# 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

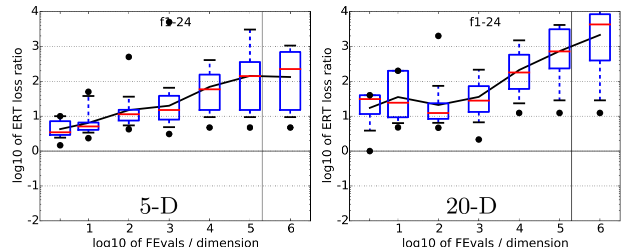
Benchmarking, Black-box optimization


## 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 $x.x$, $x.x$, $x.x$, $xx$, $xxx$, and $xxx$ milliseconds respectively.

## 2. RESULTS

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




| #FEs/D | ***f*1–*f*24 in 5-D**, maxFE/D=199951 | | | | | |
|---|---|---|---|---|---|---|
| | best | 10% | 25% | **med** | 75% | 90% |
| 2 | 1.5 | 2.4 | 2.9 | 3.5 | 7.6 | 10 |
| 10 | 2.3 | 3.3 | 4.0 | 5.1 | 6.6 | 50 |
| 100 | 4.2 | 4.8 | 7.3 | 11 | 15 | 87 |
| 1e3 | 3.1 | 4.4 | 7.7 | 15 | 40 | 94 |
| 1e4 | 4.7 | 8.7 | 15 | 59 | 1.6e2 | 4.5e2 |
| 1e5 | 4.7 | 8.7 | 15 | 1.4e2 | 3.6e2 | 4.0e3 |
| 1e6 | 4.7 | 8.7 | 15 | 2.2e2 | 7.0e2 | 1.3e3 |
| $\mathrm{RL_{US}}$/D | 2e5 | 2e5 | 2e5 | 2e5 | 2e5 | 2e5 |

| #FEs/D | ***f*1–*f*24 in 20-D**, maxFE/D=199976 | | | | | |
|---|---|---|---|---|---|---|
| | best | 10% | 25% | **med** | 75% | 90% |
| 2 | 1.0 | 3.8 | 11 | 31 | 40 | 40 |
| 10 | 4.8 | 6.0 | 8.7 | 24 | 2.0e2 | 2.0e2 |
| 100 | 4.7 | 6.4 | 8.2 | 12 | 24 | 2.7e2 |
| 1e3 | 2.1 | 5.8 | 13 | 28 | 80 | 2.4e2 |
| 1e4 | 12 | 20 | 60 | 1.8e2 | 6.0e2 | 1.7e3 |
| 1e5 | 12 | 25 | 2.3e2 | 7.2e2 | 3.3e3 | 4.3e3 |
| 1e6 | 12 | 25 | 3.4e2 | 4.3e3 | 8.7e3 | 1.4e4 |
| $\mathrm{RL_{US}}$/D | 2e5 | 2e5 | 2e5 | 2e5 | 2e5 | 2e5 |

**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 4 for results on each function subgroup.**

*Submission deadline: March 28th.


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

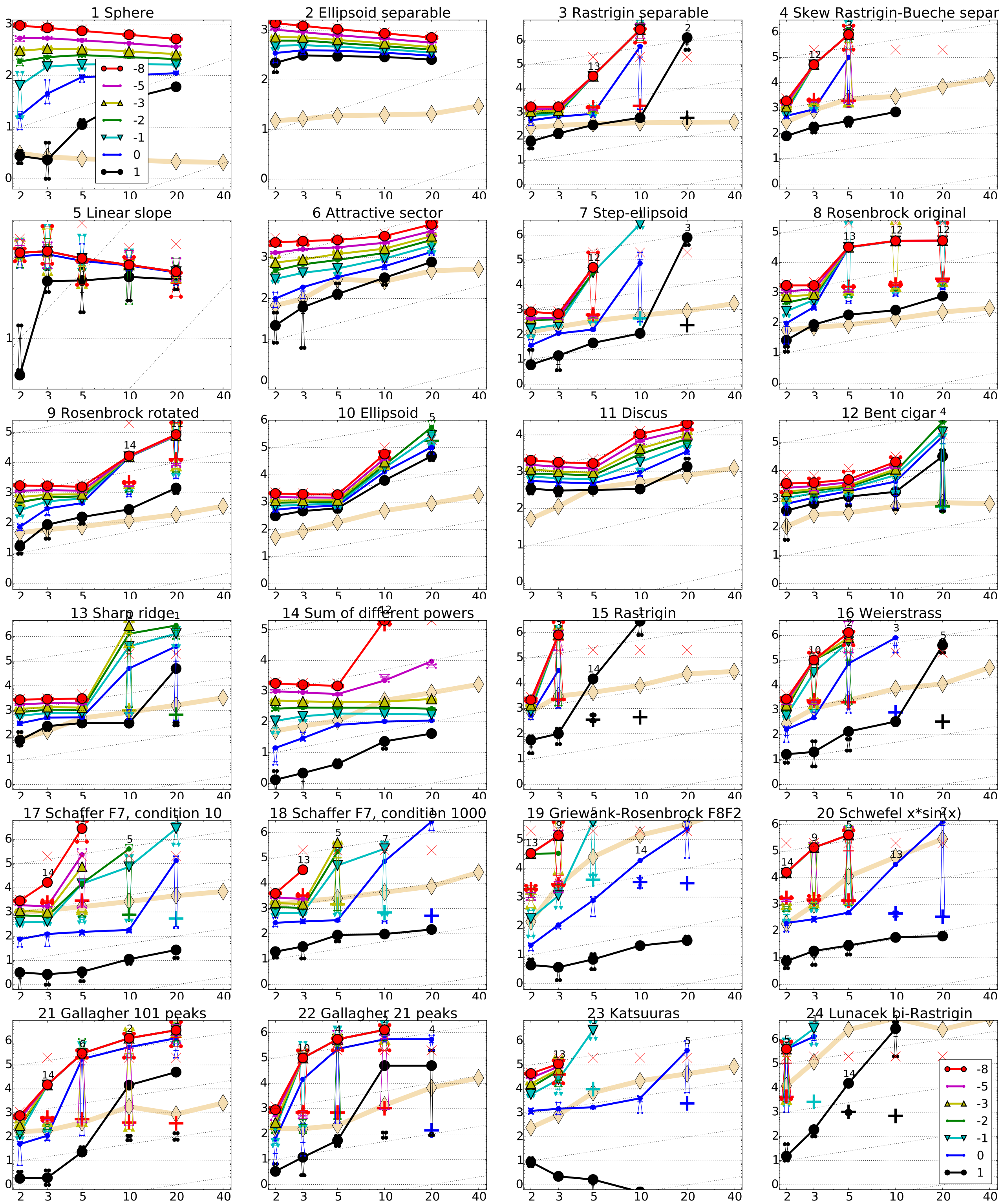



**Figure 1: Expected number of $f$-evaluations (ERT, lines) to reach $f_{\rm opt} + \Delta f$; median number of $f$-evaluations (+) to reach the most difficult target that was reached not always but at least once; maximum number of $f$-evaluations in any trial (×); interquartile range with median (notched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

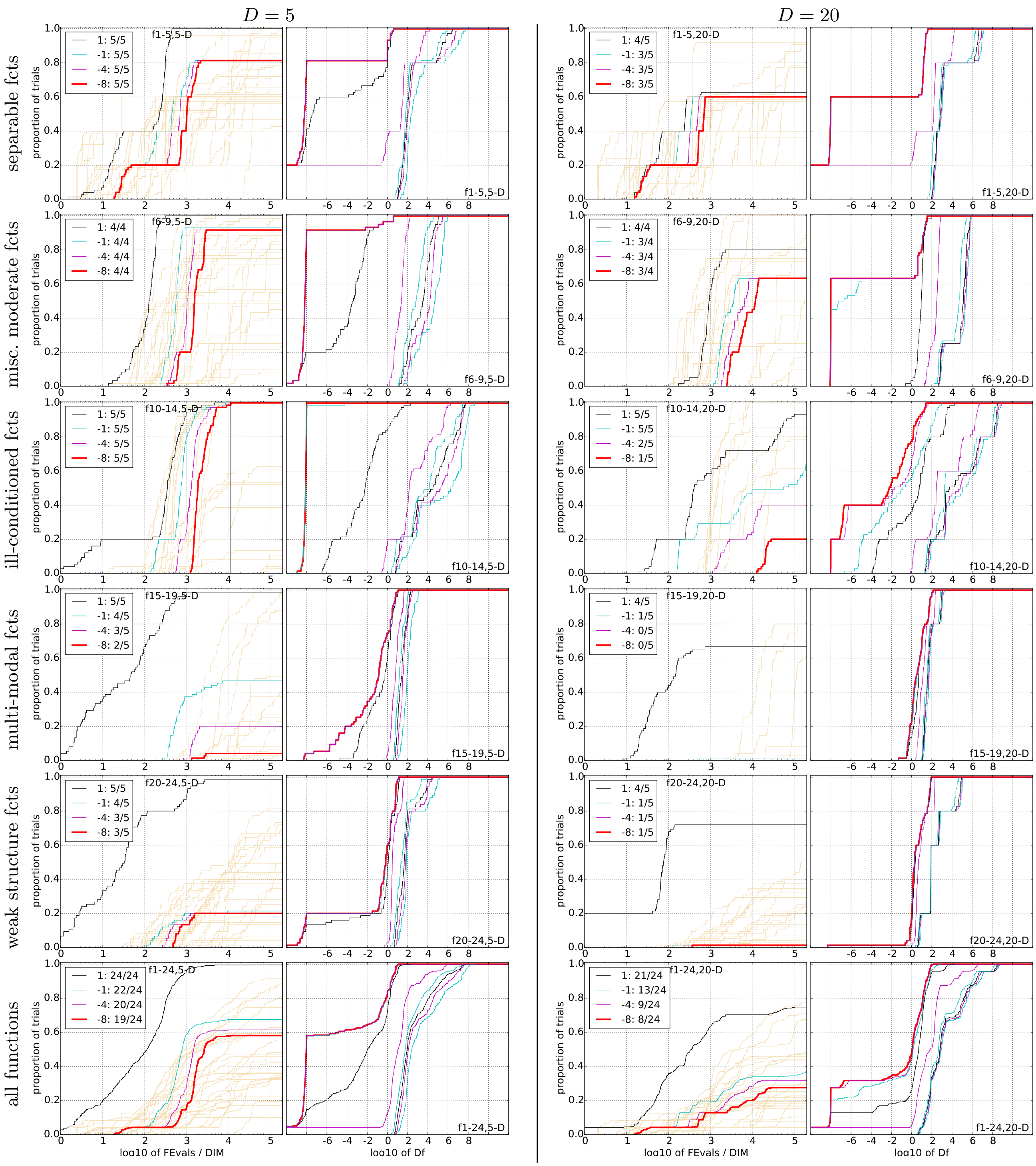


**Figure 2: Empirical cumulative distribution functions (ECDF), plotting the fraction of trials with an outcome not larger than the respective value on the $x$-axis. Left subplots: ECDF of the number of function evaluations (FEvals) divided by search space dimension $D$, to fall below $f_{\mathrm{opt}} + \Delta f$ with $\Delta f = 10^k$, where $k$ is the first value in the legend. The thick red line represents the most difficult target value $f_{\mathrm{opt}} + 10^{-8}$. 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 $\Delta f$ 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 $\Delta f$-value (red), where $\Delta f$ 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 |
|---|---|---|---|---|---|---|---|---|
| **f1** | 11 | 12 | 12 | 12 | 12 | 12 | 12 | 15/15 |
| | 5.1(3) | 39(8) | 68(14) | 103(10) | 134(10) | 199(26) | 271(21) | 15/15 |
| **f2** | 83 | 87 | 88 | 89 | 90 | 92 | 94 | 15/15 |
| | 18(2) | 22(1) | 26(2) | 31(2) | 35(2) | 43(1) | 51(2) | 15/15 |
| **f3** | 716 | 1622 | 1637 | 1642 | 1646 | 1650 | 1654 | 15/15 |
| | 2.1(0.8) | 2.7(0.4) | 97(1.0) | 97(304) | 97(304) | 97(453) | 97(152) | 13/15 |
| **f4** | 809 | 1633 | 1688 | 1758 | 1817 | 1886 | 1903 | 15/15 |
| | 1.9(0.6) | 309(611) | 2368(2365) | 2274(2129) | 2200(2334) | 2121(3572) | 2102(1180) | 3/15 |
| **f5** | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 15/15 |
| | 11(3) | 15(3) | 15(2) | 15(6) | 15(4) | 15(7) | 15(5) | 15/15 |
| **f6** | 114 | 214 | 281 | 404 | 580 | 1038 | 1332 | 15/15 |
| | 5.6(2) | 7.8(1) | 10(2) | 11(2) | 10(1) | 8.4(0.9) | 8.9(0.9) | 15/15 |
| **f7** | 24 | 324 | 1171 | 1451 | 1572 | 1572 | 1597 | 15/15 |
| | 10(7) | 2.5(0.1) | 133(214) | 108(173) | 160(318) | 160(318) | 158(470) | 12/15 |
| **f8** | 73 | 273 | 336 | 372 | 391 | 410 | 422 | 15/15 |
| | 12(4) | 572(1831) | 467(745) | 424(673) | 405(640) | 389(610) | 381(593) | 13/15 |
| **f9** | 35 | 127 | 214 | 263 | 300 | 335 | 369 | 15/15 |
| | 23(5) | 18(5) | 15(3) | 14(2) | 15(4) | 17(3) | 19(2) | 15/15 |
| **f10** | 349 | 500 | 574 | 607 | 626 | 829 | 880 | 15/15 |
| | 8.5(1) | 7.4(1) | 7.6(2) | 8.2(1) | 9.3(2) | 8.8(0.7) | 10(1.0) | 15/15 |
| **f11** | 143 | 202 | 763 | 977 | 1177 | 1467 | 1673 | 15/15 |
| | 11(4) | 12(2) | 4.2(0.3) | 4.0(0.6) | 4.0(0.3) | 4.2(0.2) | 4.5(0.6) | 15/15 |
| **f12** | 108 | 268 | 371 | 413 | 461 | 1303 | 1494 | 15/15 |
| | 55(59) | 35(33) | 31(51) | 32(41) | 32(17) | 15(3) | 15(7) | 15/15 |
| **f13** | 132 | 195 | 250 | 319 | 1310 | 1752 | 2255 | 15/15 |
| | 12(2) | 13(1) | 16(2) | 17(1) | 5.1(0.4) | 5.5(0.4) | 6.0(2) | 15/15 |
| **f14** | 10 | 41 | 58 | 90 | 139 | 251 | 476 | 15/15 |
| | 2.2(2) | 10(3) | 16(3) | 16(1) | 16(2) | 16(1) | 13(1) | 15/15 |
| **f15** | 511 | 9310 | 19369 | 19743 | 20073 | 20769 | 21359 | 14/15 |
| | 144(491) | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *1.0e6* | 0/15 |
| **f16** | 120 | 612 | 2662 | 10163 | 10449 | 11644 | 12095 | 15/15 |
| | 5.6(7) | 602(1634) | 1034(1124) | 271(368) | 383(191) | 558(407) | 538(557) | 2/15 |
| **f17** | 5.2 | 215 | 899 | 2861 | 3669 | 6351 | 7934 | 15/15 |
| | 3.3(6) | 3.5(1) | 82(0.2) | 26(87) | 100(136) | 181(79) | 820(850) | 1/15 |
| **f18** | 103 | 378 | 3968 | 8451 | 9280 | 10905 | 12469 | 15/15 |
| | 4.3(2) | 4.6(0.8) | 64(126) | 104(355) | 216(323) | ∞ | ∞ *1.0e6* | 0/15 |
| **f19** | 1 | 1 | 242 | 1.0e5 | 1.2e5 | 1.2e5 | 1.2e5 | 15/15 |
| | 35(27) | 3883(4846) | 8329(5159) | ∞ | ∞ | ∞ | ∞ *1.0e6* | 0/15 |
| **f20** | 16 | 851 | 38111 | 51362 | 54470 | 54861 | 55313 | 14/15 |
| | 8.9(7) | 2.9(0.8) | 52(39) | 39(83) | 37(23) | 36(41) | 36(81) | 5/15 |
| **f21** | 41 | 1157 | 1674 | 1692 | 1705 | 1729 | 1757 | 14/15 |
| | 2.8(3) | 756(1296) | 895(1491) | 886(1918) | 879(1025) | 868(1011) | 854(995) | 6/15 |
| **f22** | 71 | 386 | 938 | 980 | 1008 | 1040 | 1068 | 14/15 |
| | 3.9(2) | 2957(3878) | 2931(2131) | 2804(4841) | 2726(4458) | 2643(2640) | 2576(4210) | 4/15 |
| **f23** | 3.0 | 518 | 14249 | 27890 | 31654 | 33030 | 34256 | 15/15 |
| | 2.8(2) | 17(9) | 985(1297) | ∞ | ∞ | ∞ | ∞ *1.0e6* | 0/15 |
| **f24** | 1622 | 2.2e5 | 6.4e6 | 9.6e6 | 9.6e6 | 1.3e7 | 1.3e7 | 3/15 |
| | 48(3) | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *1.0e6* | 0/15 |

20-D

| Δf | 1e+1 | 1e+0 | 1e-1 | 1e-2 | 1e-3 | 1e-5 | 1e-7 | #succ |
|---|---|---|---|---|---|---|---|---|
| **f1** | 43 | 43 | 43 | 43 | 43 | 43 | 43 | 15/15 |
| | 28(4) | 52(5) | 76(8) | 98(7) | 122(5) | 168(10) | 217(6) | 15/15 |
| **f2** | 385 | 386 | 387 | 388 | 390 | 391 | 393 | 15/15 |
| | 13(0.6) | 16(0.7) | 18(1) | 21(1) | 23(1) | 29(1) | 34(2) | 15/15 |
| **f3** | 5066 | 7626 | 7635 | 7637 | 7643 | 7646 | 7651 | 15/15 |
| | 5128(5323) | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f4** | 4722 | 7628 | 7666 | 7686 | 7700 | 7758 | 1.4e5 | 9/15 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f5** | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 15/15 |
| | 11(3) | 12(5) | 12(4) | 12(5) | 12(4) | 12(4) | 12(5) | 15/15 |
| **f6** | 1296 | 2343 | 3413 | 4255 | 5220 | 6728 | 8409 | 15/15 |
| | 12(2) | 12(4) | 11(3) | 12(6) | 13(7) | 13(8) | 14(6) | 15/15 |
| **f7** | 1351 | 4274 | 9503 | 16523 | 16524 | 16524 | 16969 | 15/15 |
| | 11844(10359) | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f8** | 2039 | 3871 | 4040 | 4148 | 4219 | 4371 | 4484 | 15/15 |
| | 7.5(2) | 265(0.7) | 255(248) | 249(482) | 246(474) | 239(685) | 234(223) | 12/15 |
| **f9** | 1716 | 3102 | 3277 | 3379 | 3455 | 3594 | 3727 | 15/15 |
| | 17(2) | 483(4) | 462(604) | 454(1465) | 450(858) | 447(1101) | 446(800) | 11/15 |
| **f10** | 7413 | 8661 | 10735 | 13641 | 14920 | 17073 | 17476 | 15/15 |
| | 131(60) | 241(103) | 518(415) | 810(632) | ∞ | ∞ | ∞ *3.8e6* | 0/15 |
| **f11** | 1002 | 2228 | 6278 | 8586 | 9762 | 12285 | 14831 | 15/15 |
| | 28(17) | 32(22) | 18(4) | 18(7) | 20(4) | 23(9) | 24(2) | 15/15 |
| **f12** | 1042 | 1938 | 2740 | 3156 | 4140 | 12407 | 13827 | 15/15 |
| | 612(859) | 1664(6942) | 1679(2388) | 3316(4544) | ∞ | ∞ | ∞ *3.6e6* | 0/15 |
| **f13** | 652 | 2021 | 2751 | 3507 | 18749 | 24455 | 30201 | 15/15 |
| | 1540(1532) | 3955(8395) | 9439(17419) | 15945(21634) | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f14** | 75 | 239 | 304 | 451 | 932 | 1648 | 15661 | 15/15 |
| | 11(3) | 9.2(0.8) | 11(0.6) | 12(1.0) | 12(3) | 113(66) | 1236(831) | 0/15 |
| **f15** | 30378 | 1.5e5 | 3.1e5 | 3.2e5 | 3.2e5 | 4.5e5 | 4.6e5 | 15/15 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f16** | 1384 | 27265 | 77015 | 1.4e5 | 1.9e5 | 2.0e5 | 2.2e5 | 15/15 |
| | 5778(8661) | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f17** | 63 | 1030 | 4005 | 12242 | 30677 | 56288 | 80472 | 15/15 |
| | 8.6(6) | 2570(3830) | 13805(19985) | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f18** | 621 | 3972 | 19561 | 28555 | 67569 | 1.3e5 | 1.5e5 | 15/15 |
| | 4.8(2) | 13838(14068) | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f19** | 1 | 1 | 3.4e5 | 4.7e6 | 6.2e6 | 6.7e6 | 6.7e6 | 15/15 |
| | 628(250) | 4.5e6 (6e6) | ∞ | ∞ | ∞ | ∞ | ∞ *3.8e6* | 0/15 |
| **f20** | 82 | 46150 | 3.1e6 | 5.5e6 | 5.5e6 | 5.6e6 | 5.6e6 | 14/15 |
| | 16(4) | 563(779) | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f21** | 561 | 6541 | 14103 | 14318 | 14643 | 15567 | 17589 | 15/15 |
| | 1782(5339) | 3972(3513) | 3967(3754) | 3908(3559) | 3821(1774) | 3594(2118) | 3181(1931) | 1/15 |
| **f22** | 467 | 5580 | 23491 | 24163 | 24948 | 26847 | 1.3e5 | 12/15 |
| | 2144(6421) | 1970(2328) | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f23** | 3.2 | 1614 | 67457 | 3.7e5 | 4.9e5 | 8.1e5 | 8.4e5 | 15/15 |
| | 1.7(1) | 4993(6189) | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 0/15 |
| **f24** | 1.3e6 | 7.5e6 | 5.2e7 | 5.2e7 | 5.2e7 | 5.2e7 | 5.2e7 | 3/15 |
| | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ | ∞ *4.0e6* | 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 $\Delta f$-values are shown in the top row. #succ is the number of trials that reached the (final) target $f_{\rm opt} + 10^{-8}$. 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.**

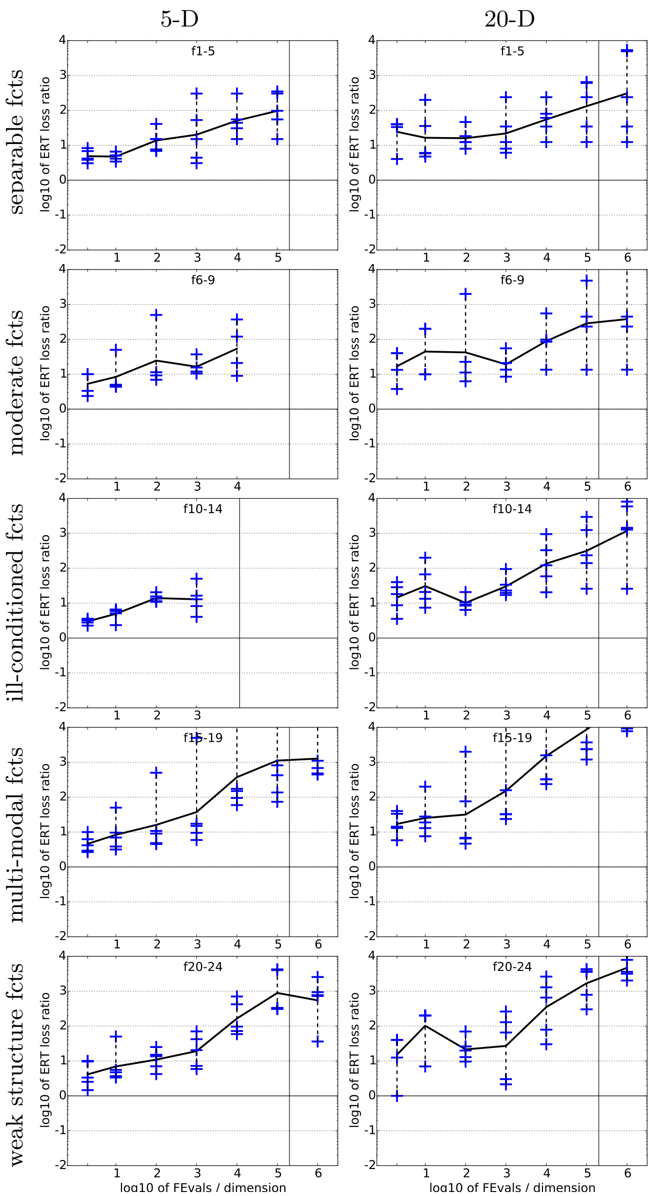



**Figure 4: ERT loss ratios (see Figure 3 for details). Each cross (+) represents a single function, the line is the geometric mean.**