Black-Box Optimization Benchmarking Template for Noiseless Function Testbed

Draft version *

Forename Name

ABSTRACT

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

General Terms

Algorithms

Keywords

Benchmarking, Black-box optimization

1. RESULTS

Results of SMBO from experiments according to [**?**] on the benchmark functions given in [**?**, **?**] are presented in Figures 1, 2 and 4 and in Tables 1 and 3.

2. REFERENCES

[1] 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.

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

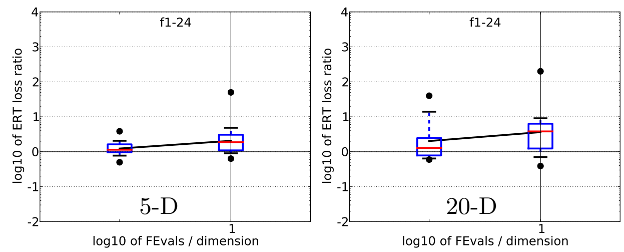
[3] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2009. Updated February 2010.

*Submission deadline: March 28th.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'13, July 6-10, 2013, Amsterdam, The Netherlands.

Copyright 2013 ACM TBA ...$15.00.



#FEs/D	best	10%	25%	**med**	75%	90%
	f_1–f_{24} in 5-D, maxFE/D=10					
2	0.50	0.75	0.95	1.1	1.6	2.2
10	0.63	0.86	1.1	1.9	3.1	5.2
100	0.86	3.1	4.7	11	21	87
$\mathrm{RL_{US}}$/D	10	10	10	10	10	10
	f_1–f_{24} in 20-D, maxFE/D=10					
#FEs/D	best	10%	25%	**med**	75%	90%
2	0.60	0.63	0.76	1.3	2.5	17
10	0.39	0.68	1.2	3.8	6.6	28
100	0.53	0.96	5.6	16	33	51
$\mathrm{RL_{US}}$/D	10	10	10	10	10	10

Figure 3: ERT **loss ratio versus the budget (both in number of f-evaluations divided by dimension). The target value f_t for a given budget** FEvals **is the best target f-value reached within the budget by the given algorithm. Shown is the** ERT **of the given algorithm divided by best** ERT **seen in GECCO-BBOB-2009 for the target f_t, or, if the best algorithm reached a better target within the budget, the budget divided by the best** ERT. **Line: geometric mean. Box-Whisker error bar: 25-75%-ile with median (box), 10-90%-ile (caps), and minimum and maximum** ERT **loss ratio (points). The vertical line gives the maximal number of function evaluations in a single trial in this function subset. See also Figure 4 for results on each function subgroup.**

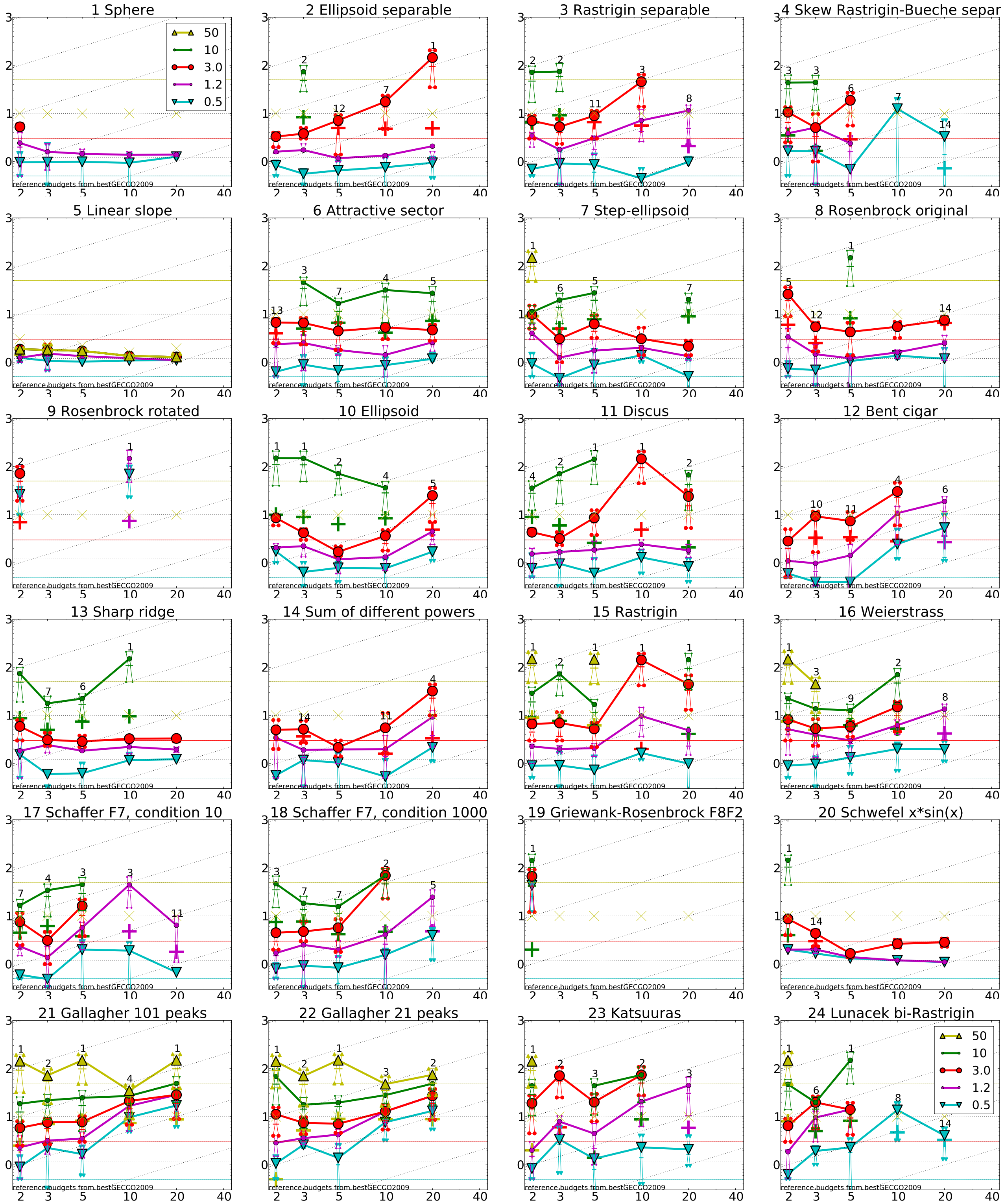


Figure 1: Expected number of f-evaluations (ERT, lines) to reach $f_{\mathrm{opt}} + \Delta f$; median number of f-evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of f-evaluations in any trial (×); interquartile range with median (notched boxes) of simulated runlengths to reach $f_{\mathrm{opt}} + \Delta f$; all values are divided by dimension and plotted as $\log_{10}$ values versus dimension. Shown is the ERT for targets just not reached by the GECCO-BBOB-2009 best algorithm within the given budget kDIM, where k is shown in the legend. Numbers above ERT-symbols indicate the number of trials reaching the respective target. Slanted grid lines indicate a scaling with $\mathcal{O}(\mathrm{DIM})$ compared to $\mathcal{O}(1)$ when using the respective 2009 best algorithm.

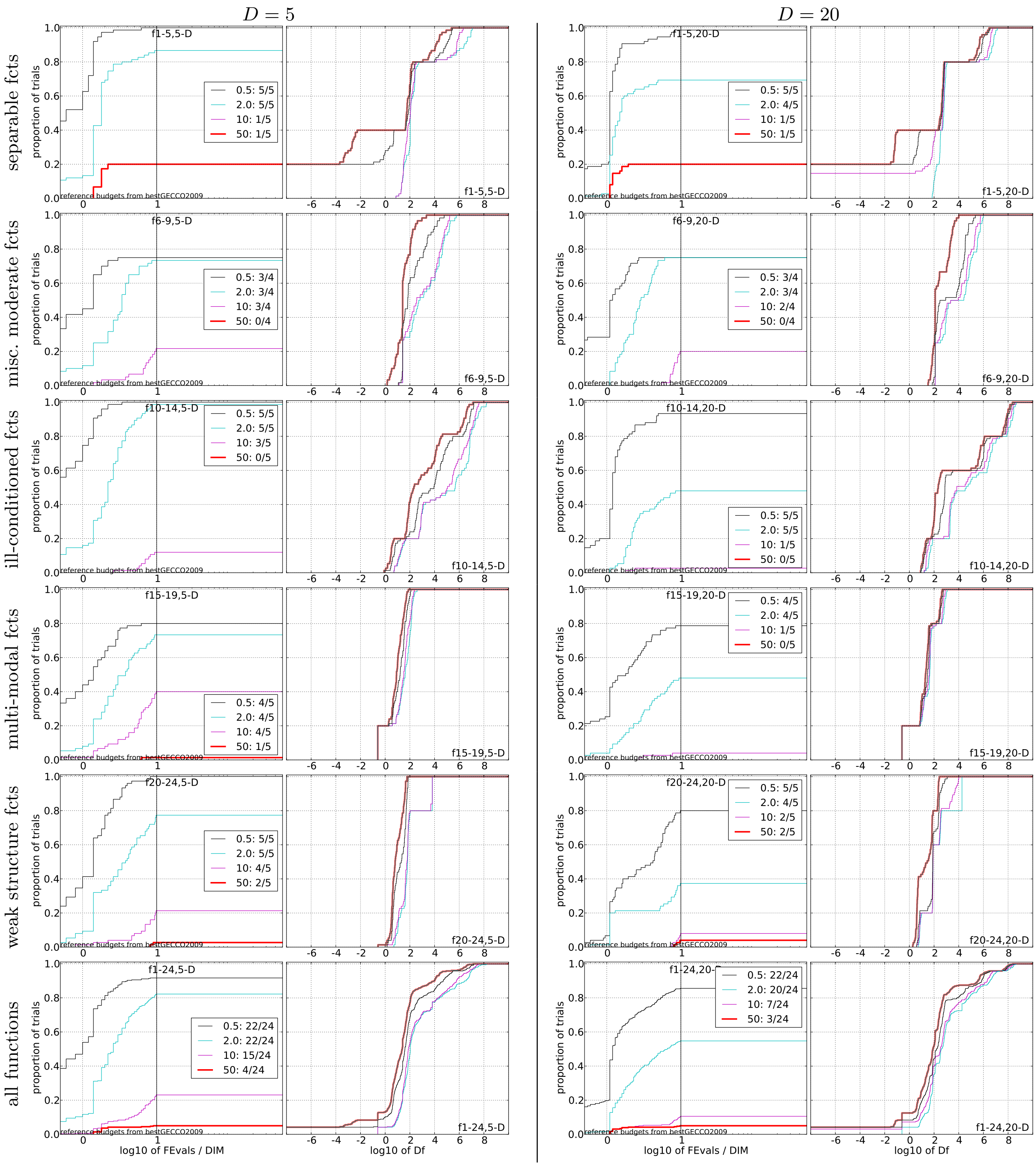


Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the x-axis. Left subplots: ECDF of number of function evaluations (FEvals) divided by search space dimension D, to fall below $f_{\mathrm{opt}} + \Delta f$ where Δf is the target just not reached by the GECCO-BBOB-2009 best algorithm within a budget of $k \times \mathrm{DIM}$ evaluations, where k is the first value in the legend. Legends indicate for each target the number of functions that were solved in at least one trial within the displayed budget. Right subplots: ECDF of the best achieved Δf for running times of $0.5D, 1.2D, 3D, 10D, 100D, 1000D, \ldots$ function evaluations (from right to left cycling cyan-magenta-black...) and final Δf-value (red), where Δf and Df denote the difference to the optimal function value.

5-D

#FEs/D	0.5	1.2	3.0	10	50	#succ
f_1	*2.5e+1*:4.8	*1.6e+1*:7.6	*1.0e-8*:12	*1.0e-8*:12	*1.0e-8*:12	15/15
	1.0(0.8)	0.96(0.5)	∞	∞	∞ *50*	0/15
f_2	*1.6e+6*:2.9	*4.0e+5*:11	*4.0e+4*:15	*6.3e+2*:58	*1.0e-8*:95	15/15
	1.1(1)	0.53(0.3)	2.4(2)	∞	∞ *50*	0/15
f_3	*1.6e+2*:4.1	*1.0e+2*:15	*6.3e+1*:23	*2.5e+1*:73	*1.0e+1*:716	15/15
	1.0(0.8)	1.0(1)	1.9(2)	∞	∞ *50*	0/15
f_4	*2.5e+2*:2.6	*1.6e+2*:10	*1.0e+2*:19	*4.0e+1*:65	*1.6e+1*:434	15/15
	1.3(1)	1.2(2)	5.0(5)	∞	∞ *50*	0/15
f_5	*6.3e+1*:4.0	*4.0e+1*:10	*1.0e-8*:10	*1.0e-8*:10	*1.0e-8*:10	15/15
	1.3(0.6)	0.68(0.4)	0.86(0.2)	0.86(0.2)	0.86(0.2)	15/15
f_6	*1.0e+5*:3.0	*2.5e+4*:8.4	*1.0e+2*:16	*2.5e+1*:54	*2.5e-1*:254	15/15
	1.1(1)	1.1(0.7)	1.4(1)	1.5(2)	∞ *50*	0/15
f_7	*1.6e+2*:4.2	*1.0e+2*:6.2	*2.5e+1*:20	*4.0e+0*:54	*1.0e+0*:324	15/15
	1.0(1)	1.4(1)	1.5(1)	2.5(2)	∞ *50*	0/15
f_8	*1.0e+4*:4.6	*6.3e+3*:6.8	*1.0e+3*:18	*6.3e+1*:54	*1.6e+0*:258	15/15
	1.1(0.9)	0.88(0.6)	1.2(1)	14(15)	∞ *50*	0/15
f_9	*2.5e+1*:20	*1.6e+1*:26	*1.0e+1*:35	*4.0e+0*:62	*1.6e-2*:256	15/15
	∞	∞	∞	∞	∞ *50*	0/15
f_{10}	*2.5e+6*:2.9	*6.3e+5*:7.0	*2.5e+5*:17	*6.3e+3*:54	*2.5e+1*:297	15/15
	1.4(1)	0.85(0.7)	0.50(0.4)	6.7(7)	∞ *50*	0/15
f_{11}	*1.0e+6*:3.0	*6.3e+4*:6.2	*6.3e+2*:16	*6.3e+1*:74	*6.3e-1*:298	15/15
	1.0(1)	1.5(1)	2.7(3)	10(12)	∞ *50*	0/15
f_{12}	*4.0e+7*:3.6	*1.6e+7*:7.6	*4.0e+6*:19	*1.6e+4*:52	*1.0e+0*:268	15/15
	0.56(0.6)	0.94(0.9)	1.9(2)	∞	∞ *50*	0/15
f_{13}	*1.0e+3*:2.8	*6.3e+2*:8.4	*4.0e+2*:17	*6.3e+1*:52	*6.3e-2*:264	15/15
	1.1(1)	1.1(0.8)	0.86(0.3)	2.2(2)	∞ *50*	0/15
f_{14}	*1.6e+1*:3.0	*1.0e+1*:10	*6.3e+0*:15	*2.5e-1*:53	*1.0e-5*:251	15/15
	1.7(2)	1(0.8)	0.70(0.5)	∞	∞ *50*	0/15
f_{15}	*1.6e+2*:3.0	*1.0e+2*:13	*6.3e+1*:24	*4.0e+1*:55	*1.6e+1*:289	5/5
	1.2(2)	0.80(0.6)	1.1(0.8)	1.5(1)	2.5(3)	1/15
f_{16}	*4.0e+1*:4.8	*2.5e+1*:16	*1.6e+1*:46	*1.0e+1*:120	*4.0e+0*:334	15/15
	1.4(1)	0.95(0.7)	**0.63**(0.6)$^{\downarrow}$	0.53(0.5)	∞ *50*	0/15
f_{17}	*1.0e+1*:5.2	*6.3e+0*:26	*4.0e+0*:57	*2.5e+0*:110	*6.3e-1*:412	15/15
	1.9(2)	1.1(1)	1.4(2)	2.1(2)	∞ *50*	0/15
f_{18}	*6.3e+1*:3.4	*4.0e+1*:7.2	*2.5e+1*:20	*1.6e+1*:58	*1.6e+0*:318	15/15
	1.2(1)	1.4(1.0)	1.4(1)	1.3(1)	∞ *50*	0/15
f_{19}	*1.6e-1*:172	*1.0e-1*:242	*6.3e-2*:675	*4.0e-2*:3078	*2.5e-2*:4946	15/15
	∞	∞	∞	∞	∞ *50*	0/15
f_{20}	*6.3e+3*:5.1	*4.0e+3*:8.4	*4.0e+1*:15	*2.5e+0*:69	*1.0e+0*:851	15/15
	1.3(0.2)	**0.83**$^{\downarrow}$	**0.55**(0.1)$^{\downarrow 4}$	∞	∞ *50*	0/15
f_{21}	*4.0e+1*:3.9	*2.5e+1*:11	*1.6e+1*:31	*6.3e+0*:73	*1.6e+0*:347	5/5
	2.1(2)	1.7(2)	1.3(1)	1.7(2)	2.1(2)	1/15
f_{22}	*6.3e+1*:3.6	*4.0e+1*:15	*2.5e+1*:32	*1.0e+1*:71	*1.6e+0*:341	5/5
	1.9(2)	1.4(0.9)	1.1(1.0)	1.4(1)	2.2(2)	1/15
f_{23}	*1.0e+1*:3.0	*6.3e+0*:9.0	*4.0e+0*:33	*2.5e+0*:84	*1.0e+0*:518	15/15
	2.2(2)	2.5(3)	3.0(4)	2.6(3)	∞ *50*	0/15
f_{24}	*6.3e+1*:15	*4.0e+1*:37	*4.0e+1*:37	*2.5e+1*:118	*1.6e+1*:692	15/15
	0.79(0.8)	1.9(2)	1.9(2)	6.3(7)	∞ *50*	0/15

20-D

#FEs/D	0.5	1.2	3.0	10	50	#succ
f_1	*6.3e+1*:24	*4.0e+1*:42	*1.0e-8*:43	*1.0e-8*:43	*1.0e-8*:43	15/15
	1.0(0.2)	**0.67**(0.1)$^{\downarrow 4}$	∞	∞	∞ *200*	0/15
f_2	*4.0e+6*:29	*2.5e+6*:42	*1.0e+5*:65	*1.0e+4*:207	*1.0e-8*:412	15/15
	0.66(0.5)	0.99(0.6)	45(49)	∞	∞ *200*	0/15
f_3	*6.3e+2*:33	*4.0e+2*:44	*1.6e+2*:109	*1.0e+2*:255	*2.5e+1*:3277	15/15
	0.60(0.4)$^{\downarrow 3}$	5.3(7)	∞	∞	∞ *200*	0/15
f_4	*6.3e+2*:22	*4.0e+2*:91	*2.5e+2*:250	*1.6e+2*:332	*6.3e+1*:1927	15/15
	3.1(4)	∞	∞	∞	∞ *200*	0/15
f_5	*2.5e+2*:19	*1.6e+2*:34	*1.0e-8*:41	*1.0e-8*:41	*1.0e-8*:41	15/15
	1.2	**0.64**$^{\downarrow 4}$	**0.63**(0.1)$^{\downarrow 4}$	**0.63**(0.1)$^{\downarrow 4}$	**0.63**(0.1)$^{\downarrow 4}$	15/15
f_6	*2.5e+5*:16	*6.3e+4*:43	*1.6e+4*:62	*1.6e+2*:353	*1.6e+1*:1078	15/15
	1.5(1)	1.2(0.6)	1.5(0.9)	1.5(1)	∞ *200*	0/15
f_7	*1.0e+3*:11	*4.0e+2*:39	*2.5e+2*:74	*6.3e+1*:319	*1.0e+1*:1351	15/15
	0.96(1)	0.68(0.3)	**0.58**(0.3)$^{\downarrow}$	1.3(1.0)	∞ *200*	0/15
f_8	*4.0e+4*:19	*2.5e+4*:35	*4.0e+3*:67	*2.5e+2*:231	*1.6e+1*:1470	15/15
	1.3(1)	1.4(0.8)	2.2(0.7)	∞	∞ *200*	0/15
f_9	*1.0e+2*:357	*6.3e+1*:560	*4.0e+1*:684	*2.5e+1*:756	*1.0e+1*:1716	15/15
	∞	∞	∞	∞	∞ *200*	0/15
f_{10}	*1.6e+6*:15	*1.0e+6*:27	*4.0e+5*:70	*6.3e+4*:231	*4.0e+3*:1015	15/15
	2.3(2)	3.3(4)	7.1(7)	∞	∞ *200*	0/15
f_{11}	*4.0e+4*:11	*2.5e+3*:27	*1.6e+2*:313	*1.0e+2*:481	*1.0e+1*:1002	15/15
	1.5(0.9)	1.3(0.5)	1.5(2)	2.8(3)	∞ *200*	0/15
f_{12}	*1.0e+8*:23	*6.3e+7*:39	*2.5e+7*:76	*4.0e+6*:209	*1.0e+1*:1042	15/15
	4.6(5)	10(10)	∞	∞	∞ *200*	0/15
f_{13}	*1.6e+3*:28	*1.0e+3*:64	*6.3e+2*:79	*4.0e+1*:211	*2.5e+0*:1724	15/15
	0.89(0.1)	**0.61**(0.2)$^{\downarrow 4}$	**0.84**(0.2)$^{\downarrow}$	∞	∞ *200*	0/15
f_{14}	*2.5e+1*:15	*1.6e+1*:42	*1.0e+1*:75	*1.6e+0*:219	*6.3e-4*:1106	15/15
	3.0(3)	4.4(5)	8.5(9)	∞	∞ *200*	0/15
f_{15}	*6.3e+2*:15	*4.0e+2*:67	*2.5e+2*:292	*1.6e+2*:846	*1.0e+2*:1671	15/15
	1.3(1)	1.5(2)	3.0(4)	3.4(4)	∞ *200*	0/15
f_{16}	*4.0e+1*:26	*2.5e+1*:127	*1.6e+1*:540	*1.6e+1*:540	*1.0e+1*:1384	15/15
	1.5(1)	2.1(2)	∞	∞	∞ *200*	0/15
f_{17}	*1.6e+1*:11	*1.0e+1*:63	*6.3e+0*:305	*4.0e+0*:468	*1.0e+0*:1030	15/15
	1.3(3)	2.0(3)	∞	∞	∞ *200*	0/15
f_{18}	*4.0e+1*:116	*2.5e+1*:252	*1.6e+1*:430	*1.0e+1*:621	*4.0e+0*:1090	15/15
	0.69(0.6)	2.0(2)	∞	∞	∞ *200*	0/15
f_{19}	*1.6e-1*:2.5e5	*1.0e-1*:3.4e5	*6.3e-2*:3.4e5	*4.0e-2*:3.4e5	*2.5e-2*:3.4e5	3/15
	∞	∞	∞	∞	∞ *200*	0/15
f_{20}	*1.6e+4*:38	*1.0e+4*:42	*2.5e+2*:62	*2.5e+0*:250	*1.6e+0*:2536	15/15
	0.57$^{\downarrow}$	**0.53**$^{\downarrow 4}$	0.91(0.3)	∞	∞ *200*	0/15
f_{21}	*6.3e+1*:36	*4.0e+1*:77	*4.0e+1*:77	*1.6e+1*:456	*4.0e+0*:1094	15/15
	10(9)	7.4(7)	7.4(7)	2.2(2)	2.7(3)	1/15
f_{22}	*6.3e+1*:45	*4.0e+1*:68	*4.0e+1*:68	*1.6e+1*:231	*6.3e+0*:1219	15/15
	5.9(5)	8.0(8)	8.0(8)	4.2(4)	1.2(1)	2/15
f_{23}	*6.3e+0*:29	*4.0e+0*:118	*2.5e+0*:306	*2.5e+0*:306	*1.0e+0*:1614	15/15
	1.5(1)	7.6(8)	∞	∞	∞ *200*	0/15
f_{24}	*2.5e+2*:208	*1.6e+2*:918	*1.0e+2*:6628	*6.3e+1*:9885	*4.0e+1*:31629	15/15
	0.39(0.2)$^{\downarrow 3}$	∞	∞	∞	∞ *200*	0/15

Table 1: Expected running time (ERT in number of function evaluations) divided by the best ERT measured during BBOB-2009. The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear in the second row of each cell, the best ERT (preceded by the target Δf-value in *italics*) in the first. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Bold entries are statistically significantly better (according to the rank-sum test) compared to the best algorithm in BBOB-2009, with $p = 0.05$ or $p = 10^{-k}$ when the number $k > 1$ is following the $\downarrow$ symbol, with Bonferroni correction by the number of functions.

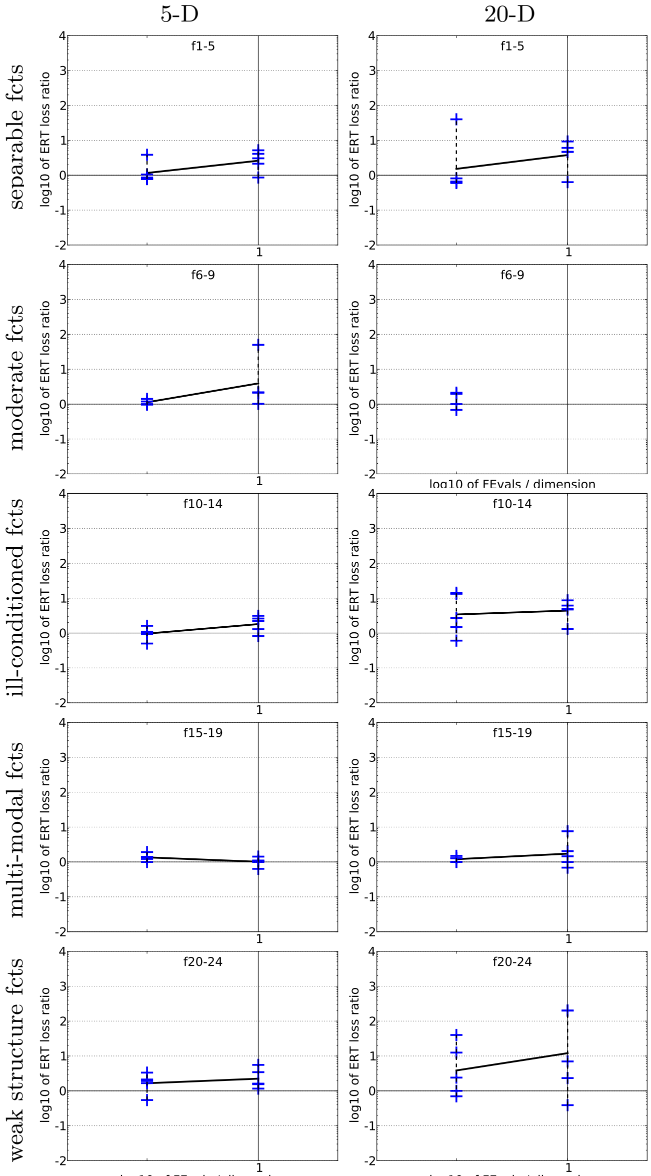


Figure 4: ERT loss ratios (see Figure 3 for details). Each cross (+) represents a single function, the line is the geometric mean.