

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

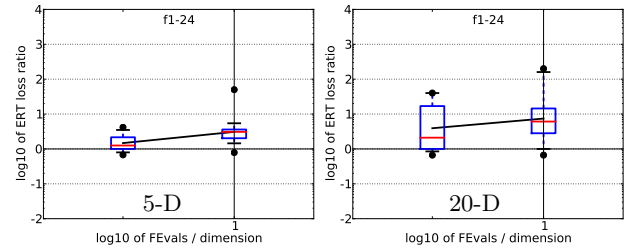
Algorithms

Keywords

Benchmarking, Black-box optimization

1. RESULTS

Results of SMBO4 from experiments according to [?] on the benchmark functions given in [?, ?] are presented in Figures ??, ?? and ?? and in Tables ?? and ??.



<i>f1–f24</i> in 5-D, maxFE/D=10						
#FEs/D	best	10%	25%	med	75%	90%
2	0.68	0.79	1.0	1.3	2.2	3.6
10	0.79	1.4	2.0	3.1	3.8	6.0
100	0.79	6.5	8.7	12	19	87
RL _{US} /D	10	10	10	10	10	10

<i>f1–f24</i> in 20-D, maxFE/D=10						
#FEs/D	best	10%	25%	med	75%	90%
2	0.67	0.84	1.0	2.1	21	40
10	0.67	0.98	2.7	6.1	16	2.0e2
100	0.67	3.6	15	26	84	2.0e3
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 ?? for results on each function subgroup.

*Submission deadline: March 28th.

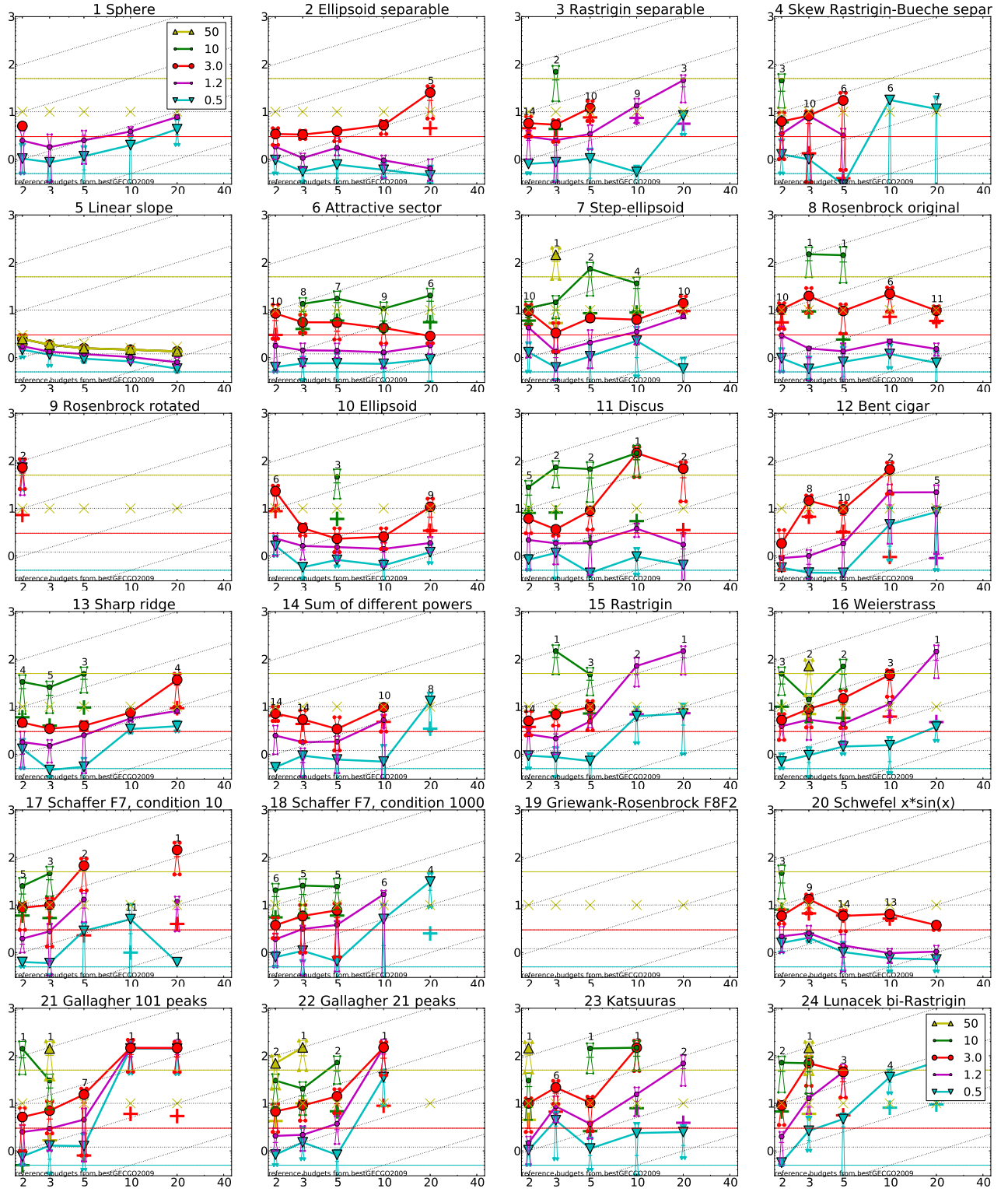


Figure 1: Expected number of f -evaluations (ERT, lines) to reach $f_{\text{opt}} + \Delta f$; median number of f -evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of f -evaluations in any trial (\times); interquartile range with median (notched boxes) of simulated runlengths to reach $f_{\text{opt}} + \Delta f$; 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 $k\text{DIM}$, 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 $\mathcal{O}(\text{DIM})$ compared to $\mathcal{O}(1)$ when using the respective 2009 best algorithm.

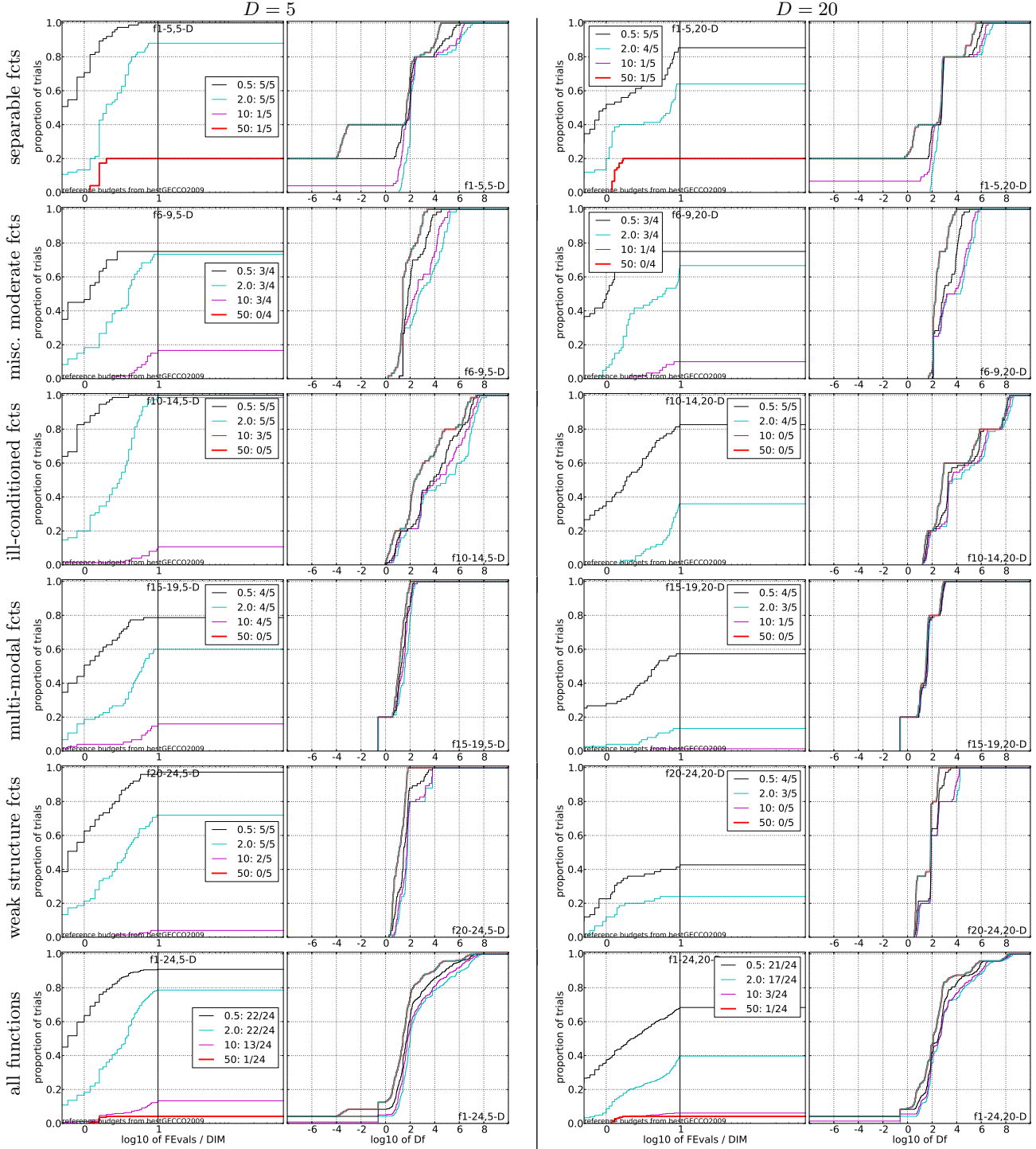


Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the x -axis. Left subplots: ECDF of number of function evaluations (FEvals) divided by search space dimension D , to fall below $f_{\text{opt}} + \Delta f$ 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 $0.5D, 1.2D, 3D, 10D, 100D, 1000D, \dots$ 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.

#FEs/D	5-D					#succ
	0.5	1.2	3.0	10	50	
f₁	2.5e+1:4.8 1.2(1)	1.6e+1:7.6 1.6(1)	1.0e-8:12 ∞	1.0e-8:12 ∞	1.0e-8:12 ∞50	15/15 0/15
f₂	1.6e+6:2.9 1.3(1)	4.0e+5:11 0.78(0.7)	4.0e+4:15 1.3(1.0)	6.3e+2:58 ∞	1.0e-8:95 ∞50	15/15 0/15
f₃	1.6e+2:4.1 1.3(1)	1.0e+2:15 1.2(1)	6.3e+1:23 2.6(2)	2.5e+1:73 ∞	1.0e+1:716 ∞50	15/15 0/15
f₄	2.5e+2:2.6 0.56(0.4)	1.6e+2:10 1.6(3)	1.0e+2:19 4.6(6)	4.0e+1:65 ∞	1.6e+1:434 ∞50	15/15 0/15
f₅	6.3e+1:4.0 1.2(0.2)	4.0e+1:10 0.60(0.1)	1.0e-8:10 0.79(0.2)	1.0e-8:10 0.79(0.2)	1.0e-8:10 0.79(0.2)	15/15 15/15
f₆	1.0e+5:3.0 1.3(1)	2.5e+4:8.4 0.83(0.4)	1.0e+2:16 1.7(2)	2.5e+1:54 1.6(2)	2.5e-1:254 ∞50	15/15 0/15
f₇	1.6e+2:4.2 1.3(2)	1.0e+2:6.2 1.7(2)	2.5e+1:20 1.7(0.6)	4.0e+0:54 6.8(7)	1.0e+0:324 ∞50	15/15 0/15
f₈	1.0e+4:4.6 0.88(1.0)	6.3e+3:6.8 1(1)	1.0e+3:18 2.7(3)	6.3e+1:54 13(16)	1.6e+0:258 ∞50	15/15 0/15
f₉	2.5e+1:20 ∞	1.6e+1:26 ∞	1.0e+1:35 ∞	4.0e+0:62 ∞	1.6e-2:256 ∞50	15/15 0/15
f₁₀	2.5e+6:2.9 1.4(1)	6.3e+5:7.0 1.1(1)	2.5e+5:17 0.68(0.6)	6.3e+3:54 4.3(5)	2.5e+1:297 ∞50	15/15 0/15
f₁₁	1.0e+6:3.0 0.73(1.0)	6.3e+4:6.2 1.5(1)	6.3e+2:16 2.8(2)	6.3e+1:74 4.5(5)	6.3e-1:298 ∞50	15/15 0/15
f₁₂	4.0e+7:3.6 0.61(0.4)	1.6e+7:7.6 1.2(0.9)	4.0e+6:19 2.5(2)	1.6e+4:52 ∞	1.0e+0:268 ∞50	15/15 0/15
f₁₃	1.0e+3:2.8 0.98(1)	6.3e+2:8.4 1.5(1)	4.0e+2:17 1.2(0.4)	6.3e+1:52 4.7(4)	6.3e-2:264 ∞50	15/15 0/15
f₁₄	1.6e+1:3.0 1.3(2)	1.0e+1:10 0.94(1.0)	6.3e+0:15 1.1(1)	2.5e-1:53 ∞	1.0e-5:251 ∞50	15/15 0/15
f₁₅	1.6e+2:3.0 1.2(2)	1.0e+2:13 1.9(2)	6.3e+1:24 2.0(2)	4.0e+1:55 4.3(5)	1.6e+1:289 ∞50	5/5 0/15
f₁₆	4.0e+1:4.8 1.5(0.9)	2.5e+1:16 1.4(2)	1.6e+1:46 1.6(2)	1.0e+1:120 2.9(3)	4.0e+0:334 ∞50	15/15 0/15
f₁₇	1.0e+1:5.2 2.7(3)	6.3e+0:26 2.5(3)	2.5e+0:57 5.9(7)	2.5e+0:110 ∞	6.3e-1:412 ∞50	15/15 0/15
f₁₈	6.3e+1:3.4 0.96(0.6)	4.0e+1:7.2 2.6(4)	2.5e+1:20 1.9(3)	1.6e+1:58 2.1(2)	1.6e+0:318 ∞50	15/15 0/15
f₁₉	1.6e-1:172 ∞	1.0e-1:242 ∞	6.3e-2:675 ∞	4.0e-2:3078 ∞	2.5e-2:4946 ∞50	15/15 0/15
f₂₀	6.3e+3:5.1 1(1.0)	4.0e+3:8.4 0.84(0.7)	4.0e+1:15 1.9(1)	2.5e+0:69 ∞	1.0e+0:851 ∞50	15/15 0/15
f₂₁	4.0e+1:3.9 1.6(2)	2.5e+1:11 2.2(3)	1.6e+1:31 2.5(3)	6.3e+0:73 ∞	1.6e+0:347 ∞50	5/5 0/15
f₂₂	6.3e+1:3.6 1.1(0.7)	4.0e+1:15 1.3(1)	2.5e+1:32 2.2(2)	1.0e+1:71 5.1(5)	1.6e+0:341 ∞50	5/5 0/15
f₂₃	1.0e+1:3.0 1.9(2)	6.3e+0:9.0 2.1(2)	4.0e+0:33 1.5(2)	2.5e+0:84 8.5(10)	1.0e+0:518 ∞50	15/15 0/15
f₂₄	6.3e+1:15 1.6(2)	4.0e+1:37 6.2(7)	4.0e+1:37 6.2(7)	2.5e+1:118 ∞	1.6e+1:692 ∞50	15/15 0/15

#FEs/D	20-D					#succ
	0.5	1.2	3.0	10	50	
f₁	6.3e+1:24 3.5(3)	4.0e+1:42 3.7(0.7)	1.0e-8:43 ∞	1.0e-8:43 ∞	1.0e-8:43 ∞200	15/15 0/15
f₂	4.0e+6:29 0.32(0.3)↓	2.5e+6:42 0.31(0.2)↓ ⁴	1.0e+5:65 7.8(9)	1.0e+4:207 ∞	1.0e-8:412 ∞200	15/15 0/15
f₃	6.3e+2:33 5.1(6)	4.0e+2:44 21(22)	1.6e+2:109 ∞	1.0e+2:255 ∞	2.5e+1:3277 ∞200	15/15 0/15
f₄	6.3e+2:22 11(14)	4.0e+2:91 ∞	2.5e+2:250 ∞	1.6e+2:332 ∞	6.3e+1:1927 ∞200	15/15 0/15
f₅	2.5e+2:19 0.62(0.2)	1.6e+2:34 0.47(0.1)↓ ⁴	1.0e-8:41 0.67(0.1)↓ ⁴	1.0e-8:41 0.67(0.1)↓ ⁴	1.0e-8:41 0.67(0.1)↓ ⁴	15/15 15/15
f₆	2.5e+5:16 1.1(1)	6.3e+4:43 0.84(0.3)	1.6e+4:62 0.92(0.4)	1.6e+2:353 1.1(1)	1.6e+1:1078 ∞200	15/15 0/15
f₇	1.0e+3:11 1.1(2)	4.0e+2:39 3.8(2)	2.5e+2:74 3.8(3)	6.3e+1:319 ∞	1.0e+1:1351 ∞200	15/15 0/15
f₈	4.0e+4:19 0.83(0.8)	2.5e+4:35 0.86(0.7)	4.0e+3:67 2.9(2)	2.5e+2:231 ∞	1.6e+1:1470 ∞200	15/15 0/15
f₉	1.0e+2:357 ∞	6.3e+1:560 ∞	4.0e+1:684 ∞	2.5e+1:756 ∞	1.0e+1:1716 ∞200	15/15 0/15
f₁₀	1.6e+6:15 1.6(0.9)	1.0e+6:27 1.4(0.7)	4.0e+5:70 3.1(3)	6.3e+4:231 ∞	4.0e+3:1015 ∞200	15/15 0/15
f₁₁	4.0e+4:11 1.1(1)	2.5e+3:27 1.3(2)	1.6e+2:313 4.4(5)	1.0e+2:481 ∞	1.0e+1:1002 ∞200	15/15 0/15
f₁₂	1.0e+8:23 7.2(9)	6.3e+7:39 11(13)	2.5e+7:76 ∞	4.0e+6:209 ∞	1.0e+1:1042 ∞200	15/15 0/15
f₁₃	1.6e+3:28 2.8(1)	1.0e+3:64 2.5(0.4)	6.3e+2:79 9.3(9)	4.0e+1:211 ∞	2.5e+0:1724 ∞200	15/15 0/15
f₁₄	2.5e+1:15 18(21)	1.6e+1:42 ∞	1.0e+1:75 ∞	1.6e+0:219 ∞	6.3e-4:1106 ∞200	15/15 0/15
f₁₅	6.3e+2:15 9.3(13)	4.0e+2:67 44(50)	2.5e+2:292 ∞	1.6e+2:846 ∞	1.0e+2:1671 ∞200	15/15 0/15
f₁₆	4.0e+1:26 2.9(3)	2.5e+1:127 23(25)	1.6e+1:540 ∞	1.6e+1:540 ∞	1.0e+1:1384 ∞200	15/15 0/15
f₁₇	1.6e+1:11 1.2(3)	1.0e+1:63 3.7(5)	6.3e+0:305 9.4(10)	4.0e+0:468 ∞	1.0e+0:1030 ∞200	15/15 0/15
f₁₈	4.0e+1:116 5.3(6)	2.5e+1:252 ∞	1.6e+1:430 ∞	1.0e+1:621 ∞	4.0e+0:1090 ∞200	15/15 0/15
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 ∞200	3/15 0/15
f₂₀	1.6e+4:38 0.37(0.2)↓	1.0e+4:42 0.50(0.2)↓ ⁴	2.5e+2:62 1.2(0.6)	2.5e+0:250 ∞	1.6e+0:2536 ∞200	15/15 0/15
f₂₁	6.3e+1:36 79(80)	4.0e+1:77 38(42)	4.0e+1:77 38(38)	1.6e+1:456 ∞	4.0e+0:1094 ∞200	15/15 0/15
f₂₂	6.3e+1:45 ∞	4.0e+1:68 ∞	4.0e+1:68 ∞	1.6e+1:231 ∞	6.3e+0:1219 ∞200	15/15 0/15
f₂₃	6.3e+0:29 1.7(2)	4.0e+0:118 12(13)	2.5e+0:306 ∞	2.5e+0:306 ∞	1.0e+0:1614 ∞200	15/15 0/15
f₂₄	2.5e+2:208 7.2(7)	1.6e+2:918 ∞	1.0e+2:6628 ∞	6.3e+1:9885 ∞	4.0e+1:31629 ∞200	15/15 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.

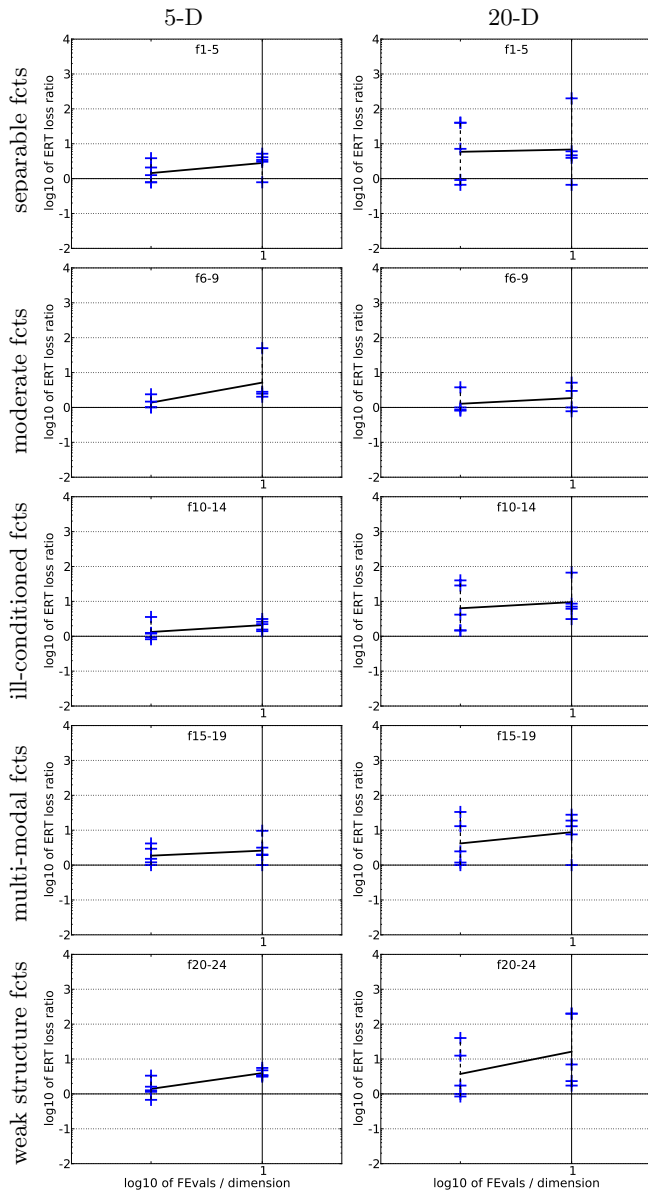


Figure 4: ERT loss ratios (see Figure ?? for details). Each cross (+) represents a single function, the line is the geometric mean.