Scaling Up Indicator-based MOEAs by Approximating the Least Hypervolume Contributor: A Preliminary Study

Thomas Voß
Institut für Neuroinformatik
Ruhr-Universität Bochum
Bochum, Germany

Tobias Friedrich
Max-Planck-Institut
für Informatik
Saarbrücken, Germany

Karl Bringmann
Universität des Saarlandes
Saarbrücken, Germany

Christian Igel
Institut für Neuroinformatik
Ruhr-Universität Bochum
Bochum, Germany

Categories and Subject Descriptors

F.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity; G.1.6 [**Optimization**]: Global Optimization; G.3 [**Probability and Statistics**]: Probabilistic algorithms (including Monte Carlo); I.2.8 [**Problem Solving, Control Methods, and Search**]: Heuristic methods

General Terms

Algorithms, Measurement, Performance

Keywords

multi-objective optimization, hypervolume, approximation, covariance matrix adaptation, evolution strategy, MO-CMA-ES

1. INTRODUCTION

It is known that the performance of multi-objective evolutionary algorithms (MOEAs) in general deteriorates with increasing number of objectives. For few objectives, MOEAs relying on the *contributing hypervolume* as (second-level) sorting criterion are the methods of choice. These include the *multi-objective covariance matrix adaptation evolution strategy* (MO-CMA-ES, [10, 11, 14, 16]) the *SMS-EMOA* [5], and variants of the *indicator-based evolutionary algorithm* (IBEA, [18]). However, the computational complexity of calculating the contributing hypervolume prevents the broad application of these powerful MOEAs to objective functions with many (say, more than 4) objectives.

Recently, a Monte-Carlo algorithm for approximately determining the least hypervolume contributor of a given Pareto-front approximation has been presented in [8]. We hypothesize that using this approximation instead of the exact contributing hypervolume will make the aforementioned EMOAs applicable to problems with many objectives and that the resulting algorithms will push the boundaries of today's EMOAs for many-objective optimization.

In theory, the approximation allows for the application of hypervolume-based MOEAs to optimization problems with an arbitrary number of objectives. However, the effects of replacing the exact hypervolume calculation with an approximation algorithm on the overall performance of recent MOEAs are unknown. In this study, we employ the approximation within the steady-state MO-CMA-ES (termed $(\mu + 1)$-MO-CMA-ES and using the recent improvements presented in [16]) and the SMS-EMOA to empirically investigate whether the Monte-Carlo approximation is indeed useful in practice.

The remainder of the document is structured as follows. We first introduce the problem of determining the least hypervolume contributor. Then, we summarize the $(\mu + 1)$-MO-CMA-ES. We conclude with the results of a preliminary experiment and an outlook on future work.

2. IDENTIFYING THE LEAST HYPERVOLUME CONTRIBUTOR

Measures based on the hypervolume indicator are the only known Pareto-compliant unary performance indicators for comparing different Pareto-front approximations [20]. For this reason, they serve as second-level sorting criterion in many recent multi-objective evolutionary algorithms (MOEAs).

The hypervolume measure or $\mathcal{S}$-metric (see [19]) of a population A is defined as

$$\mathrm{HYP}_{\boldsymbol{f}^{\mathrm{ref}}}(A) = \Lambda\left(\bigcup_{a\in A}\left[f_1(a), f_1^{\mathrm{ref}}\right]\times\cdots\times\left[f_m(a), f_m^{\mathrm{ref}}\right]\right),$$

with $\boldsymbol{f}^{\mathrm{ref}} \in \mathbb{R}^m$ referring to an appropriately chosen reference point, $\Lambda(\cdot)$ being the Lebesgue measure and the multi-objective optimization problem $\boldsymbol{f}\colon A \to \mathbb{R}^m$, $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x}))$. The contributing hypervolume of a

GECCO'10, July 7–11, 2010, Portland, Oregon, USA.

point $a \in A$ is given by

$$\mathrm{CON}(a, A) = \mathrm{HYP}_{\boldsymbol{f}^{\mathrm{ref}}}(A) - \mathrm{HYP}_{\boldsymbol{f}^{\mathrm{ref}}}(A \setminus \{a\}) \enspace .$$

Note that the contributing hypervolume of a dominated individual is zero.

Most hypervolume based algorithms like the steady-state MO-CMA-ES [10, 11, 14] or the SMS-EMOA [5] remove the individual

$$a^* := \operatorname*{argmin}_{a \in A} \mathrm{CON}(a, A),$$

contributing the least hypervolume to the population A.

The definition of the least-contributing individual can be used to rank the individuals in a population. Such a ranking is required if more than one individual has to be selected or discarded from a Pareto set A of *non-dominated* solutions. The *contribution rank* $\mathrm{cont}(a, A)$ of $a \in A$ can then be defined as follows.

The element $a^* \in A$ with the smallest contributing hypervolume is assigned lowest contribution rank 1. The next rank is assigned by considering $A \setminus \{a^*\}$ etc. More precisely, let $c_0(A) = \mathrm{argmin}_{a \in A} \mathrm{CON}(a^*, A)$ and

$$c_i(A) = c_0 \left(A \setminus \bigcup_{j=0}^{i-1} \{c_j(A)\} \right)$$

for $i > 0$ and we assume that argmin breaks ties at random. For $a \in A$ we define the contribution rank $\mathrm{cont}(a, A)$ to be $1 + i$ iff $a = c_i(A)$.

Several algorithms for calculating the hypervolume have been developed. The first ones was the Hypervolume by Slicing Objectives (HSO) algorithm which was suggested independently by Zitzler [17] and Knowles [12]. Another popular algorithm to calculate HYP is by Beume and Rudolph [3, 4] which is based on Overmars and Yap's algorithm for Klee's Measure Problem [13]. It is known that exact calculation of HYP is **#P**-hard in the number of objectives m [6]. Therefore, unless $\mathbf{P} = \mathbf{NP}$ there is no hypervolume algorithms with a runtime polynomial in m. Also $\mathrm{CON}(a, A)$ is **#P**-hard to solve exactly [8] and even **NP**-hard to approximate by a factor of $2^{m^{1-\varepsilon}}$ for all $\varepsilon > 0$.

In the context of our algorithm we are only interested in determining $c_0(A)$. The classical way to do this would be $\mu+1$ calculations of CON. We use Bringmann and Friedrich's adaption [7] of [3, 4, 13] which calculates all $\mathrm{CON}(a, A)$, $a \in A$, in a single run. This achieves a speed-up of μ compared to the algorithm of Beume and Rudolph [3, 4] and is the fastest known exact algorithm.

Recently, Monte-Carlo approximation algorithms for determining $c_0(A)$ appeared in the literature [1, 2, 8]. We decided to implement the algorithm of Bringmann and Friedrich [8], which returns for a population A and arbitrarily given $\varepsilon, \delta > 0$, a solution $a \in A$ with contribution $\mathrm{CON}(a)$ at most $(1 + \varepsilon)\,\mathrm{CON}(a^*, A)$ with probability at least $(1 - \delta)$. As this problem is **NP**-hard [8], this algorithm cannot run in polynomial time for all instances, but it has been shown to perform very fast on artificial data-sets [8]. To the best of our knowledge this is the first paper which examines the use of an approximate hypervolume contribution for MOEAs.

3. THE $(\mu + 1)$-MO-CMA-ES

In the following, we briefly outline the MO-CMA-ES according to [10, 11, 14], see Algorithm 1. For a detailed description and an in-depth performance evaluation, we refer to [10, 15, 16].

In the MO-CMA-ES, a candidate solution $a_i^{(g)}$ in generation g is a tuple $\left[\boldsymbol{x}_i^{(g)}, \bar{p}_{\mathrm{succ},i}^{(g)}, \sigma_i^{(g)}, \boldsymbol{p}_{i,c}^{(g)}, \boldsymbol{C}_i^{(g)}\right]$, where $\boldsymbol{x}_i^{(g)} \in \mathbb{R}^n$ is the current search point, $\bar{p}_{\mathrm{succ},i}^{(g)} \in [0,1]$ is the smoothed success probability, $\sigma_i^{(g)} \in \mathbb{R}_0^+$ is the global step size, $\boldsymbol{p}_{i,c}^{(g)} \in \mathbb{R}^n$ is the cumulative evolution path, and $\boldsymbol{C}_i^{(g)} \in \mathbb{R}^{n \times n}$ is the covariance matrix of the search distribution.

The success indicator $\mathrm{succ}_{Q^{(g)}}\left(a_i^{(g)}, a'^{(g)}\right)$ evaluates to one if the mutation that has created $a'^{(g)}$ is considered to be successful and to zero otherwise.

Algorithm 1: $(\mu$ +1)-MO-CMA-ES

1 $g \leftarrow 0$, initialize parent population $Q^{(0)}$
2 **repeat**
3 $\quad i \sim \mathcal{U}\left(1, \ldots, \left|\mathrm{ndom}\left(Q^{(g)}\right)\right|\right)$
4 $\quad a'^{(g+1)} \leftarrow a_i^{(g)}$
5 $\quad \boldsymbol{x}'^{(g+1)} \sim \boldsymbol{x}_i^{(g)} + \sigma_i^{(g)} \mathcal{N}\left(\mathbf{0}, \boldsymbol{C}_i^{(g)}\right)$
6 $\quad Q^{(g)} \leftarrow Q^{(g)} \cup \left\{a'^{(g+1)}\right\}$
7 $\quad \sigma\text{-update}\left(a'^{(g+1)}, \mathrm{succ}_{Q^{(g)}}\left(a_i^{(g)}, a'^{(g)}\right)\right)$
8 $\quad \boldsymbol{C}\text{-update}\left(a'^{(g+1)}, \mathrm{succ}_{Q^{(g)}}\left(a_i^{(g)}, a'^{(g)}\right)\right)$
9 $\quad \sigma\text{-update}\left(a_i^{(g)}, \mathrm{succ}_{Q^{(g)}}\left(a_i^{(g)}, a'^{(g)}\right)\right)$
10 $\quad Q^{(g+1)} \leftarrow Q^{(g)} \setminus \left\{ \mathrm{argmin}_{a \in Q^{(g)}} \mathrm{CON}(a, Q^{(g)}) \right\}$
11 $\quad g \leftarrow g + 1$
until stopping criterion is met

In each generation, the algorithm samples an offspring individual $a'^{(g+1)}$. The parent individual $a_i^{(g)}$ is chosen uniformly at random from the set of non-dominated individuals $\mathrm{ndom}\left(Q^{(g)}\right)$ (lines 3–5). Next, the strategy parameters of both the parent and offspring individual are adapted (lines 7–9). The decision whether a new candidate solution is better than its parent is made in the context of the population $Q^{(g)}$ of parent and offspring individuals subject to the indicator-based selection strategy implemented in the algorithm. The step sizes and the covariance matrix of the offspring individual are updated (lines 7–8). Subsequently, the step size $\sigma_i^{(g)}$ of the parent individual $a_i^{(g)}$ is adapted (line 9). Finally, the new parent population is selected from the set of parent and offspring individuals according to the indicator-based selection scheme (line 10).

4. PRELIMINARY EMPIRICAL EVALUATION

This section presents the intended setup of the empirical performance evaluation of the $(\mu + 1)$-MO-CMA-ES relying on the approximation of the least hypervolume contributor. We compare the $(\mu + 1)$-MO-CMA-ES using the approximation of the least hypervolume contributor (referred to as

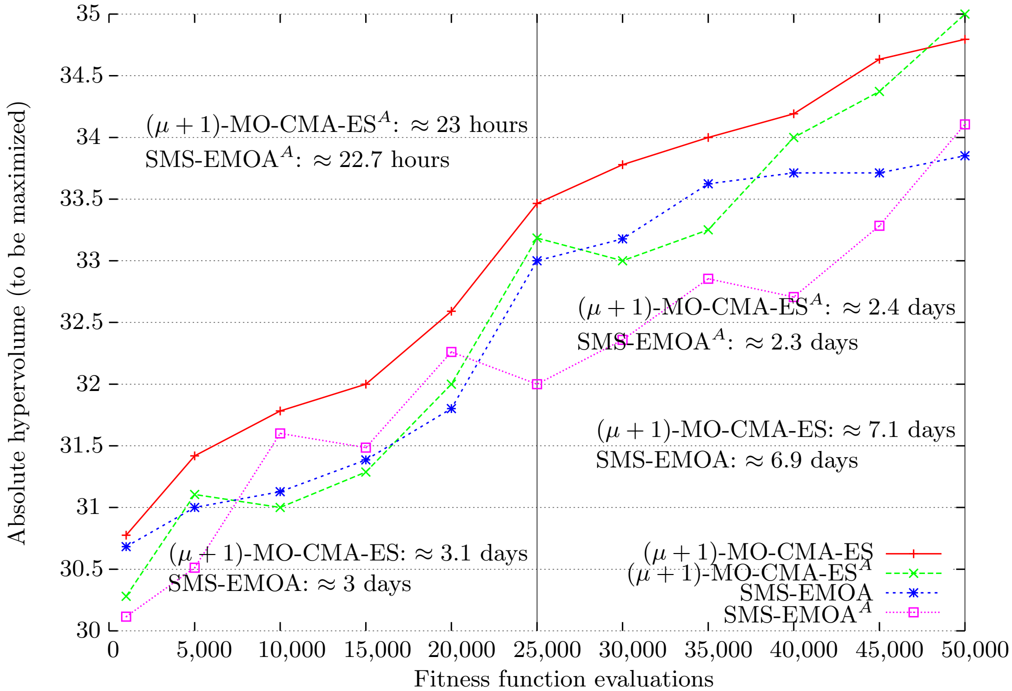

Figure 1: Evolution of the absolute hypervolume for the fitness function DTLZ2 with five objectives. Vertical lines mark the actual time that the algorithms required to carry out 25,000 and 50,000 fitness function evaluations, respectively.

(μ+1)-MO-CMA-ES^A) to the results of the original (μ+1)-MO-CMA-ES. The results of both algorithms are compared to both the results of the original SMS-EMOA (see [5]) and the results of the SMS-EMOA relying on the approximation of the least hypervolume contributor (SMS-EMOA^A). Additionally, we compare the running time of the different algorithms. We adhere to the suggestions given in [16] for statistical testing purposes.

4.1 Experimental Setup

We compare the algorithms on the well-known benchmark function DTLZ2 (see [9]). The number of objectives was set to 5, search space dimension was 30, and the number of parent individuals was $\mu = 50$. For the approximation algorithm, we used the parameters $\varepsilon = 10^{-6}$ and $\delta = 10^{-4}$.

We conducted 10 independent trials with 50,000 fitness function evaluations each. We sampled the performance of the algorithms after every 1,000th fitness function evaluations.

The experiments were carried out on a workstation with an 2.93 GHz Intel Quad-Core Xeon processor running Linux. We used the GNU compiler chain and enabled compiler optimizations according to the following command line arguments: `-O3 -ffast-math -msse4 -mtune=core2`.

4.2 Results

The results are shown in Fig. 1. For both the (μ+1)-MO-CMA-ES as well as the SMS-EMOA, the variants relying on the approximation of the least hypervolume contributor did not perform worse than the variants employing the exact hypervolume indicator. In general, both variants of the (μ+1)-MO-CMA-ES outperformed the respective variants of the SMS-EMOA.

For both the (μ + 1)-MO-CMA-ES^A and the SMS-EMOA^A, the running time of the experiments was significantly reduced (see markers in Fig. 1) when compared to the variants of the algorithms relying on the exact hypervolume indicator.

5. OUTLOOK

We presented an overview of our current work regarding the hypervolume indicator. We are empirically investigating how replacing the exact hypervolume indicator with a Monte-Carlo approximation of the least hypervolume contributor affects the performance of the multi-objective evolutionary algorithms relying on the well-known indicator-based selection strategy. The preliminary results shown here are promising and suggest that the performance of the algorithms does not suffer from the additional noise introduced by the Monte-Carlo approximation of the least hypervolume contributor. Additionally, we observe a reduction of the running time that is significant for more than four objectives.

Nevertheless, the performance analysis needs to be carried out across a broad range of fitness functions to ensure that the algorithms' performance remains robust despite possible failures in the selection procedure. Moreover, we will examine different settings of the error bound ε and the error probability δ as well as the scaling with respect to the number of objectives. Future work should also include a comparison with other approximation schemes such as HypE [1, 2].

Acknowledgments

CI acknowledges support from the German Federal Ministry of Education and Research within the National Network Computational Neuroscience under grant number 01GQ0951.

References

[1] J. Bader and E. Zitzler. HypE: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary Computation*, 2010+. To appear.

[2] J. Bader, K. Deb, and E. Zitzler. Faster Hypervolume-Based Search Using Monte Carlo Sampling. In *Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems (MCDM)*, volume 636 of *Lecture Notes in Economics and Mathematical Systems*, pages 313–326. Springer-Verlag, 2010.

[3] N. Beume. S-metric calculation by considering dominated hypervolume as Klee's measure problem. *Evolutionary Computation*, 17(4):477–492, 2009.

[4] N. Beume and G. Rudolph. Faster S-metric calculation by considering dominated hypervolume as Klee's measure problem. In *Proc. International Conference on Computational Intelligence (IASTED)*, pages 231–236. ACTA Press, 2006.

[5] N. Beume, B. Naujoks, and M. Emmerich. SMS-EMOA: Multiobjective selection based on dominated hypervolume. *European Journal of Operational Research*, 181(3):1653–1669, 2007.

[6] K. Bringmann and T. Friedrich. Approximating the volume of unions and intersections of high-dimensional geometric objects. In *Proc. 19th International Symposium on Algorithms and Computation (ISAAC)*, volume 5369 of *Lecture Notes in Computer Science*, pages 436–447, 2008.

[7] K. Bringmann and T. Friedrich. Don't be greedy when calculating hypervolume contributions. In *Proc. 10th International Workshop on Foundations of Genetic Algorithms (FOGA)*, pages 103–112, 2009.

[8] K. Bringmann and T. Friedrich. Approximating the least hypervolume contributor: NP-hard in general, but fast in practice. In *Proc. 5th International Conference on Evolutionary Multi-Criterion Optimization (EMO)*, pages 6–20, 2009.

[9] K. Deb, L. Thiele, M. Laumanns, and E. Zitzler. Scalable multi-objective optimization test problems. In *Proc. Congress on Evolutionary Computation (CEC)*, pages 825–830. IEEE Press, 2002.

[10] C. Igel, N. Hansen, and S. Roth. Covariance matrix adaptation for multi-objective optimization. *Evolutionary Computation*, 15(1):1–28, 2007.

[11] C. Igel, T. Suttorp, and N. Hansen. Steady-state selection and efficient covariance matrix update in the multi-objective CMA-ES. In *Proc. 4th International Conference on Evolutionary Multi-Criterion Optimization (EMO)*, volume 4403 of *Lecture Notes in Computer Science*. Springer-Verlag, 2007.

[12] J. D. Knowles. *Local-Search and Hybrid Evolutionary Algorithms for Pareto Optimization*. PhD thesis, Department of Computer Science, University of Reading, UK, 2002.

[13] M. H. Overmars and C.-K. Yap. New upper bounds in Klee's measure problem. *SIAM Journal on Computing*, 20(6):1034–1045, 1991.

[14] T. Suttorp, N. Hansen, and C. Igel. Efficient covariance matrix update for variable metric evolution strategies. *Machine Learning*, 75(2):167–197, 2009.

[15] T. Voß, N. Beume, G. Rudolph, and C. Igel. Scalarization versus indicator-based selection in multi-objective CMA evolution strategies. In *Proc. IEEE Congress on Evolutionary Computation (CEC)*, pages 3041–3048. IEEE Press, 2008.

[16] T. Voß, N. Hansen, and C. Igel. Improved step size adaptation for the MO-CMA-ES. In *Proc. 12th Annual Conference on Genetic and Evolutionary Computation Conference (GECCO)*. ACM Press, 2010.

[17] E. Zitzler. Hypervolume metric calculation, 2001. Computer Engineering and Networks Laboratory (TIK), ETH Zurich, Switzerland, see ftp://ftp.tik.ee.ethz.ch/pub/people/zitzler/hypervol.c.

[18] E. Zitzler and S. Künzli. Indicator-based selection in multiobjective search. In X. Yao et al., editors, *Parallel Problem Solving from Nature (PPSN)*, volume 3242 of *Lecture Notes in Computer Science*, pages 832–842, Berlin, Germany, 2004. Springer-Verlag.

[19] E. Zitzler and L. Thiele. Multiobjective optimization using evolutionary algorithms – a comparative case study. In *Proc. 5th International Conference on Parallel Problem Solving from Nature (PPSN)*, volume 1498 of *Lecture Notes in Computer Science*, pages 292–301. Springer-Verlag, 1998.

[20] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. Grunert da Fonseca. Performance assesment of multiobjective optimizers: An analysis and review. *IEEE Transactions on Evolutionary Computation*, 7 (2):117–132, 2003.