



HAL
open science

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

Raymond Ros

► **To cite this version:**

Raymond Ros. Comparison of NEWUOA with Different Numbers of Interpolation Points on the BBOB Noiseless Testbed. Genetic and Evolutionary Computation Conference 2010, Jul 2010, Portland, OR, United States. inria-00473774

HAL Id: inria-00473774

<https://inria.hal.science/inria-00473774>

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 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, $f_t = f_{\text{opt}} +$

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.

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.

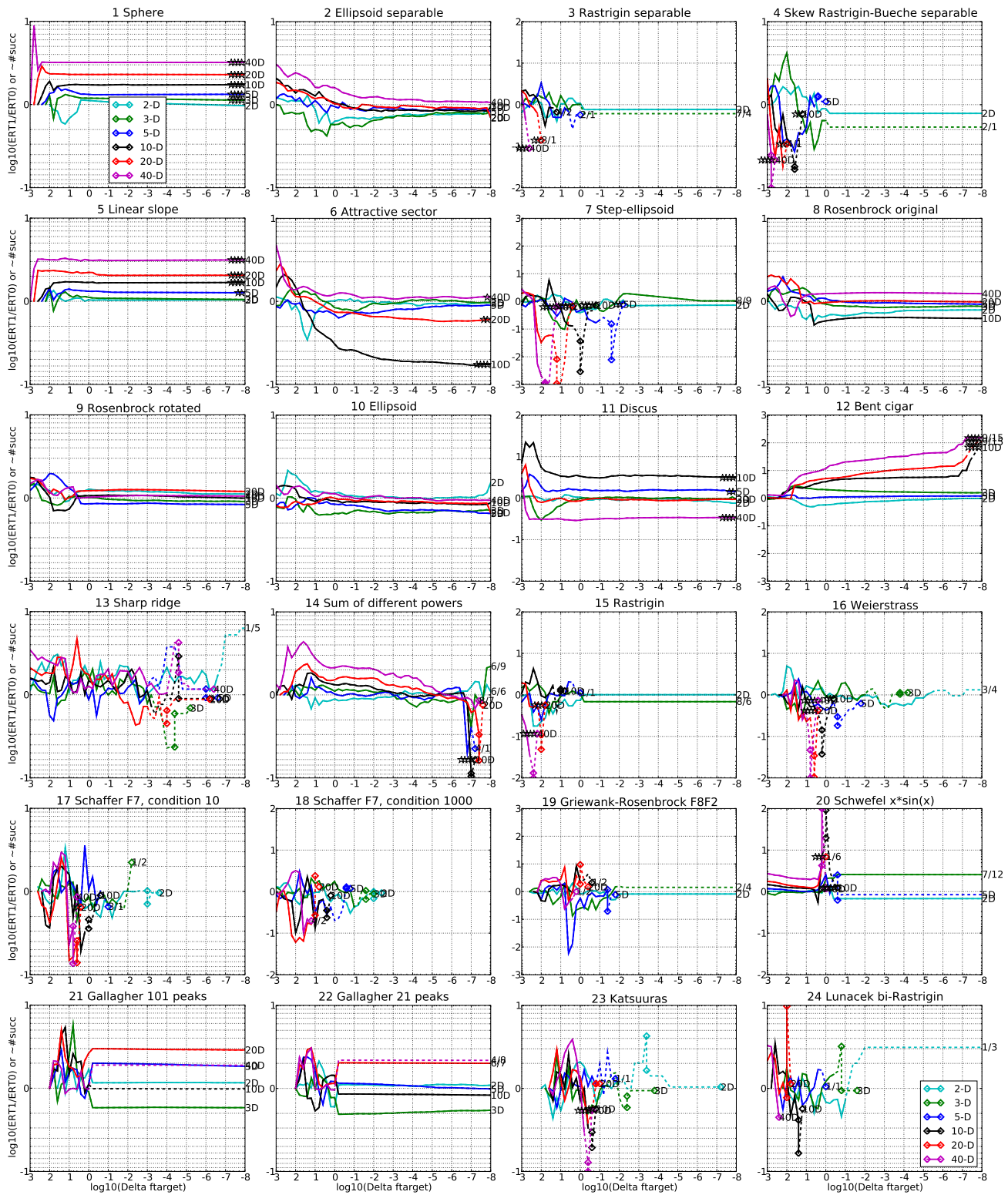


Figure 1: ERT ratio of avg-NEWUOA divided by NEWUOA versus $\log_{10}(\Delta f)$ for f_1 – f_{24} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds thrice the median of the trial-wise 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 $p = 10^{-\#\ast}$ otherwise, with Bonferroni correction within each figure.

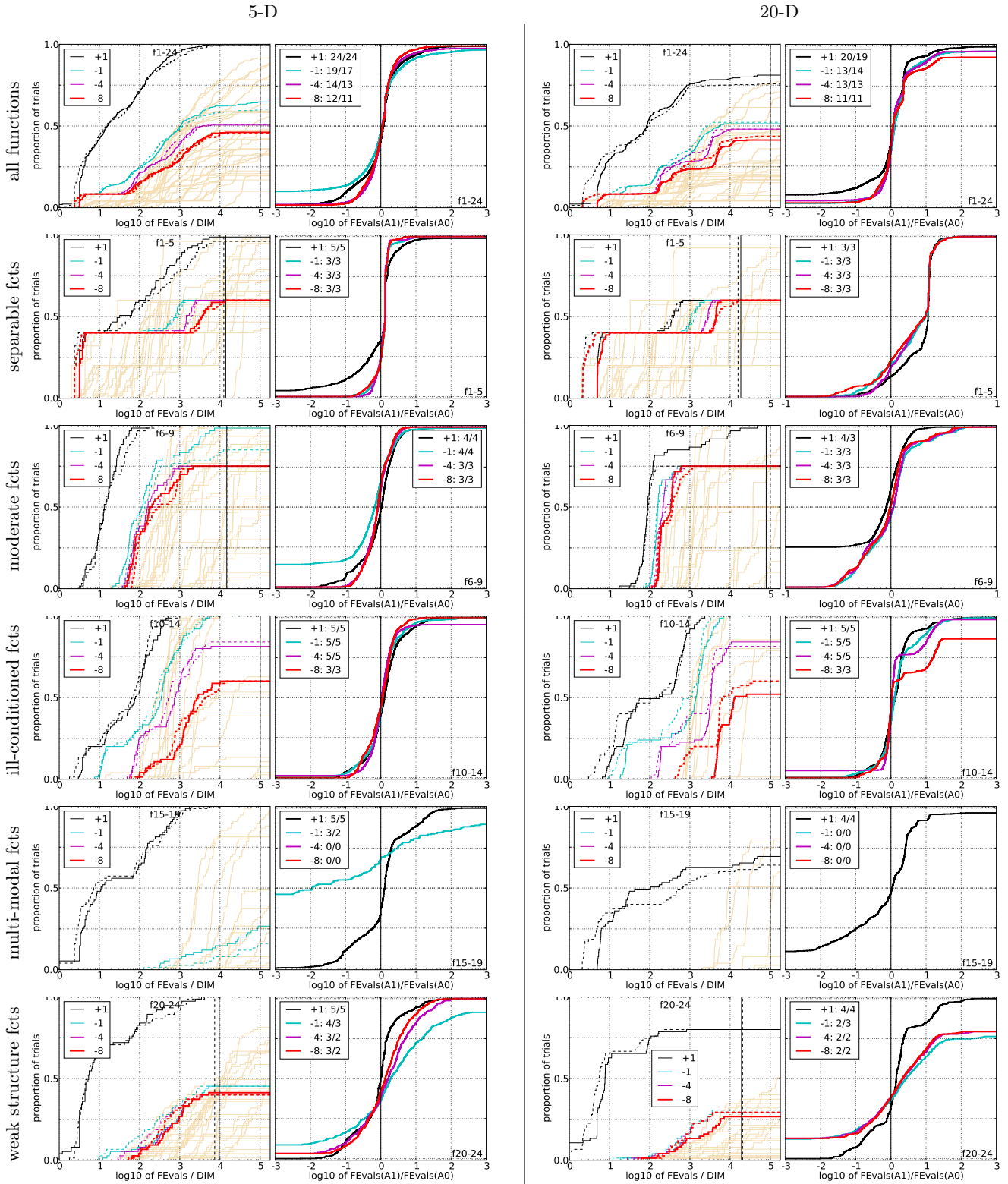


Figure 2: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/ D) to reach a target value $f_{\text{opt}} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ 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).

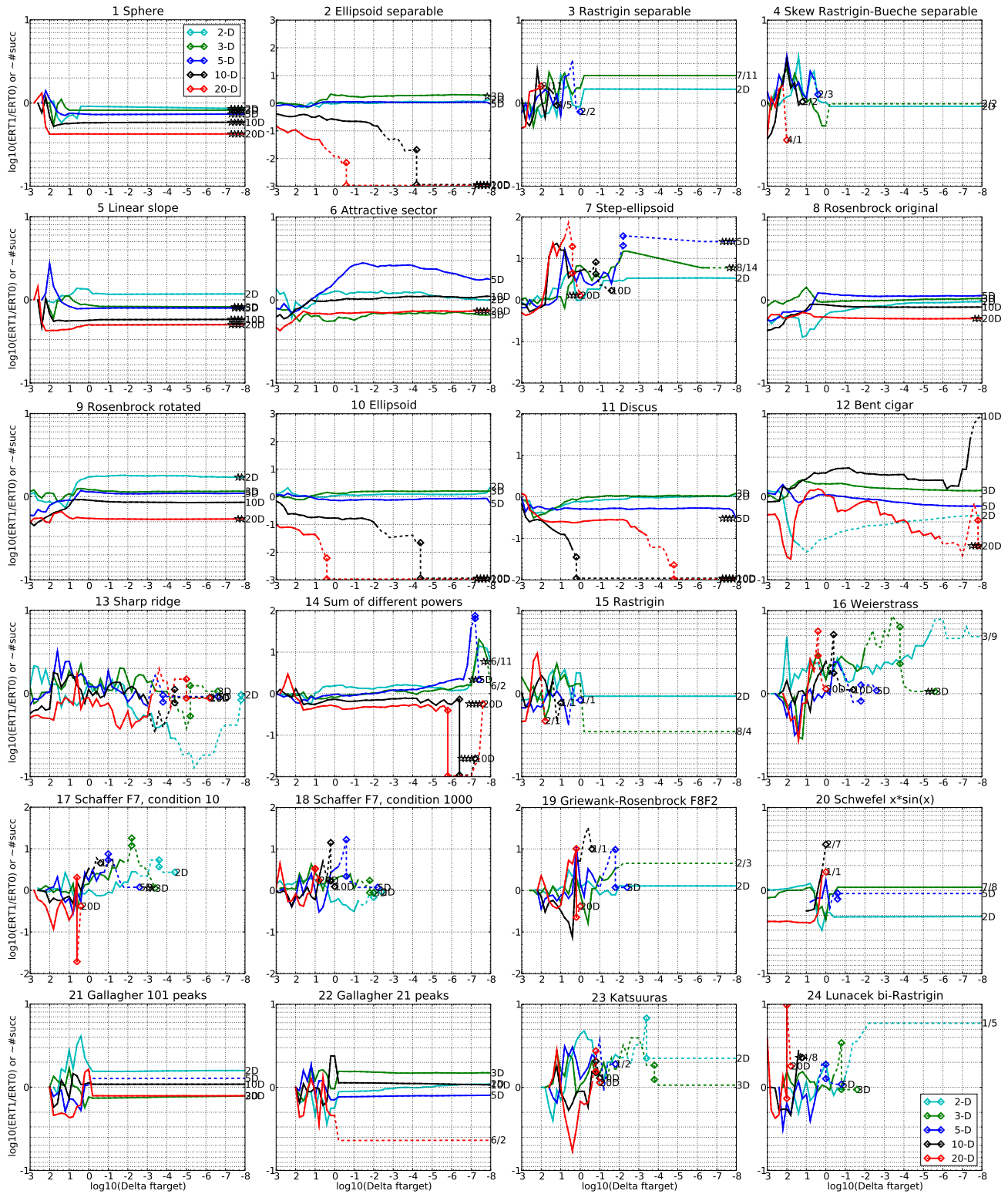


Figure 3: ERT ratio of avg-NEWUOA divided by full-NEWUOA versus $\log_{10}(\Delta f)$ for f_1 – f_{24} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of avg-NEWUOA, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds thrice the median of the 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 $\{1 \dots 9\}$ for avg-NEWUOA (1st number) and non-zero for full-NEWUOA (2nd number). Results are significant with $p = 0.05$ for one star and $p = 10^{-\#*}$ otherwise, with Bonferroni correction within each figure.

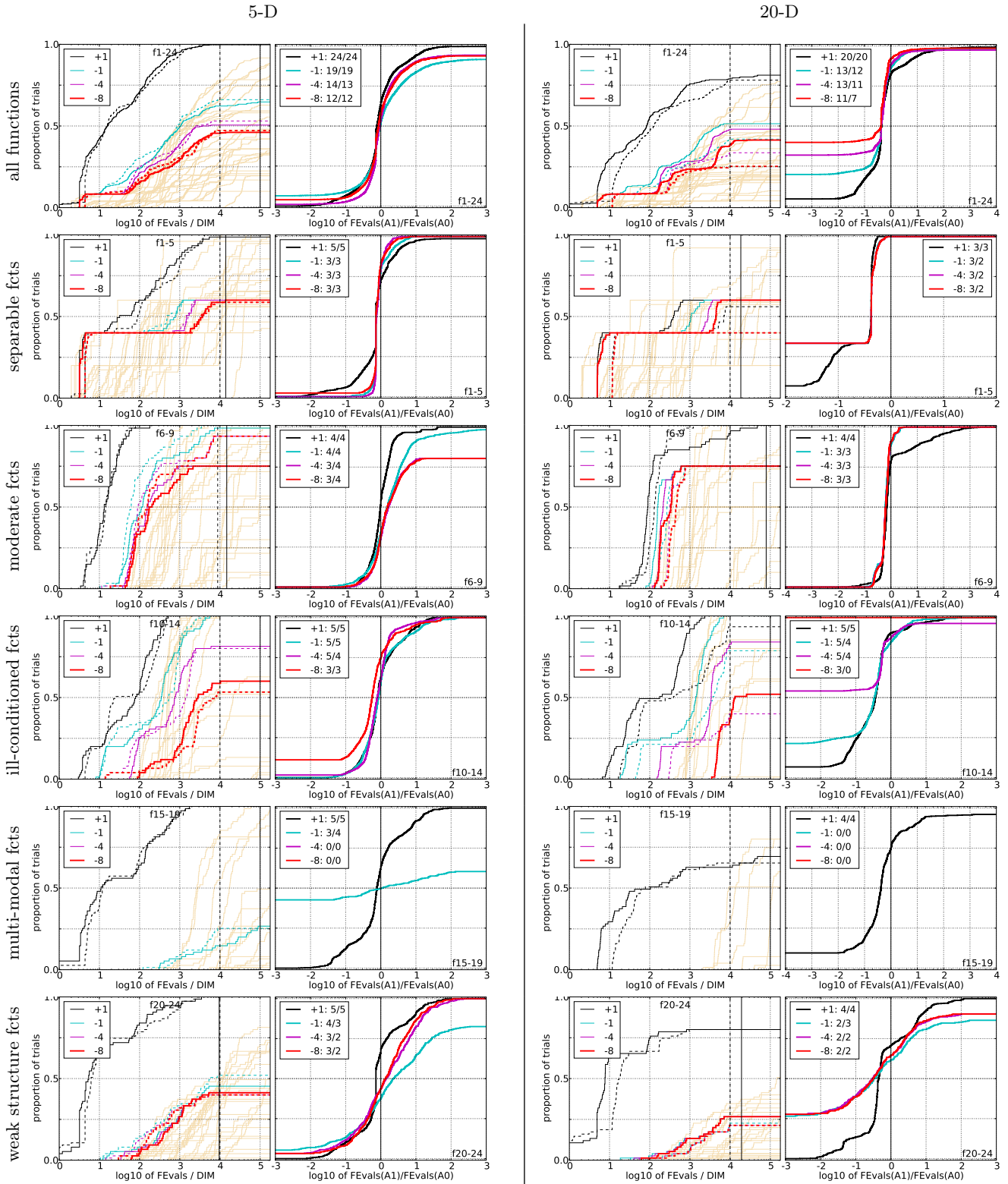


Figure 4: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/ D) to reach a target value $f_{\text{opt}} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ 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).

