Black-Box Optimization Benchmarking Template for Noiseless Function Testbed

Draft version [*]

Forename Name

ABSTRACT

to be written

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. CPU TIMING

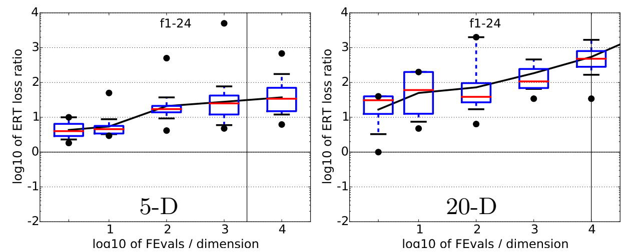
In order to evaluate the CPU timing of the algorithm, we have run the MY-ALGORITHM-NAME on the function f_8 with restarts for at least 30 seconds and until a maximum budget equal to $400(D+2)$ is reached. The code was run on a Mac Intel(R) Core(TM) i5-2400S CPU @ 2.50GHz with 1 processor and 4 cores. The time per function evaluation for dimensions 2, 3, 5, 10, 20, 40 equals $1.8{*}e^{-8}$, $1.8{*}e^{-7}$, $6.0{*}e^{-3}$, $1.9{*}e$, $2.1{*}e^{2}$, $4.4{*}e^{3}$,$1.1{*}e^{5}$ milliseconds respectively.

2. RESULTS

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

[*]Submission deadline: March 28th.

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



	f_1–f_{24} **in 5-D**, maxFE/D=2491					
#FEs/D	best	10%	25%	**med**	75%	90%
2	1.8	2.2	2.9	4.0	6.7	10
10	3.0	3.2	3.4	4.6	5.8	14
100	4.2	8.8	14	17	21	88
1e3	4.8	5.8	12	25	42	1.1e2
1e4	6.2	11	15	34	83	1.9e2
$\mathrm{RL_{US}}$/D	2e3	2e3	2e3	2e3	2e3	2e3

	f_1–f_{24} **in 20-D**, maxFE/D=9998					
#FEs/D	best	10%	25%	**med**	75%	90%
2	1.0	2.9	12	31	40	40
10	4.8	7.4	12	61	2.0e2	2.0e2
100	6.4	17	27	39	99	2.0e3
1e3	34	63	69	1.1e2	2.7e2	2.4e3
1e4	34	1.4e2	2.7e2	4.8e2	8.0e2	2.1e3
1e5	34	2.9e2	8.5e2	3.1e3	6.8e3	2.1e4
$\mathrm{RL_{US}}$/D	1e4	1e4	1e4	1e4	1e4	1e4

Figure 3: ERT loss ratio versus the budget in number of f-evaluations divided by dimension. For each given budget FEvals, the target value f_t is computed as the best target f-value reached within the budget by the given algorithm. Shown is then the ERT to reach f_t for the given algorithm or the budget, if the GECCO-BBOB-2009 best algorithm reached a better target within the budget, divided by the best ERT seen in GECCO-BBOB-2009 to reach f_t. 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.

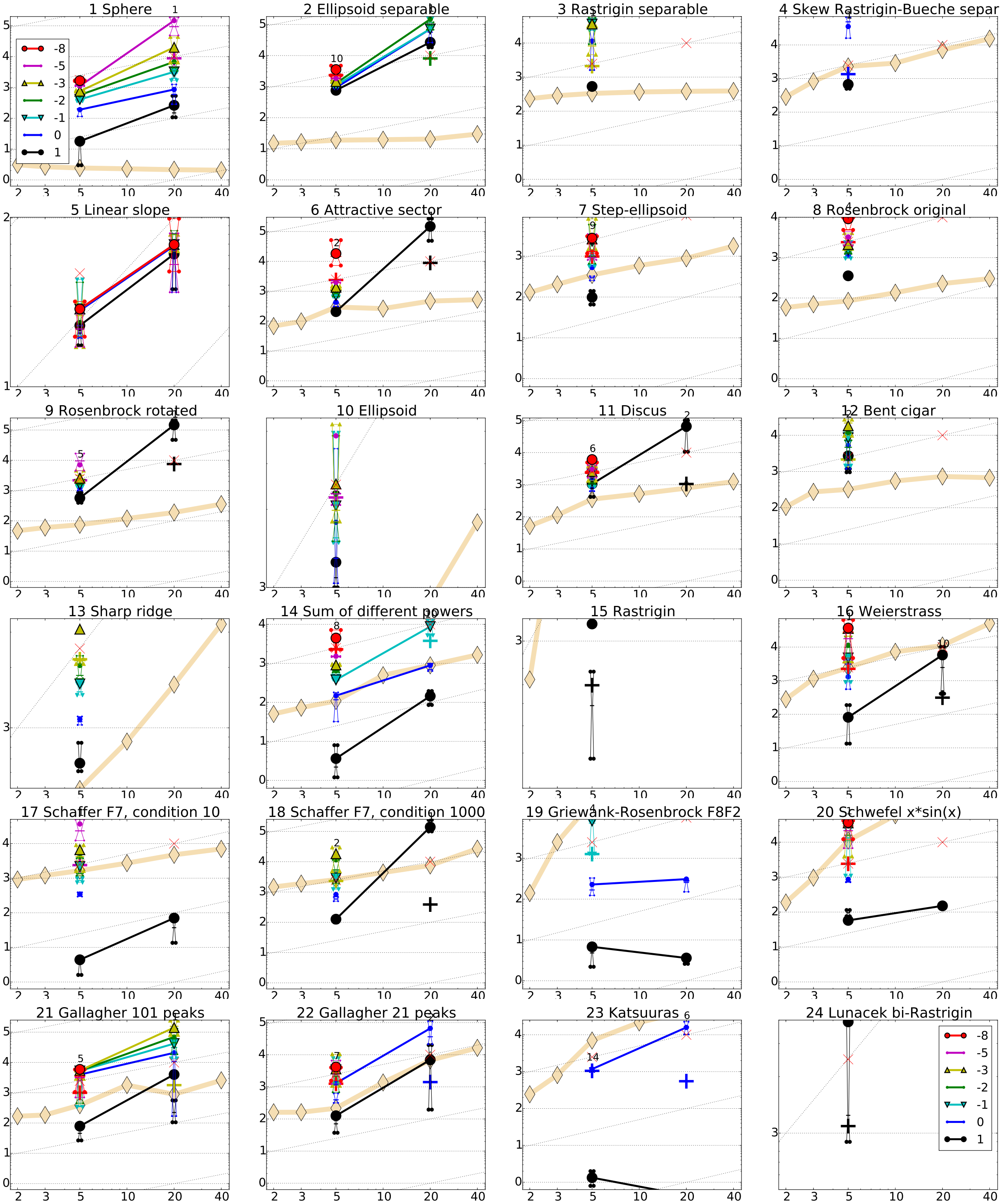


Figure 1: Expected number of f-evaluations (ERT, lines) to reach $f_{\mathrm{opt}} + \Delta f$; median number of f-evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of f-evaluations in any trial (×); interquartile range with median (notched 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 are $\Delta f = 10^{\{1,0,-1,-2,-3,-5,-8\}}$. Numbers above ERT-symbols (if appearing) indicate the number of trials reaching the respective target. The light thick line with diamonds indicates the respective best result from BBOB-2009 for $\Delta f = 10^{-8}$. Horizontal lines mean linear scaling, slanted grid lines depict quadratic scaling.

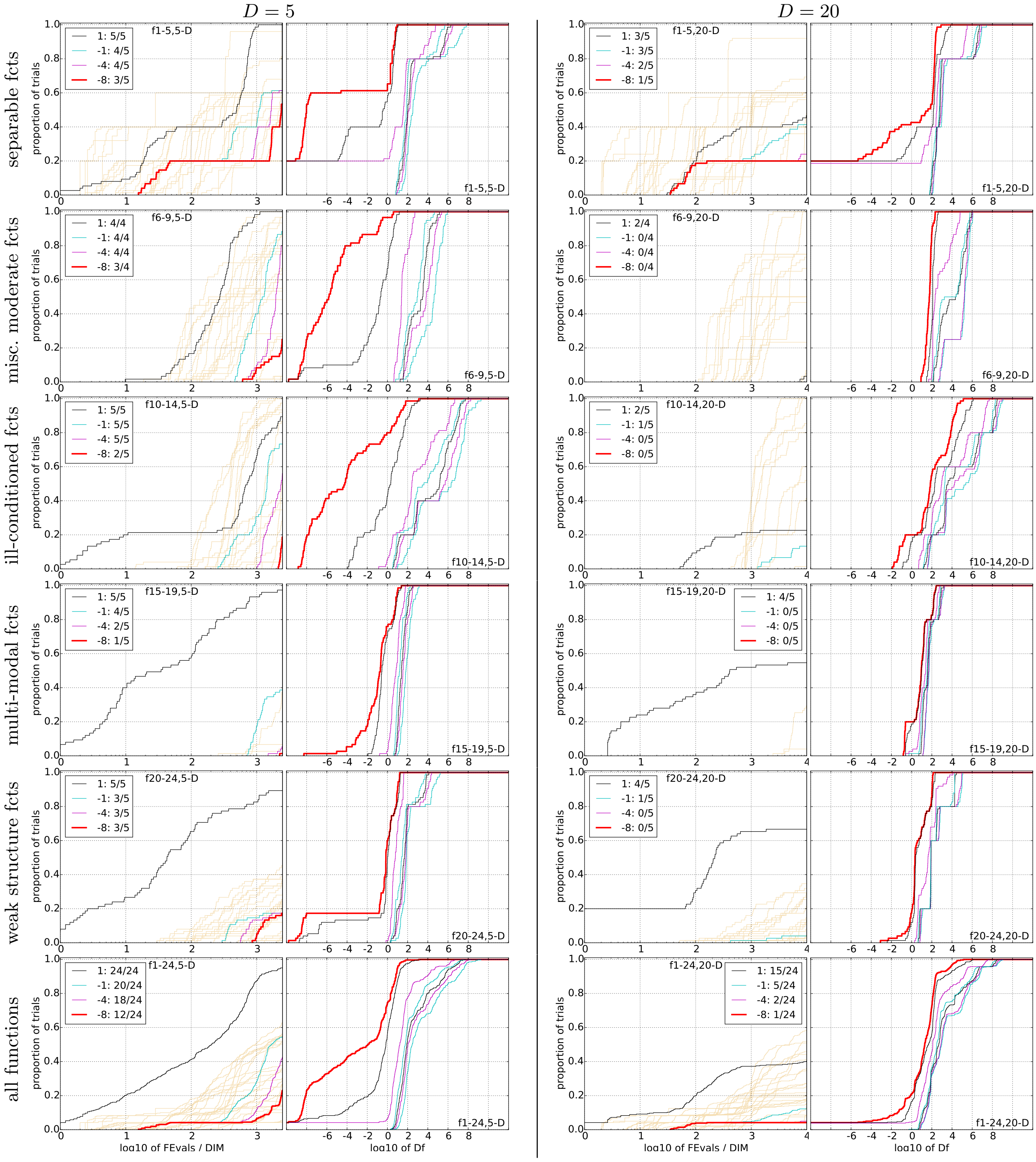


Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the x-axis. Left subplots: ECDF of the number of function evaluations (FEvals) divided by search space dimension D, to fall below $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where k is the first value in the legend. The thick red line represents the most difficult target value $f_{\rm opt} + 10^{-8}$. 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. Light brown lines in the background show ECDFs for the most difficult target of all algorithms benchmarked during BBOB-2009.

5-D

Δf	1e+1	1e+0	1e-1	1e-2	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	11	12	12	12	12	12	12	15/15
	8.2(9)	78(28)	167(18)	233(21)	310(37)	456(22)	601(26)	15/15
$\mathbf{f_2}$	83	87	88	89	90	92	94	15/15
	46(7)	53(4)	62(4)	72(5)	81(5)	99(6)	117(7)	10/15
$\mathbf{f_3}$	716	1622	1637	1642	1646	1650	1654	15/15
	3.7(2)	35(28)	110(104)	110(171)	110(297)	∞	∞ *1.2e4*	0/15
$\mathbf{f_4}$	809	1633	1688	1758	1817	1886	1903	15/15
	4.1(0.8)	108(86)	∞	∞	∞	∞	∞ *1.2e4*	0/15
$\mathbf{f_5}$	10	10	10	10	10	10	10	15/15
	12(3)	14(8)	14(7)	14(7)	14(6)	14(8)	14(5)	15/15
$\mathbf{f_6}$	114	214	281	404	580	1038	1332	15/15
	9.1(5)	10(3)	13(3)	12(4)	12(3)	10(1)	19(12)	2/15
$\mathbf{f_7}$	24	324	1171	1451	1572	1572	1597	15/15
	21(10)	8.2(21)	10(11)	8.9(5)	8.5(14)	8.5(15)	8.5(7)	9/15
$\mathbf{f_8}$	73	273	336	372	391	410	422	15/15
	24(4)	22(5)	23(14)	26(13)	27(12)	40(27)	85(110)	4/15
$\mathbf{f_9}$	35	127	214	263	300	335	369	15/15
	82(56)	49(18)	37(11)	35(5)	43(22)	105(119)	∞ *1.2e4*	0/15
$\mathbf{f_{10}}$	349	500	574	607	626	829	880	15/15
	18(3)	18(10)	18(11)	19(12)	20(16)	23(12)	211(183)	0/15
$\mathbf{f_{11}}$	143	202	763	977	1177	1467	1673	15/15
	37(44)	39(26)	14(13)	12(0.5)	11(8)	10(7)	13(8)	6/15
$\mathbf{f_{12}}$	108	268	371	413	461	1303	1494	15/15
	125(144)	98(106)	114(75)	145(148)	197(249)	∞	∞ *1.2e4*	0/15
$\mathbf{f_{13}}$	132	195	250	319	1310	1752	2255	15/15
	25(8)	28(3)	33(13)	32(12)	12(5)	∞	∞ *1.2e4*	0/15
$\mathbf{f_{14}}$	10	41	58	90	139	251	476	15/15
	1.9(2)	18(10)	33(10)	36(8)	33(4)	30(3)	26(15)	8/15
$\mathbf{f_{15}}$	511	9310	19369	19743	20073	20769	21359	14/15
	11(10)	∞	∞	∞	∞	∞	∞ *1.2e4*	0/15
$\mathbf{f_{16}}$	120	612	2662	10163	10449	11644	12095	15/15
	3.4(4)	11(8)	8.9(7)	5.6(7)	17(18)	16(18)	15(11)	1/15
$\mathbf{f_{17}}$	5.2	215	899	2861	3669	6351	7934	15/15
	4.2(4)	8.1(3)	12(11)	7.3(4)	9.0(11)	29(28)	∞ *1.2e4*	0/15
$\mathbf{f_{18}}$	103	378	3968	8451	9280	10905	12469	15/15
	6.1(3)	11(2)	3.7(4)	6.8(5)	10(7)	∞	∞ *1.2e4*	0/15
$\mathbf{f_{19}}$	1	1	242	1.0e5	1.2e5	1.2e5	1.2e5	15/15
	34(20)	1143(779)	170(121)	∞	∞	∞	∞ *1.2e4*	0/15
$\mathbf{f_{20}}$	16	851	38111	51362	54470	54861	55313	14/15
	18(15)	5.1(0.8)	4.6(5)	3.5(4)	3.3(4)	3.3(3)	3.3(4)	1/15
$\mathbf{f_{21}}$	41	1157	1674	1692	1705	1729	1757	14/15
	10(5)	17(24)	16(16)	16(18)	16(16)	16(16)	16(19)	5/15
$\mathbf{f_{22}}$	71	386	938	980	1008	1040	1068	14/15
	8.9(17)	17(24)	18(27)	18(7)	18(24)	18(8)	19(22)	7/15
$\mathbf{f_{23}}$	3.0	518	14249	27890	31654	33030	34256	15/15
	2.2(1)	11(7)	∞	∞	∞	∞	∞ *1.2e4*	0/15
$\mathbf{f_{24}}$	1622	2.2e5	6.4e6	9.6e6	9.6e6	1.3e7	1.3e7	3/15
	12(11)	∞	∞	∞	∞	∞	∞ *1.2e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-2	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	43	43	43	43	43	43	43	15/15
	122(71)	400(292)	1498(813)	3160(2499)	9316(15057)	68996(96179)	∞ *2.0e5*	0/15
$\mathbf{f_2}$	385	386	387	388	390	391	393	15/15
	1394(703)	3697(7117)	3710(2840)	7619(2958)	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_3}$	5066	7626	7635	7637	7643	7646	7651	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_4}$	4722	7628	7666	7686	7700	7758	1.4e5	9/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_5}$	41	41	41	41	41	41	41	15/15
	30(17)	34(20)	34(10)	34(10)	34(11)	34(14)	34(8)	15/15
$\mathbf{f_6}$	1296	2343	3413	4255	5220	6728	8409	15/15
	2295(3353)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_7}$	1351	4274	9503	16523	16524	16524	16969	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_8}$	2039	3871	4040	4148	4219	4371	4484	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_9}$	1716	3102	3277	3379	3455	3594	3727	15/15
	1713(1480)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{10}}$	7413	8661	10735	13641	14920	17073	17476	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{11}}$	1002	2228	6278	8586	9762	12285	14831	15/15
	1317(1791)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{12}}$	1042	1938	2740	3156	4140	12407	13827	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{13}}$	652	2021	2751	3507	18749	24455	30201	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{14}}$	75	239	304	451	932	1648	15661	15/15
	39(21)	74(35)	576(772)	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{15}}$	30378	1.5e5	3.1e5	3.2e5	3.2e5	4.5e5	4.6e5	15/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{16}}$	1384	27265	77015	1.4e5	1.9e5	2.0e5	2.2e5	15/15
	84(252)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{17}}$	63	1030	4005	12242	30677	56288	80472	15/15
	22(14)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{18}}$	621	3972	19561	28555	67569	1.3e5	1.5e5	15/15
	4511(5463)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{19}}$	1	1	3.4e5	4.7e6	6.2e6	6.7e6	6.7e6	15/15
	72(19)	6185(9153)	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{20}}$	82	46150	3.1e6	5.5e6	5.5e6	5.6e6	5.6e6	14/15
	37(13)	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{21}}$	561	6541	14103	14318	14643	15567	17589	15/15
	142(267)	64(99)	59(96)	97(205)	193(133)	∞	∞ *2.0e5*	0/15
$\mathbf{f_{22}}$	467	5580	23491	24163	24948	26847	1.3e5	12/15
	297(534)	237(420)	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{23}}$	3.2	1614	67457	3.7e5	4.9e5	8.1e5	8.4e5	15/15
	2.8(2)	196(280)	∞	∞	∞	∞	∞ *2.0e5*	0/15
$\mathbf{f_{24}}$	1.3e6	7.5e6	5.2e7	5.2e7	5.2e7	5.2e7	5.2e7	3/15
	∞	∞	∞	∞	∞	∞	∞ *2.0e5*	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 in the first. 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. 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.

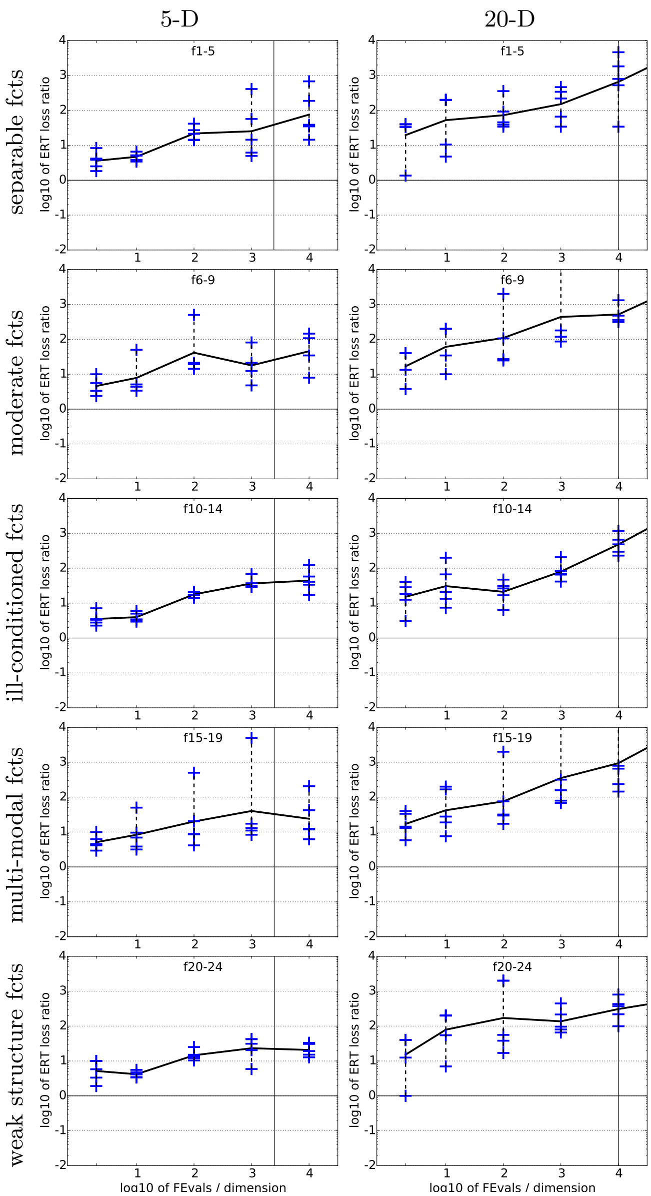


Figure 4: ERT **loss ratios (see Figure ?? for details). Each cross (**$+$**) represents a single function, the line is the geometric mean.**