Comparison of NEWUOA with Different Numbers of Interpolation Points on the BBOB Noiseless Testbed

Raymond Ros

Comparison of NEWUOA with Different Numbers of Interpolation Points on the BBOB Noiseless Testbed

Raymond Ros
TAO Team-Project – INRIA Saclay
LRI, Bat 490, Univ. Paris-Sud
F-91405 Orsay Cedex, France
raymond.ros@inria.fr

ABSTRACT

In this paper, we study the performances of the NEW Unconstrained Optimization Algorithm (NEWUOA) with different numbers of interpolation points. NEWUOA is a trust region method, the number of points used to build the surrogate model is an input parameter of the algorithm. We compare the performances of NEWUOA using three different number of points in search spaces of dimension from two to forty on problems from the BBOB 2009 noiseless function testbed.

In particular we study the performances of an 'average' number of interpolation points that scales like the dimension of the search space to the power $3/2$. Using this number of interpolation points is expectedly faster than using the maximum number of interpolation points (scaling like the square of the dimension), though it does not grant better performances than using a number of interpolation points scaling like the dimension.

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. INTRODUCTION

The NEWUOA, for NEW Unconstrained Optimization Algorithm was introduced in [5] as a method for unconstrained derivative-free optimization. NEWUOA is a trust-region method which uses m points to build a quadratic approximation of the objective function. The approximation is considered reliable within the radius of the current trust region. In this paper, we study the effect of the number of interpolation points m on the performances of NEWUOA on a testbed of noiseless functions.

We use three different values for m which will be denoted **NEWUOA**, **avg-NEWUOA** and **full-NEWUOA**. These variants are sorted by ascending numbers of interpolation points. The number of interpolation points of these variants depends on the dimension of the search space n. The variant denoted NEWUOA uses $2n + 1$ interpolation points as recommended in [5]. The avg-NEWUOA uses the rounded value of $\sqrt{(n + 1/2)(n + 1)(n + 2)}$ interpolation points which is intermediate. The full-NEWUOA uses the maximum number $\frac{(n+1)(n+2)}{2}$. These three settings were already compared on a few test problems in [5].

The performances of the avg-NEWUOA are obtained on the BBOB 2009 testbed of noiseless functions. The avg-NEWUOA is successively compared to NEWUOA and full-NEWUOA. The performances of both NEWUOA and full-NEWUOA on the BBOB 2009 noiseless functions were presented in [7].

2. EXPERIMENTAL PROCEDURE

To benchmark the avg-NEWUOA, we use the exact same experimental procedure that was presented in [7]. In particular the algorithm uses an independent multi-start procedure, as do NEWUOA and full-NEWUOA. The crafting effort [3] is equal to CrE $= 0$ for all three variants of the NEWUOA.

3. CPU TIMING EXPERIMENT

According to [5], the complexity of the NEWUOA variants we consider is at worst $\Theta(mn)$, where n is the dimension of the search space. The algorithms were run on f_8 for at least 30 seconds, according to [3].

Results for the the different variants of NEWUOA are given in Table 1. These figures were obtained on a Intel Core 2 6700 processor (2.66 GHz) with Linux 2.6.24.7.

4. RESULTS

Results from experiments according to [3] on the benchmark functions given in [1, 4] are presented in this section. The Figures 1 and 2 and the Table 2 compare the avg-NEWUOA to NEWUOA. The Figures 3 and 4 and the Table 3 compare the avg-NEWUOA to full-NEWUOA. The **expected running time (ERT)**, used in the figures and tables, depends on a given target function value, $f_t = f_{opt} +$

GECCO'10, July 7–11, 2010, Portland, Oregon, USA.

Table 1: CPU-Time per function evaluations of f_8 in microseconds for variants of NEWUOA with different number of interpolation points

	2-D	3-D	5-D	10-D	20-D	40-D
NEWUOA	8.1	11	21	58	170	620
avg	8.0	13	27	100	580	3900
full	9.0	15	38	240	2400	32000

Δf, and is computed over all relevant trials as the number of function evaluations executed during each trial while the best function value did not reach f_t, summed over all trials and divided by the number of trials that actually reached f_t [3, 6]. **Statistical significance** is tested with the rank-sum test for a given target Δf_t (10^{-8} in Figure 1) using, for each trial, either the number of needed function evaluations to reach Δf_t (inverted and multiplied by -1), or, if the target was not reached, the best Δf-value achieved, measured only up to the smallest number of overall function evaluations for any unsuccessful trial under consideration.

The performances of all variants of NEWUOA are rather similar over all test functions of the BBOB 2009 noiseless testbed as shown in the top sub-figures of Figures 2 and 4. Some differences can be spotted. Figure 1 shows NEWUOA is faster than avg-NEWUOA on functions f_1 and f_5 by a factor growing as the dimension of the search space increases. There is a factor of two in 20-D, a factor of three in 40-D. Also, on the Bent Cigar function f_{12}, NEWUOA is faster than avg-NEWUOA by a factor of up to a hundred.

In a rather similar way, avg-NEWUOA is faster than full-NEWUOA on functions f_1 and f_5 by a factor growing as the dimension of the search space increases. The avg-NEWUOA is faster than full-NEWUOA in 20-D on functions f_2, f_6, f_8, f_9, f_{10}, f_{11}, f_{12}.

The performances of all three NEWUOA variants on f_{14} show strange behaviour since the precision of 10^{-7} is reached consistently but 10^{-8} is not. None of the three variants reached the target precision 10^{-8} on functions f_{15} to f_{19} except in 2-D and 3-D.

Acknowledgment

The author would like to acknowledge Dr Mohamed Jebalia for his insights on this paper.

5. REFERENCES

[1] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Presentation of the noiseless functions. Technical Report 2009/20, Research Center PPE, 2009. Updated February 2010.

[2] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2009: Experimental setup. Technical Report RR-6828, INRIA, 2009.

[3] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2010: Experimental setup. Technical Report RR-7215, INRIA, 2010.

[4] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2009. Updated February 2010.

[5] M. J. D. Powell. The NEWUOA software for unconstrained optimization without derivatives. *Large Scale Nonlinear Optimization*, pages 255–297, 2006.

[6] K. Price. Differential evolution vs. the functions of the second ICEO. In *Proceedings of the IEEE International Congress on Evolutionary Computation*, pages 153–157, 1997.

[7] R. Ros. Benchmarking the NEWUOA on the BBOB-2009 function testbed. In F. Rothlauf, editor, *GECCO (Companion)*, pages 2421–2428. ACM, 2009.

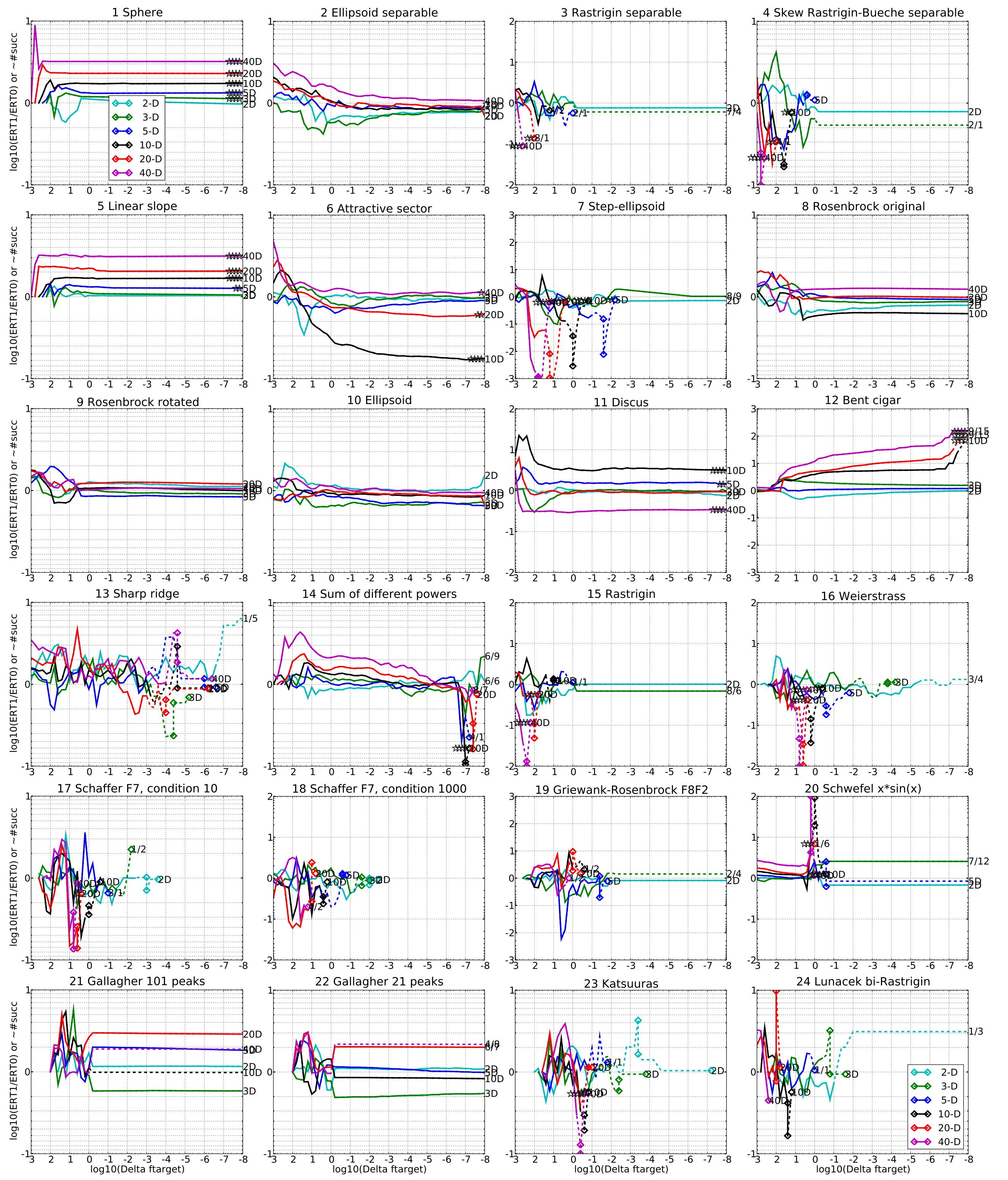

Figure 1: ERT ratio of avg-NEWUOA divided by NEWUOA versus $\log_{10}(\Delta f)$ for f_1–f_{24} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds thrice the median of the trial-wise overall number of f-evaluations for the same algorithm on this function. Symbols indicate the best achieved Δf-value of one algorithm (ERT gets undefined to the right). The dashed line continues as the fraction of successful trials of the other algorithm, where 0 means 0% and the y-axis limits mean 100%, values below zero for avg-NEWUOA. The line ends when no algorithm reaches Δf anymore. The number of successful trials is given, only if it was in $\{1 \dots 9\}$ for avg-NEWUOA (1st number) and non-zero for NEWUOA (2nd number). Results are significant with $p = 0.05$ for one star and $p = 10^{-\#\star}$ otherwise, with Bonferroni correction within each figure.

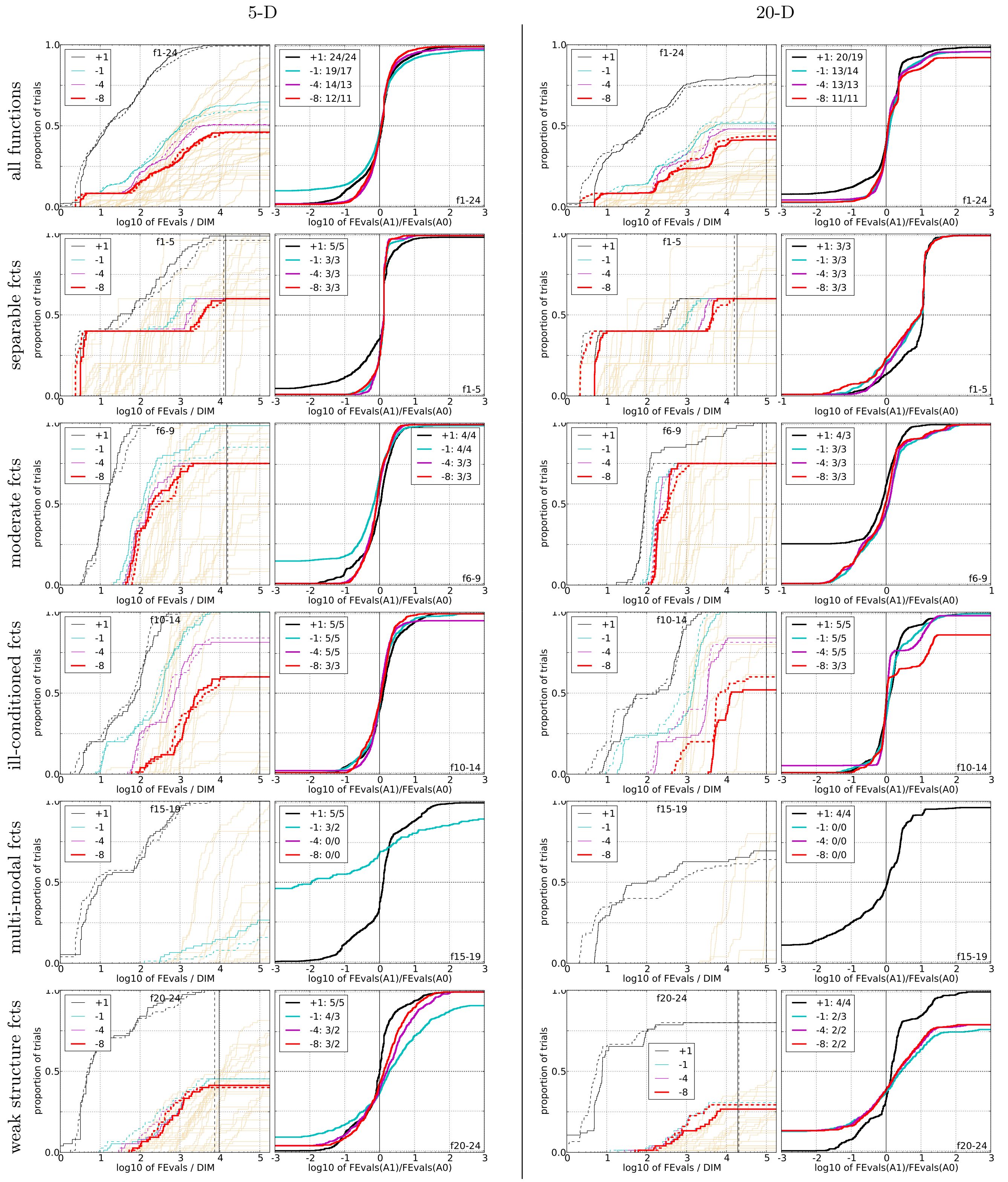

Figure 2: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/D) to reach a target value $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ is given by the first value in the legend, for avg-NEWUOA (solid) and NEWUOA (dashed). Light beige lines show the ECDF of FEvals for target value $\Delta f = 10^{-8}$ of algorithms benchmarked during BBOB-2009. Right sub-columns: ECDF of FEval ratios of avg-NEWUOA divided by NEWUOA, all trial pairs for each function. Pairs where both trials failed are disregarded, pairs where one trial failed are visible in the limits being > 0 or < 1. The legends indicate the number of functions that were solved in at least one trial (avg-NEWUOA first).

5-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f1	11	12	12	12	12	12	15/15
0: NEW	**1.1**$^{\star 3}$	**1**$^{\star 3}$	**1**$^{\star 3}$	**1**$^{\star 3}$	**1**$^{\star 3}$	**1**$^{\star 3}$	15/15
1: AVG	1.5	1.3	1.3	1.3	1.3	1.3	15/15
f2	83	87	88	90	92	94	15/15
0: NEW	5.7	22	45	85	129	166	15/15
1: AVG	6.4	21	41	75	108	145	15/15
f3	716	1622	1637	1646	1650	1654	15/15
0: NEW	6.1	229	∞	∞	∞	∞ *2.5e4*	0/15
1: AVG	3.0	130	∞	∞	∞	∞ *3.0e4*	0/15
f4	809	1633	1688	1817	1886	1903	15/15
0: NEW	27	305	∞	∞	∞	∞ *3.4e4*	0/15
1: AVG	14	∞	∞	∞	∞	∞ *4.1e4*	0/15
f5	10	10	10	10	10	10	15/15
0: NEW	**1.3**$^{\star 3}$	**1.5**$^{\star 2}$	**1.5**$^{\star 2}$	**1.5***	**1.5***	**1.5***	15/15
1: AVG	1.8	1.9	1.9	1.9	1.9	1.9	15/15
f6	114	214	281	580	1038	1332	15/15
0: NEW	1.7	2.4	3.6	3.3	2.7	2.9	15/15
1: AVG	1.3	1.6	2.6	2.6	2.4	2.5	15/15
f7	24	324	1171	1572	1572	1597	15/15
0: NEW	10	13	60	∞	∞	∞ *2.9e4*	0/15
1: AVG	4.4	5.9	**13***	∞	∞	∞ *3.8e4*	0/15
f8	73	273	336	391	410	422	15/15
0: NEW	1	1.1	1.2	1.2	1.2	1.2	15/15
1: AVG	1.2	1.2	1.1	1.1	1.1	1.1	15/15
f9	35	127	214	300	335	369	15/15
0: NEW	1.8	3.6	2.5	1.9	1.9	1.7	15/15
1: AVG	2.4	3.1	2.1	1.7	1.6	1.5	15/15
f10	349	500	574	626	829	880	15/15
0: NEW	3.1	5.5	8.1	14	16	21	15/15
1: AVG	3.1	4.6	6.6	10	11	14	15/15
f11	143	202	763	1177	1467	1673	15/15
0: NEW	3.5	4.7	1.8	**1.8***	**2.0**$^{\star 2}$	**2.2**$^{\star 2}$	15/15
1: AVG	5.4	7.2	2.7	2.8	3.1	3.4	15/15
f12	108	268	371	461	1303	1494	15/15
0: NEW	3.5	2.6	2.5	2.6	1.1	1.1	15/15
1: AVG	3.5	2.8	2.8	3.0	1.3	1.4	15/15
f13	132	195	250	1310	1752	2255	15/15
0: NEW	3.1	9.3	35	54	335	∞ *4.0e4*	0/15
1: AVG	4.5	8.1	42	68	391	∞ *4.7e4*	0/15
f14	10	41	58	139	251	476	15/15
0: NEW	1.7	1	1	1.2	5.5	2525	0/15
1: AVG	2.1	1.0	1.0	1.2	5.0	1029	0/15
f15	511	9310	19369	20073	20769	21359	14/15
0: NEW	5.8	41	∞	∞	∞	∞ *2.5e4*	0/15
1: AVG	5.8	46	∞	∞	∞	∞ *2.9e4*	0/15
f16	120	612	2662	10449	11644	12095	15/15
0: NEW	2.1	29	∞	∞	∞	∞ *3.6e4*	0/15
1: AVG	2.6	12	47	∞	∞	∞ *4.0e4*	0/15
f17	5.2	215	899	3669	6351	7934	15/15
0: NEW	2.3	40	617	∞	∞	∞ *3.4e4*	0/15
1: AVG	3.1	42	405	∞	∞	∞ *5.5e4*	0/15
f18	103	378	3968	9280	10905	12469	15/15
0: NEW	31	1351	∞	∞	∞	∞ *9.2e4*	0/15
1: AVG	10	272	∞	∞	∞	∞ *1.5e5*	0/15
f19	1	1	242	1.20e5	1.21e5	1.22e5	15/15
0: NEW	**14***	26728	1415	∞	∞	∞ *5.0e5*	0/15
1: AVG	24	15619	995	∞	∞	∞ *5.0e5*	0/15
f20	16	851	38111	54470	54861	55313	14/15
0: NEW	1	3.3	∞	∞	∞	∞ *3.2e4*	0/15
1: AVG	1	8.4	12	8.2	8.1	8.0	1/15
f21	41	1157	1674	1705	1729	1757	14/15
0: NEW	1.1	2.2	1.8	1.8	1.8	1.9	15/15
1: AVG	1.7	2.5	3.6	3.5	3.5	3.5	15/15
f22	71	386	938	1008	1040	1068	14/15
0: NEW	2.1	2.1	2.0	2.1	2.3	2.4	15/15
1: AVG	3.4	2.6	2.3	2.3	2.4	2.4	15/15
f23	3.0	518	14249	31654	33030	34256	15/15
0: NEW	6.2	2.4	7.1	∞	∞	∞ *3.3e4*	0/15
1: AVG	6.0	2.5	14	∞	∞	∞ *4.5e4*	0/15
f24	1622	2.16e5	6.36e6	9.62e6	1.28e7	1.28e7	3/15
0: NEW	2.9	2.1	∞	∞	∞	∞ *3.0e4*	0/15
1: AVG	2.0	2.2	∞	∞	∞	∞ *3.3e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f1	43	43	43	43	43	43	15/15
0: NEW	**1.0**$^{\star 3}$	**1.0**$^{\star 3}$	**1.0**$^{\star 3}$	**1.0**$^{\star 3}$	**1.0**$^{\star 3}$	**1.0**$^{\star 3}$	15/15
1: AVG	2.3	2.3	2.3	2.3	2.3	2.3	15/15
f2	385	386	387	390	391	393	15/15
0: NEW	18	42	71	125	174	219	15/15
1: AVG	21	43	63	116	161	199	15/15
f3	5066	7626	7635	7643	7646	7651	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *1.3e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.1e5*	0/15
f4	4722	7628	7666	7700	7758	1.41e5	9/15
0: NEW	∞	∞	∞	∞	∞	∞ *2.2e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *3.2e5*	0/15
f5	41	41	41	41	41	41	15/15
0: NEW	**1.2**$^{\star 3}$	**1.5**$^{\star 3}$	**1.6**$^{\star 3}$	**1.6**$^{\star 3}$	**1.6**$^{\star 3}$	**1.6**$^{\star 3}$	15/15
1: AVG	2.7	3.2	3.3	3.3	3.3	3.3	15/15
f6	1296	2343	3413	5220	6728	8409	15/15
0: NEW	1	1	1	1.1	1.3	1.3	15/15
1: AVG	1.00	0.74	0.72	0.70$^{\downarrow 4}$	0.74$^{\downarrow 4}$	**0.73**$^{\star 2 \downarrow 4}$	15/15
f7	1351	4274	9503	16524	16524	16969	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *4.8e5*	0/15
1: AVG	**101**$^{\star 3}$	∞	∞	∞	∞	∞ *4.5e5*	0/15
f8	2039	3871	4040	4219	4371	4484	15/15
0: NEW	1	1	1	1	1	1	15/15
1: AVG	0.96	0.97	1.0	1.0	1.0	0.99	15/15
f9	1716	3102	3277	3455	3594	3727	15/15
0: NEW	1.0	1	1	1	1	1	15/15
1: AVG	1.0	1.2	1.2	1.3	1.2	1.2	15/15
f10	7413	8661	10735	14920	17073	17476	15/15
0: NEW	1.7	2.6	3.3	4.0	4.7	5.8	15/15
1: AVG	1.5	2.6	3.1	3.6	4.2	5.0	15/15
f11	1002	2228	6278	9762	12285	14831	15/15
0: NEW	15	13	5.8	6.1	6.6	6.5	15/15
1: AVG	15	11	5.7	5.6	5.8	6.1	15/15
f12	1042	1938	2740	4140	12407	13827	15/15
0: NEW	3.0	3.0	3.0	**2.5**$^{\star 2}$	**1**$^{\star 3}$	**1**$^{\star 3}$	15/15
1: AVG	11	15	18	24	12	21	9/15
f13	652	2021	2751	18749	24455	30201	15/15
0: NEW	**1***	3.0	9.3	19	∞	∞ *1.8e5*	0/15
1: AVG	1.5	5.3	14	14	172	∞ *3.0e5*	0/15
f14	75	239	304	932	1648	15661	15/15
0: NEW	**1.5**$^{\star 2}$	**1***	**1**$^{\star 3}$	**1**$^{\star 2}$	9.1	43	0/15
1: AVG	2.7	1.5	1.6	1.3	9.3	**26**$^{\star 3}$	0/15
f15	30378	1.47e5	3.12e5	3.20e5	4.49e5	4.59e5	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *1.3e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
f16	1384	27265	77015	1.88e5	1.98e5	2.20e5	15/15
0: NEW	16	∞	∞	∞	∞	∞ *2.3e5*	0/15
1: AVG	3.6	∞	∞	∞	∞	∞ *3.2e5*	0/15
f17	63	1030	4005	30677	56288	80472	15/15
0: NEW	16	∞	∞	∞	∞	∞ *1.5e6*	0/15
1: AVG	2.4	∞	∞	∞	∞	∞ *8.7e5*	0/15
f18	621	3972	19561	67569	1.31e5	1.47e5	15/15
0: NEW	11930	∞	∞	∞	∞	∞ *1.6e6*	0/15
1: AVG	3217	∞	∞	∞	∞	∞ *1.2e6*	0/15
f19	1	1	3.43e5	6.22e6	6.69e6	6.74e6	15/15
0: NEW	**76***	4.29e6	∞	∞	∞	∞ *2.0e6*	0/15
1: AVG	210	8.03e6	∞	∞	∞	∞ *2.0e6*	0/15
f20	82	46150	3.10e6	5.54e6	5.59e6	5.64e6	14/15
0: NEW	1	**15**$^{\star 2}$	∞	∞	∞	∞ *3.8e5*	0/15
1: AVG	1.3	107	∞	∞	∞	∞ *3.4e5*	0/15
f21	561	6541	14103	14643	15567	17589	15/15
0: NEW	1.7	2.2	1.2	1.2	1.1	1	15/15
1: AVG	3.2	5.7	3.5	3.4	3.3	2.9	14/15
f22	467	5580	23491	24948	26847	1.35e5	12/15
0: NEW	1	4.9	6.8	6.4	6.0	1.2	7/15
1: AVG	2.0	5.6	14	13	12	2.4	6/15
f23	3.2	1614	67457	4.89e5	8.11e5	8.38e5	15/15
0: NEW	12	3.5	32	∞	∞	∞ *1.5e5*	0/15
1: AVG	15	4.7	∞	∞	∞	∞ *3.7e5*	0/15
f24	1.34e6	7.48e6	5.19e7	5.20e7	5.20e7	5.20e7	3/15
0: NEW	∞	∞	∞	∞	∞	∞ *1.7e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.1e5*	0/15

Table 2: Expected running time (ERT in number of function evaluations) divided by the best ERT measured during BBOB-2009 (given in the respective first row) for different Δf values for functions f_1–f_{24}. The median number of conducted function evaluations is additionally given in *italics*, if $\mathrm{ERT}(10^{-7}) = \infty$. #succ is the number of trials that reached the final target $f_{\mathrm{opt}} + 10^{-8}$. 0: NEW is NEWUOA and 1: AVG is avg-NEWUOA. Bold entries are statistically significantly better compared to the other algorithm, with $p = 0.05$ or $p = 10^{-k}$ where $k > 1$ is the number following the $\star$ symbol, with Bonferroni correction of 48.

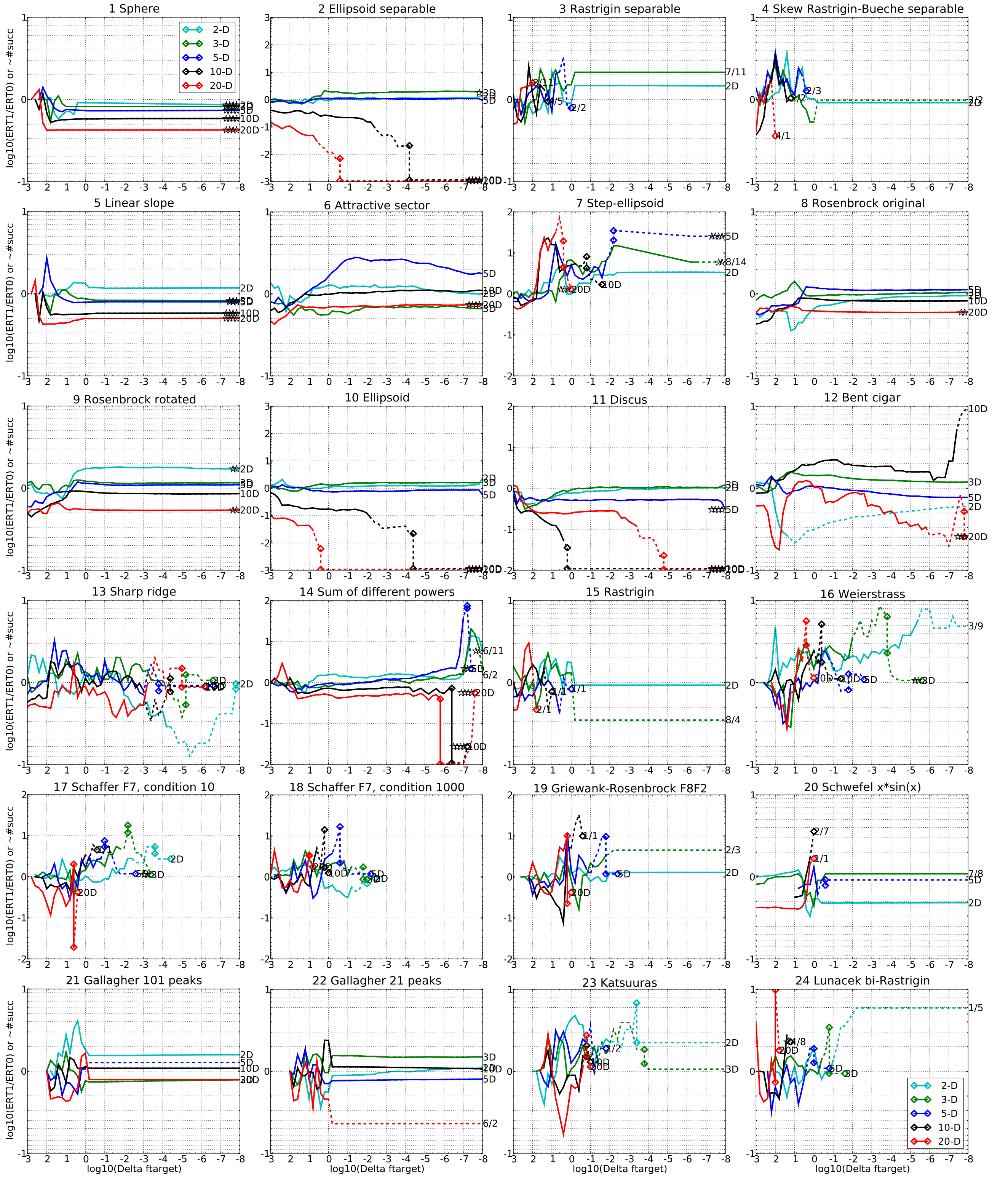

Figure 3: ERT ratio of avg-NEWUOA divided by full-NEWUOA versus $\log_{10}(\Delta f)$ for f_1–f_{24} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds thrice the median of the trial-wise overall number of f-evaluations for the same algorithm on this function. Symbols indicate the best achieved Δf-value of one algorithm (ERT gets undefined to the right). The dashed line continues as the fraction of successful trials of the other algorithm, where 0 means 0% and the y-axis limits mean 100%, values below zero for avg-NEWUOA. The line ends when no algorithm reaches Δf anymore. The number of successful trials is given, only if it was in $\{1\ldots9\}$ for avg-NEWUOA (1st number) and non-zero for full-NEWUOA (2nd number). Results are significant with $p = 0.05$ for one star and $p = 10^{-\#\star}$ otherwise, with Bonferroni correction within each figure.

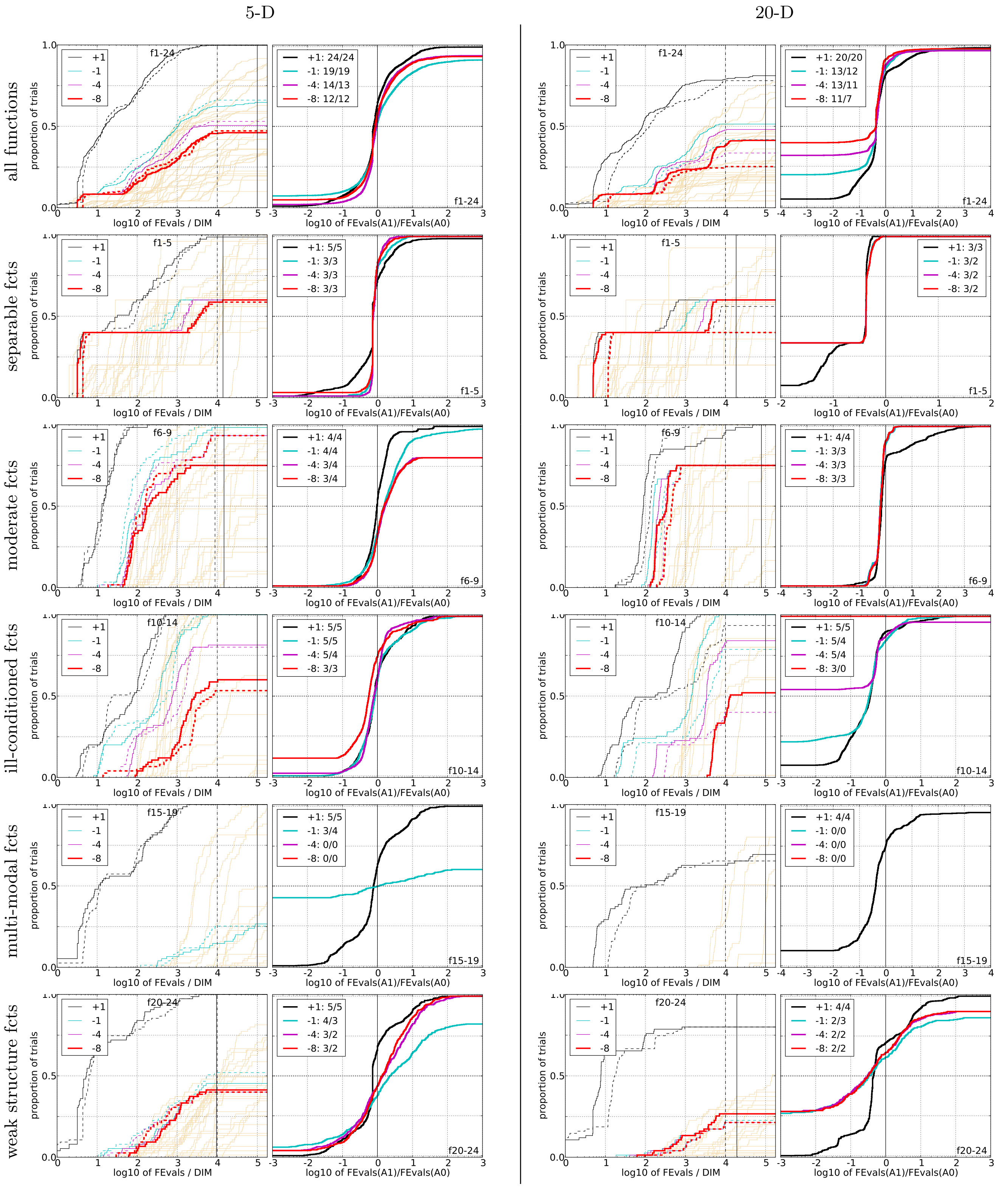

Figure 4: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/D) to reach a target value $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ is given by the first value in the legend, for avg-NEWUOA (solid) and full-NEWUOA (dashed). Light beige lines show the ECDF of FEvals for target value $\Delta f = 10^{-8}$ of algorithms benchmarked during BBOB-2009. Right sub-columns: ECDF of FEval ratios of avg-NEWUOA divided by full-NEWUOA, all trial pairs for each function. Pairs where both trials failed are disregarded, pairs where one trial failed are visible in the limits being > 0 or < 1. The legends indicate the number of functions that were solved in at least one trial (avg-NEWUOA first).

5-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f1	11	12	12	12	12	12	15/15
0: FUL	1.9	1.8	1.8	1.8	1.8	1.8	15/15
1: AVG	**1.5**$^{\star 2}$	**1.3**$^{\star 3}$	**1.3**$^{\star 3}$	**1.3**$^{\star 3}$	**1.3**$^{\star 3}$	**1.3**$^{\star 3}$	15/15
f2	83	87	88	90	92	94	15/15
0: FUL	6.9	19	36	69	103	134	14/15
1: AVG	6.4	21	41	75	108	145	15/15
f3	716	1622	1637	1646	1650	1654	15/15
0: FUL	4.2	164	∞	∞	∞	∞ *3.7e4*	0/15
1: AVG	3.0	130	∞	∞	∞	∞ *3.0e4*	0/15
f4	809	1633	1688	1817	1886	1903	15/15
0: FUL	12	∞	∞	∞	∞	∞ *5.0e4*	0/15
1: AVG	14	∞	∞	∞	∞	∞ *4.1e4*	0/15
f5	10	10	10	10	10	10	15/15
0: FUL	2.2	2.4	2.4	2.4	2.4	2.4	15/15
1: AVG	**1.8**$^{\star 2}$	**1.9**$^{\star 3}$	**1.9**$^{\star 2}$	**1.9**$^{\star 2}$	**1.9**$^{\star 2}$	**1.9**$^{\star 2}$	15/15
f6	114	214	281	580	1038	1332	15/15
0: FUL	1.2	1	1	1	1	1.4	15/15
1: AVG	1.3	1.6	2.6	2.6	2.4	2.5	15/15
f7	24	324	1171	1572	1572	1597	15/15
0: FUL	1	1.2	4.0	**24***	**24***	**23***	11/15
1: AVG	4.4	5.9	13	∞	∞	∞ *3.8e4*	0/15
f8	73	273	336	391	410	422	15/15
0: FUL	1.6	1	1	1	1	1	15/15
1: AVG	1.2	1.2	1.1	1.1	1.1	1.1	15/15
f9	35	127	214	300	335	369	15/15
0: FUL	2.6	2.7	1.9	1.5	1.4	1.3	15/15
1: AVG	2.4	3.1	2.1	1.7	1.6	1.5	15/15
f10	349	500	574	626	829	880	15/15
0: FUL	3.6	6.2	8.7	13	13	17	12/15
1: AVG	3.1	4.6	6.6	10	11	14	15/15
f11	143	202	763	1177	1467	1673	15/15
0: FUL	11	14	5.4	5.3	5.9	6.5	13/15
1: AVG	**5.4**$^{\star 2}$	**7.2**$^{\star 3}$	**2.7**$^{\star 3}$	**2.8**$^{\star 3}$	**3.1**$^{\star 3}$	**3.4**$^{\star 3}$	15/15
f12	108	268	371	461	1303	1494	15/15
0: FUL	3.7	2.6	2.7	3.3	1.6	1.7	15/15
1: AVG	3.5	2.8	2.8	3.0	1.3	1.4	15/15
f13	132	195	250	1310	1752	2255	15/15
0: FUL	1.8	6.7	23	97	∞	∞ *5.0e4*	0/15
1: AVG	4.5	8.1	42	68	391	∞ *4.7e4*	0/15
f14	10	41	58	139	251	476	15/15
0: FUL	2.7	1.1	1.1	1	**3.2***	**26**$^{\star 3}$	0/15
1: AVG	**2.1***	1.0	1.0	1.2	5.0	1029	0/15
f15	511	9310	19369	20073	20769	21359	14/15
0: FUL	6.3	55	∞	∞	∞	∞ *3.4e4*	0/15
1: AVG	5.8	46	∞	∞	∞	∞ *2.9e4*	0/15
f16	120	612	2662	10449	11644	12095	15/15
0: FUL	2.7	12	29	∞	∞	∞ *4.8e4*	0/15
1: AVG	2.6	12	47	∞	∞	∞ *4.0e4*	0/15
f17	5.2	215	899	3669	6351	7934	15/15
0: FUL	4.9	25	76	∞	∞	∞ *5.0e4*	0/15
1: AVG	3.1	42	405	∞	∞	∞ *5.5e4*	0/15
f18	103	378	3968	9280	10905	12469	15/15
0: FUL	10	84	90	∞	∞	∞ *5.0e4*	0/15
1: AVG	10	272	∞	∞	∞	∞ *1.5e5*	0/15
f19	1	1	242	1.20e5	1.21e5	1.22e5	15/15
0: FUL	31	10526	865	∞	∞	∞ *5.0e4*	0/15
1: AVG	24	15619	995	∞	∞	∞ *5.0e5*	0/15
f20	16	851	38111	54470	54861	55313	14/15
0: FUL	1.4	6.4	∞	∞	∞	∞ *3.2e4*	0/15
1: AVG	**1**$^{\star 3}$	8.4	12	8.2	8.1	8.0	1/15
f21	41	1157	1674	1705	1729	1757	14/15
0: FUL	2.4	2.3	2.8	2.8	2.7	2.7	15/15
1: AVG	1.7	2.5	3.6	3.5	3.5	3.5	15/15
f22	71	386	938	1008	1040	1068	14/15
0: FUL	4.3	3.7	3.0	2.9	3.0	3.1	15/15
1: AVG	3.4	2.6	2.3	2.3	2.4	2.4	15/15
f23	3.0	518	14249	31654	33030	34256	15/15
0: FUL	5.4	2.0	3.8	∞	∞	∞ *5.0e4*	0/15
1: AVG	6.0	2.5	14	∞	∞	∞ *4.5e4*	0/15
f24	1622	2.16e5	6.36e6	9.62e6	1.28e7	1.28e7	3/15
0: FUL	2.5	1.1	∞	∞	∞	∞ *3.4e4*	0/15
1: AVG	2.0	2.2	∞	∞	∞	∞ *3.3e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f1	43	43	43	43	43	43	15/15
0: FUL	5.4	5.4	5.5	5.5	5.5	5.5	15/15
1: AVG	**2.3**$^{\star 3}$	**2.3**$^{\star 3}$	**2.3**$^{\star 3}$	**2.3**$^{\star 3}$	**2.3**$^{\star 3}$	**2.3**$^{\star 3}$	15/15
f2	385	386	387	390	391	393	15/15
0: FUL	448	3769	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	**21**$^{\star 3}$	**43**$^{\star 3}$	**63**$^{\star 3}$	**116**$^{\star 3}$	**161**$^{\star 3}$	**199**$^{\star 3}$	15/15
f3	5066	7626	7635	7643	7646	7651	15/15
0: FUL	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.1e5*	0/15
f4	4722	7628	7666	7700	7758	1.41e5	9/15
0: FUL	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *3.2e5*	0/15
f5	41	41	41	41	41	41	15/15
0: FUL	6.2	6.5	6.6	6.6	6.6	6.6	15/15
1: AVG	**2.7**$^{\star 3}$	**3.2**$^{\star 3}$	**3.3**$^{\star 3}$	**3.3**$^{\star 3}$	**3.3**$^{\star 3}$	**3.3**$^{\star 3}$	15/15
f6	1296	2343	3413	5220	6728	8409	15/15
0: FUL	1.4	1.1	1.0	1	1	1	15/15
1: AVG	**1.00**$^{\star 3}$	**0.74**$^{\star 3}$	**0.72**$^{\star 3}$	**0.70**$^{\star 3 \downarrow 4}$	**0.74**$^{\star 3 \downarrow 4}$	**0.73**$^{\star 3 \downarrow 4}$	15/15
f7	1351	4274	9503	16524	16524	16969	15/15
0: FUL	4.6	**700**$^{\star 2}$	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	101	∞	∞	∞	∞	∞ *4.5e5*	0/15
f8	2039	3871	4040	4219	4371	4484	15/15
0: FUL	1.4	1.6	1.6	1.7	1.7	1.7	15/15
1: AVG	**0.96***	0.97	1.0	**1.0***	**1.0***	**0.99***	15/15
f9	1716	3102	3277	3455	3594	3727	15/15
0: FUL	1.8	2.2	2.3	2.3	2.3	2.2	15/15
1: AVG	**1.0**$^{\star 3}$	1.2	1.2	1.3	1.2	1.2	15/15
f10	7413	8661	10735	14920	17073	17476	15/15
0: FUL	34	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	**1.5**$^{\star 3}$	**2.6**$^{\star 3}$	**3.1**$^{\star 3}$	**3.6**$^{\star 3}$	**4.2**$^{\star 3}$	**5.0**$^{\star 3}$	15/15
f11	1002	2228	6278	9762	12285	14831	15/15
0: FUL	57	45	21	36	∞	∞ *2.0e5*	0/15
1: AVG	**15**$^{\star 3}$	**11**$^{\star 3}$	**5.7**$^{\star 3}$	**5.6**$^{\star 3}$	**5.8**$^{\star 3}$	**6.1**$^{\star 3}$	15/15
f12	1042	1938	2740	4140	12407	13827	15/15
0: FUL	11	15	26	38	35	105	0/15
1: AVG	11	15	18	24	12	21	9/15
f13	652	2021	2751	18749	24455	30201	15/15
0: FUL	1.8	6.0	18	26	116	∞ *2.0e5*	0/15
1: AVG	**1.5***	5.3	14	14	172	∞ *3.0e5*	0/15
f14	75	239	304	932	1648	15661	15/15
0: FUL	5.9	3.0	3.6	2.6	20	∞ *2.0e5*	0/15
1: AVG	**2.7**$^{\star 3}$	**1.5**$^{\star 3}$	**1.6**$^{\star 3}$	**1.3**$^{\star 3}$	**9.3**$^{\star 3}$	**26**$^{\star 3}$	0/15
f15	30378	1.47e5	3.12e5	3.20e5	4.49e5	4.59e5	15/15
0: FUL	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
f16	1384	27265	77015	1.88e5	1.98e5	2.20e5	15/15
0: FUL	4.6	108	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	3.6	∞	∞	∞	∞	∞ *3.2e5*	0/15
f17	63	1030	4005	30677	56288	80472	15/15
0: FUL	13	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	**2.4**$^{\star 2}$	∞	∞	∞	∞	∞ *8.7e5*	0/15
f18	621	3972	19561	67569	1.31e5	1.47e5	15/15
0: FUL	948	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	3217	∞	∞	∞	∞	∞ *1.2e6*	0/15
f19	1	1	3.43e5	6.22e6	6.69e6	6.74e6	15/15
0: FUL	475	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	210	8.03e6	∞	∞	∞	∞ *2.0e6*	0/15
f20	82	46150	3.10e6	5.54e6	5.59e6	5.64e6	14/15
0: FUL	3.1	64	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	**1.3**$^{\star 3}$	107	∞	∞	∞	∞ *3.4e5*	0/15
f21	561	6541	14103	14643	15567	17589	15/15
0: FUL	7.4	3.4	4.5	4.4	4.1	3.7	14/15
1: AVG	3.2	5.7	3.5	3.4	3.3	2.9	14/15
f22	467	5580	23491	24948	26847	1.35e5	12/15
0: FUL	2.2	12	60	57	53	11	2/15
1: AVG	2.0	5.6	14	13	12	2.4	6/15
f23	3.2	1614	67457	4.89e5	8.11e5	8.38e5	15/15
0: FUL	14	7.0	44	∞	∞	∞ *2.0e5*	0/15
1: AVG	15	4.7	∞	∞	∞	∞ *3.7e5*	0/15
f24	1.34e6	7.48e6	5.19e7	5.20e7	5.20e7	5.20e7	3/15
0: FUL	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.1e5*	0/15

Table 3: Expected running time (ERT in number of function evaluations) divided by the best ERT measured during BBOB-2009 (given in the respective first row) for different Δf values for functions f_1–f_{24}. The median number of conducted function evaluations is additionally given in *italics*, if $\mathrm{ERT}(10^{-7}) = \infty$. #succ is the number of trials that reached the final target $f_{\mathrm{opt}} + 10^{-8}$. 0: FUL is full-NEWUOA and 1: AVG is avg-NEWUOA. Bold entries are statistically significantly better compared to the other algorithm, with $p = 0.05$ or $p = 10^{-k}$ where $k > 1$ is the number following the $\star$ symbol, with Bonferroni correction of 48.