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

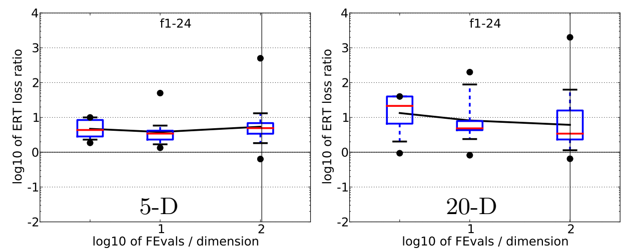
Algorithms

Keywords

Benchmarking, Black-box optimization

1. RESULTS

Results of CMA-ES from experiments according to [?] on the benchmark functions given in [?, ?] are presented in Figures ??, ?? and ?? and in Tables ?? and ??.



	$f1$–$f24$ **in 5-D**, maxFE/D=103					
#FEs/D	best	10%	25%	**med**	75%	90%
2	1.9	2.3	2.7	4.4	8.5	10
10	1.3	1.6	2.3	3.4	4.2	10
100	0.64	1.6	3.1	4.9	7.0	18
1e3	0.66	2.0	4.8	17	29	96
$\mathrm{RL}_{\mathrm{US}}$/D	1e2	1e2	1e2	1e2	1e2	1e2

	$f1$–$f24$ **in 20-D**, maxFE/D=100					
#FEs/D	best	10%	25%	**med**	75%	90%
2	0.94	1.9	6.4	21	40	40
10	0.82	2.2	4.3	4.9	8.3	1.1e2
100	0.65	1.1	2.3	3.4	16	70
1e3	1.1	2.7	4.4	11	35	4.8e2
$\mathrm{RL}_{\mathrm{US}}$/D	1e2	1e2	1e2	1e2	1e2	1e2

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 ?? for results on each function subgroup.**

*Submission deadline: March 28th.

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

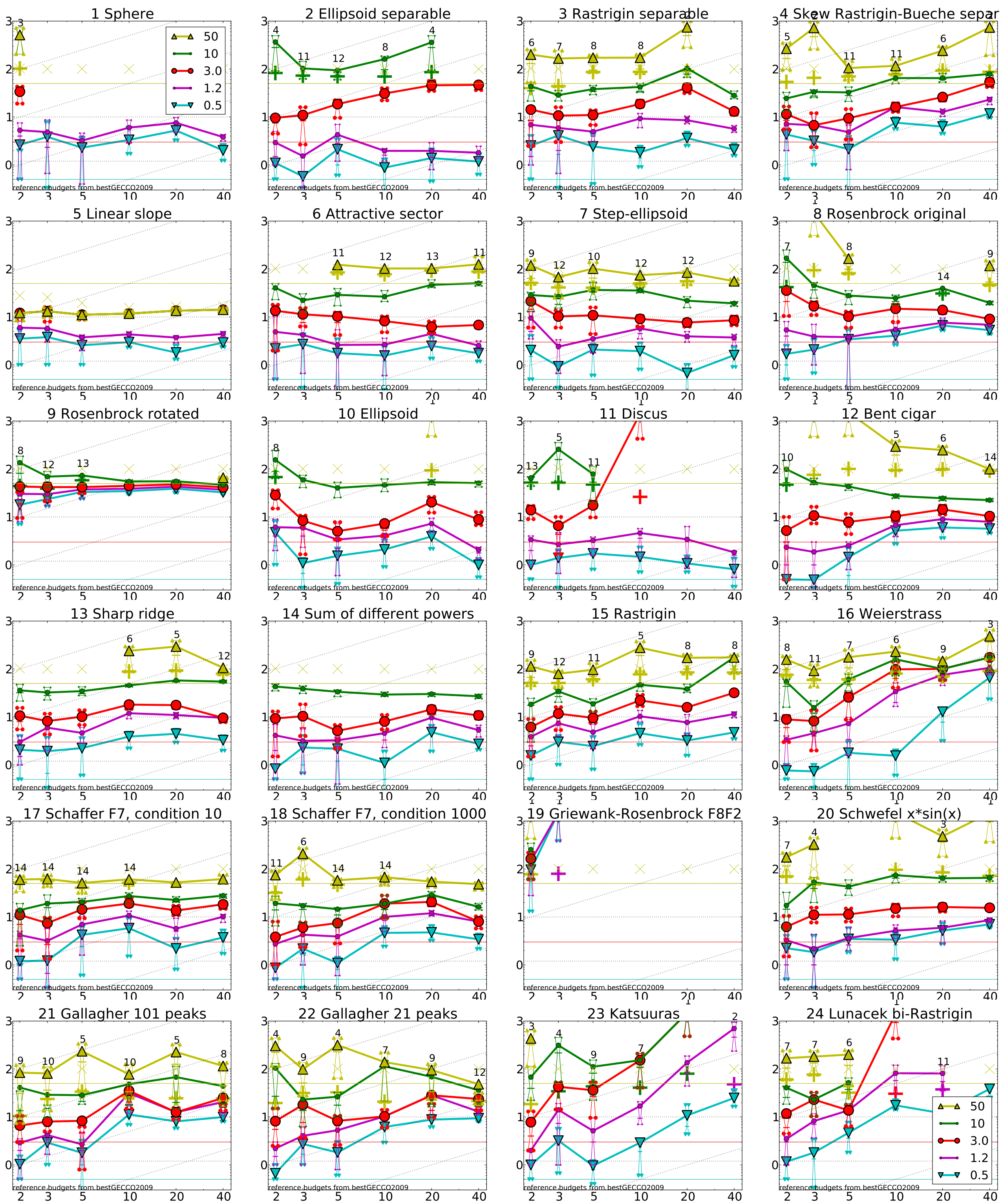


Figure 1: Expected number of f-evaluations (ERT, lines) to reach $f_{\mathrm{opt}} + \Delta f$; median number of f-evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of f-evaluations in any trial (×); interquartile range with median (notched 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.

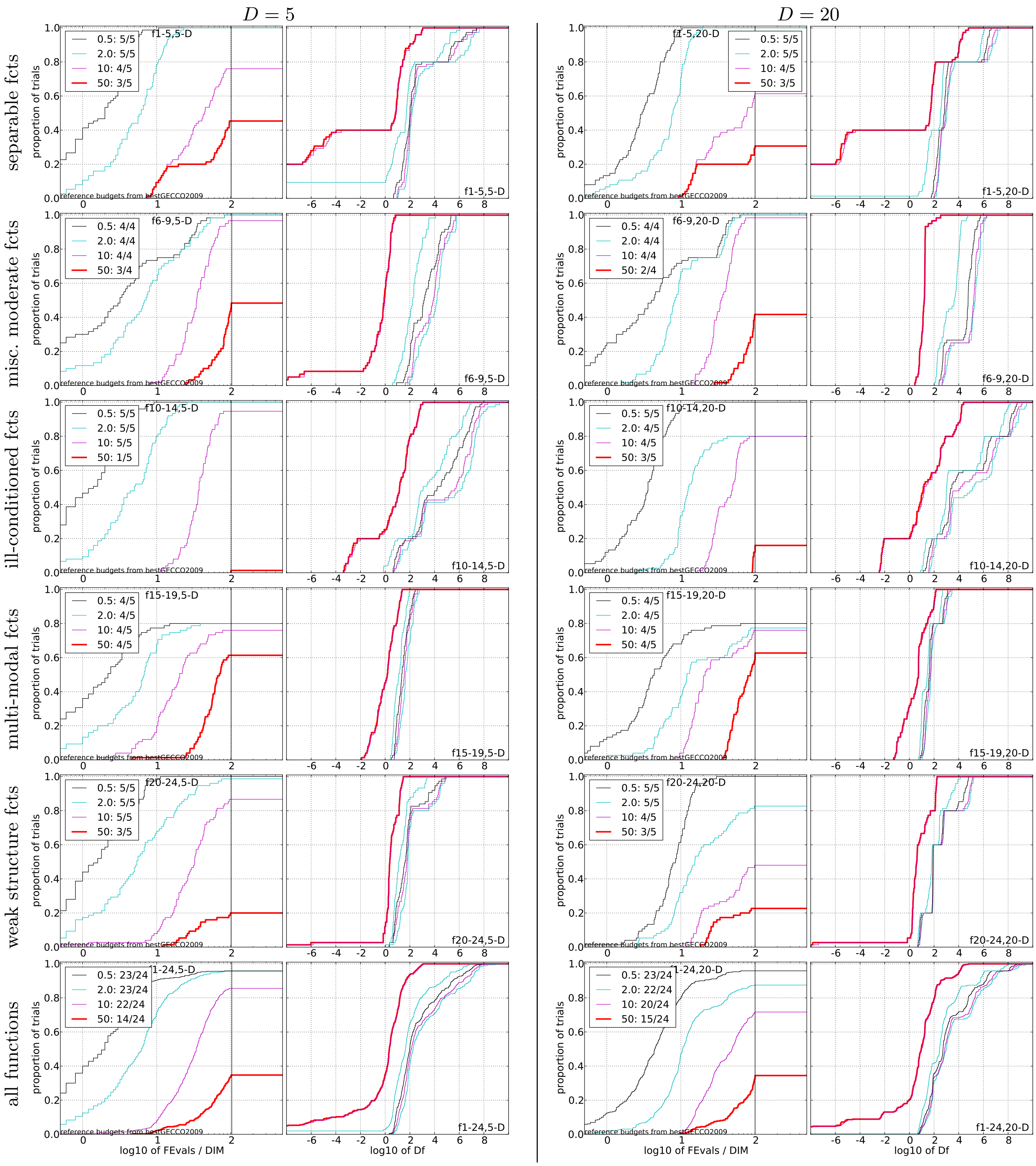


Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the x-axis. Left subplots: ECDF of number of function evaluations (FEvals) divided by search space dimension D, to fall below $f_{\text{opt}} + \Delta f$ where Δf is the target just not reached by the GECCO-BBOB-2009 best algorithm within a budget of $k \times \text{DIM}$ 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
f1	*2.5e+1*:4.8	*1.6e+1*:7.6	*1.0e-8*:12	*1.0e-8*:12	*1.0e-8*:12	15/15
	2.4(2)	2.1(2)	∞	∞	∞*510*	0/15
f2	*1.6e+6*:2.9	*4.0e+5*:11	*4.0e+4*:15	*6.3e+2*:58	*1.0e-8*:95	15/15
	3.7(5)	1.9(2)	6.1(3)	8.2(5)	∞*510*	0/15
f3	*1.6e+2*:4.1	*1.0e+2*:15	*6.3e+1*:23	*2.5e+1*:73	*1.0e+1*:716	15/15
	2.9(3)	1.7(1)	2.3(2)	2.6(1)	1.2(1)	8/15
f4	*2.5e+2*:2.6	*1.6e+2*:10	*1.0e+2*:19	*4.0e+1*:65	*1.6e+1*:434	15/15
	4.1(4)	2.5(4)	2.5(3)	2.5(1)	1.2(0.8)	11/15
f5	*6.3e+1*:4.0	*4.0e+1*:10	*1.0e-8*:10	*1.0e-8*:10	*1.0e-8*:10	15/15
	3.2(2)	1.9(2)	5.6(1)	5.6(1)	5.6(1)	15/15
f6	*1.0e+5*:3.0	*2.5e+4*:8.4	*1.0e+2*:16	*2.5e+1*:54	*2.5e-1*:254	15/15
	2.9(4)	1.6(2)	3.2(4)	2.7(2)	2.4(1)	11/15
f7	*1.6e+2*:4.2	*1.0e+2*:6.2	*2.5e+1*:20	*4.0e+0*:54	*1.0e+0*:324	15/15
	2.5(3)	2.8(3)	2.7(2)	3.4(2)	1.6(2)	10/15
f8	*1.0e+4*:4.6	*6.3e+3*:6.8	*1.0e+3*:18	*6.3e+1*:54	*1.6e+0*:258	15/15
	3.7(3)	2.8(3)	2.9(2)	2.6(1)	3.2(3)	8/15
f9	*2.5e+1*:20	*1.6e+1*:26	*1.0e+1*:35	*4.0e+0*:62	*1.6e-2*:256	15/15
	8.4(3)	7.3(3)	6.1(3)	6.0(5)	∞*510*	0/15
f10	*2.5e+6*:2.9	*6.3e+5*:7.0	*2.5e+5*:17	*6.3e+3*:54	*2.5e+1*:297	15/15
	2.7(3)	2.4(3)	1.5(1)	3.8(2)	∞*510*	0/15
f11	*1.0e+6*:3.0	*6.3e+4*:6.2	*6.3e+2*:16	*6.3e+1*:74	*6.3e-1*:298	15/15
	2.9(2)	2.6(3)	5.5(5)	5.3(5)	∞*510*	0/15
f12	*4.0e+7*:3.6	*1.6e+7*:7.6	*4.0e+6*:19	*1.6e+4*:52	*1.0e+0*:268	15/15
	2.0(3)	1.7(2)	2.1(2)	4.2(1)	28(30)	1/15
f13	*1.0e+3*:2.8	*6.3e+2*:8.4	*4.0e+2*:17	*6.3e+1*:52	*6.3e-2*:264	15/15
	4.0(4)	2.8(2)	3.0(2)	3.3(1.0)	∞*510*	0/15
f14	*1.6e+1*:3.0	*1.0e+1*:10	*6.3e+0*:15	*2.5e-1*:53	*1.0e-5*:251	15/15
	3.6(6)	1.7(2)	1.7(2)	3.2(1)	∞*510*	0/15
f15	*1.6e+2*:3.0	*1.0e+2*:13	*6.3e+1*:24	*4.0e+1*:55	*1.6e+1*:289	5/5
	4.1(4)	1.9(2)	2.0(0.9)	1.7(0.9)	1.7(1)	11/15
f16	*4.0e+1*:4.8	*2.5e+1*:16	*1.6e+1*:46	*1.0e+1*:120	*4.0e+0*:334	15/15
	1.9(2)	2.3(4)	2.8(2)	2.5(3)	2.6(2)	7/15
f17	*1.0e+1*:5.2	*6.3e+0*:26	*4.0e+0*:57	*2.5e+0*:110	*6.3e-1*:412	15/15
	4.1(6)	1.3(1)	1.3(1.0)	0.93(0.4)	**0.62**(0.3)↓	14/15
f18	*6.3e+1*:3.4	*4.0e+1*:7.2	*2.5e+1*:20	*1.6e+1*:58	*1.6e+0*:318	15/15
	1.6(2)	2.7(3)	1.9(1)	1.2(0.7)	0.92(0.3)	14/15
f19	*1.6e-1*:172	*1.0e-1*:242	*6.3e-2*:675	*4.0e-2*:3078	*2.5e-2*:4946	15/15
	∞	∞	∞	∞	∞*510*	0/15
f20	*6.3e+3*:5.1	*4.0e+3*:8.4	*4.0e+1*:15	*2.5e+0*:69	*1.0e+0*:851	15/15
	3.4(3)	2.2(2)	3.7(2)	3.1(1)	∞*510*	0/15
f21	*4.0e+1*:3.9	*2.5e+1*:11	*1.6e+1*:31	*6.3e+0*:73	*1.6e+0*:347	5/5
	2.3(3)	1.3(1)	1.3(2)	2.0(1)	3.4(4)	5/15
f22	*6.3e+1*:3.6	*4.0e+1*:15	*2.5e+1*:32	*1.0e+1*:71	*1.6e+0*:341	5/5
	2.5(3)	1.8(2)	1.3(2)	1.9(1)	4.6(5)	4/15
f23	*1.0e+1*:3.0	*6.3e+0*:9.0	*4.0e+0*:33	*2.5e+0*:84	*1.0e+0*:518	15/15
	1.6(1)	2.9(3)	5.4(5)	6.7(7)	∞*510*	0/15
f24	*6.3e+1*:15	*4.0e+1*:37	*4.0e+1*:37	*2.5e+1*:118	*1.6e+1*:692	15/15
	1.6(2)	1.9(2)	1.9(2)	2.2(2)	1.4(1)	6/15

20-D

#FEs/D	0.5	1.2	3.0	10	50	#succ
f1	*6.3e+1*:24	*4.0e+1*:42	*1.0e-8*:43	*1.0e-8*:43	*1.0e-8*:43	15/15
	4.3(2)	3.6(2)	∞	∞	∞*2000*	0/15
f2	*4.0e+6*:29	*2.5e+6*:42	*1.0e+5*:65	*1.0e+4*:207	*1.0e-8*:412	15/15
	0.97(0.9)	0.94(0.7)	14(5)	35(32)	∞*2000*	0/15
f3	*6.3e+2*:33	*4.0e+2*:44	*1.6e+2*:109	*1.0e+2*:255	*2.5e+1*:3277	15/15
	2.2(1)	3.9(2)	7.5(3)	8.1(4)	4.5(5)	2/15
f4	*6.3e+2*:22	*4.0e+2*:91	*2.5e+2*:250	*1.6e+2*:332	*6.3e+1*:1927	15/15
	5.8(3)	2.8(1)	2.1(0.7)	3.9(1)	2.5(2)	6/15
f5	*2.5e+2*:19	*1.6e+2*:34	*1.0e-8*:41	*1.0e-8*:41	*1.0e-8*:41	15/15
	1.9(1)	2.2(0.7)	6.7(1)	6.7(1)	6.7(1)	15/15
f6	*2.5e+5*:16	*6.3e+4*:43	*1.6e+4*:62	*1.6e+2*:353	*1.6e+1*:1078	15/15
	3.1(2)	2.1(0.9)	2.0(0.9)	2.6(1)	1.9(1)	13/15
f7	*1.0e+3*:11	*4.0e+2*:39	*2.5e+2*:74	*6.3e+1*:319	*1.0e+1*:1351	15/15
	1.3(1)	2.0(1)	2.1(0.8)	1.4(0.4)	1.3(0.9)	12/15
f8	*4.0e+4*:19	*2.5e+4*:35	*4.0e+3*:67	*2.5e+2*:231	*1.6e+1*:1470	15/15
	7.0(4)	4.3(2)	4.2(0.8)	3.4(2)	∞*2000*	0/15
f9	*1.0e+2*:357	*6.3e+1*:560	*4.0e+1*:684	*2.5e+1*:756	*1.0e+1*:1716	15/15
	2.2(0.5)	1.5(0.3)	1.4(0.2)	1.5(0.3)	∞*2000*	0/15
f10	*1.6e+6*:15	*1.0e+6*:27	*4.0e+5*:70	*6.3e+4*:231	*4.0e+3*:1015	15/15
	5.2(4)	5.3(3)	5.8(4)	4.6(1)	30(34)	1/15
f11	*4.0e+4*:11	*2.5e+3*:27	*1.6e+2*:313	*1.0e+2*:481	*1.0e+1*:1002	15/15
	1.9(1)	2.5(2)	∞	∞	∞*2000*	0/15
f12	*1.0e+8*:23	*6.3e+7*:39	*2.5e+7*:76	*4.0e+6*:209	*1.0e+1*:1042	15/15
	5.2(3)	4.6(2)	3.8(1)	2.4(0.4)	4.8(4)	6/15
f13	*1.6e+3*:28	*1.0e+3*:64	*6.3e+2*:79	*4.0e+1*:211	*2.5e+0*:1724	15/15
	3.2(2)	3.4(1)	4.5(0.9)	5.5(0.5)	3.4(3)	5/15
f14	*2.5e+1*:15	*1.6e+1*:42	*1.0e+1*:75	*1.6e+0*:219	*6.3e-4*:1106	15/15
	6.6(4)	4.6(3)	3.8(1)	2.8(0.6)	∞*2000*	0/15
f15	*6.3e+2*:15	*4.0e+2*:67	*2.5e+2*:292	*1.6e+2*:846	*1.0e+2*:1671	15/15
	4.2(4)	2.3(1)	1.1(0.2)	0.90(0.3)	2.1(1)	8/15
f16	*4.0e+1*:26	*2.5e+1*:127	*1.6e+1*:540	*1.6e+1*:540	*1.0e+1*:1384	15/15
	10(9)	12(8)	3.7(3)	3.7(3)	2.1(2)	9/15
f17	*1.6e+1*:11	*1.0e+1*:63	*6.3e+0*:305	*4.0e+0*:468	*1.0e+0*:1030	15/15
	4.2(4)	1.8(1)	0.89(0.3)	0.96(0.3)	1.0(0.5)	15/15
f18	*4.0e+1*:116	*2.5e+1*:252	*1.6e+1*:430	*1.0e+1*:621	*4.0e+0*:1090	15/15
	0.82(0.8)	0.95(0.4)	0.96(0.3)	0.94(0.3)	1.0(0.3)	15/15
f19	*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
	∞	∞	∞	∞	∞*2000*	0/15
f20	*1.6e+4*:38	*1.0e+4*:42	*2.5e+2*:62	*2.5e+0*:250	*1.6e+0*:2536	15/15
	2.7(1)	2.8(2)	5.2(1)	5.1(1)	3.8(4)	3/15
f21	*6.3e+1*:36	*4.0e+1*:77	*4.0e+1*:77	*1.6e+1*:456	*4.0e+0*:1094	15/15
	4.5(2)	3.3(1)	3.3(1)	3.0(4)	4.1(5)	5/15
f22	*6.3e+1*:45	*4.0e+1*:68	*4.0e+1*:68	*1.6e+1*:231	*6.3e+0*:1219	15/15
	3.9(2)	8.4(15)	8.4(15)	6.0(9)	1.6(2)	9/15
f23	*6.3e+0*:29	*4.0e+0*:118	*2.5e+0*:306	*2.5e+0*:306	*1.0e+0*:1614	15/15
	7.5(8)	23(25)	97(105)	97(105)	∞*2000*	0/15
f24	*2.5e+2*:208	*1.6e+2*:918	*1.0e+2*:6628	*6.3e+1*:9885	*4.0e+1*:31629	15/15
	1.1(0.3)	1.8(1)	∞	∞	∞*2000*	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 ↓ symbol, with Bonferroni correction by the number of functions.

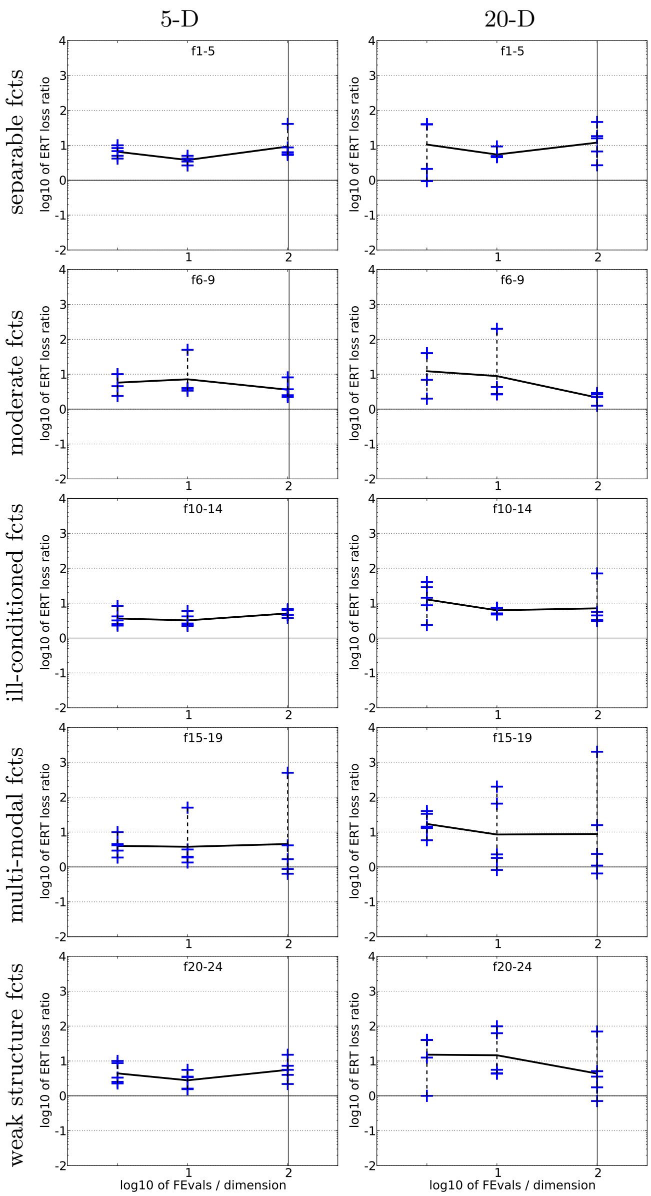


Figure 4: ERT loss ratios (see Figure ?? for details). Each cross (+) represents a single function, the line is the geometric mean.