HAL archives-ouvertes.fr

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

Raymond Ros

To cite this version:

Raymond Ros. Comparison of NEWUOA with Different Numbers of Interpolation Points on the BBOB Noisy Testbed. Genetic and Evolutionary Computation Conference 2010, Jul 2010, Portland, OR, United States. inria-00473776

HAL Id: inria-00473776
https://hal.inria.fr/inria-00473776

Submitted on 16 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of NEWUOA with Different Numbers of Interpolation Points on the BBOB Noisy 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 noisy function testbed. Using the maximum number of interpolation points grants the better results in this noisy setting.

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 noisy functions.

We use three different values for m which will be denoted **NEWUOA**, **avg-NEWUOA** and **full-NEWUOA**.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
GECCO'10, July 7–11, 2010, Portland, Oregon, USA.
Copyright 2010 ACM 978-1-4503-0073-5/10/07 ...$10.00.

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. RESULTS

Results of the CPU-timing experiments are given in the paper benchmarking NEWUOA with all three settings of m on the BBOB noiseless testbed submitted to the same workshop.

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 1 compare the avg-NEWUOA to NEWUOA. The Figures 3 and 4 and the Table 2 compare the avg-NEWUOA to full-NEWUOA. The **expected running time (ERT)**, used in the figures and table, depends on a given target function value, $f_\mathrm{t} = f_\mathrm{opt} + \Delta f_\mathrm{t}$, 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 success probability of avg-NEWUOA reaching the precision 10^{-8} is only slightly larger than that of NEWUOA: in 5-D avg-NEWUOA is successful on 52 function instances (out of 450) in 5-D and NEWUOA only on 42; in 20-D avg-NEWUOA is successful on 30 function instances and NEWUOA only on 15. On functions f_{101}, f_{102} and f_{103} which both NEWUOA and avg-NEWUOA solve in 5-D, the NEWUOA is slower faster than avg-NEWUOA.

To reach the precision 10^{-8}, full-NEWUOA has the best success probability out of the three variants of NEWUOA: it solves 77 function instances in 5-D, 31 in 20-D. Also, the full-NEWUOA solves in 5-D the Rosenbrock function f_{106} and the Sphere function f_{109}, both with moderate Cauchy noise, whereas avg-NEWUOA does not.

In 20-D, the NEWUOA only solves the Sphere function with moderate noise. The full-NEWUOA solves the Sphere function with all three noise models and is the fastest out of the three NEWUOA variants. In 20-D again, avg-NEWUOA does not solve f_{103}, NEWUOA does not solve neither f_{102} nor f_{103} and is the slowest out of the three variants of NEWUOA on f_{101}.

Overall the full-NEWUOA variant which uses the largest number of interpolation points performs better than the avg-NEWUOA and the NEWUOA on the BBOB 2009 noisy testbed.

4. REFERENCES

[1] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2010: Presentation of the noisy functions. Technical Report 2009/21, Research Center PPE, 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: Noisy functions definitions. Technical Report RR-6869, 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 noisy testbed. In F. Rothlauf, editor, *GECCO (Companion)*, pages 2429–2434. ACM, 2009.

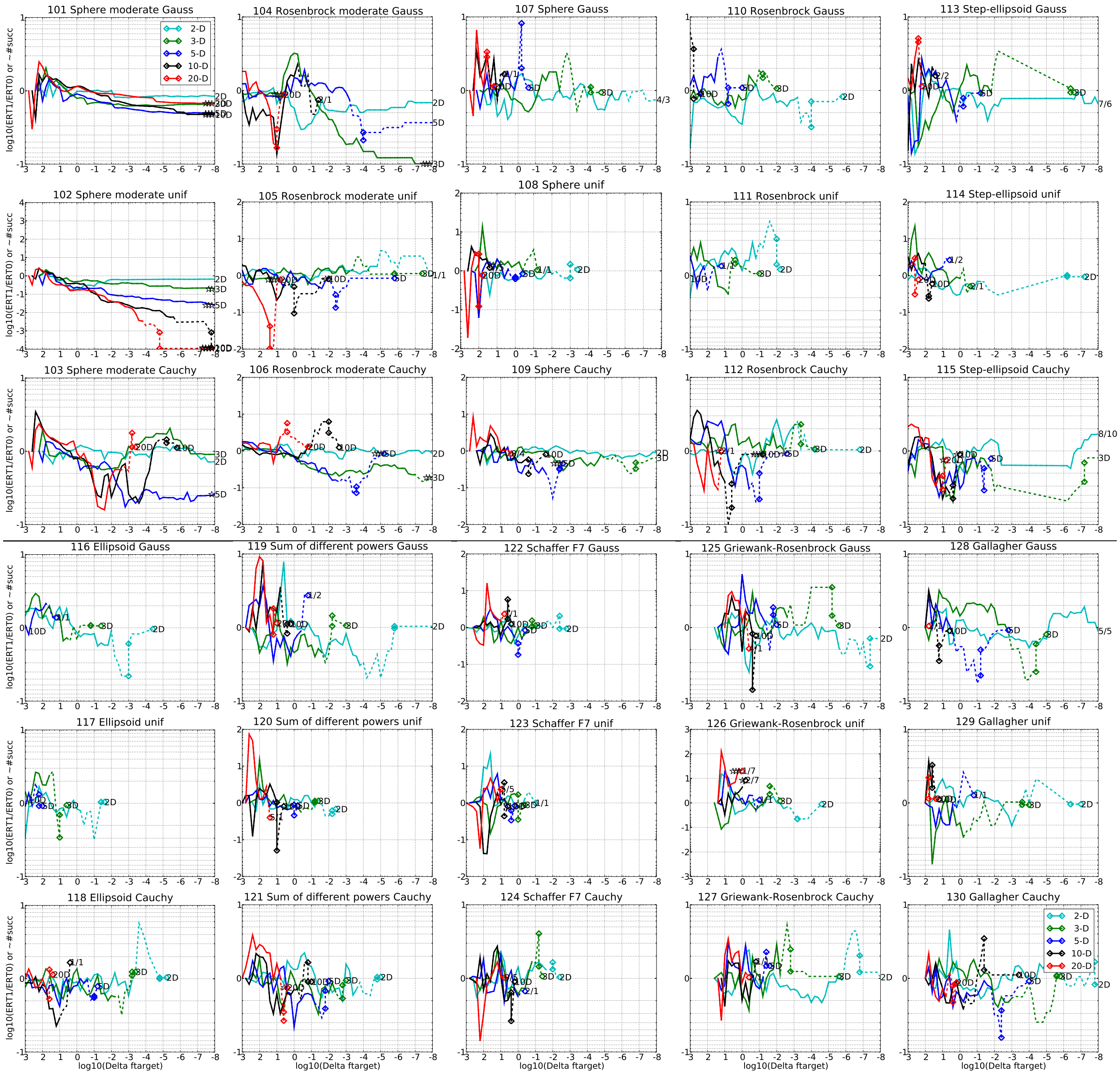

Figure 1: Ratio of the expected running times (ERT) of avg-NEWUOA divided by NEWUOA versus $\log_{10}(\Delta f)$ for f_{101}–f_{130} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds 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 statistically significant with $p = 0.05$ for one star and $p = 10^{-\#\star}$ otherwise, with Bonferroni correction within each figure.

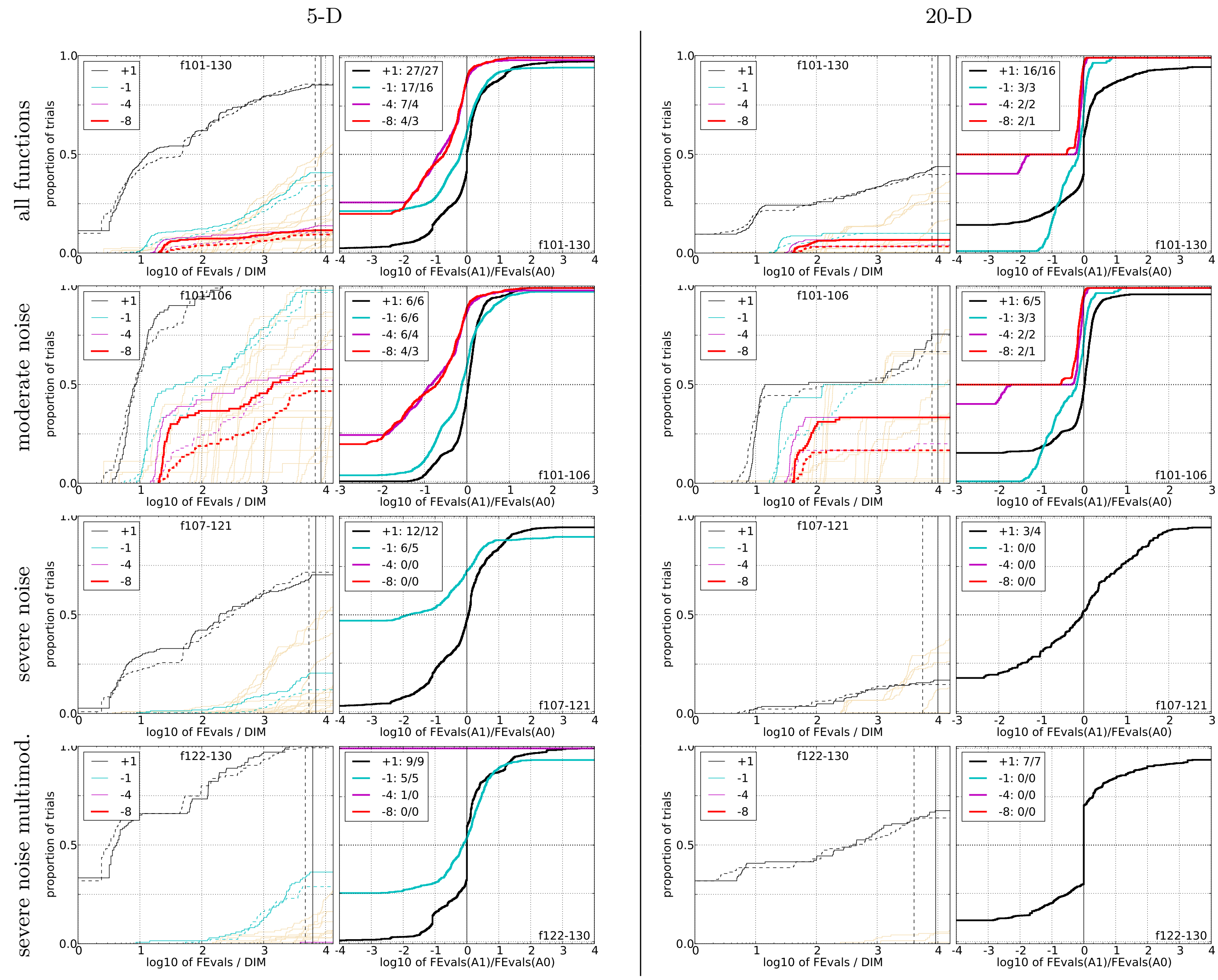

Figure 2: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of necessary function evaluations divided by dimension D (FEvals/D) to reached a target value $f_{\mathrm{opt}} + \Delta f$ with $\Delta f = 10^k$, 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 all 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
f_{101}	11	37	44	62	69	75	15/15
0: NEW	2.5	1.6	2.1	2.6	3.0	3.1	15/15
1: AVG	2.9	1.5	1.6	1.5	**1.5***	**1.5**$^{\star 2}$	15/15
f_{102}	11	35	50	72	86	99	15/15
0: NEW	6.3	6.0	7.0	20	33	41	15/15
1: AVG	2.7	1.4	1.5	**1.5**$^{\star 2}$	**1.6**$^{\star 2}$	**1.5**$^{\star 2}$	15/15
f_{103}	11	28	30	31	35	115	15/15
0: NEW	2.4	1.9	5.7	60	178	136	12/15
1: AVG	2.5	1.6	3.6	13	42	34	15/15
f_{104}	173	773	1287	1768	2040	2284	15/15
0: NEW	1.2	3.4	6.0	24	∞	∞ *2.7e4*	0/15
1: AVG	1.0	5.0	7.6	23	20	24	7/15
f_{105}	167	1436	5174	10388	10824	11202	15/15
0: NEW	1.7	2.7	3.3	∞	∞	∞ *2.5e4*	0/15
1: AVG	1.7	2.4	1.9	14	46	∞ *3.4e4*	0/15
f_{106}	86	529	1050	2666	2887	3087	15/15
0: NEW	1	2.2	5.0	59	∞	∞ *3.3e4*	0/15
1: AVG	0.91	1.5	2.4	**9.1***	**200***	∞ *4.1e4*	0/15
f_{107}	40	228	453	940	1376	1850	15/15
0: NEW	60	194	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	68	317	∞	∞	∞	∞ *3.1e4*	0/15
f_{108}	87	5144	14469	30935	58628	80667	15/15
0: NEW	77	64	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	155	44	∞	∞	∞	∞ *3.1e4*	0/15
f_{109}	11	57	216	572	873	946	15/15
0: NEW	4.8	13	83	∞	∞	∞ *2.3e4*	0/15
1: AVG	4.3	3.6	26	∞	∞	∞ *3.0e4*	0/15
f_{110}	949	33625	1.20e5	5.93e5	6.03e5	6.11e5	15/15
0: NEW	118	10	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	241	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{111}	6856	6.12e5	8.83e6	2.30e7	3.10e7	3.13e7	3/15
0: NEW	∞	∞	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{112}	107	1684	3421	4502	5132	5596	15/15
0: NEW	1.9	7.7	105	∞	∞	∞ *2.5e4*	0/15
1: AVG	3.1	4.9	23	∞	∞	∞ *3.3e4*	0/15
f_{113}	133	1883	8081	24128	24128	24402	15/15
0: NEW	13	44	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	14	31	55	∞	∞	∞ *3.1e4*	0/15
f_{114}	767	14720	56311	83272	83272	84949	15/15
0: NEW	43	∞	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	74	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{115}	64	485	1829	2550	2550	2970	15/15
0: NEW	2.9	14	42	∞	∞	∞ *2.2e4*	0/15
1: AVG	1.1	4.2	28	∞	∞	∞ *3.0e4*	0/15
f_{116}	5730	14472	22311	26868	30329	31661	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{117}	26686	76052	1.10e5	1.37e5	1.73e5	1.92e5	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{118}	429	1217	1555	1998	2430	2913	15/15
0: NEW	4.3	10	116	∞	∞	∞ *2.6e4*	0/15
1: AVG	1.8	8.3	64	∞	∞	∞ *3.4e4*	0/15
f_{119}	12	657	1136	10372	35296	49747	15/15
0: NEW	26	35	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	19	23	∞	∞	∞	∞ *3.1e4*	0/15
f_{120}	16	2900	18698	72438	3.33e5	5.48e5	15/15
0: NEW	130	55	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	94	49	∞	∞	∞	∞ *3.1e4*	0/15
f_{121}	8.6	111	273	1583	3870	6195	15/15
0: NEW	4.8	15	76	∞	∞	∞ *2.2e4*	0/15
1: AVG	4.3	3.3	45	∞	∞	∞ *3.0e4*	0/15
f_{122}	10	1727	9190	30087	53743	1.11e5	15/15
0: NEW	14	91	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	6.1	34	∞	∞	∞	∞ *3.1e4*	0/15
f_{123}	11	16066	81505	3.36e5	6.71e5	2.22e6	15/15
0: NEW	65	∞	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	80	∞	∞	∞	∞	∞ *3.1e4*	0/15
f_{124}	10	202	1040	20478	45337	95200	15/15
0: NEW	3.0	158	∞	∞	∞	∞ *2.2e4*	0/15
1: AVG	6.1	89	∞	∞	∞	∞ *2.9e4*	0/15
f_{125}	1	1	1	2.39e5	2.43e5	2.46e5	15/15
0: NEW	3.9	15	6088	∞	∞	∞ *2.3e4*	0/15
1: AVG	2.0	81	9966	∞	∞	∞ *3.1e4*	0/15
f_{126}	1	1	1	∞	∞	∞	0
0: NEW	1.2	1053	3.47e5	∞	∞	∞	0/15
1: AVG	22	1636	4.54e5	∞	∞	∞	0/15
f_{127}	1	1	1	3.42e5	3.89e5	3.95e5	15/15
0: NEW	2.5	14	7248	∞	∞	∞ *2.2e4*	0/15
1: AVG	2.0	18	5233	∞	∞	∞ *2.9e4*	0/15
f_{128}	111	4248	7808	12447	17217	21162	15/15
0: NEW	12	17	43	∞	∞	∞ *2.3e4*	0/15
1: AVG	11	9.3	7.5	∞	∞	∞ *3.1e4*	0/15
f_{129}	64	10710	59443	2.85e5	5.11e5	5.80e5	15/15
0: NEW	124	16	∞	∞	∞	∞ *2.3e4*	0/15
1: AVG	68	20	∞	∞	∞	∞ *3.1e4*	0/15
f_{130}	55	812	3034	32823	33889	34528	10/15
0: NEW	2.3	11	10	∞	∞	∞ *2.2e4*	0/15
1: AVG	1.3	6.4	5.9	3.0	∞	∞ *3.0e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f_{101}	59	361	513	700	739	783	15/15
0: NEW	3.1	1	1	1.1	1.5	1.6	15/15
1: AVG	3.3	1.2	0.95	0.94	**1.0**$^{\star 2}$	**1.1**$^{\star 2}$	15/15
f_{102}	231	399	579	921	1157	1407	15/15
0: NEW	2.9	6.1	6.3	45	∞	∞ *1.0e5*	0/15
1: AVG	0.92	1.1	**0.93***	**0.88**$^{\star 3}$	**1.1**$^{\star 3}$	**1.3**$^{\star 3}$	15/15
f_{103}	65	417	629	1313	1893	2464	14/15
0: NEW	2.3	1	5.9	1231	∞	∞ *1.1e5*	0/15
1: AVG	3.0	0.95	2.0	655	∞	∞ *2.0e5*	0/15
f_{104}	23690	85656	1.71e5	1.82e5	1.89e5	1.96e5	15/15
0: NEW	68	∞	∞	∞	∞	∞ *1.1e5*	0/15
1: AVG	11	∞	∞	∞	∞	∞ *2.0e5*	0/15
f_{105}	1.92e5	6.11e5	6.32e5	6.49e5	6.60e5	6.70e5	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *9.4e4*	0/15
1: AVG	**7.2**$^{\star 3}$	∞	∞	∞	∞	∞ *2.0e5*	0/15
f_{106}	11480	21668	23746	25470	26492	27360	15/15
0: NEW	7.0	31	∞	∞	∞	∞ *1.5e5*	0/15
1: AVG	8.1	∞	∞	∞	∞	∞ *2.0e5*	0/15
f_{107}	8571	13582	16226	27357	52486	65052	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{108}	58063	97228	2.03e5	4.46e5	6.30e5	8.98e5	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{109}	333	632	1138	2287	3583	4952	15/15
0: NEW	17	∞	∞	∞	∞	∞ *8.4e4*	0/15
1: AVG	17	∞	∞	∞	∞	∞ *1.9e5*	0/15
f_{110}	∞	∞	∞	∞	∞	∞	0
0: NEW	∞	∞	∞	∞	∞	∞	0/15
1: AVG	∞	∞	∞	∞	∞	∞	0/15
f_{111}	∞	∞	∞	∞	∞	∞	0
0: NEW	∞	∞	∞	∞	∞	∞	0/15
1: AVG	∞	∞	∞	∞	∞	∞	0/15
f_{112}	25552	64124	69621	73557	76137	78238	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *9.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
f_{113}	50123	3.64e5	5.60e5	5.88e5	5.88e5	5.91e5	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{114}	2.08e5	1.12e6	1.45e6	1.57e6	1.57e6	1.58e6	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{115}	2405	30268	91749	1.27e5	1.27e5	1.29e5	15/15
0: NEW	236	∞	∞	∞	∞	∞ *8.5e4*	0/15
1: AVG	108	∞	∞	∞	∞	∞ *1.9e5*	0/15
f_{116}	4.98e5	6.94e5	8.93e5	1.03e6	1.08e6	1.12e6	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.7e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{117}	1.79e6	2.46e6	2.60e6	2.91e6	3.24e6	3.62e6	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.7e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{118}	6908	11786	17514	26342	30062	32659	15/15
0: NEW	∞	∞	∞	∞	∞	∞ *1.1e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *2.0e5*	0/15
f_{119}	2771	29365	35930	4.11e5	1.40e6	1.90e6	15/15
0: NEW	398	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{120}	36040	1.79e5	2.81e5	1.59e6	6.74e6	1.35e7	13/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{121}	249	769	1426	9304	34434	57404	15/15
0: NEW	31	∞	∞	∞	∞	∞ *8.0e4*	0/15
1: AVG	49	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{122}	692	52008	1.40e5	7.93e5	2.00e6	5.82e6	15/15
0: NEW	82	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	125	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{123}	1063	5.30e5	1.49e6	5.29e6	2.71e7	1.58e8	0
0: NEW	174	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	405	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{124}	192	1959	40840	1.27e5	3.89e5	7.99e5	15/15
0: NEW	91	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	95	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{125}	1	1	1	2.50e7	8.03e7	8.06e7	4/15
0: NEW	1	414	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	1	493	∞	∞	∞	∞ *1.8e5*	0/15
f_{126}	1	1	1	∞	∞	∞	0
0: NEW	4.2	**1.32e5**$^{\star 2}$	∞	∞	∞	∞	0/15
1: AVG	122	2.60e6	∞	∞	∞	∞	0/15
f_{127}	1	1	1	4.43e6	7.27e6	7.43e6	15/15
0: NEW	3.7	253	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	7.7	219	∞	∞	∞	∞ *1.8e5*	0/15
f_{128}	1.40e5	1.34e7	1.72e7	1.72e7	1.72e7	1.72e7	9/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{129}	7.81e6	4.13e7	4.15e7	4.18e7	4.21e7	4.24e7	5/15
0: NEW	∞	∞	∞	∞	∞	∞ *7.8e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞ *1.8e5*	0/15
f_{130}	4904	93149	2.52e5	2.54e5	2.55e5	2.57e5	7/15
0: NEW	9.1	∞	∞	∞	∞	∞ *8.0e4*	0/15
1: AVG	6.6	∞	∞	∞	∞	∞ *1.8e5*	0/15

Table 1: 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_{101}–f_{130}. #succ is the number of trials that reached the final target $f_{\rm opt} + 10^{-8}$. 0: NEW is NEWUOA and 1: AVG is avg-NEWUOA. #succ is the number of trials that reached the final target $f_{\rm opt} + 10^{-8}$. 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 60.

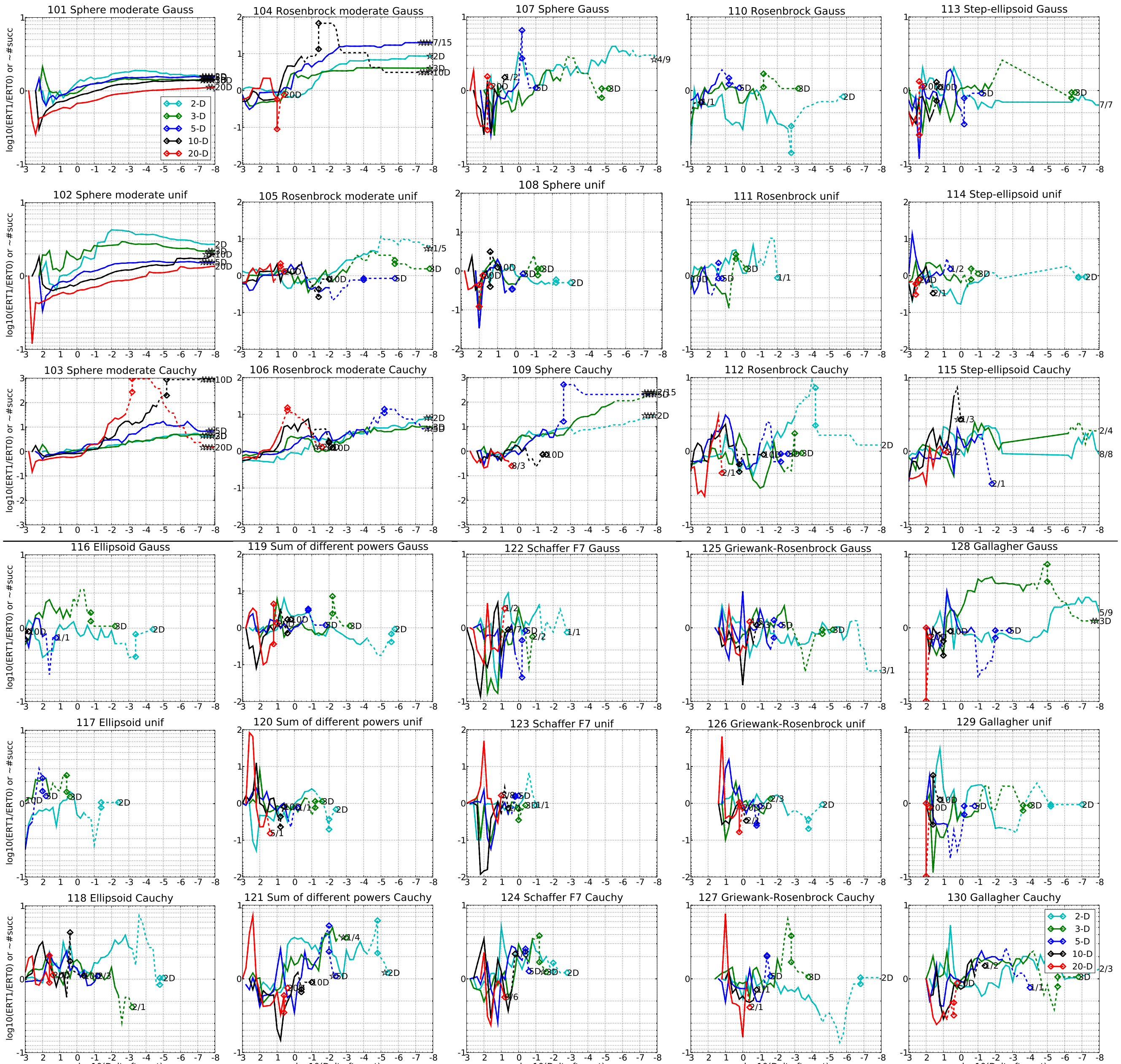

Figure 3: Ratio of the expected running times (ERT) of avg-NEWUOA divided by full-NEWUOA versus $\log_{10}(\Delta f)$ for f_{101}–f_{130} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT 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 $\{1 \ldots 9\}$ for avg-NEWUOA (1st number) and non-zero for full-NEWUOA (2nd number). Results are statistically significant with $p = 0.05$ for one star and $p = 10^{-\#\star}$ otherwise, with Bonferroni correction within each figure.

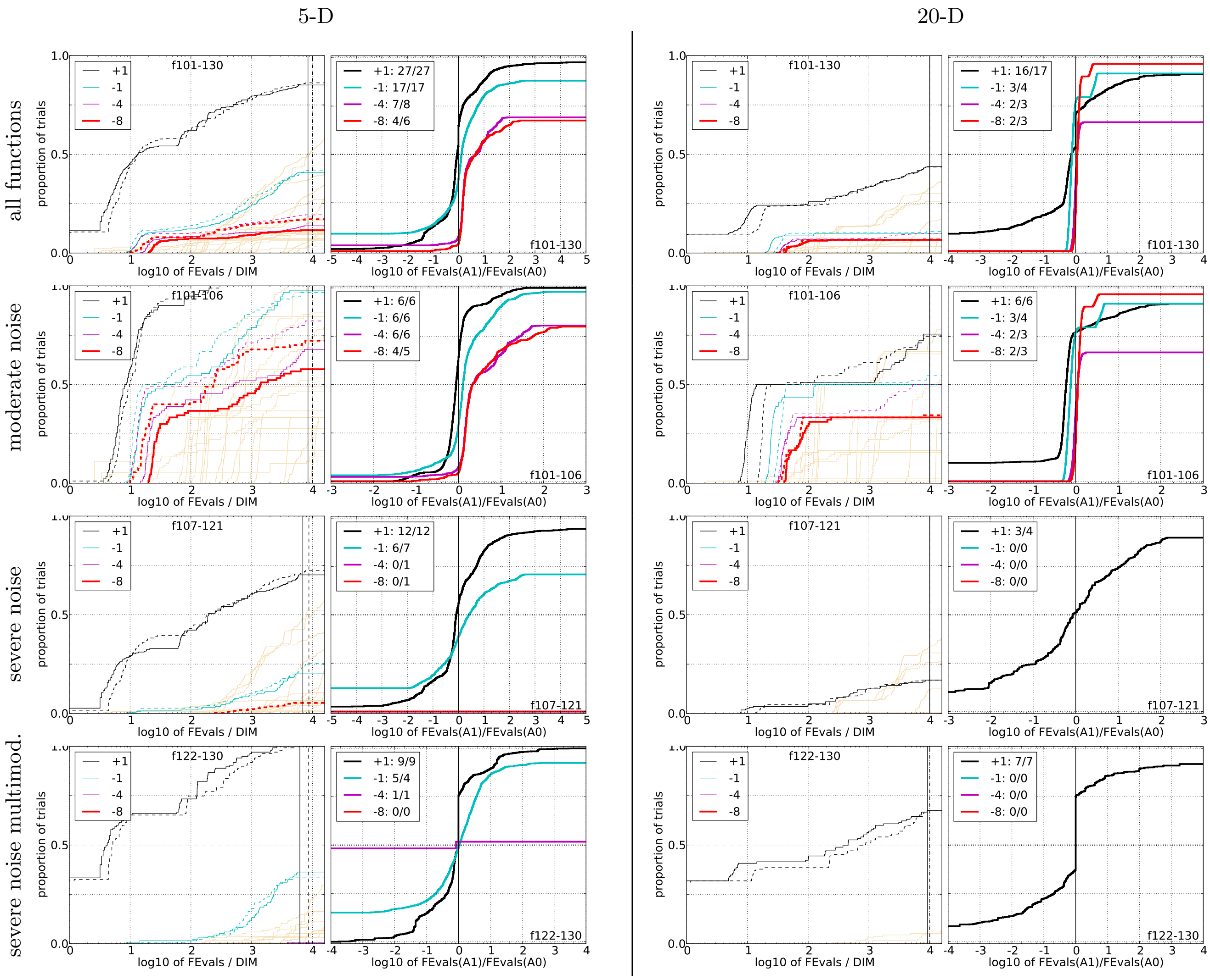

Figure 4: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of necessary function evaluations divided by dimension D (FEvals/D) to reached a target value $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ 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 all 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
f101	11	37	44	62	69	75	15/15
0: FUL	3.3	1.3	**1.3***	**1**$^{\star 3}$	**1**$^{\star 3}$	**1**$^{\star 3}$	15/15
1: AVG	2.9	1.5	1.6	1.5	1.5	1.5	15/15
f102	11	35	50	72	86	99	15/15
0: FUL	3.2	1.4	1.2	**1**$^{\star 2}$	**1**$^{\star 2}$	**1**$^{\star 2}$	15/15
1: AVG	2.7	1.4	1.5	1.5	1.6	1.5	15/15
f103	11	28	30	31	35	115	15/15
0: FUL	2.8	1.6	1.8	**2.6**$^{\star 2}$	**2.4**$^{\star 3}$	**4.8***	15/15
1: AVG	2.5	1.6	3.6	13	42	34	15/15
f104	173	773	1287	1768	2040	2284	15/15
0: FUL	1.1	1.8	**1.7***	**1.4**$^{\star 3}$	**1.3**$^{\star 2}$	**1.2**$^{\star 2}$	15/15
1: AVG	1.0	5.0	7.6	23	20	24	7/15
f105	167	1436	5174	10388	10824	11202	15/15
0: FUL	1.7	1.6	4.5	31	∞	∞*4.6e4*	0/15
1: AVG	1.7	2.4	1.9	14	46	∞*3.4e4*	0/15
f106	86	529	1050	2666	2887	3087	15/15
0: FUL	1.1	1	1	3.1	18	31	5/15
1: AVG	0.91	1.5	2.4	9.1	200	∞*4.1e4*	0/15
f107	40	228	453	940	1376	1850	15/15
0: FUL	85	247	1373	∞	∞	∞*4.2e4*	0/15
1: AVG	68	317	∞	∞	∞	∞*3.1e4*	0/15
f108	87	5144	14469	30935	58628	80667	15/15
0: FUL	78	∞	∞	∞	∞	∞*4.3e4*	0/15
1: AVG	155	44	∞	∞	∞	∞*3.1e4*	0/15
f109	11	57	216	572	873	946	15/15
0: FUL	5.6	5.0	6.1	**21**$^{\star 3}$	**22**$^{\star 3}$	**20**$^{\star 3}$	12/15
1: AVG	4.3	3.6	26	∞	∞	∞*3.0e4*	0/15
f110	949	33625	1.20e5	5.93e5	6.03e5	6.11e5	15/15
0: FUL	144	∞	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	241	∞	∞	∞	∞	∞*3.1e4*	0/15
f111	6856	6.12e5	8.83e6	2.30e7	3.10e7	3.13e7	3/15
0: FUL	∞	∞	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*3.1e4*	0/15
f112	107	1684	3421	4502	5132	5596	15/15
0: FUL	1	6.3	18	∞	∞	∞*4.3e4*	0/15
1: AVG	3.1	4.9	23	∞	∞	∞*3.3e4*	0/15
f113	133	1883	8081	24128	24128	24402	15/15
0: FUL	19	57	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	14	31	55	∞	∞	∞*3.1e4*	0/15
f114	767	14720	56311	83272	83272	84949	15/15
0: FUL	60	∞	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	74	∞	∞	∞	∞	∞*3.1e4*	0/15
f115	64	485	1829	2550	2550	2970	15/15
0: FUL	1	2.8	17	∞	∞	∞*4.0e4*	0/15
1: AVG	1.1	4.2	28	∞	∞	∞*3.0e4*	0/15
f116	5730	14472	22311	26868	30329	31661	15/15
0: FUL	∞	∞	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*3.1e4*	0/15
f117	26686	76052	1.10e5	1.37e5	1.73e5	1.92e5	15/15
0: FUL	∞	∞	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*3.1e4*	0/15
f118	429	1217	1555	1998	2430	2913	15/15
0: FUL	1.3	5.9	70	∞	∞	∞*4.3e4*	0/15
1: AVG	1.8	8.3	64	∞	∞	∞*3.4e4*	0/15
f119	12	657	1136	10372	35296	49747	15/15
0: FUL	12	22	261	∞	∞	∞*4.2e4*	0/15
1: AVG	19	23	∞	∞	∞	∞*3.1e4*	0/15
f120	16	2900	18698	72438	3.33e5	5.48e5	15/15
0: FUL	151	63	∞	∞	∞	∞*4.3e4*	0/15
1: AVG	94	49	∞	∞	∞	∞*3.1e4*	0/15
f121	8.6	111	273	1583	3870	6195	15/15
0: FUL	3.2	10	33	∞	∞	∞*4.0e4*	0/15
1: AVG	4.3	3.3	45	∞	∞	∞*3.0e4*	0/15
f122	10	1727	9190	30087	53743	1.11e5	15/15
0: FUL	30	110	∞	∞	∞	∞*4.2e4*	0/15
1: AVG	6.1	34	∞	∞	∞	∞*3.1e4*	0/15
f123	11	16066	81505	3.36e5	6.71e5	2.22e6	15/15
0: FUL	113	18	∞	∞	∞	∞*4.3e4*	0/15
1: AVG	80	∞	∞	∞	∞	∞*3.1e4*	0/15
f124	10	202	1040	20478	45337	95200	15/15
0: FUL	5.3	45	∞	∞	∞	∞*3.9e4*	0/15
1: AVG	6.1	89	∞	∞	∞	∞*2.9e4*	0/15
f125	1	1	1	2.39e5	2.43e5	2.46e5	15/15
0: FUL	3.9	26	5566	∞	∞	∞*4.2e4*	0/15
1: AVG	2.0	81	9966	∞	∞	∞*3.1e4*	0/15
f126	1	1	1	∞	∞	∞	0
0: FUL	2.6	1743	∞	∞	∞	∞	0/15
1: AVG	22	1636	4.54e5	∞	∞	∞	0/15
f127	1	1	1	3.42e5	3.89e5	3.95e5	15/15
0: FUL	4.1	25	10666	∞	∞	∞*3.9e4*	0/15
1: AVG	2.0	**18***	5233	∞	∞	∞*2.9e4*	0/15
f128	111	4248	7808	12447	17217	21162	15/15
0: FUL	15	11	36	∞	∞	∞*4.2e4*	0/15
1: AVG	11	9.3	7.5	∞	∞	∞*3.1e4*	0/15
f129	64	10710	59443	2.85e5	5.11e5	5.80e5	15/15
0: FUL	128	58	∞	∞	∞	∞*4.3e4*	0/15
1: AVG	68	20	∞	∞	∞	∞*3.1e4*	0/15
f130	55	812	3034	32823	33889	34528	10/15
0: FUL	3.0	7.0	3.0	2.4	∞	∞*4.0e4*	0/15
1: AVG	1.3	6.4	5.9	3.0	∞	∞*3.0e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f101	59	361	513	700	739	783	15/15
0: FUL	5.8	1.7	1.3	1	1	1	15/15
1: AVG	**3.3**$^{\star 3}$	**1.2**$^{\star 3}$	**0.95**$^{\star 3}$	0.94	1.0	1.1	15/15
f102	231	399	579	921	1157	1407	15/15
0: FUL	1.6	1.7	1.3	1	1	1	15/15
1: AVG	**0.92**$^{\star 3}$	**1.1**$^{\star 3}$	**0.93**$^{\star 2}$	0.88	1.1	1.3	15/15
f103	65	417	629	1313	1893	2464	14/15
0: FUL	5.2	1.5	1.1	**7.8**$^{\star 2}$	**97**$^{\star 3}$	**554**$^{\star 3}$	1/15
1: AVG	**3.0**$^{\star 3}$	**0.95**$^{\star 3}$	2.0	655	∞	∞*2.0e5*	0/15
f104	23690	85656	1.71e5	1.82e5	1.89e5	1.96e5	15/15
0: FUL	20	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	11	∞	∞	∞	∞	∞*2.0e5*	0/15
f105	1.92e5	6.11e5	6.32e5	6.49e5	6.60e5	6.70e5	15/15
0: FUL	7.2	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	7.2	∞	∞	∞	∞	∞*2.0e5*	0/15
f106	11480	21668	23746	25470	26492	27360	15/15
0: FUL	3.3	12	28	∞	∞	∞*2.0e5*	0/15
1: AVG	8.1	∞	∞	∞	∞	∞*2.0e5*	0/15
f107	8571	13582	16226	27357	52486	65052	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f108	58063	97228	2.03e5	4.46e5	6.30e5	8.98e5	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f109	333	632	1138	2287	3583	4952	15/15
0: FUL	23	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	17	∞	∞	∞	∞	∞*1.9e5*	0/15
f110	∞	∞	∞	∞	∞	∞	0
0: FUL	∞	∞	∞	∞	∞	∞	0/15
1: AVG	∞	∞	∞	∞	∞	∞	0/15
f111	∞	∞	∞	∞	∞	∞	0
0: FUL	∞	∞	∞	∞	∞	∞	0/15
1: AVG	∞	∞	∞	∞	∞	∞	0/15
f112	25552	64124	69621	73557	76137	78238	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
f113	50123	3.64e5	5.60e5	5.88e5	5.88e5	5.91e5	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f114	2.08e5	1.12e6	1.45e6	1.57e6	1.57e6	1.58e6	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f115	2405	30268	91749	1.27e5	1.27e5	1.29e5	15/15
0: FUL	116	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	108	∞	∞	∞	∞	∞*1.9e5*	0/15
f116	4.98e5	6.94e5	8.93e5	1.03e6	1.08e6	1.12e6	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f117	1.79e6	2.46e6	2.60e6	2.91e6	3.24e6	3.62e6	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f118	6908	11786	17514	26342	30062	32659	15/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
f119	2771	29365	35930	4.11e5	1.40e6	1.90e6	15/15
0: FUL	1046	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f120	36040	1.79e5	2.81e5	1.59e6	6.74e6	1.35e7	13/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f121	249	769	1426	9304	34434	57404	15/15
0: FUL	73	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	49	∞	∞	∞	∞	∞*1.8e5*	0/15
f122	692	52008	1.40e5	7.93e5	2.00e6	5.82e6	15/15
0: FUL	182	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	125	∞	∞	∞	∞	∞*1.8e5*	0/15
f123	1063	5.30e5	1.49e6	5.29e6	2.71e7	1.58e8	0
0: FUL	248	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	405	∞	∞	∞	∞	∞*1.8e5*	0/15
f124	192	1959	40840	1.27e5	3.89e5	7.99e5	15/15
0: FUL	117	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	95	∞	∞	∞	∞	∞*1.8e5*	0/15
f125	1	1	1	2.50e7	8.03e7	8.06e7	4/15
0: FUL	1.7	856	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	1	493	∞	∞	∞	∞*1.8e5*	0/15
f126	1	1	1	∞	∞	∞	0
0: FUL	311	∞	∞	∞	∞	∞	0/15
1: AVG	122	2.60e6	∞	∞	∞	∞	0/15
f127	1	1	1	4.43e6	7.27e6	7.43e6	15/15
0: FUL	1	1360	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	7.7	**219**$^{\star 2}$	∞	∞	∞	∞*1.8e5*	0/15
f128	1.40e5	1.34e7	1.72e7	1.72e7	1.72e7	1.72e7	9/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f129	7.81e6	4.13e7	4.15e7	4.18e7	4.21e7	4.24e7	5/15
0: FUL	∞	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	∞	∞	∞	∞	∞	∞*1.8e5*	0/15
f130	4904	93149	2.52e5	2.54e5	2.55e5	2.57e5	7/15
0: FUL	19	∞	∞	∞	∞	∞*2.0e5*	0/15
1: AVG	6.6	∞	∞	∞	∞	∞*1.8e5*	0/15

Table 2: 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_{101}–f_{130}. #succ is the number of trials that reached the final target $f_{\rm opt} + 10^{-8}$. 0: FUL is full-NEWUOA and 1: AVG is avg-NEWUOA. #succ is the number of trials that reached the final target $f_{\rm opt} + 10^{-8}$. 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 60.