Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB Noiseless Testbed

Asma Atamna

Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB Noiseless Testbed

Asma Atamna
Projet TAO, Inria
Saclay–Île-de-France
LRI, Bât. 660, Univ. Paris-Sud
91405 Orsay Cedex, France
atamna@lri.fr

ABSTRACT

We benchmark IPOP-CMA-ES, a restart Covariance Matrix Adaptation Evolution Strategy with increasing population size, with two step-size adaptation mechanisms, Two-Point Step-Size Adapation (TPA) and Median Success Rule (MSR), on the BBOB noiseless testbed. We then compare IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR to IPOP-CMA-ES with the standard step-size adaptation mechanism, Cumulative Step-size Adaptation (CSA). We conduct experiments for a budget of 10^5 times the dimension of the search space. As expected, the algorithms perform alike on most functions. However, we observe some relevant differences, the most significant being on the attractive sector function where IPOP-CMA-TPA and IPOP-CMA-CSA outperform IPOP-CMA-MSR, and on the Rastrigin function where IPOP-CMA-MSR is the only algorithm to solve the function in all tested dimensions. We also observe that at least one of the three algorithms is comparable to the best BBOB-09 artificial algorithm on 13 functions.

Categories and Subject Descriptors

G.1.6 [**Numerical Analysis**]: Optimization—*global optimization, unconstrained optimization*; F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems

Keywords

Benchmarking, Black-box optimization

1. INTRODUCTION

This paper compares three step-size adaptation methods coupled with IPOP-CMA-ES [2], a restarted version of the state-of-the-art Evolution Strategy (ES), the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [8], where the population size is increased for each restart, on the BBOB noiseless testbed [3, 7]. The step-size adaptation algorithms under consideration are Two-Point Step-Size Adaptation (TPA) [5], Median Success Rule (MSR) [1], and Cumulative Step-Size Adaptation (CSA) [8], the latter being the default step-size adaptation method in CMA-ES. We first recall the general principle of the considered ES, we then describe the studied step-size adaptation algorithms, with a particular focus on TPA and MSR, and evaluate them empirically.

GECCO '15, July 11 - 15, 2015, Madrid, Spain

ISBN 978-1-4503-3488-4/15/07…$15.00
DOI: http://dx.doi.org/10.1145/2739482.2768467

2. THE $(\mu/\mu, \lambda)$-ES

In this paper, we consider the $(\mu/\mu, \lambda)$-ES with weighted recombination, where λ is the population size, μ is the number of parents, and ',' denotes non-elitist selection [4]. At iteration t, λ offspring, $\mathbf{X}_t^1, \ldots, \mathbf{X}_t^\lambda$, are sampled independently from a multivariate normal distribution according to

$$\mathbf{X}_t^i = \mathbf{X}_t + \sigma_t \mathcal{N}_t(0, \mathbf{C}_t) \ \ , i = 1, \ldots, \lambda \tag{1}$$

where $\mathcal{N}_t(0, \mathbf{C}_t)$ is the multivariate normal distribution with mean 0 and covariance matrix $\mathbf{C}_t$, σ_t is the step-size and defines the width of the sampling distribution. The μ best offspring are recombined to form the new solution

$$\mathbf{X}_{t+1} = \sum_{i=1}^{\mu} w_i \, \mathbf{X}_t^{i:\lambda} \tag{2}$$

where $\mathbf{X}_t^{i:\lambda}$ is the ith best offspring fitness-wise, $w_i > 0$ and $\sum_{i=1}^{\mu} w_i = 1$. In adaptive ES, σ_t and $\mathbf{C}_t$ are updated during the search process in order to achieve fast convergence.

3. IPOP-CMA-ES

IPOP-CMA-ES consists in launching independent restarts of CMA-ES by increasing the population size by a factor of two for each restart. Increasing the population size allows for a better covering of the search space and improves the performance of CMA-ES on multimodal functions [2]. The principle of the algorithm can be summed up in two steps:

1. run CMA-ES
2. if CMA-ES stops before reaching the target value and before exceeding the budget, double the population size and go to step 1

For a detailed description of the algorithm, see [2].

CMA-ES. In this paper, we consider the $(\mu/\mu, \lambda)$-CMA-ES with weighted recombination, fully described in [8].

4. STEP-SIZE ADAPTATION METHODS

This section describes the three step-size adaptation methods under investigation.

4.1 TPA

In Two-Point Step-Size Adaptation, the first two offspring are sampled along the shift vector from the previous solution, $\mathbf{X}_{t-1}$, to the current solution $\mathbf{X}_t$, as a mirrored pair, symmetric to $\mathbf{X}_t$.

$$\mathbf{X}_t^{1,2} = \mathbf{X}_t \pm \sigma_t \times \|\mathcal{N}_t(0, \mathbf{I})\| \frac{\mathbf{X}_t - \mathbf{X}_{t-1}}{\|\mathbf{X}_t - \mathbf{X}_{t-1}\|} \tag{3}$$

where $\mathbf{I}$ is the identity matrix. We decide whether to increase or decrease the step-size σ_t depending on the fitness of $\mathbf{X}_t^1$ and $\mathbf{X}_t^2$: if $\mathbf{X}_t^1$ is better than $\mathbf{X}_t^2$, σ_t is increased as this indicates that there are better solutions in the direction of the latest solution shift. Otherwise, it is decreased. The following equations give the step-size update.

$$s_1 = (1 - c_\sigma)\, s_{t-1} + c_\sigma \frac{\mathrm{rank}(\mathbf{X}_t^2) - \mathrm{rank}(\mathbf{X}_t^1)}{\lambda - 1} \tag{4}$$

$$\sigma_{t+1} = \sigma_t \exp\left(\frac{s_t}{d_\sigma}\right) \tag{5}$$

where $\mathrm{rank}(\mathbf{X}_t^i)$ is the fitness ranking of the ith individual among the entire population, $s_0 = 0$, $c_\sigma = 0.3$, and $d_\sigma = \sqrt{D}$ where D is the dimension of the search space. A more thorough description of the algorithm can be found in [5].

4.2 MSR

The Median Success Rule Step-Size Adaptation can be seen as a generalization of the 1/5th success rule [10] to the case of $(\mu/\mu, \lambda)$-ES. The success is defined as the median individual (fitness-wise) of the current population, $\mathbf{X}_t^{m(\lambda)}$, being better than the jth best individual of the previous population, $\mathbf{X}_{t-1}^{j:\lambda}$. In practice, j is chosen such that the median success probability is 1/2 with optimal step-size on the sphere function [1]. The idea is then to increase the step-size if $\mathbf{X}_t^{m(\lambda)}$ is fitter than $\mathbf{X}_{t-1}^{j:\lambda}$ and decrease it otherwise. The step-size σ_t is updated as

$$s_1 = (1 - c_\sigma)\, s_{t-1} + c_\sigma \frac{2}{\lambda}\left(K_{\mathrm{succ}} - \frac{\lambda}{2}\right) \tag{6}$$

$$\sigma_{t+1} = \sigma_t \exp\left(\frac{s_t}{d_\sigma}\right) \tag{7}$$

where K_{succ} is the number of successful individuals, $s_0 = 0$, $c_\sigma = 0.3$, and $d_\sigma = 2 - 2/D$.

4.3 CSA

The Cumulative Step-Size Adaptation is the standard step-size adaptation method in CMA-ES. A detailed description of the method can be found in [8].

5. EXPERIMENTAL PROCEDURE

We ran the algorithms with a budget of $10^5 \times D$ on the BBOB noiseless functions in six different dimensions. We used the python implementation of CMA-ES, cma 1.1.06. The source code can be found at [11]. TPA, MSR, and CSA are implemented in cma 1.1.06 as well as the IPOP restart strategy. For each run of the algorithms, the initial solution $\mathbf{X}_0$ is sampled uniformly in $[-4, 4]^D$ and the initial step-size σ_0 is set to 2.5. The maximum number of restarts is set to 9. For all other parameters, default values are used (for instance, the population size $\lambda = 4 + \lfloor 3 \ln D \rfloor$ and the number of parents $\mu = \lambda/2$).

6. RESULTS

Results from experiments according to [6] on the benchmark functions given in [3, 7] are presented in Figures 1, 3 and 4 and in Tables 1 and 2. The **expected running time (ERT)**, used in the figures and tables, depends on a given target function value, $f_{\mathrm{t}} = f_{\mathrm{opt}} + \Delta f$, and is computed over all relevant trials as the number of function evaluations executed during each trial while the best function value did not reach f_{t}, summed over all trials and divided by the number of trials that actually reached f_{t} [6, 9]. **Statistical significance** is tested with the rank-sum test for a given target Δf_{t} using, for each trial, either the number of needed function evaluations to reach Δf_{t} (inverted and multiplied by -1), or, if the target was not reached, the best Δf-value achieved, measured only up to the smallest number of overall function evaluations for any unsuccessful trial under consideration.

For the sake of simplicity, we will refer to IPOP-CMA-ES-TPA, IPOP-CMA-ES-MSR, and IPOP-CMA-ES-CSA as TPA, MSR, and CSA respectively in the following.

ERT versus dimension. Figure 1 shows that in 5-D (respectively 20-D), TPA, MSR, and CSA solve 22 (respectively 19), 20 (respectively 20), and 22 (respectively 20) out of 24 functions. For unsolved functions (mainly multi-modal and weakly structured multi-modal functions), a larger budget is required (at least $10^6 \times D$ function evaluations). The algorithms have a comparable performance on most of the functions and scale similarly with the dimension. This corresponds to our expectations, as the three algorithms are very similar. On some functions, however, we observe relevant differences in the performance: on function 1 (sphere), TPA performs significantly better than MSR and CSA in at least one dimension. We also observe a significant difference on function 6 (attractive sector) where TPA and CSA outperform MSR in large dimensions. Single runs on function 6 show that MSR generates smaller step-sizes than TPA and CSA, which leads to its larger ERT. Figure 2 displays single runs of MSR (left) and CSA (right) in 20-D (due to space limitations, results for TPA are not presented). On function 3 (separable Rastrigin), MSR has the best performance. Our explanation is that having small step-sizes avoids getting stuck in local optima. On functions 16 (Weierstrass) and 19 (Griewank-Rosenbrock), TPA and CSA perform very similarly and better than MSR. On function 20 (Schwefel), CSA performs slightly better than TPA in small dimensions. The gap we see in 10-D between TPA and CSA is due to insufficient budget and should disappear by increasing the budget. Another significant difference is observed on function 23 (Katsuuras) where MSR solves the function within the maximum budget and performs better than TPA and CSA. On function 21 (Gallagher 101 peaks), a larger budget is necessary to decide whether the observed difference is significant, since the ERTs are close to the maximum budget. Another observation is that each algorithm performs similarly on the original/rotated ellipsoid and Rosenbrock due to their rotational invariance. On Rastrigin functions, however, this is not the case, likely because the rotated function does not correspond to the original one.

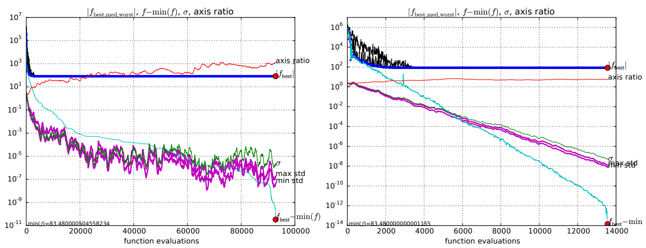

Figure 2: Single runs of IPOP-CMA-MSR (left) and IPOP-CMA-CSA (right) on one instance of the attractive sector function in 20-D. x-axis shows function evaluations. Line with dots (blue): best f-value of the iteration in absolute value, median and worst displayed in thin black lines; cyan line: difference between current f-value and f_{opt}; green line: step-size σ_t, largest and smallest coordinate-wise standard deviation of the sample distribution in purple; red line: square root of the condition number of the covariance matrix.

Empirical cumulative distribution functions. Figures 3 and 4 show the empirical cumulative distribution functions (ECDFs) of the number of function evaluations for 50 targets in dimensions 5 and 20 respectively. In 5-D, the ECDFs are quite similar for moderate and ill-conditioned functions. On separable functions, MSR solves about 82% of the problems for the fixed budget ($10^5 \times D$) while TPA and CSA solve about 73%. On multi-modal functions, TPA and CSA manage to solve all problems while MSR solves about 88% of the problems. While no algorithm solves all weakly structured multi-modal problems, TPA and CSA solve up to 76% of the problems for the maximum budget while MSR only solves about 56%. On the overall set of functions, TPA, MSR, and CSA solve roughly the same proportion of problems up to $10^4 \times D$ function evaluations. For the maximum budget, however, TPA and CSA solve about 90% of the problems while MSR only solves about 84%. In 20-D, two main differences are observed: firstly, TPA and CSA solve about 8% (respectively 10%) less separable (respectively multi-modal) problems than in 5-D (none of them managed to solve function 3 in 20-D). Secondly, CSA is better than MSR and TPA on weakly structured multi-modal problems and solves about 50% of the problems, being 10% more than MSR and 13% more than TPA.

7. DISCUSSION

We evaluated IPOP-CMA-ES with two different and relatively new step-size adaptation schemes, TPA and MSR, on the BBOB noiseless continuous functions. We then compared them to IPOP-CMA-ES with the standard step-size adaptation method, CSA. As expected, empirical results showed that the three algorithms need nearly the same number of function evaluations in average to solve the target $f_{\mathrm{t}} = f_{\mathrm{opt}} + 10^{-8}$ on a large number of functions. However, significant differences were observed, the most notable were on the attractive sector function where TPA and CSA outperformed MSR in large dimensions and on Rastrigin where MSR was the best. 16 functions out of 24 were solved by all the algorithms in all dimensions while some multi-modal and weekly structured multi-modal functions remained unsolved because the chosen budget ($10^5 \times D$ function evaluations) was insufficient. On the other hand, the performance was comparable to the best BBOB-09 results on 13 functions for at least one algorithm, generally in large dimensions.

Acknowledgments

This work was supported by the NumBBO project (grant ANR-2012-MONU-0009) of the French National Research Agency. The author would like to thank Nikolaus Hansen and Anne Auger for their comments on this work.

8. REFERENCES

[1] O. Ait Elhara, A. Auger, and N. Hansen. A median success rule for non-elitist evolution strategies: Study of feasibility. In *Genetic and Evolutionary Computation Conference (GECCO 2013), Proceedings*, pages 415–422. ACM, 2013.

[2] A. Auger and N. Hansen. A restart CMA evolution strategy with increasing population size. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2005)*, pages 1769–1776. IEEE Press, 2005.

[3] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Presentation of the noiseless functions. Technical Report 2009/20, Research Center PPE, 2009. Updated February 2010.

[4] N. Hansen, D. Arnold, and A. Auger. Evolution strategies. To appear in Janusz Kacprzyk and Witold Pedrycz (Eds.): *Handbook of Computational Intelligence*, Springer.

[5] N. Hansen, A. Atamna, and A. Auger. How to assess step-size adaptation mechanisms in randomised search. In *Parallel Problem Solving from Nature - PPSN XIII, Proceedings*, pages 60–69. Springer, 2014.

[6] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2012: Experimental setup. Technical report, INRIA, 2012.

[7] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2009. Updated February 2010.

[8] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.

[9] K. Price. Differential evolution vs. the functions of the second ICEO. In *Proceedings of the IEEE International Congress on Evolutionary Computation*, pages 153–157, 1997.

[10] I. Rechenberg. *Evolutionsstrategie '94*. Frommann-Holzboog Verlag, 1994.

[11] http://coco.gforge.inria.fr/doku.php?id=bbob-2015-results.

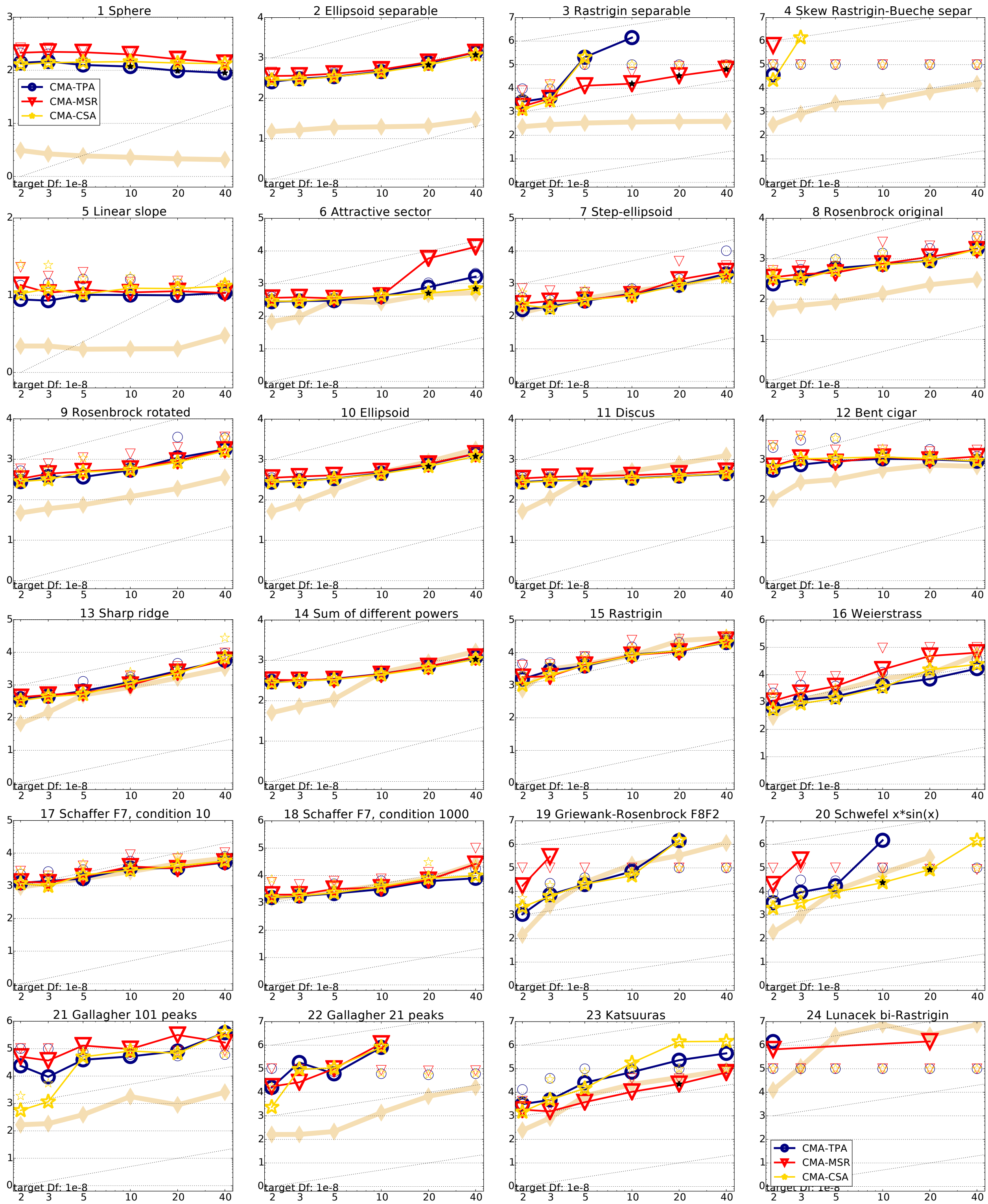

Figure 1: Expected running time (ERT in number of f-evaluations as $\log_{10}$ value), divided by dimension for target function value 10^{-8} versus dimension. Slanted grid lines indicate quadratic scaling with the dimension. Different symbols correspond to different algorithms given in the legend of f_1 and f_{24}. Light symbols give the maximum number of function evaluations from the longest trial divided by dimension. Black stars indicate a statistically better result compared to all other algorithms with $p < 0.01$ and Bonferroni correction number of dimensions (six). Legend: ∘:CMA-TPA, ▽:CMA-MSR, ⋆:CMA-CSA

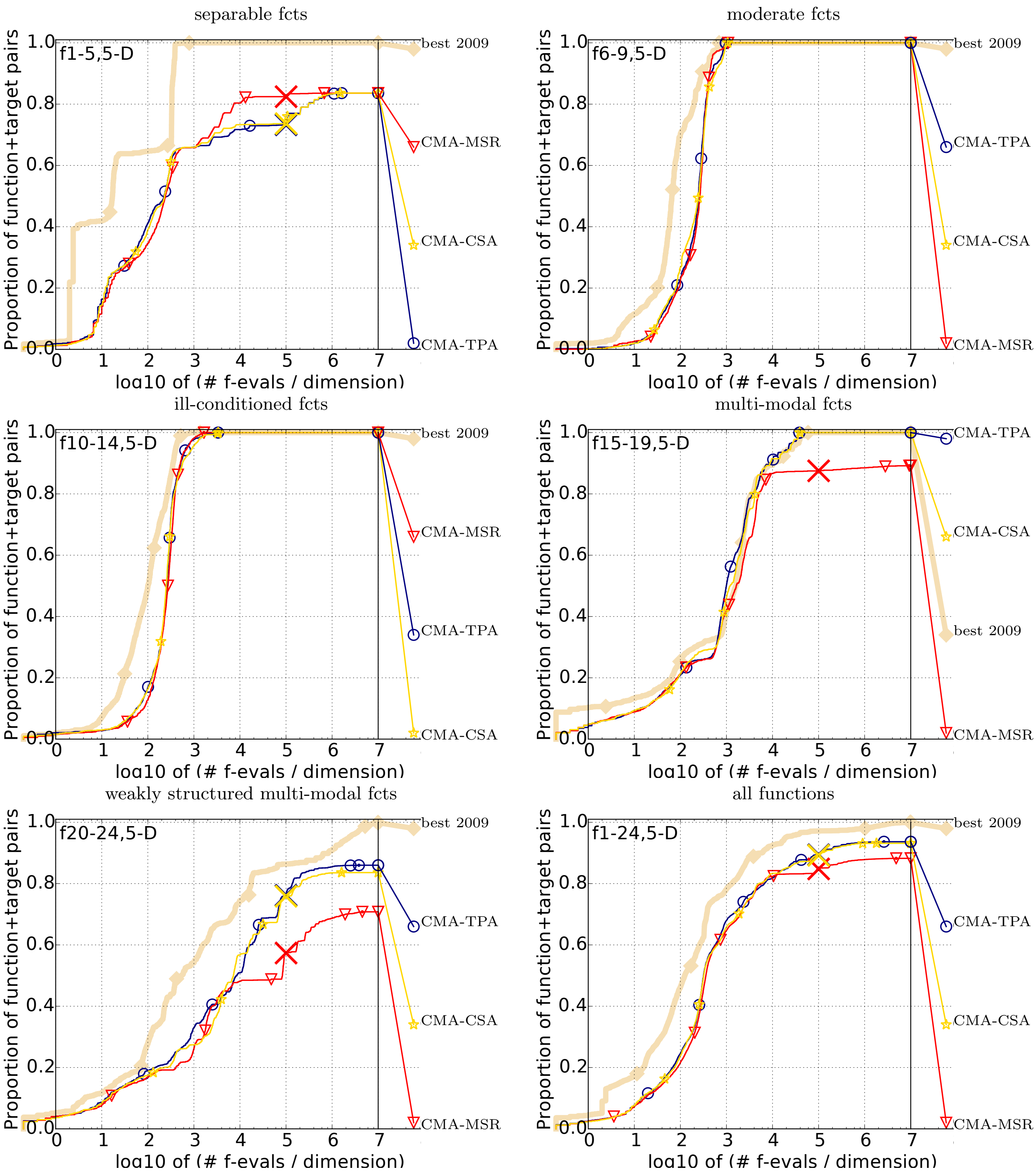

Figure 3: Bootstrapped empirical cumulative distribution of the number of objective function evaluations divided by dimension (FEvals/DIM) for 50 targets in $10^{[-8..2]}$ for all functions and subgroups in 5-D. The "best 2009" line corresponds to the best ERT observed during BBOB 2009 for each single target.

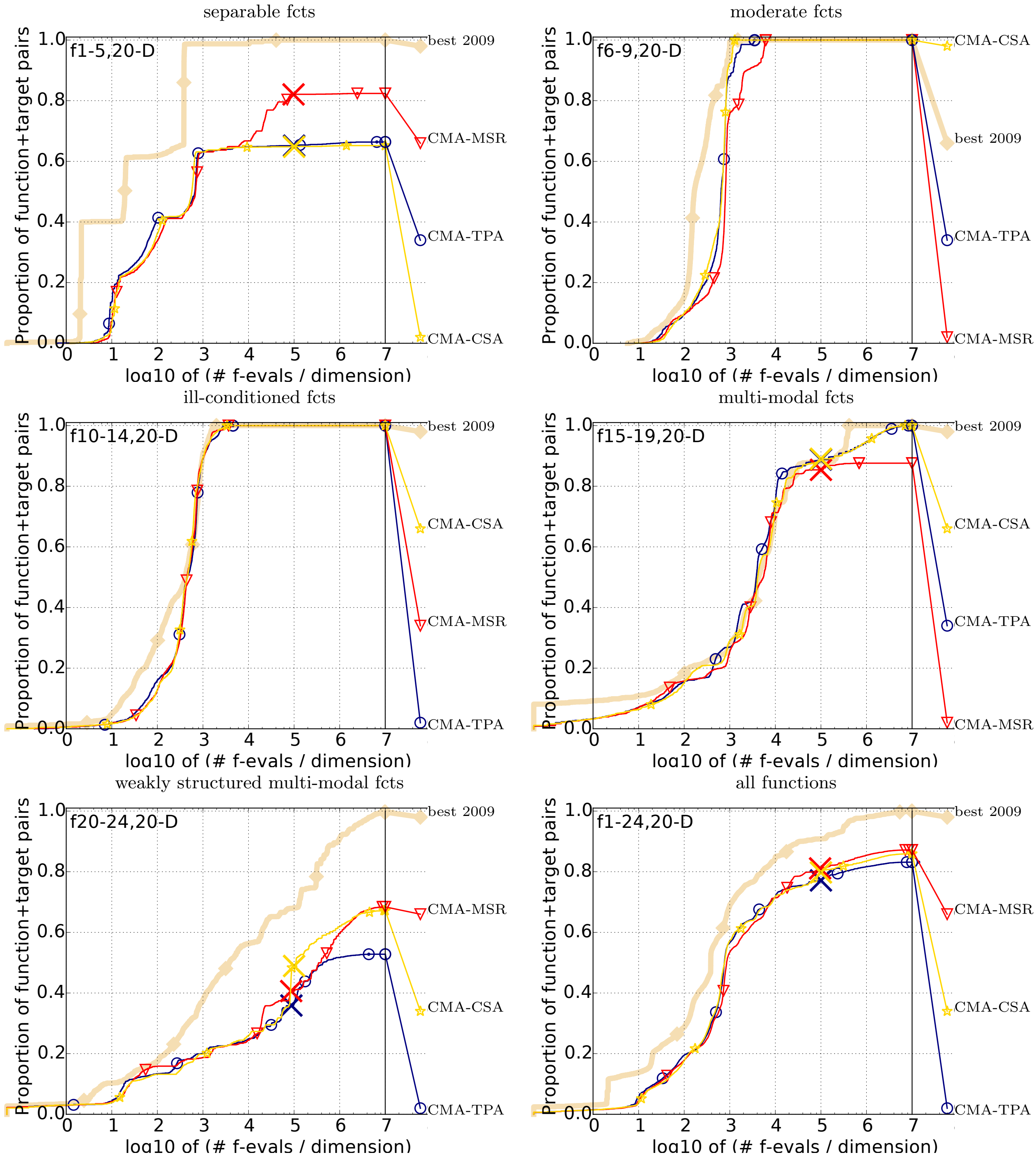

Figure 4: Bootstrapped empirical cumulative distribution of the number of objective function evaluations divided by dimension (FEvals/DIM) for 50 targets in $10^{[-8..2]}$ for all functions and subgroups in 20-D. The "best 2009" line corresponds to the best ERT observed during BBOB 2009 for each single target.

Δf_{opt}	1e1	1e0	1e-1	1e-2	1e-3	1e-5	1e-7	#succ
f1	11	12	12	12	12	12	12	15/15
CMA-TPA	**3.2**(2)	**9.2**(4)	**14**(6)	**20**(5)	**24**(4)	**36**(4)	**47**(8)	15/15
CMA-MSR	3.6(4)	12(3)	21(5)	31(2)	41(6)	62(9)	82(6)	15/15
CMA-CSA	3.8(3)	10(4)	16(3)	22(4)	28(3)	40(3)	52(5)	15/15
f2	83	87	88	89	90	92	94	15/15
CMA-TPA	**10**(2)	**12**(2)	14(0.5)	15(3)	15(2)	17(3)	18(3)	15/15
CMA-MSR	12(3)	13(2)	14(2)	15(3)	16(2)	18(2)	20(1)	15/15
CMA-CSA	11(2)	13(2)	**14**(1)	**14**(2)	**15**(1)	**16**(1)	**17**(2)	15/15
f3	716	1622	1637	1642	1646	1650	1654	15/15
CMA-TPA	**0.81**(0.7)	9.3(10)	632(925)	630(1153)	629(926)	628(766)	627(458)	5/15
CMA-MSR	1.7(2)	**5.7**(2)	**36**(86)	**36**(154)	**36**(155)	**37**(164)	**38**(12)	14/15
CMA-CSA	1.4(1)	32(82)	623(1075)	622(460)	621(837)	619(840)	618(607)	5/15
f4	809	1633	1688	1758	1817	1886	1903	15/15
CMA-TPA	2.7(4)	∞	∞	∞	∞	∞	∞ *5e5*	0/15
CMA-MSR	2.2(1)	∞	∞	∞	∞	∞	∞ *5e5*	0/15
CMA-CSA	**2.2**(3)	∞	∞	∞	∞	∞	∞ *5e5*	0/15
f5	10	10	10	10	10	10	10	15/15
CMA-TPA	4.0(2)	**5.0**(2)	**5.1**(2)	**5.1**(2)	**5.1**(2)	**5.1**(2)	**5.1**(2)	15/15
CMA-MSR	4.2(2)	5.8(3)	5.9(2)	5.9(3)	5.9(2)	5.9(2)	5.9(2)	15/15
CMA-CSA	**3.6**(0.9)	5.0(2)	5.2(2)	5.2(2)	5.2(2)	5.2(3)	5.2(2)	15/15
f6	114	214	281	404	580	1038	1332	15/15
CMA-TPA	2.2(0.9)	1.9(0.2)	**1.9**(0.7)	**1.7**(0.5)	**1.4**(0.3)	**1.0**(0.2)	**1.0**(0.1)	15/15
CMA-MSR	2.5(0.6)	2.0(0.5)	2.1(0.4)	1.9(0.3)	1.6(0.1)	1.2(0.1)	1.2(0.2)	15/15
CMA-CSA	**2.0**(0.9)	**1.9**(0.3)	2.0(0.4)	1.8(0.1)	1.5(0.3)	1.2(0.2)	1.1(0.2)	15/15
f7	24	324	1171	1451	1572	1572	1597	15/15
CMA-TPA	**4.1**(2)	**0.98**(1)	0.93(0.5)	0.86(0.4)	0.82(0.3)	0.82(0.3)	**0.83**(0.7)	15/15
CMA-MSR	5.3(5)	1.1(0.7)	0.94(0.4)	0.90(0.4)	0.90(0.6)	0.90(0.6)	0.92(0.5)	15/15
CMA-CSA	4.8(2)	1.3(1)	**0.87**(0.8)	**0.80**(0.9)	**0.80**(0.8)	**0.80**(0.7)	0.86(0.6)	15/15
f8	73	273	336	372	391	410	422	15/15
CMA-TPA	4.0(2)	6.0(4)	6.1(3)	6.2(2)	6.3(3)	6.5(3)	6.7(3)	15/15
CMA-MSR	4.6(3)	**3.6**(2)	**4.1**(1)	**4.3**(1)	**4.3**(1)	**4.7**(0.7)	**5.1**(0.5)	15/15
CMA-CSA	**3.0**(0.8)	5.1(5)	5.3(4)	5.4(4)	5.5(3)	5.7(2)	6.0(4)	15/15
f9	35	127	214	263	300	335	369	15/15
CMA-TPA	**5.4**(2)	**5.8**(3)	**5.2**(2)	**5.0**(1)	**4.8**(1)	**4.9**(1)	**4.8**(0.9)	15/15
CMA-MSR	7.2(0.7)	9.4(7)	7.5(2)	6.8(6)	6.3(3)	6.3(5)	6.4(5)	15/15
CMA-CSA	5.7(0.7)	10(11)	7.7(7)	7.1(4)	6.7(0.5)	6.5(5)	6.4(4)	15/15
f10	349	500	574	607	626	829	880	15/15
CMA-TPA	**2.5**(0.4)	2.2(0.2)	2.1(0.2)	2.1(0.2)	2.2(0.1)	1.8(0.1)	1.8(0.1)	15/15
CMA-MSR	2.6(0.6)	**2.1**(0.5)	2.1(0.3)	2.2(0.2)	2.3(0.2)	2.0(0.2)	2.2(0.1)	15/15
CMA-CSA	2.5(0.4)	2.1(0.2)	**2.0**(0.2)	**2.0**(0.1)	**2.1**(0.2)	**1.8**(0.1)	**1.8**(0.1)	15/15
f11	143	202	763	977	1177	1467	1673	15/15
CMA-TPA	5.1(0.9)	4.6(0.7)	1.3(0.1)	1.1(0.1)	1.0(0.1)	**0.91**(0.1)	0.89(0.1)	15/15
CMA-MSR	5.9(0.7)	5.0(0.3)	1.5(0.2)	1.3(0.2)	1.2(0.1)	1.1(0.1)	1.1(0.1)	15/15
CMA-CSA	**4.9**(1.0)	**4.3**(0.6)	**1.3**(0.2)	**1.1**(0.2)	**1.00**(0.1)	0.91(0.1)	**0.88**(0.1)	15/15
f12	108	268	371	413	461	1303	1494	15/15
CMA-TPA	8.3(5)	6.1(8)	6.0(3)	6.2(8)	6.2(1)	2.7(3)	2.9(3)	15/15
CMA-MSR	**7.7**(4)	**5.4**(6)	**5.5**(2)	**5.8**(3)	**6.0**(4)	**2.7**(1)	**2.8**(1)	15/15
CMA-CSA	10(12)	7.1(7)	6.9(9)	7.2(8)	7.4(7)	3.5(1)	3.5(5)	15/15

Δf_{opt}	1e1	1e0	1e-1	1e-2	1e-3	1e-5	1e-7	#succ
f13	132	195	250	319	1310	1752	2255	15/15
CMA-TPA	**2.9**(0.9)	3.8(1)	4.2(2)	4.0(1)	1.2(0.2)	1.3(0.4)	1.2(0.2)	15/15
CMA-MSR	3.2(0.5)	3.6(0.7)	**3.8**(0.7)	4.0(0.7)	1.2(0.1)	1.2(0.1)	1.1(0.1)	15/15
CMA-CSA	3.3(0.8)	**3.4**(2)	4.1(1)	**3.9**(0.9)	**1.1**(0.2)	**1.1**(0.2)	**1.1**(0.2)	15/15
f14	10	41	58	90	139	251	476	15/15
CMA-TPA	2.1(1)	3.3(2)	3.7(1)	3.9(1)	3.9(0.9)	4.0(0.5)	3.1(0.5)	15/15
CMA-MSR	2.5(1)	3.4(2)	4.7(0.7)	5.0(1)	4.4(0.9)	4.1(0.4)	3.1(0.3)	15/15
CMA-CSA	**1.7**(2)	**2.7**(1)	**3.6**(0.8)	**3.7**(0.8)	**3.8**(0.7)	**3.9**(0.6)	**3.0**(0.3)	15/15
f15	511	9310	19369	19743	20073	20769	21359	15/15
CMA-TPA	1.9(2)	**0.90**(0.5)	**0.87**(0.6)	**0.88**(0.6)	**0.88**(0.7)	**0.88**(0.6)	**0.89**(0.5)	15/15
CMA-MSR	1.9(2)	0.95(0.8)	0.89(0.7)	0.89(0.6)	0.91(0.6)	0.93(0.6)	0.95(0.8)	15/15
CMA-CSA	**1.1**(0.9)	1.1(0.8)	0.91(0.3)	0.92(0.4)	0.92(0.2)	0.92(0.5)	0.92(0.3)	15/15
f16	120	612	2662	10163	10449	11644	12095	15/15
CMA-TPA	**1.7**(1)	3.1(3)	1.8(1)	0.56(0.3)	0.62(0.8)	0.62(0.6)	0.65(0.3)	15/15
CMA-MSR	5.9(7)	5.8(5)	4.7(4)	1.6(1)	1.6(1)	1.5(2)	1.5(2)	15/15
CMA-CSA	2.2(1)	**1.9**(1)	**1.4**(1)	**0.49**(0.3)	**0.54**(0.2)	**0.55**(0.3)	**0.56**(0.3)	15/15
f17	5.2	215	899	2861	3669	6351	7934	15/15
CMA-TPA	24(3)	2.6(2)	1.6(2)	0.97(0.4)	0.94(0.3)	**0.88**(0.3)	**1.0**(0.6)	15/15
CMA-MSR	**4.2**(2)	**0.93**(0.2)	0.97(0.6)	**0.83**(0.6)	**0.82**(0.5)	0.96(0.8)	1.1(0.5)	15/15
CMA-CSA	4.2(6)	0.98(0.3)	**0.53**(0.3)	1.0(0.2)	1.2(0.5)	1.1(0.6)	1.3(0.4)	15/15
f18	103	378	3968	8451	9280	10905	12469	15/15
CMA-TPA	**0.92**(0.5)	**1.8**(2)	0.67(1)	0.59(0.4)	**0.69**(0.4)	**0.70**(0.3)	**0.85**(0.3)	15/15
CMA-MSR	1.1(0.7)	5.0(6)	1.0(2)	0.70(0.3)	1.0(0.8)	1.2(0.8)	1.3(1)	15/15
CMA-CSA	1.3(2)	2.4(0.1)	**0.61**(0.4)	**0.54**(0.6)	0.74(0.5)	0.77(0.4)	0.90(0.8)	15/15
f19	1	1	242	1.0e5	1.2e5	1.2e5	1.2e5	15/15
CMA-TPA	25(21)	**959**(777)	**84**(62)	**0.68**(0.7)	**0.78**(0.5)	**0.80**(0.7)	**0.80**(0.6)	15/15
CMA-MSR	31(60)	2573(3243)	306(78)	67(63)	∞	∞	∞ *5e5*	0/15
CMA-CSA	**19**(12)	2971(3103)	153(107)	0.86(0.7)	0.83(0.7)	0.83(0.7)	0.84(0.6)	15/15
f20	16	851	38111	51362	54470	54861	55313	15/15
CMA-TPA	3.9(2)	17(17)	2.0(0.5)	1.5(0.5)	1.5(0.5)	1.5(0.7)	1.5(0.9)	15/15
CMA-MSR	4.8(0.8)	1666(2186)	∞	∞	∞	∞	∞ *5e5*	0/15
CMA-CSA	**3.7**(1)	**9.2**(9)	**1.1**(0.5)	**0.83**(0.4)	**0.80**(0.6)	**0.82**(0.5)	**0.84**(0.5)	15/15
f21	41	1157	1674	1692	1705	1729	1757	14/15
CMA-TPA	2.2(0.5)	88(75)	**116**(213)	**115**(425)	**114**(216)	**113**(332)	**112**(177)	10/15
CMA-MSR	5.3(23)	206(6)	388(196)	384(439)	382(577)	377(491)	371(430)	6/15
CMA-CSA	**1.9**(1)	**55**(150)	119(226)	148(319)	147(155)	145(278)	143(207)	9/15
f22	71	386	938	980	1008	1040	1068	14/15
CMA-TPA	**2.5**(6)	223(4)	**323**(820)	**310**(534)	**301**(348)	**292**(409)	**285**(305)	8/15
CMA-MSR	14(13)	457(1052)	531(574)	508(663)	494(951)	479(519)	467(1081)	7/15
CMA-CSA	4.1(1)	**135**(138)	345(479)	426(534)	535(782)	519(629)	507(413)	6/15
f23	3.0	518	14249	27890	31654	33030	34256	15/15
CMA-TPA	3.2(2)	16(23)	8.1(4)	4.2(2)	3.8(5)	3.8(8)	3.7(18)	13/15
CMA-MSR	2.5(2)	**3.2**(6)*	**0.91**(0.6)	**0.52**(0.6)	**0.48**(0.4)	**0.51**(0.5)	**0.53**(0.6)	15/15
CMA-CSA	**2.3**(3)	13(15)	4.7(0.8)	2.5(2)	2.2(2)	2.2(2)	2.1(1)	15/15
f24	1622	2.2e5	6.4e6	9.6e6	9.6e6	1.3e7	1.3e7	3/15
CMA-TPA	**1.3**(2)	**10**(20)	∞	∞	∞	∞	∞ *5e5*	0/15
CMA-MSR	1.3(1)	33(42)	**1.1**(1)	∞	∞	∞	∞ *5e5*	0/15
CMA-CSA	2.0(2)	∞	∞	∞	∞	∞	∞ *5e5*	0/15

Table 1: Expected running time (ERT in number of function evaluations) divided by the respective best ERT measured during BBOB-2009 in dimension 5. The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and target, the corresponding best ERT in the first row. The different target Δf-values are shown in the top row. #succ is the number of trials that reached the (final) target $f_{\mathrm{opt}} + 10^{-8}$. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries, succeeded by a star, are statistically significantly better (according to the rank-sum test) when compared to all other algorithms of the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

Δf_{opt}	1e1	1e0	1e-1	1e-2	1e-3	1e-5	1e-7	#succ
f1	43	43	43	43	43	43	43	15/15
CMA-TPA	**6.4**(1)*	**11**(1)$^{\star 3}$	**15**(2)$^{\star 3}$	**19**(1)$^{\star 4}$	**24**(2)$^{\star 4}$	**32**(2)$^{\star 4}$	**41**(2)$^{\star 4}$	15/15
CMA-MSR	9.2(1)	16(1.0)	23(3)	30(3)	38(3)	53(3)	68(4)	15/15
CMA-CSA	7.7(1)	14(2)	20(1)	26(2)	32(2)	45(3)	57(4)	15/15
f2	385	386	387	388	390	391	393	15/15
CMA-TPA	25(3)	30(2)	33(2)	35(1)	36(1)	37(2)	37(1)	15/15
CMA-MSR	27(4)	32(4)	35(2)	36(2)	37(2)	38(3)	39(2)	15/15
CMA-CSA	**23**(2)	**27**(2)*	**29**(0.9)$^{\star 3}$	**30**(1)$^{\star 3}$	**31**(1.0)$^{\star 3}$	**32**(2)$^{\star 3}$	**33**(1)$^{\star 3}$	15/15
f3	5066	7626	7635	7637	7643	7646	7651	15/15
CMA-TPA	8.8(5)	1756(2177)	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-MSR	**6.4**(1)	**38**(20)$^{\star 3}$	**70**(45)$^{\star 4}$	**73**(56)$^{\star 4}$	**76**(41)$^{\star 4}$	**81**(36)$^{\star 4}$	**86**(68)$^{\star 4}$	15/15
CMA-CSA	10(8)	∞	∞	∞	∞	∞	∞ *2e6*	0/15
f4	4722	7628	7666	7686	7700	7758	1.4e5	9/15
CMA-TPA	∞	∞	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-MSR	**5792**(3817)	∞	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-CSA	∞	∞	∞	∞	∞	∞	∞ *2e6*	0/15
f5	41	41	41	41	41	41	41	15/15
CMA-TPA	**4.3**(0.9)	**4.9**(2)	**4.9**(1)	**4.9**(0.8)	**4.9**(1)	**4.9**(0.9)	**4.9**(0.8)	15/15
CMA-MSR	5.0(1)	5.5(2)	5.6(0.6)	5.6(1)	5.6(1)	5.6(1)	5.6(0.8)	15/15
CMA-CSA	4.9(1)	5.8(0.9)	6.0(1)	6.0(1)	6.0(1)	6.0(0.9)	6.0(1)	15/15
f6	1296	2343	3413	4255	5220	6728	8409	15/15
CMA-TPA	1.6(0.4)	1.3(0.2)	1.2(0.3)	1.3(0.3)	1.4(0.3)	1.5(0.4)	1.6(0.5)	15/15
CMA-MSR	**1.5**(0.7)	1.9(2)	2.4(2)	3.9(4)	5.7(4)	11(6)	13(1)	15/15
CMA-CSA	1.6(0.3)	**1.3**(0.2)	**1.1**(0.2)	**1.1**(0.2)	**1.1**(0.1)*	**1.1**(0.1)$^{\star 2}$	**1.1**(0.1)$^{\star 2}$	15/15
f7	1351	4274	9503	16523	16524	16524	16969	15/15
CMA-TPA	2.1(1)	2.7(0.7)	**1.6**(0.6)	**1.0**(0.4)	**1.0**(0.4)	**1.0**(0.4)	**1.0**(0.4)	15/15
CMA-MSR	2.1(0.7)	4.2(1)	2.4(1)	1.6(0.5)	1.6(2)	1.6(0.3)	1.5(0.6)	15/15
CMA-CSA	**1.7**(1)	**2.3**(1)	1.7(0.6)	1.1(0.4)	1.1(0.3)	1.1(0.4)	1.0(0.3)	15/15
f8	2039	3871	4040	4148	4219	4371	4484	15/15
CMA-TPA	**3.1**(0.7)	3.5(0.1)	3.8(1)	3.9(1)	3.9(0.2)	3.9(1)	3.9(0.4)	15/15
CMA-MSR	3.6(0.8)	4.6(3)	4.8(3)	4.8(3)	4.8(0.4)	4.8(0.5)	4.9(3)	15/15
CMA-CSA	3.4(0.7)	**3.4**(0.4)	**3.6**(0.2)	**3.7**(0.3)	**3.8**(0.5)	**3.8**(0.2)	**3.8**(0.2)	15/15
f9	1716	3102	3277	3379	3455	3594	3727	15/15
CMA-TPA	3.8(0.7)	5.5(8)	5.8(0.4)	5.8(8)	5.8(2)	5.8(7)	5.8(1)	15/15
CMA-MSR	**3.8**(0.9)	4.5(0.5)	4.8(3)	4.8(0.4)	4.8(2)	4.8(0.3)	4.8(2)	15/15
CMA-CSA	3.8(0.4)	**4.1**(0.3)	**4.3**(0.3)	**4.4**(0.2)	**4.4**(0.2)	**4.5**(0.2)	**4.5**(0.4)	15/15
f10	7413	8661	10735	13641	14920	17073	17476	15/15
CMA-TPA	1.4(0.1)	1.4(0.1)	1.2(0.1)	1.0(0.1)	0.95(0.1)	0.86(0.0)	0.86(0.0)	15/15
CMA-MSR	1.3(0.2)	1.3(0.1)	1.2(0.1)	0.99(0.1)	0.93(0.1)	0.86(0.0)	0.88(0.0)	15/15
CMA-CSA	**1.2**(0.2)	**1.2**(0.1)	**1.0**(0.1)$^{\star 2}$	**0.86**(0.0)$^{\star 3}$	**0.81**(0.0)$^{\star 3}$	**0.74**(0.0)$^{\star 4}$	**0.76**(0.0)$^{\star 3}$	15/15
f11	1002	2228	6278	8586	9762	12285	14831	15/15
CMA-TPA	**4.5**(0.3)	2.3(0.1)	0.89(0.0)	0.69(0.0)	0.65(0.0)	0.57(0.0)	0.51(0.0)	15/15
CMA-MSR	4.7(0.5)	2.6(0.2)	1.0(0.1)	0.80(0.0)	0.74(0.0)	0.65(0.0)	0.58(0.0)	15/15
CMA-CSA	4.6(0.2)	**2.3**(0.1)	**0.86**(0.0)	**0.67**(0.0)	**0.63**(0.0)*	**0.55**(0.0)	**0.50**(0.0)	15/15
f12	1042	1938	2740	3156	4140	12407	13827	15/15
CMA-TPA	3.8(3)	4.1(2)	3.8(2)	3.9(1)	3.3(2)	1.4(0.3)	1.4(0.5)	15/15
CMA-MSR	3.7(4)	**3.3**(2)	**3.5**(1)	**3.6**(2)	**3.2**(2)	**1.3**(0.5)	**1.4**(0.4)	15/15
CMA-CSA	**3.6**(2)	3.5(2)	3.8(2)	3.9(1)	3.5(1)	1.4(0.3)	1.5(0.4)	15/15
f13	652	2021	2751	3507	18749	24455	30201	15/15
CMA-TPA	4.7(5)	4.7(3)	5.0(3)	5.4(2)	1.1(0.7)	1.3(0.6)	1.5(0.5)	15/15
CMA-MSR	4.4(4)	**3.3**(4)	4.9(3)	**4.2**(2)	**0.87**(0.4)	**1.0**(0.4)	1.5(0.8)	15/15
CMA-CSA	**3.2**(4)	4.2(3)	**4.0**(3)	4.5(1)	0.93(0.4)	1.1(0.6)	**1.3**(0.7)	15/15
f14	75	239	304	451	932	1648	15661	15/15
CMA-TPA	**3.5**(1)	**2.3**(0.6)	**2.8**(0.8)$^{\star 2}$	**3.1**(0.4)*	**2.8**(0.3)	**3.8**(0.4)	0.71(0.1)	15/15
CMA-MSR	4.2(1)	2.8(0.5)	3.4(0.5)	3.6(0.2)	2.9(0.2)	3.9(0.4)	0.73(0.0)	15/15
CMA-CSA	4.2(1)	2.9(0.5)	3.7(0.2)	4.1(0.3)	3.3(0.3)	3.9(0.3)	**0.67**(0.1)	15/15
f15	30378	1.5e5	3.1e5	3.2e5	3.2e5	4.5e5	4.6e5	15/15
CMA-TPA	0.94(0.2)	1.1(0.2)	0.63(0.1)	0.64(0.2)	0.64(0.2)	0.48(0.0)	0.49(0.2)	15/15
CMA-MSR	0.98(0.3)	**0.95**(0.2)	**0.54**(0.5)	**0.55**(0.1)	**0.56**(0.3)	**0.43**(0.2)	**0.45**(0.3)	15/15
CMA-CSA	**0.83**(0.6)	0.99(0.3)	0.64(0.2)	0.65(0.1)	0.65(0.2)	0.49(0.2)	0.49(0.2)	15/15
f16	1384	27265	77015	1.4e5	1.9e5	2.0e5	2.2e5	15/15
CMA-TPA	1.2(0.4)	0.78(0.4)	**0.80**(0.4)	**0.67**(0.4)	**0.63**(0.5)	**0.66**(0.4)	**0.62**(0.2)	15/15
CMA-MSR	**0.80**(0.1)$^{\star 2}$	0.84(0.6)	1.1(0.4)	1.3(1)	3.3(6)	4.7(4)	4.3(5)	12/15
CMA-CSA	1.9(0.5)	**0.64**(0.6)	0.84(0.4)	1.2(1)	1.4(0.6)	1.5(1)	1.4(1)	15/15
f17	63	1030	4005	12242	30677	56288	80472	15/15
CMA-TPA	**2.7**(1)	1.4(0.6)	1.5(0.9)	**0.94**(0.6)	**0.74**(0.6)	**0.71**(0.4)	**0.80**(0.3)	15/15
CMA-MSR	2.7(0.5)	6.5(3)	3.5(2)	1.9(2)	0.97(0.4)	0.88(0.2)	0.81(0.3)	15/15
CMA-CSA	3.0(2)	**1.0**(0.2)	**1.4**(2)	1.2(0.7)	0.74(0.4)	0.88(0.3)	0.88(0.2)	15/15
f18	621	3972	19561	28555	67569	1.3e5	1.5e5	15/15
CMA-TPA	1.6(3)	1.3(0.8)	**0.77**(0.4)	**0.96**(0.2)	**0.57**(0.3)	**0.58**(0.5)	**0.74**(0.3)	15/15
CMA-MSR	2.8(14)	2.8(2)	1.4(0.4)	2.0(0.8)	1.2(0.5)	0.83(0.3)	0.87(0.3)	15/15
CMA-CSA	**0.96**(0.3)	**0.72**(0.1)	0.81(0.4)	1.1(0.6)	0.83(0.4)	1.1(2)	1.0(0.3)	15/15
f19	1	1	3.4e5	4.7e6	6.2e6	6.7e6	6.7e6	15/15
CMA-TPA	**177**(45)	**1.9e4**(8426)	1.6(0.8)	1.2(1)	4.7(6)	**4.3**(5)	**4.3**(4)	1/15
CMA-MSR	212(72)	3.5e4(2e4)	1.2(0.4)	∞	∞	∞	∞ *2e6*	0/15
CMA-CSA	221(81)	3.3e4(9280)	**0.82**(0.4)	**0.56**(0.4)	**2.4**(2)	4.5(3)	4.5(3)	1/15
f20	82	46150	3.1e6	5.5e6	5.5e6	5.6e6	5.6e6	15/15
CMA-TPA	**4.0**(0.7)	∞	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-MSR	5.1(1)	∞	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-CSA	5.0(1)	**2.5**(0.2)$^{\star 4}$	**0.35**(0.1)$^{\star 4}$	**0.29**(9e-3)$^{\star 4}$	**0.29**(0.0)$^{\star 4}$	**0.29**(4e-3)$^{\star 4}$	**0.30**(0.1)$^{\star 4}$	15/15
f21	561	6541	14103	14318	14643	15567	17589	15/15
CMA-TPA	63(187)	248(674)	115(240)	114(95)	111(133)	105(157)	93(165)	6/15
CMA-MSR	**24**(86)	278(305)	449(364)	442(533)	433(508)	407(453)	360(361)	3/15
CMA-CSA	113(403)	**159**(110)	**95**(61)	**94**(182)	**92**(59)	**87**(89)	**77**(110)	7/15
f22	467	5580	23491	24163	24948	26847	1.3e5	12/15
CMA-TPA	162(11)	216(94)	∞	∞	∞	∞	∞ *1e6*	0/15
CMA-MSR	254(876)	249(632)	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-CSA	**22**(26)	**145**(279)	∞	∞	∞	∞	∞ *1e6*	0/15
f23	3.2	1614	67457	3.7e5	4.9e5	8.1e5	8.4e5	15/15
CMA-TPA	6.5(5)	23(25)	4.8(12)	3.0(3)	9.3(10)	5.6(13)	5.5(5)	5/15
CMA-MSR	6.8(5)	**2.0**(2)$^{\star 2}$	**0.79**(0.5)*	**0.74**(0.3)	**0.73**(0.2)$^{\star 2}$	**0.49**(0.2)$^{\star 2}$	**0.51**(0.1)*	15/15
CMA-CSA	**6.1**(5)	93(33)	13(13)	16(18)	58(59)	35(27)	34(53)	1/15
f24	1.3e6	7.5e6	5.2e7	5.2e7	5.2e7	5.2e7	5.2e7	15/15
CMA-TPA	6.2(8)	3.9(3)	∞	∞	∞	∞	∞ *2e6*	0/15
CMA-MSR	**4.1**(5)	**1.8**(2)	**0.55**(0.7)	**0.55**(0.4)	**0.55**(1)	**0.55**(0.6)	**0.55**(0.5)	1/15
CMA-CSA	6.0(7)	∞	∞	∞	∞	∞	∞ *2e6*	0/15

Table 2: Expected running time (ERT in number of function evaluations) divided by the respective best ERT measured during BBOB-2009 in dimension 20. The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and target, the corresponding best ERT in the first row. The different target Δf-values are shown in the top row. #succ is the number of trials that reached the (final) target $f_{\mathrm{opt}} + 10^{-8}$. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries, succeeded by a star, are statistically significantly better (according to the rank-sum test) when compared to all other algorithms of the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.