

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

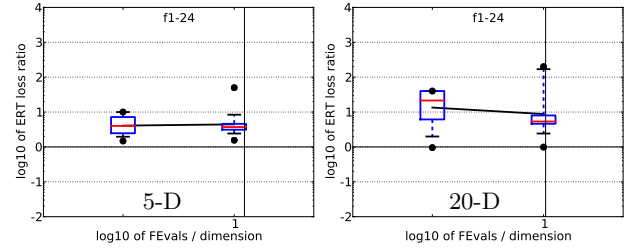
Algorithms

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 ??.



<i>f1-f24 in 5-D, maxFE/D=11</i>						
#FEs/D	best	10%	25%	med	75%	90%
2	1.5	1.9	2.4	4.0	7.6	10
10	1.6	2.3	3.1	3.7	4.6	14
100	3.1	4.1	12	16	21	88
RL _{US} /D	10	10	10	10	10	10

<i>f1-f24 in 20-D, maxFE/D=10</i>						
#FEs/D	best	10%	25%	med	75%	90%
2	0.96	1.8	6.0	21	40	40
10	0.99	2.3	4.6	5.4	8.4	2.0e2
100	7.9	13	22	28	46	2.7e2
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 subgroup. See also Figure ?? for results on each function subgroup.

*Submission deadline: March 28th.

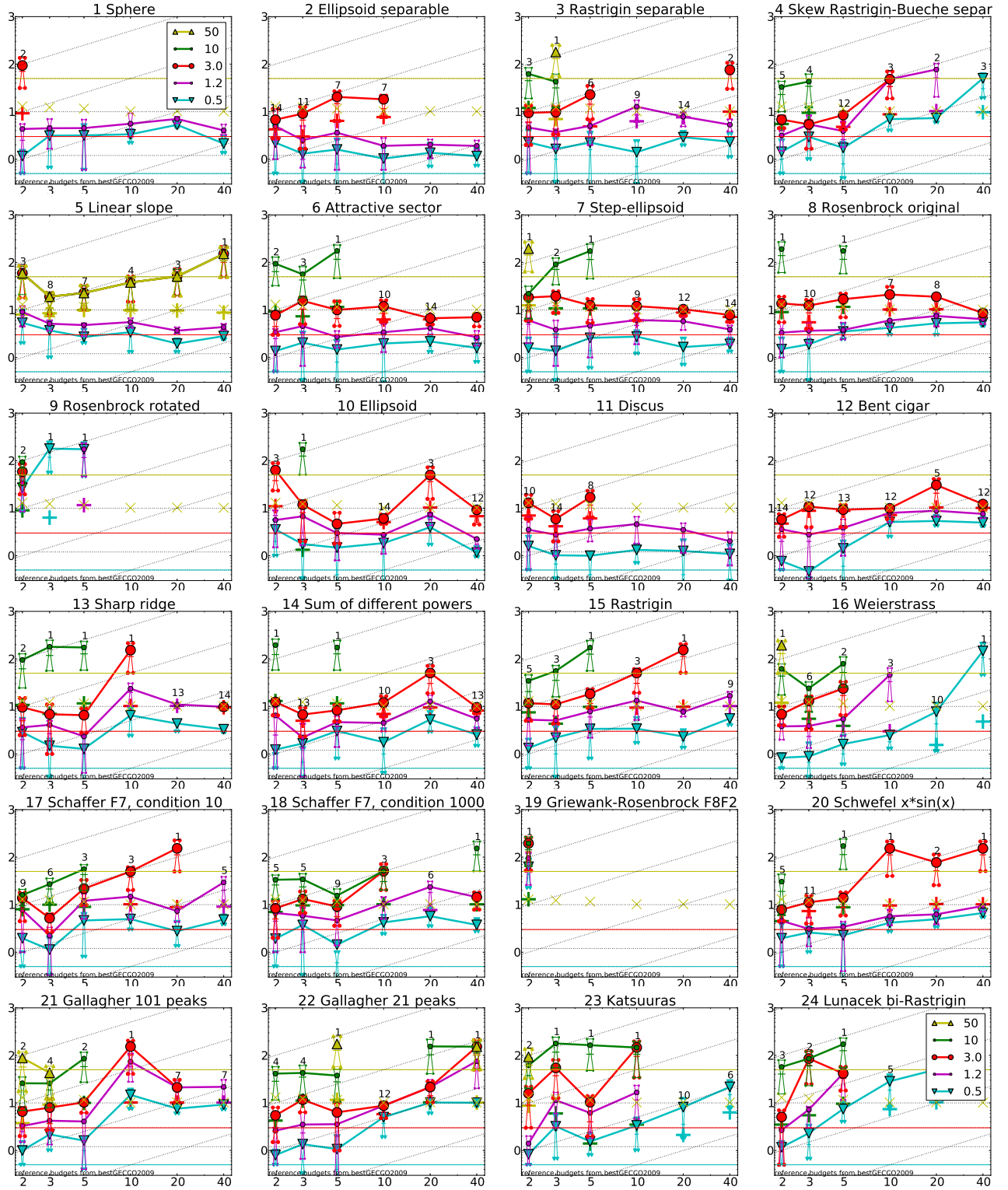


Figure 1: Expected number of f -evaluations (ERT, lines) to reach $f_{\text{opt}} + \Delta f$; median number of f -evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of f -evaluations in any trial (\times); interquartile range with median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

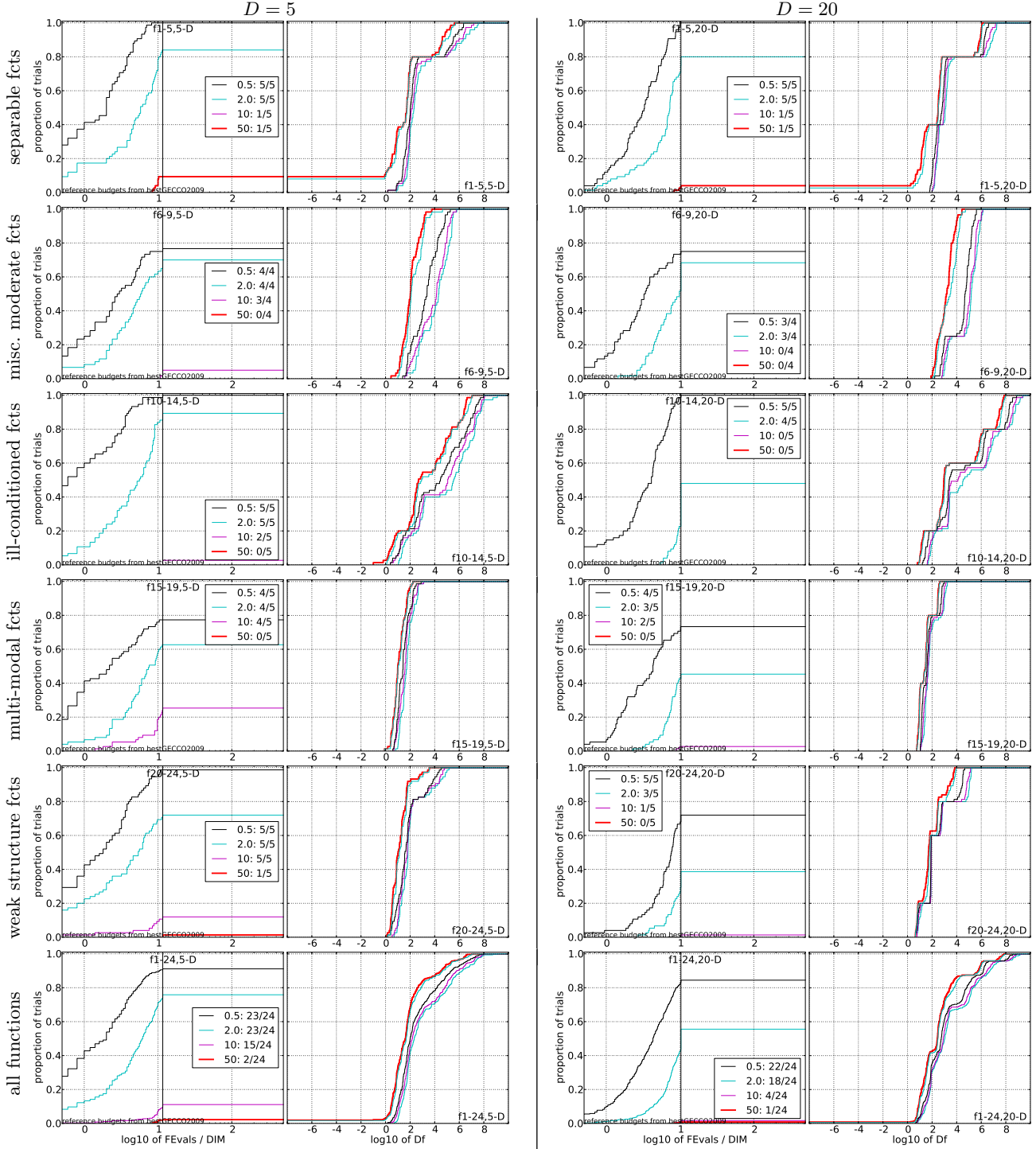


Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the x -axis. Left subplots: ECDF of number of function evaluations (FEvals) divided by search space dimension D , to fall below $f_{\text{opt}} + \Delta f$ 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	<i>2.5e+1:4.8</i> 3.3(3)	<i>1.6e+1:7.6</i> 3.0(3)	<i>1.0e-8:12</i> ∞	<i>1.0e-8:12</i> ∞	<i>1.0e-8:12</i> ∞58	15/15 0/15
f₂	<i>1.6e+6:2.9</i> 2.8(3)	<i>4.0e+5:11</i> 1.6(2)	<i>4.0e+4:15</i> 6.7(6)	<i>6.3e+2:58</i> ∞	<i>1.0e-8:95</i> ∞58	15/15 0/15
f₃	<i>1.6e+2:4.1</i> 2.7(3)	<i>1.0e+2:15</i> 1.7(2)	<i>6.3e+1:23</i> 4.9(5)	<i>2.5e+1:73</i> ∞	<i>1.0e+1:716</i> ∞58	15/15 0/15
f₄	<i>2.5e+2:2.6</i> 3.3(4)	<i>1.6e+2:10</i> 2.0(2)	<i>1.0e+2:19</i> 2.3(2)	<i>4.0e+1:65</i> ∞	<i>1.6e+1:434</i> ∞58	15/15 0/15
f₅	<i>6.3e+1:4.0</i> 3.4(4)	<i>4.0e+1:10</i> 2.4(1)	<i>1.0e-8:10</i> 11(9)	<i>1.0e-8:10</i> 11(9)	<i>1.0e-8:10</i> 11(9)	15/15 7/15
f₆	<i>1.0e+5:3.0</i> 2.4(3)	<i>2.5e+4:8.4</i> 1.6(1)	<i>1.0e+2:16</i> 3.1(3)	<i>2.5e+1:54</i> 16(16)	<i>2.5e-1:254</i> ∞58	15/15 0/15
f₇	<i>1.6e+2:4.2</i> 3.1(4)	<i>1.0e+2:6.2</i> 3.7(5)	<i>2.5e+1:20</i> 3.1(3)	<i>4.0e+0:54</i> 16(18)	<i>1.0e+0:324</i> ∞58	15/15 0/15
f₈	<i>1.0e+4:4.6</i> 3.8(2)	<i>6.3e+3:6.8</i> 2.8(2)	<i>1.0e+3:18</i> 4.7(4)	<i>6.3e+1:54</i> 16(18)	<i>1.6e+0:258</i> ∞58	15/15 0/15
f₉	<i>2.5e+1:20</i> 44(44)	<i>1.6e+1:26</i> 34(36)	<i>1.0e+1:35</i> ∞	<i>4.0e+0:62</i> ∞	<i>1.6e-2:256</i> ∞58	15/15 0/15
f₁₀	<i>2.5e+6:2.9</i> 2.6(3)	<i>6.3e+5:7.0</i> 2.1(2)	<i>2.5e+5:17</i> 1.4(1)	<i>6.3e+3:54</i> ∞	<i>2.5e+1:297</i> ∞58	15/15 0/15
f₁₁	<i>1.0e+6:3.0</i> 1.7(2)	<i>6.3e+4:6.2</i> 3.0(3)	<i>6.3e+2:16</i> 5.3(5)	<i>6.3e+1:74</i> ∞	<i>6.3e-1:298</i> ∞58	15/15 0/15
f₁₂	<i>4.0e+7:3.6</i> 2.0(3)	<i>1.6e+7:7.6</i> 2.5(3)	<i>4.0e+6:19</i> 2.4(2)	<i>1.6e+4:52</i> ∞	<i>1.0e+0:268</i> ∞58	15/15 0/15
f₁₃	<i>1.0e+3:2.8</i> 2.3(2)	<i>6.3e+2:8.4</i> 1.4(2)	<i>6.3e+2:17</i> 2.0(1)	<i>6.3e+1:52</i> 17(19)	<i>6.3e-2:264</i> ∞58	15/15 0/15
f₁₄	<i>1.6e+1:3.0</i> 5.1(5)	<i>1.0e+1:10</i> 2.4(2)	<i>6.3e+0:15</i> 2.7(2)	<i>2.5e-1:53</i> 17(17)	<i>1.0e-5:251</i> ∞58	15/15 0/15
f₁₅	<i>1.6e+2:3.0</i> 5.6(8)	<i>1.0e+2:13</i> 3.0(3)	<i>6.3e+1:24</i> 2.6(3)	<i>4.0e+1:55</i> 3.3(4)	<i>1.6e+1:289</i> ∞58	5/5 0/15
f₁₆	<i>4.0e+1:4.8</i> 1.7(2)	<i>2.5e+1:16</i> 1.7(1)	<i>1.6e+1:46</i> 2.6(3)	<i>1.0e+1:120</i> 3.3(4)	<i>4.0e+0:334</i> ∞58	15/15 0/15
f₁₇	<i>1.0e+1:5.2</i> 4.5(4)	<i>6.3e+0:26</i> 2.3(2)	<i>4.0e+0:57</i> 1.9(2)	<i>2.5e+0:110</i> 2.5(3)	<i>6.3e-1:412</i> ∞58	15/15 0/15
f₁₈	<i>6.3e+1:3.4</i> 2.1(2)	<i>4.0e+1:7.2</i> 3.2(3)	<i>2.5e+1:20</i> 2.3(2)	<i>1.6e+1:58</i> 1.3(1)	<i>1.6e+0:318</i> ∞58	15/15 0/15
f₁₉	<i>1.6e-1:172</i> ∞	<i>1.0e-1:242</i> ∞	<i>6.3e-2:675</i> ∞	<i>4.0e-2:3078</i> ∞	<i>2.5e-2:4946</i> ∞58	15/15 0/15
f₂₀	<i>6.3e+3:5.1</i> 2.2(2)	<i>4.0e+3:8.4</i> 2.0(3)	<i>4.0e+1:15</i> 4.5(4)	<i>2.5e+0:69</i> 12(13)	<i>1.0e+0:851</i> ∞58	15/15 0/15
f₂₁	<i>4.0e+1:3.9</i> 2.1(3)	<i>2.5e+1:11</i> 1.9(2)	<i>1.6e+1:31</i> 1.7(1)	<i>6.3e+0:73</i> 5.9(6)	<i>1.6e+0:347</i> ∞58	5/5 0/15
f₂₂	<i>6.3e+1:3.6</i> 1.5(2)	<i>4.0e+1:15</i> 1.2(1)	<i>2.5e+1:32</i> 0.98(0.8)	<i>1.0e+1:71</i> 2.7(3)	<i>1.6e+0:341</i> 2.6(3)	5/5 1/15
f₂₃	<i>1.0e+1:3.0</i> 2.6(2)	<i>6.3e+0:9.0</i> 3.4(4)	<i>4.0e+0:33</i> 1.6(2)	<i>2.5e+0:84</i> 10(11)	<i>1.0e+0:518</i> ∞58	15/15 0/15
f₂₄	<i>6.3e+1:15</i> 2.6(1)	<i>4.0e+1:37</i> 5.7(6)	<i>4.0e+1:37</i> 5.7(6)	<i>2.5e+1:118</i> 7.3(8)	<i>1.6e+1:692</i> ∞58	15/15 0/15

#FEs/D	20-D					#succ
	0.5	1.2	3.0	10	50	
f₁	<i>6.3e+1:24</i> 4.4(1)	<i>4.0e+1:42</i> 3.4(1)	<i>1.0e-8:43</i> ∞	<i>1.0e-8:43</i> ∞	<i>1.0e-8:43</i> ∞210	15/15 0/15
f₂	<i>4.0e+6:29</i> 0.96(0.9)	<i>2.5e+6:42</i> 0.97(0.8)	<i>1.0e+5:65</i> ∞	<i>1.0e+4:207</i> ∞	<i>1.0e-8:412</i> ∞210	15/15 0/15
f₃	<i>6.3e+2:33</i> 1.8(0.7)	<i>4.0e+2:44</i> 3.6(2)	<i>1.6e+2:109</i> ∞	<i>1.0e+2:255</i> ∞	<i>2.5e+1:3277</i> ∞210	15/15 0/15
f₄	<i>6.3e+2:22</i> 6.8(2)	<i>4.0e+2:91</i> 17(19)	<i>2.5e+2:250</i> ∞	<i>1.6e+2:332</i> ∞	<i>6.3e+1:1927</i> ∞210	15/15 0/15
f₅	<i>2.5e+2:19</i> 2.1(1)	<i>1.6e+2:34</i> 2.1(0.6)	<i>1.0e-8:41</i> 25(23)	<i>1.0e-8:41</i> 25(23)	<i>1.0e-8:41</i> 25(25)	15/15 3/15
f₆	<i>2.5e+5:16</i> 2.7(2)	<i>6.3e+4:43</i> 1.9(0.6)	<i>1.6e+4:62</i> 2.2(1)	<i>1.6e+2:353</i> ∞	<i>1.6e+1:1078</i> ∞210	15/15 0/15
f₇	<i>1.0e+3:11</i> 3.1(3)	<i>4.0e+2:39</i> 3.0(2)	<i>2.5e+2:74</i> 2.8(2)	<i>6.3e+1:319</i> ∞	<i>1.0e+1:1351</i> ∞210	15/15 0/15
f₈	<i>4.0e+4:19</i> 5.6(3)	<i>2.5e+4:35</i> 4.2(2)	<i>4.0e+3:67</i> 5.6(5)	<i>2.5e+2:231</i> ∞	<i>1.6e+1:1470</i> ∞210	15/15 0/15
f₉	<i>1.0e+2:357</i> ∞	<i>6.3e+1:560</i> ∞	<i>4.0e+1:684</i> ∞	<i>2.5e+1:756</i> ∞	<i>1.0e+1:1716</i> ∞210	15/15 0/15
f₁₀	<i>1.6e+6:15</i> 5.3(4)	<i>1.0e+6:27</i> 5.3(4)	<i>4.0e+5:70</i> 14(14)	<i>6.3e+4:231</i> ∞	<i>4.0e+3:1015</i> ∞210	15/15 0/15
f₁₁	<i>4.0e+4:11</i> 2.2(3)	<i>2.5e+3:27</i> 2.6(3)	<i>1.6e+2:313</i> ∞	<i>1.0e+2:481</i> ∞	<i>1.0e+1:1002</i> ∞210	15/15 0/15
f₁₂	<i>1.0e+8:23</i> 4.6(3)	<i>6.3e+7:39</i> 4.5(1)	<i>2.5e+7:76</i> 8.1(7)	<i>4.0e+6:209</i> ∞	<i>1.0e+1:1042</i> ∞210	15/15 0/15
f₁₃	<i>1.6e+3:28</i> 3.2(2)	<i>1.0e+3:64</i> 3.4(2)	<i>6.3e+2:79</i> ∞	<i>4.0e+1:211</i> ∞	<i>2.5e+0:1724</i> ∞210	15/15 0/15
f₁₄	<i>2.5e+1:15</i> 7.2(4)	<i>1.6e+1:42</i> 6.0(3)	<i>1.0e+1:75</i> 13(14)	<i>1.6e+0:219</i> ∞	<i>6.3e-4:1106</i> ∞210	15/15 0/15
f₁₅	<i>6.3e+2:15</i> 3.0(2)	<i>4.0e+2:67</i> 2.3(0.6)	<i>2.5e+2:292</i> 11(12)	<i>1.6e+2:846</i> ∞	<i>1.0e+2:1671</i> ∞210	15/15 0/15
f₁₆	<i>4.0e+1:26</i> 5.9(8)	<i>2.5e+1:127</i> ∞	<i>1.6e+1:540</i> ∞	<i>1.6e+1:540</i> ∞	<i>1.0e+1:1384</i> ∞210	15/15 0/15
f₁₇	<i>1.6e+1:11</i> 5.2(4)	<i>1.0e+1:63</i> 2.3(1)	<i>6.3e+0:305</i> 10(11)	<i>4.0e+0:468</i> ∞	<i>1.0e+0:1030</i> ∞210	15/15 0/15
f₁₈	<i>4.0e+1:116</i> 0.99(0.7)	<i>2.5e+1:252</i> 1.9(2)	<i>1.6e+1:430</i> ∞	<i>1.0e+1:621</i> ∞	<i>4.0e+0:1090</i> ∞210	15/15 0/15
f₁₉	<i>1.6e-1:2.5e5</i> ∞	<i>1.0e-1:3.4e5</i> ∞	<i>6.3e-2:3.4e5</i> ∞	<i>4.0e-2:3.4e5</i> ∞	<i>2.5e-2:3.4e5</i> ∞210	3/15 0/15
f₂₀	<i>1.6e+4:38</i> 2.6(1)	<i>1.0e+4:42</i> 3.0(1)	<i>2.5e+2:62</i> 25(25)	<i>2.5e+0:250</i> ∞	<i>1.6e+0:2536</i> ∞210	15/15 0/15
f₂₁	<i>6.3e+1:36</i> 4.2(1)	<i>4.0e+1:77</i> 5.5(4)	<i>4.0e+1:77</i> 5.5(4)	<i>1.6e+1:456</i> ∞	<i>4.0e+0:1094</i> ∞210	15/15 0/15
f₂₂	<i>6.3e+1:45</i> 4.5(3)	<i>4.0e+1:68</i> 6.4(5)	<i>4.0e+1:68</i> 6.4(5)	<i>1.6e+1:231</i> 13(13)	<i>6.3e+0:1219</i> ∞210	15/15 0/15
f₂₃	<i>6.3e+0:29</i> 5.7(8)	<i>4.0e+0:118</i> ∞	<i>2.5e+0:306</i> ∞	<i>2.5e+0:306</i> ∞	<i>1.0e+0:1614</i> ∞210	15/15 0/15
f₂₄	<i>2.5e+2:208</i> 4.9(5)	<i>1.6e+2:918</i> ∞	<i>1.0e+2:6628</i> ∞	<i>6.3e+1:9885</i> ∞	<i>4.0e+1:31629</i> ∞210	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.

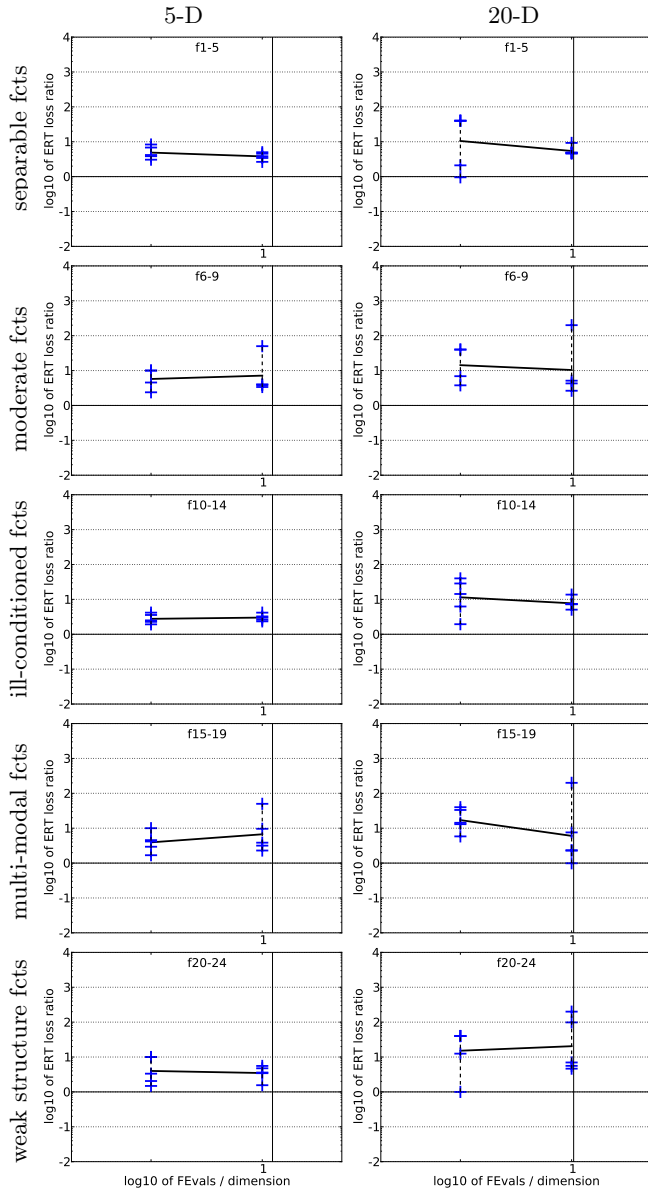


Figure 4: ERT loss ratios (see Figure ?? for details). Each cross (+) represents a single function, the line is the geometric mean.