



**HAL**  
open science

## Investigating the Impact of Sequential Selection in the (1,4)-CMA-ES on the Noisy BBOB-2010 Testbed

Anne Auger, Dimo Brockhoff, Nikolaus Hansen

► **To cite this version:**

Anne Auger, Dimo Brockhoff, Nikolaus Hansen. Investigating the Impact of Sequential Selection in the (1,4)-CMA-ES on the Noisy BBOB-2010 Testbed. GECCO workshop on Black-Box Optimization Benchmarking (BBOB'2010), Jul 2010, Portland, OR, United States. pp.1611-1616, 10.1145/1830761.1830780 . inria-00502434

**HAL Id: inria-00502434**

**<https://inria.hal.science/inria-00502434v1>**

Submitted on 14 Jul 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.

# Investigating the Impact of Sequential Selection in the (1,4)-CMA-ES on the Noisy BBOB-2010 Testbed

[Black-Box Optimization Benchmarking Workshop]

Anne Auger, Dimo Brockhoff, and Nikolaus Hansen  
Projet TAO, INRIA Saclay—Ile-de-France  
LRI, Bât 490, Univ. Paris-Sud  
91405 Orsay Cedex, France  
firstname.lastname@inria.fr

## ABSTRACT

Sequential selection, introduced for Evolution Strategies (ESs) with the aim of accelerating their convergence, consists in performing the evaluations of the different offspring sequentially, stopping the sequence of evaluations as soon as an offspring is better than its parent and updating the new parent to this offspring solution. This paper investigates the impact of the application of sequential selection to the (1,4)-CMA-ES on the BBOB-2010 noisy benchmark testbed. The performance of the (1,4<sup>s</sup>)-CMA-ES, where sequential selection is implemented, is compared to the baseline algorithm (1,4)-CMA-ES. Independent restarts for the two algorithms are conducted till a maximum of  $10^4 D$  function evaluations per trial was reached, where  $D$  is the dimension of the search space.

The results show that the sequential selection within the (1,4<sup>s</sup>)-CMA-ES clearly outperforms the baseline algorithm (1,4)-CMA-ES by at least 12% on 7 functions in 20D whereas no statistically significant worsening can be observed. Moreover, the (1,4<sup>s</sup>)-CMA-ES shows shorter expected running times on 6 functions of up to 32% compared to the function-wise best algorithm of the BBOB-2009 benchmarking (in 20D and for a target value of  $10^{-7}$ ).

## 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

Evolution Strategies (ESs) are robust stochastic search algorithms for numerical optimization where the objective function to be minimized,  $f$ , maps the continuous search space  $\mathbb{R}^D$  into  $\mathbb{R}$ . In ESs, a population of  $\lambda$  candidate solutions is sampled at each iteration by adding to a current solution  $\lambda$  random vectors following a multivariate normal distribution. In the local search (1,  $\lambda$ )-ES we are interested in, the best of the  $\lambda$  solutions, i.e., the solution having the smallest objective function value, is selected to become the new current solution.

Sequential selection has been recently introduced for Evolution Strategies with the aim of accelerating their convergence [2]. When sequential selection is applied in a (1,  $\lambda$ )-ES, the evaluations are carried out sequentially and the sequence of evaluations is stopped as soon as an offspring turns out to be better than its parent. The parent for the next iteration is then set to this offspring. In this paper, we evaluate the impact of sequential selection on the (1,4)-Covariance-Matrix-Adaptation Evolution-Strategy (CMA-ES) using the BBOB-2010 noisy testbed. The performance of the (1,4<sup>s</sup>)-CMA-ES implementing sequential selection is compared to the performance of the (1,4)-CMA-ES. The algorithms as well as the CPU timing experiments are described in a complementing paper in the same proceedings [1].

## 2. COMPARING THE (1,4) AND THE (1,4<sup>s</sup>)-CMA-ES

Results from experiments comparing (1,4)-CMA-ES and (1,4<sup>s</sup>)-CMA-ES according to [4] on the benchmark functions given in [3, 5] are presented in Figures 1, 2 and 3 and in Table 1. The **expected running time (ERT)**, used in the figures and table, depends on a given target function value,  $f_t = f_{\text{opt}} + \Delta f_t$ , 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  $f_t$  [4, 6]. **Statistical significance** is tested with the rank-sum test for a given target  $\Delta f_t$  ( $10^{-8}$  in Figure 1) using, for each trial, either the number of needed function evaluations to reach  $\Delta f_t$  (inverted and multiplied by  $-1$ ), or, if the target was not reached, the best  $\Delta f$ -value achieved, measured only up to the smallest number of overall function evaluations for any unsuccessful trial under consideration.

First of all, it is to mention that already the simple (1,4)-CMA-ES outperforms the function-wise best algorithm of the BBOB-2009 benchmarking in 20D on the Gallagher function with Cauchy noise ( $f_{130}$ ) by about 40% (although only 11 of the 15 runs are successful) and that it shows the same expected running time than the BBOB-2009 function-wise best algorithm on the sphere function with moderate Cauchy noise ( $f_{103}$ ).

Moreover, the sequential selection in the (1,4<sup>s</sup>)-CMA-ES further improves over the (1,4)-CMA-ES on seven functions statistically significant in 20D and for a target value of  $10^{-7}$ : on  $f_{101-103}$ , the improvement is between 12% and 20%, on  $f_{106}$  and  $f_{118}$ , the improvement is 40% and on  $f_{121}$  and  $f_{112}$ , the running time of the (1,4<sup>s</sup>)-CMA-ES is smaller than the one of the (1,4)-CMA-ES by a factor of about 2 and 3 respectively (all results statistically significant). No statistically significant worsening on any function in 5D and 20D can be observed although the expected running times on  $f_{130}$  are approximately 50% higher for the (1,4<sup>s</sup>)-CMA-ES than for the (1,4)-CMA-ES and also the success probability of the (1,4<sup>s</sup>)-CMA-ES is smaller on this function (8 versus 11 instances solved).

Despite this result on  $f_{130}$ , the (1,4<sup>s</sup>)-CMA-ES shows, in comparison to the function-wise best algorithm of the BBOB-2009 benchmarking, better results on all functions that are solved except for the comparably easy functions  $f_{101}$  and  $f_{102}$  as well as on  $f_{118}$  (in 20D and for a target value of  $10^{-7}$ ): on  $f_{106}$ ,  $f_{121}$ , and  $f_{130}$ , the improvements are rather small ( $\leq 10\%$ ) but on  $f_{103}$ , the (1,4<sup>s</sup>)-CMA-ES is 18% faster, on  $f_{109}$  32% faster, and on  $f_{112}$  26% faster than the function-wise best algorithm of BBOB-2009, which was in all those cases the IPOP-SEP-CMA-ES of [7]—showing that incorporating the sequential selection idea into the separable CMA-ES of [7] might even further improve the results.

### 3. CONCLUSIONS

The idea behind the sequential selection scheme introduced in [2] is to finish the iteration as soon as an offspring is evaluated which is better than the current solution and thereby save some of the  $\lambda$  function evaluations per iteration in a (1  $\dagger$   $\lambda$ )-ES. Here, the concept of sequential selection has been integrated into a comma-strategy, the so-called (1,4<sup>s</sup>)-CMA-ES, and compared with the corresponding baseline (1,4)-CMA-ES on the noisy BBOB-2010 testbed.

The results show that the (1,4)-CMA-ES and its improved version (1,4<sup>s</sup>)-CMA-ES with sequential selection solve 9 of the 30 functions overall. No statistically significantly worse results can be observed for the (1,4<sup>s</sup>)-CMA-ES although the expected running times on  $f_{130}$  are approximately 50% higher for the (1,4<sup>s</sup>)-CMA-ES than for the (1,4)-CMA-ES. Instead, the sequential selection in the (1,4<sup>s</sup>)-CMA-ES improves over the (1,4)-CMA-ES on seven functions statistically significantly in 20D with improvements of at least 12%.

Moreover, the (1,4<sup>s</sup>)-CMA-ES even shows an improved performance over the overall best algorithm from the BBOB-2009 benchmarking on 6 functions (in 20D and for a target value of  $10^{-7}$ ). Interestingly, all those 6 functions belong to the class of functions with additional Cauchy noise. The largest improvements are obtained on  $f_{103}$  (18% faster than the best algorithm of the BBOB-2009 benchmarking on that function), on  $f_{109}$  (32% faster), and on  $f_{112}$  (26% faster).

### 4. REFERENCES

- [1] A. Auger, D. Brockhoff, and N. Hansen. Investigating the impact of sequential selection in the (1,4)-CMA-ES on the noiseless BBOB-2010 testbed. In *GECCO (Companion)*, 2010.
- [2] A. Auger, D. Brockhoff, and N. Hansen. Mirrored sampling and sequential selection for evolution strategies. Rapport de Recherche RR-7249, INRIA Saclay—Île-de-France, April 2010.
- [3] 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.
- [4] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2010: Experimental setup. Technical Report RR-7215, INRIA, 2010.
- [5] 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.
- [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 sep-CMA-ES on the BBOB-2009 noisy testbed. In F. Rothlauf, editor, *GECCO (Companion)*, pages 2441–2446. ACM, 2009.

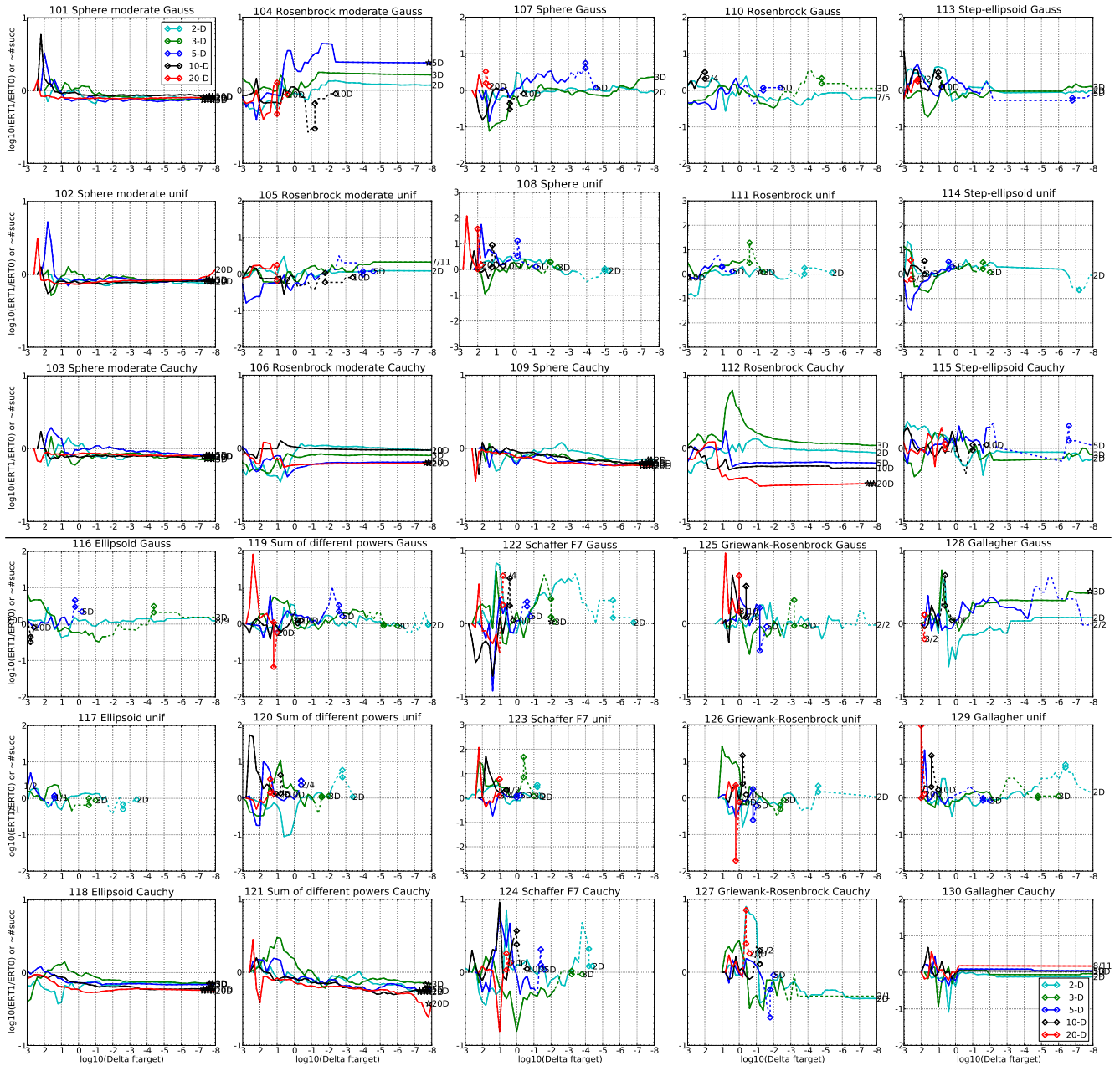


Figure 1: Ratio of the expected running times (ERT) of  $(1,4^s)$ -CMA-ES divided by  $(1,4)$ -CMA-ES versus  $\log_{10}(\Delta f)$  for  $f_{101}-f_{130}$  in 2, 3, 5, 10, 20. Ratios  $< 10^0$  indicate an advantage of  $(1,4^s)$ -CMA-ES, smaller values are always better. The line gets dashed when for 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  $\Delta f$ -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  $(1,4^s)$ -CMA-ES. The line ends when no algorithm reaches  $\Delta f$  anymore. The number of successful trials is given, only if it was in  $\{1 \dots 9\}$  for  $(1,4^s)$ -CMA-ES (1st number) and non-zero for  $(1,4)$ -CMA-ES (2nd number). Results are statistically significant with  $p = 0.05$  for one star and  $p = 10^{-\#\star}$  otherwise, with Bonferroni correction within each figure.

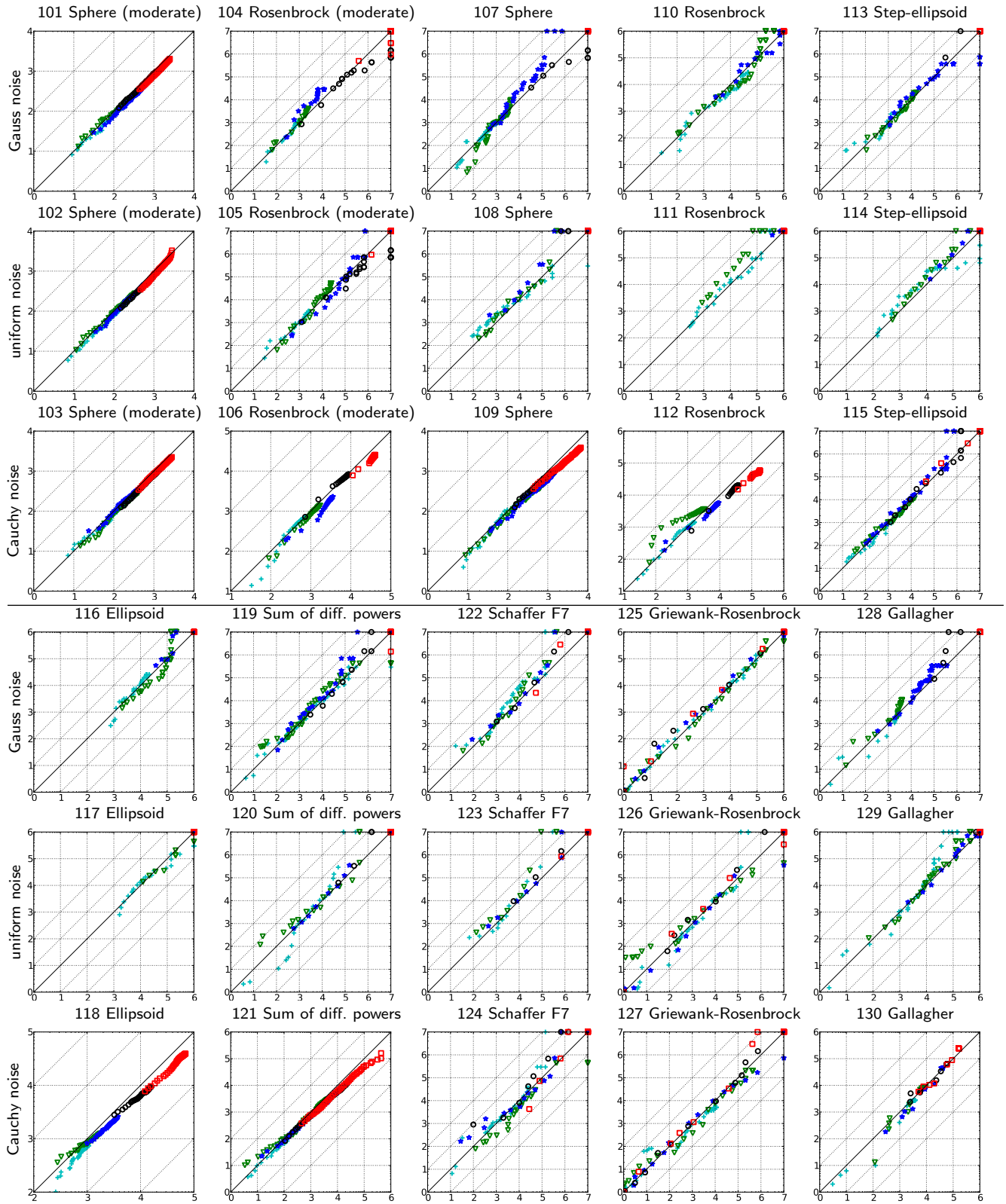
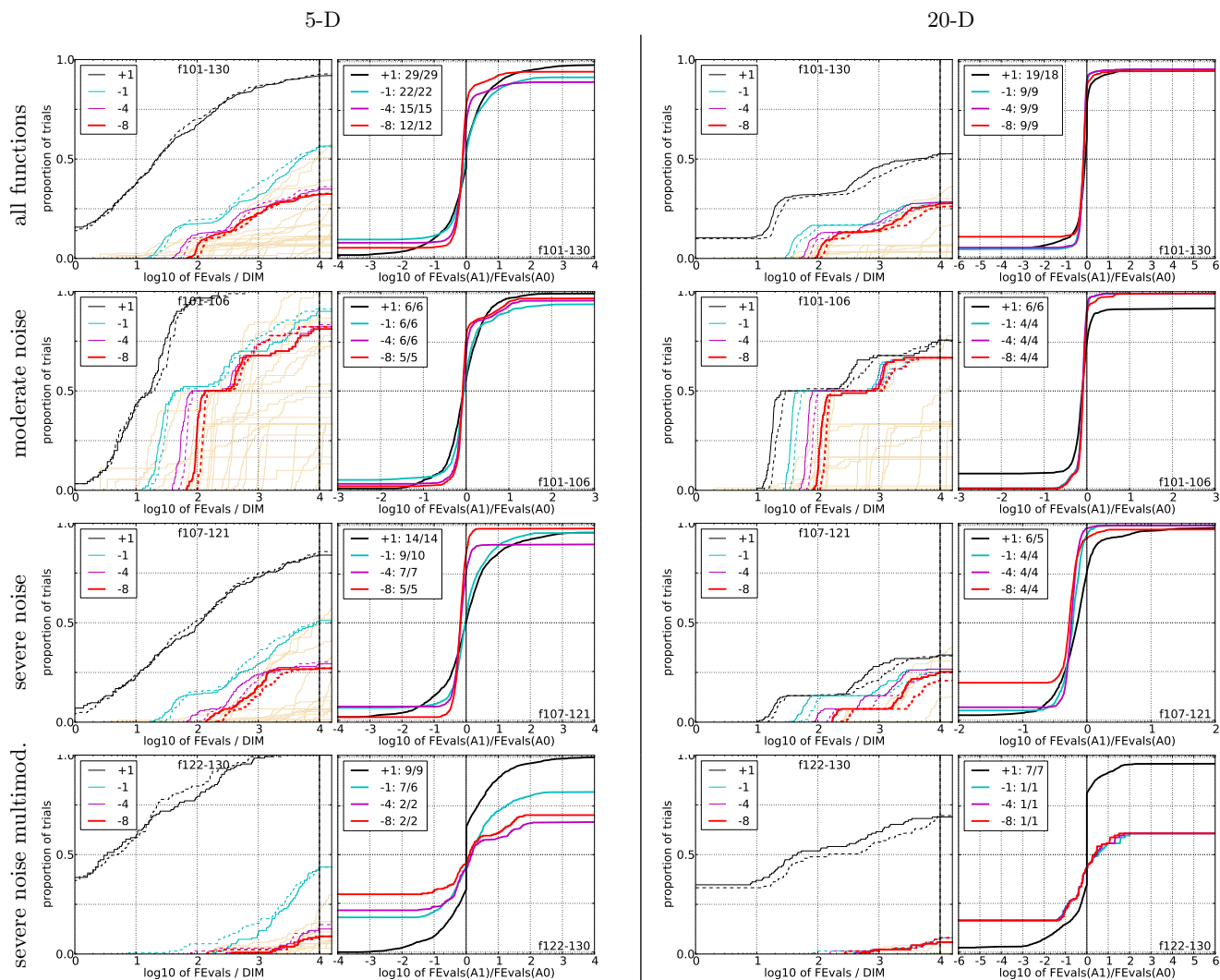


Figure 2: Expected running time (ERT in log10 of number of function evaluations) of  $(1,4^s)$ -CMA-ES versus  $(1,4)$ -CMA-ES for 46 target values  $\Delta f \in [10^{-8}, 10]$  in each dimension for functions  $f_{101}$ - $f_{130}$ . Markers on the upper or right edge indicate that the target value was never reached by  $(1,4^s)$ -CMA-ES or  $(1,4)$ -CMA-ES respectively. Markers represent dimension: 2:+, 3: $\nabla$ , 5:\*, 10:○, 20:□.



**Figure 3: 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  $D$  ( $\text{FEvals}/D$ ) 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  $(1,4^s)$ -CMA-ES (solid) and  $(1,4)$ -CMA-ES (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  $(1,4^s)$ -CMA-ES divided by  $(1,4)$ -CMA-ES, 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 ( $(1,4^s)$ -CMA-ES first).

