial u_k}$ with $\frac{\partial y_k}{\partial u_k} = \varphi'(u_k)$. Note that in our setting the gradient does not aid in finding the Pareto front, since the front is already implicitly encoded in the parameterization φ . Instead it indicates how to improve the spread of the set S (e.g., by means of a gradient ascent step).

3.2 Decomposition into Cuboids

In this section we emphasize the possibility to represent the dominated hypervolume explicitly as a disjoint¹ union of simple volumes, in this case m -dimensional axis-aligned cuboids (rectangles in two dimensions, cuboids in three dimensions). We restrict the following consideration to $m = 2$ and $m = 3$. A crucial observation is that the cuboids are not only aligned to the coordinate axes but also start and end at objective values that appear in the set $S^+ = S \cup \{r\}$. Each rectangle can thus be represented as a $2m$ -tuple of values² $(y_1^-, y_1^+, \dots, y_m^-, y_m^+)$ representing the cuboid $[(y_1^-)_1, (y_1^+)_1] \times \dots \times [(y_m^-)_m, (y_m^+)_m] \subset \mathbb{R}^m$.

For $m = 2$ the set S is sorted by one objective, which results in the reverse order in the other objective. Then the hypervolume is computed by splitting the dominated set into disjoint rectangles and summing their areas (see figure 2 (a)). Thus the decomposition into exactly n rectangles is a cheap by-product of the hypervolume computation: sorting the points requires $\mathcal{O}(n \log(n))$ operations, while there are only n rectangles.

¹ For simplicity of presentation the cuboids' boundaries may overlap. However, the corresponding open cuboids (the topological interiors) are disjoint. This does not impact the hypervolume computation.

² For efficiency reasons, indices or pointers may be used in actual implementations.

For $m = 3$ we adopt the sweep-based algorithm from [10]. Instead of computing the hypervolume on the fly the modified algorithm stores and reports a list of n to $2n - 1$ and hence $\Theta(n)$ cuboids. So again the collection of the cuboids is a rather cheap by-product of the $\mathcal{O}(n \log(n))$ hypervolume computation. The result is illustrated in figure 2 (b).

The decomposition into cuboids is not more costly than the computation of the hypervolume itself, and for $m \leq 3$ the number of cuboids is only linear in n . This proceeding has the advantage that subsequent tasks such as the computation of the hypervolume or its partial derivatives (possibly higher derivatives) [8], dependence of hypervolume contributions on other points [9], etc. do not need to be incorporated into the bookkeeping-heavy sweeping algorithm. Instead they can be realized in a unified scheme, namely by first invoking the cuboid decomposition algorithm and a subsequent loop over the list of cuboids.

4 Hypervolume Optimization

4.1 Gradient-based Optimization

In the following we assume an algorithm that represents the dominated hypervolume as a disjoint union of m -dimensional axis aligned cuboids as discussed in section 3. This allows for the trivial computation of $H_r(S)$ as the sum of the elementary volumes.

The coordinate wise lower and upper bounds of the cuboids are given by coordinates of points from the set $S^+ = S \cup \{r\}$. This representation allows for trivial differentiation of $H_r(S)$ w.r.t. points in S , yielding n derivatives $\frac{\partial H_r(S)}{\partial y_k}$ (see also section 3.1). Thus, starting from any initial configuration S can be iteratively refined with a gradient-based optimization procedure.

We propose simple gradient ascent steps $u_k \leftarrow u_k + \eta \cdot \nabla_{u_k} H_r(S)$ with learning rate $\eta > 0$. If a step happens to decrease the hypervolume then backtracking is applied: the previous set S is restored and the learning rate η is halved. Otherwise the learning rate is optimistically increased (here by a factor of 1.05). This heuristic is analog to success-based step-size control in elitist evolution strategies.

4.2 Dealing with Multi-modality

Gradient ascent is an efficient technique for the localization of a local maximum. It turns out that for $m > 2$ multi-modality of the hypervolume indicator is a practically relevant issue. We address this problem with two simple yet effective techniques: proper initialization and a multi-start strategy.

Uniform initialization of the parameters $\{u_1, \dots, u_n\} \subset U$ may lead to a highly distorted distribution of the actual values $\{\varphi(u_1), \dots, \varphi(u_n)\} \subset Y^*$. Furthermore, at least in the limit $n \rightarrow \infty$ the density of values should depend on the slope of the front. For $m = 2$ objectives this was derived in [3]. Here we present a heuristic (inexact yet practical) extension of these ideas to $m = 3$ objectives.

Importantly, we do not claim to solve the problem of asymptotically optimal distributions but rather aim for a procedure generating suitable initial solutions for gradient-based optimization.

For simplicity let us assume that the hypervolume contribution of a point $y = \varphi(u)$ consists of the volume of a cuboid with side lengths b_1, b_2, b_3 and volume $V = b_1 \cdot b_2 \cdot b_3$. Furthermore assume that the point is not close to the boundary of the front and that the front surface is regular at y . This implies that locally the front can be approximated by a hyperplane (here a plane), spanned by the derivative vectors $\frac{\partial \varphi}{\partial u_1}$ and $\frac{\partial \varphi}{\partial u_2}$. Its orientation is characterized by the normal vector $w(u) = \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2}$, which can be obtained as the cross product of the tangent vectors. The norm $\|w\|$ measures the (local) volume growth of φ .

For large n the normal vector determines the local distribution of values y which—in the optimum—locally have equal hypervolume contributions $V \approx \text{const.}$ At the same time it should hold $w_1 b_1 \approx w_2 b_2 \approx w_3 b_3$ for the shape of the cuboid. Putting these together we obtain that b_i is proportional to $\sqrt[3]{w_1 w_2 w_3} / w_i$. The “area of the front” covered by the cuboid can be approximated by the area of the triangle spanned by the cuboid vertices $y + (b_1, 0, 0)$, $y + (0, b_2, 0)$, and $y + (0, 0, b_3)$. It is computed as $A = \frac{1}{2} \cdot \sqrt{b_1^2 b_2^2 + b_1^2 b_3^2 + b_2^2 b_3^2}$. Under the above considerations the optimal density of parameters u is proportional to $\|w\|/A$.

We sample an initial set of size n from the above density by means of rejection sampling. For this purpose $\kappa = 100$ random points are drawn from the uniform distribution on $u \in U$. The maximal value of $\|w\|/A$ over these points, multiplied by a safety margin of two, is kept as a tentative upper bound B on the unnormalized density. Then parameters u are sampled uniformly and rejected with probability $\min\{1, 1 - \|w\|/(A \cdot B)\}$ until n samples are accepted.

Our multi-start procedure generates N independent initial sets of size n as described above. Each initial set serves as a starting point for the gradient-based optimization algorithm.

4.3 Implementation

We provide an efficient C++ implementation of the above described hypervolume maximization algorithm with rejection sampling initialization and restart strategy. The program can be downloaded from <http://www.ini.rub.de/PEOPLE/glasmtbl/code/opt-hv/>.

5 Optimized Sets for the ZDT and DTLZ Problems

In this section we present the close to optimal sets obtained by our optimization algorithm for the ZDT and DTLZ problems. In analogy to [1] we have conducted all experiments for cardinalities $n \in \{2, 3, 4, 5, 10, 20, 50, 100, 1000\}$.

5.1 The Bi-objective Case

The ZDT and DTLZ problems have been optimized with specifically tailored algorithms. The results are found on the website [1], which is a valuable resource

when experimenting with these benchmark problems. We stick to the original reference point $r = (11, 11)$. This test aims to validate our optimization algorithm.

n	ZDT1,4	ZDT2	ZDT3	ZDT6	DTLZ1	DTLZ2-4
2	120.0248764	120.0000000	128.0147714	117.2489467	120.7500000	120.0000000
3	120.3877279	120.1481481	128.4523400	117.3723140	120.8125000	120.0857864
4	120.4915975	120.2041588	128.5997409	117.4178988	120.8333333	120.1215851
5	120.5397291	120.2339071	128.6671568	117.4417417	120.8437500	120.1415358
10	120.6137609	120.2868199	128.7459431	117.4832459	120.8611111	120.1789660
20	120.6423963	120.3106986	128.7632012	117.5014399	120.8684211	120.1968576
50	120.6574465	120.3243978	128.7707848	117.5116580	120.8724490	120.2074851
100	120.6621372	120.3288807	128.7739496	117.5149559	120.8737374	120.2110337
1000	120.6662212	120.3328889	128.7774084	117.5178796	120.8748749	120.2142433

Table 1. Maximal dominated hypervolume covered by sets of cardinalities $n \in \{2, 3, 4, 5, 10, 20, 50, 100, 1000\}$ for bi-objective problems with reference point $(11, 11)$.

We have run the gradient-based optimization procedure $N = 100$ times with random initial configurations. The results are presented concisely in table 1. Standard deviations across repetitions are extremely small (usually below 10^{-10}), indicating that the global optimum is obtained in each single run. Most of our results reproduce the optimized fronts obtained in [1]. For large values of n we observe slight improvements. For example, for problem DTLZ2 with $n = 100$ our gradient-based procedure obtains a dominated hypervolume of 120.2110337 instead of the previously reported value of 120.210644. The improvement in itself may seem minor, however, for $n = 100$ and $n = 1000$ our optimization procedure improves on most of the existing numbers.

The ZDT3 problem is an exception. Here we observe improved values for small n . This is because the left extreme point should not be fixed for the optimization, see also Theorem 2 in [3]. On the other hand our results for large n are significantly worse than those reported at [1] since gradient-ascent cannot deal well with the disconnected front of the ZDT3 problem and the resulting discontinuous parameterization φ .

5.2 The Tri-objective Case

A major motivation for the present work is to obtain optimized fronts for the DTLZ problems in their standard form, which is with three objectives. Results of our gradient-based optimizer are presented in table 2. The optimized sets are available for download at <http://www.ini.rub.de/PEOPLE/glasmtbl/code/opt-hv/> in csv format, and as eps and png figures.

The variance in the results is significantly higher than in the bi-objective case. This is because of the multi-modality of the problem. Hence we have increased

front	n	mean	stddev	25%	50%	75%	max	
DTLZ1	2	7.5281031	0.0094041	7.5312500	7.5312500	7.5312500	7.5312500	
	3	7.8750000	0.1102255	7.6445649	7.8750000	7.8750000	7.8750000	
	4	7.8946157	0.0526134	7.9062500	7.9062500	7.9062500	7.9120370	
	5	7.9222298	0.0212193	7.9238281	7.9242346	7.9259728	7.9260397	
	10	7.9532612	0.0005410	7.9529850	7.9532053	7.9537283	7.9539787	
	20	7.9644361	0.0001552	7.9643638	7.9644671	7.9645441	7.9647401	
	50	7.9712554	0.0000490	7.9712280	7.9712615	7.9712901	7.9713876	
	100	7.9739706	0.0000267	7.9739557	7.9739739	7.9739892	7.9740466	
	1000	7.9776989	0.0000039	7.9776965	7.9776992	7.9777016	7.9777110	
	DTLZ2-4	2	6.0000000	0.0000000	6.0000000	6.0000000	6.0000000	6.0000000
3		6.8272558	0.3365748	7.0000000	7.0000000	7.0000000	7.0000000	
4		7.0694591	0.1090694	7.0857864	7.0857864	7.0857864	7.0857864	
5		7.1467484	0.0206817	7.1493061	7.1493061	7.1493061	7.1493061	
DTLZ2		10	7.2809948	0.0049230	7.2780682	7.2795647	7.2860090	7.2874732
DTLZ3		20	7.3485703	0.0022931	7.3472972	7.3488734	7.3501758	7.3545152
DTLZ4		50	7.3994118	0.0010188	7.3987853	7.3995002	7.4001307	7.4022754
100		7.4228644	0.0006244	7.4224787	7.4229145	7.4232955	7.4246456	
1000		7.4597704	0.0001156	7.4596963	7.4597782	7.4598501	7.4601203	

Table 2. Characteristics (mean, standard deviation, quantiles, and maximum) of the empirical distributions of dominated hypervolume for the DTLZ1 front (upper half) and the DTLZ2-4 front (lower half) with $m = 3$ objectives, reference point $r = (2, 2, 2)$, and cardinalities $n \in \{2, 3, 4, 5, 10, 20, 50, 100, 1000\}$.

the number of runs to $N = 10,000$. Running the procedure with even more repetitions will most probably give slightly higher hypervolumes. However, most reasonable optimization procedures may get stuck in local optima. Therefore not only the global optimum is of interest but also the distribution of local optima. The descriptive statistics in table 2 provide such data. This allows to judge the performance of algorithm on an absolute scale w.r.t. a reference distribution, e.g., by measuring how often a certain quantile of the empirical distribution of local optima is reached.

6 Conclusion

We have presented an efficient algorithm for the maximization of dominated hypervolume of sets of fixed cardinality when a parametric form of the Pareto front is known. Such sets are of practical relevance when comparing multi-objective optimizers on benchmark problems. While existing studies have been restricted to relative comparisons we are now in the position to relate differences to an absolute scale given by the best known hypervolume and by the empirical distribution of local optima as identified by our multi-start procedure. This also allows to report the performance of a single (e.g., novel) algorithm on an absolute scale rather than relative to (arbitrarily chosen) competitors. Our gradient-based pro-

cedure is computationally efficient. This algorithm has been integrated into a standalone software with easy-to-use command line interface.

References

1. ZDT and DTLZ test problems. <http://people.ee.ethz.ch/~sop/download/supplementary/testproblems/>. Accessed: 2014-03-17.
2. A. Auger, J. Bader, and D. Brockhoff. Theoretically Investigating Optimal μ -Distributions for the Hypervolume Indicator: First Results For Three Objectives. In *Parallel Problem Solving from Nature, PPSN XI*, pages 586–596. Springer, 2010.
3. 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 *Proceedings of the tenth ACM SIGEVO workshop on Foundations of genetic algorithms*, pages 87–102. ACM, 2009.
4. A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Hypervolume-based Multiobjective Optimization: Theoretical Foundations and Practical Implications. *Theoretical Computer Science*, 425:75–103, 2012.
5. K. Bringmann. Klee’s measure problem on fat boxes in time $\mathcal{O}(n(d+2)/3)$. In *Proceedings of the twenty-sixth annual symposium on Computational geometry*, pages 222–229. ACM, 2010.
6. K. Bringmann and T. Friedrich. Approximating the least hypervolume contributor: NP-hard in general, but fast in practice. *Theoretical Computer Science*, 425:104–116, 2012.
7. K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable Multi-Objective Optimization Test Problems. In *Congress on Evolutionary Computation (CEC 2002)*, pages 825–830. IEEE Press, 2002.
8. M. Emmerich and A. Deutz. Time complexity and zeros of the hypervolume indicator gradient field. In O. Schuetze, C. A. Coello Coello, A.-A. Tantar, E. Tantar, P. Bouvry, P. Del Moral, and P. Legrand, editors, *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation III*, volume 500 of *Studies in Computational Intelligence*, pages 169–193. Springer, 2014.
9. M. Emmerich and C. M. Fonseca. Computing hypervolume contributions in low dimensions: asymptotically optimal algorithm and complexity results. In *EMO’11: Proceedings of the 6th international conference on Evolutionary multi-criterion optimization*, pages 121–135. Springer-Verlag, 2011.
10. C. M. Fonseca, L. Paquete, and M. Lopez-Ibanez. An improved dimension-sweep algorithm for the hypervolume indicator. In *IEEE Congress on Evolutionary Computation, CEC 2006*, pages 1157–1163, 2006.
11. T. Friedrich, F. Neumann, and C. Thyssen. Multiplicative Approximations, Optimal Hypervolume Distributions, and the Choice of the Reference Point. Technical Report arXiv:1309.3816, arXiv.org, 2013.
12. E. Zitzler, D. Brockhoff, and L. Thiele. The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In *Evolutionary Multi-Criterion Optimization*, pages 862–876. Springer, 2007.
13. E. Zitzler, K. Deb, and L. Thiele. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evolutionary Computation*, 8(2):173–195, 2000.
14. E. Zitzler, L. Thiele, M. Laumanns, C. Fonseca, and V. Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *Evolutionary Computation, IEEE Transactions on*, 7(2):117–132, 2003.