Open access • Proceedings Article • DOI:10.1145/1570256.1570333

Benchmarking a BI-population CMA-ES on the BBOB-2009 function testbed

— Source link

Nikolaus Hansen

Institutions: French Institute for Research in Computer Science and Automation

Published on: 08 Jul 2009 - Genetic and Evolutionary Computation Conference

Topics: Population and Testbed

Benchmarking a BI-Population CMA-ES on the BBOB-2009 Function Testbed

Nikolaus Hansen

Benchmarking a BI-Population CMA-ES on the BBOB-2009 Function Testbed

Nikolaus Hansen
Microsoft Research–INRIA Joint Centre
28 rue Jean Rostand
91893 Orsay Cedex, France
Nikolaus.Hansen@inria.fr

ABSTRACT

We propose a multistart CMA-ES with equal budgets for two interlaced restart strategies, one with an increasing population size and one with varying small population sizes. This BI-population CMA-ES is benchmarked on the BBOB-2009 noiseless function testbed and could solve 23, 22 and 20 functions out of 24 in search space dimensions 10, 20 and 40, respectively, within a budget of less than $10^6 D$ function evaluations per trial.

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, Evolutionary computation, CMA-ES

1. INTRODUCTION

The *covariance matrix adaptation evolution strategy* (CMA-ES) is a stochastic, population-based search method in continuous search spaces, aiming at minimizing an objective function $f : \mathbb{R}^D \to \mathbb{R}$ in a black-box scenario. In this paper, the $(\mu/\mu_\mathrm{W}, \lambda)$-CMA-ES [3] is applied in a multistart strategy and benchmarked on 24 functions. Comprehensive results for the number of function evaluations to reach a target function value are given.

2. THE $(\mu/\mu_\mathrm{W}, \lambda)$-CMA-ES

In the standard $(\mu/\mu_\mathrm{W}, \lambda)$-CMA-ES [3, 6, 8], in each iteration step t, λ new solutions $\boldsymbol{x}_i \in \mathbb{R}^D$ are generated by sampling a multi-variate normal distribution, $\mathcal{N}\left(\mathbf{0}, \boldsymbol{C}^t\right)$, with mean $\mathbf{0}$ and $n \times n$ covariance matrix $\boldsymbol{C}^t$, see Eq. (1). The μ best solutions are selected to update the distribution parameters for the next iteration $t+1$. The complete algorithm is depicted in Table 1. We set $\mu = \lfloor \frac{\lambda}{2} \rfloor$, $w_i = \frac{\ln(\mu+1) - \ln i}{\sum_{j=1}^{\mu}(\ln(\mu+1) - \ln j)}$, $\mu_\mathrm{W}^{-1} = \sum_{i=1}^{\mu} w_i^2$, and $d_\sigma = 1 + c_\sigma + 2\max\left(0, \sqrt{\frac{\mu_\mathrm{W}-1}{D+1}} - 1\right)$ usually close to one, $\boldsymbol{C}^{t-\frac{1}{2}}$ is symmetric and satisfies $\boldsymbol{C}^{t-\frac{1}{2}}\boldsymbol{C}^{t-\frac{1}{2}} = \left(\boldsymbol{C}^t\right)^{-1}$. The remaining learning parameters have been slightly modified, without connection to the BBOB-2009 testbed, and are chosen as $c_\sigma = \frac{\mu_\mathrm{W}+2}{D+\mu_\mathrm{W}+5}$, $c_\mathrm{c} = \frac{4+\mu_\mathrm{W}/D}{D+4+2\mu_\mathrm{W}/D}$, $c_1 = \frac{2}{(D+1.3)^2+\mu_\mathrm{W}}$ and $c_\mu = \min\left(1 - c_1,\, 2\frac{\mu_\mathrm{W}-2+1/\mu_\mathrm{W}}{(D+2)^2+\mu_\mathrm{W}}\right)$.

2.1 BI-Population Multistart Scheme

The $(\mu/\mu_\mathrm{W}, \lambda)$-CMA-ES with the default population size $\lambda_\mathrm{def} = 4 + \lfloor 3 \ln D \rfloor$ is a robust and fast *local* search method [9]. With a large(r) population size a more global search can be accomplished successfully [6, 7]. After a first single run with default population size, we apply two interlaced multistart regimes, each equipped with a function evaluation budget accounting for the so far conducted function evaluations. Depending on which budget value is smaller, a complete run of either one or the other strategy is launched. The first and last restart are conducted under the first regime.

Under the first regime, we restart with **increasing population size**, where before each restart the population size λ is increased by a factor of two [1]. At most nine restarts are conducted, *i.e.*, the largest population size is $\lambda = 2^9 \lambda_\mathrm{def} = 512\,\lambda_\mathrm{def}$. The initial $\sigma^0 = 2$ (*i.e.*, 1/5 of the domain width). The budget is loaded from the first restart, *i.e.*, the first single run with population size λ_def is disregarded.

Second, a multistart regime with **small population size** is applied, where the population size λ is set to

$$\lambda_\mathrm{s} = \left\lfloor \lambda_\mathrm{def} \left(\frac{1}{2}\frac{\lambda_\ell}{\lambda_\mathrm{def}}\right)^{\mathcal{U}[0,1]^2} \right\rfloor ,$$

where λ_ℓ is the latest population size from the first regime with increasing (large) λ. Here $\mathcal{U}[0,1]$ denote independent uniformly distributed numbers in $[0,1]$ and $\lambda_\mathrm{s} \in [\lambda_\mathrm{def}, \lambda/2]$. The initial step-size is set to $\sigma^0 = 2 \times 10^{-2\,\mathcal{U}[0,1]}$. A maximum number of function evaluations of half of the recent large budget is enforced, but probably of minor relevance. The second multistart regime is launched, if and only if its recent budget is smaller than the one for the first regime with increasing populations.

GECCO'09, July 8–12, 2009, Montréal Québec, Canada.

Table 1: Update equations for the state variables in the $(\mu/\mu_{\mathbf{w}},\lambda)$-CMA-ES with iteration index $t=0,1,2,\dots$, where $\boldsymbol{p}_\sigma^{t=0}=\boldsymbol{p}_\mathrm{c}^{t=0}=\mathbf{0}$ and $\boldsymbol{C}^{t=0}=\boldsymbol{I}$. Here, $\boldsymbol{x}_{i:\lambda}$ is the i-th best of the solutions $\boldsymbol{x}_1,\dots,\boldsymbol{x}_\lambda$ and $h_\sigma=1$ if $\|\boldsymbol{p}_\sigma^{t+1}\|<\sqrt{1-(1-c_\sigma)^{2(t+1)}}\left(1.4+\frac{2}{D+1}\right)\mathsf{E}\|\mathcal{N}(\mathbf{0},\boldsymbol{I})\|$ and zero otherwise. Further symbols and constants and $\boldsymbol{m}^{t=0}$ and $\sigma^{t=0}$ are given in the text. The chosen ordering of equations allows to remove the time index in all variables but $\boldsymbol{m}^t$

Given $t\in\mathbb{N}$, $\boldsymbol{m}^t\in\mathbb{R}^D$, $\sigma^t\in\mathbb{R}_+$, $\boldsymbol{C}^t\in\mathbb{R}^{D\times D}$ positive definite, $\boldsymbol{p}_\sigma^t\in\mathbb{R}^D$, and $\boldsymbol{p}_\mathrm{c}^t\in\mathbb{R}^D$

$$\boldsymbol{x}_i \sim \boldsymbol{m}^t+\sigma^t\times\mathcal{N}_i\left(\mathbf{0},\boldsymbol{C}^t\right) \quad \text{is normally distributed for } i=1,\dots,\lambda \text{ and evaluated on } f \tag{1}$$

$$\boldsymbol{m}^{t+1} = \sum_{i=1}^{\mu} w_i\,\boldsymbol{x}_{i:\lambda} \quad \text{where } f(\boldsymbol{x}_{1:\lambda})\le\dots\le f(\boldsymbol{x}_{\mu:\lambda})\le f(\boldsymbol{x}_{\mu+1:\lambda})\dots \tag{2}$$

$$\boldsymbol{p}_\sigma^{t+1} = (1-c_\sigma)\,\boldsymbol{p}_\sigma^t+\sqrt{c_\sigma(2-c_\sigma)\mu_\mathrm{w}}\,\boldsymbol{C}^{t\,-\frac{1}{2}}\,\frac{\boldsymbol{m}^{t+1}-\boldsymbol{m}^t}{\sigma^t} \tag{3}$$

$$\boldsymbol{p}_\mathrm{c}^{t+1} = (1-c_\mathrm{c})\,\boldsymbol{p}_\mathrm{c}^t+h_\sigma\sqrt{c_\mathrm{c}(2-c_\mathrm{c})\mu_\mathrm{w}}\,\frac{\boldsymbol{m}^{t+1}-\boldsymbol{m}^t}{\sigma^t} \tag{4}$$

$$\boldsymbol{C}^{t+1} = (1-c_1-c_\mu+(1-h_\sigma)\,c_1c_\mathrm{c}(2-c_\mathrm{c}))\,\boldsymbol{C}^t+c_1\,\boldsymbol{p}_\mathrm{c}^{t+1}\boldsymbol{p}_\mathrm{c}^{t+1\,\mathrm{T}}+c_\mu\sum_{i=1}^{\mu}w_i\,\frac{\boldsymbol{x}_{i:\lambda}-\boldsymbol{m}^t}{\sigma^t}\times\frac{\left(\boldsymbol{x}_{i:\lambda}-\boldsymbol{m}^t\right)^\mathrm{T}}{\sigma^t} \tag{5}$$

$$\sigma^{t+1} = \sigma^t\times\exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma^{t+1}\|}{\mathsf{E}\|\mathcal{N}(\mathbf{0},\boldsymbol{I})\|}-1\right)\right) \tag{6}$$

2.2 Initial and Termination Criteria

The initial mean $\boldsymbol{m}^0$ is sampled uniformly distributed in $[-4,4]^D$. A single run of the $(\mu/\mu_\mathrm{w},\lambda)$-CMA-ES is terminated, when the final target function value is reached or one of the following termination conditions is satisfied.

`MaxIter` $=100+50(D+3)^2/\sqrt{\lambda}$ is the maximal number of iterations in each run of CMA-ES

`TolHistFun` $=10^{-12}$: the range of the best function values during the last $10+\lceil 30D/\lambda\rceil$ iterations is smaller than `TolHistFun`.

`EqualFunVals`: in more than $1/3^{\mathrm{rd}}$ of the last D iterations the objective function value of the best and the k-th best solution are identical, that is $f(\boldsymbol{x}_{1:\lambda})=f(\boldsymbol{x}_{k:\lambda})$, where $k=1+\lceil 0.1+\lambda/4\rceil$.

`TolX` $=10^{-12}$: all components of $\boldsymbol{p}_\mathrm{c}^t$ and all square roots of diagonal components of $\boldsymbol{C}^t$, multiplied by σ^t/σ^0, are smaller than `TolX`.

`TolUpSigma` $=10^{20}$: $\sigma^t/\sigma^0>$ `TolUpSigma` $\sqrt{l^t}$, where l^t is the largest eigenvalue of $\boldsymbol{C}^t$, indicates a mismatch between σ increase and decrease of all eigenvalues in $\boldsymbol{C}$. In this, rather untypical, case the progression of the strategy is usually very low and a restart is indicated.

`Stagnation`: the median of the 20 newest values is not smaller than the median of the 20 oldest values, respectively, in the two arrays containg the best function values and the median function values of the last $\lceil 0.2\,t+120+30D/\lambda\rceil$ iterations.

`ConditionCov`: the condition number of $\boldsymbol{C}^t$ exceeds 10^{14}.

`NoEffectAxis`: $\boldsymbol{m}^t$ remains numerically constant when adding $0.1\sigma^t\sqrt{l^t}\,\boldsymbol{v}^t$, where l^t is the $1+(t \mod D)$-largest eigenvalue of $\boldsymbol{C}^t$ and $\boldsymbol{v}^t$ is the corresponding normalized eigenvector.

`NoEffectCoor`: any element of $\boldsymbol{m}^t$ remains numerically constant when adding $0.2\sigma^t\boldsymbol{l}^t$, where elements of $\boldsymbol{l}^t$ are the square root of the diagonal elements of $\boldsymbol{C}^t$. Condition `NoEffectCoor` was never satisfied.

Most criteria are standard part of our production codes of $(\mu/\mu_\mathrm{w},\lambda)$-CMA-ES (see also next section). Restarts are launched until the final target function value or the largest, final population size is reached (see above). In neither case more than 10^6D function evaluations were conducted.

3. PARAMETER TUNING

No thorough parameter study has been done. We have experimented with restarts from a so-far best found solution point but had comparatively little success. The parameters for the first multistart scheme are taken from [1], those for the second are ad-hoc settings. We reckon that even smaller population sizes λ_s could be useful. The maximum number of iterations `MaxIter` has been set to prevent excessive long runs and is chosen such that most functions should be solvable within this limit. Most other termination criteria are standard, while `TolUpSigma` and `Stagnation` have been only recently added to the set of standard termination criteria. The former indicates a problem in acquiring the functions topography and seems only effective up to $D=10$. The latter is of major relevance for noisy functions. The same D-dependent parameter setting is used on all functions and therefore the crafting effort [4] computes to $\mathrm{CrE}=0$.

4. CPU TIMING EXPERIMENT

For the timing experiment the complete algorithm was run on f_8 and restarted until at least 30 seconds had passed (according to Figure 2 in [4]). These experiments have been conducted with an Intel dual core T5600 processor with 1.8 GHz under Linux 2.6.27-11 using Matlab R2008a. The results are shown in the following table.

D	2	3	5	10	20	40	80
seconds $\times 10^{-4}$	2.8	2.4	2.0	1.8	1.8	2.0	6.0

Up to 10-D, the necessary CPU time even reduces with increasing dimension, presumbably due to a larger number of initialization procedures for the restarts until 30 seconds have passed.

Equations (1) and (3) require a decomposition of $\boldsymbol{C}^t$. An eigendecomposition with time complexity $\propto D^3$ is applied and for computational efficiency reasons only conducted until after

$$\frac{(c_1 + c_\mu)^{-1}}{10D} \tag{7}$$

iterations have passed. Therefore, a slightly outdated decomposition is used in case. This policy results in a quadratic scaling of the internal time complexity with the dimension. For larger dimension, a computational burden between 10^{-8} and $10^{-7} \times D^2$ seconds per function evaluation is the typical outcome of timing experiments (for $D = 80$ the table reveals $9 \times 10^{-8} D^2$ seconds).

5. RESULTS AND DISCUSSION

Results from experiments according to [4] on the benchmark functions given in [2, 5] are presented in Figures 1 and 2 and in Table 2.

The number of solved functions amounts to 24, 24, 24, 23, 22, 20 out of 24 for dimension 2, 3, 5, 10, 20, 40. Two functions, f_3 and f_4, seem to become practically unsolvable with increasing dimension. The scaling of the running time (expected number of function evaluations, ERT) with the problem dimension is linear for f_1, f_5 and f_{12} and clearly sub-quadratic for most unimodal functions. For the multi-modal functions the scaling is typically quadratic, in some cases worse, but never better. Running times to reach the final target function value in 20-D range between D and somewhat above $3 \times 10^5 D$. They are typically above $300D$ and below $30\,000D$.

The failure on f_3 for larger dimensions is unexpected and caused by the introduced deformation of the Rastrigin function (see [2, 5]). We suspect that a local minimum with a larger attraction basin has been generated, while this seems not to be the case for f_{15}.

Functions f_4 and f_{24} had been designed to be deceptive for evolution strategies. Nevertheless, f_{24} can be solved, but only with a very large budget of $3 \times 10^5 D^2$ function evaluations, also due to a small success probability.

6. SUCCESSFUL POPULATION SIZE

We investigate the population sizes of the final successful runs whenever at least one restart was executed. In Table 3 minimal, median (the larger in case of even data) and maximal population size are given. For the functions not listed, no restarts were necessary in 20-D (with one exception with a single restart in one trial on f_9). On all multi-modal functions $f_{15\text{–}24}$ restarts are applied. Functions 20 and 24 require a population size above 1000. Functions 19, 21 and 23 are solved with the largest range of different population sizes.

Table 4 tabulates minimal, median (the larger in case of even data) and maximal initial step-size σ^0 of the final successful runs, whenever $\sigma^0 < 2$ in at least one case. Only for functions 23 and 24, the smaller initial step-size appears to be beneficial, while for f_{22} the data are not conclusive. The multi-modal functions f_{17}, f_{18} and f_{20} were never solved with an initially small step-size.

Table 3: Final population sizes in 20-D, where $\lambda_{\text{def}} = 12$, when at least one restart was executed

f	min	med	max
7	96	96	96
13	24	48	96
15	200	384	768
16	56	115	384
17	48	96	192
18	96	192	192
19	236	6144	6144
20	3072	6144	6144
21	12	101	1678
22	14	45	202
23	114	381	1441
24	2137	4456	4675

Table 4: Initial step-size σ^0 of successful restarts in 20-D for functions, where $\sigma^0 < 2$ was successful at least once

f	min	med	max
15	0.04	2	2
16	0.066	2	2
17	1.06	2	2
19	0.054	2	2
21	0.044	1.66	2
22	0.024	0.4	0.6
23	0.02	0.032	0.1
24	0.036	0.068	0.166

7. CONCLUSION

The BI-population CMA-ES performs satisfactorily on many functions of the BBOB-2009 testbed and exhibits a reasonable scaling behavior: between linear and quadratic on unimodal functions, between quadratic and cubic on multi-modal functions. Yet, it can be considerably outperformed at least (a) on functions that are smooth, "regular" and only moderately ill-conditioned (f_1, f_5, f_8, f_9), (b) on separable functions (in particular f_3 and f_4) and (c) on the multi-modal functions f_{21} and f_{22}. The former two cases are intrinsic and connected to invariance properties of the algorithm, namely (a) invariance to order-preserving transformations of the function value and (b) rotational invariance. Case (c) might be successfully addressed by an improved restart schedule.

Acknowledgments

The author would like to acknowledge the great and hard work of the BBOB team with particular kudos to Raymond Ros, Steffen Finck and Anne Auger, and Anne Auger and Marc Schoenauer for their kind and persistent support.

8. REFERENCES

[1] A. Auger and N. Hansen. A restart CMA evolution strategy with increasing population size. In *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2005)*, pages 1769–1776. IEEE Press, 2005.

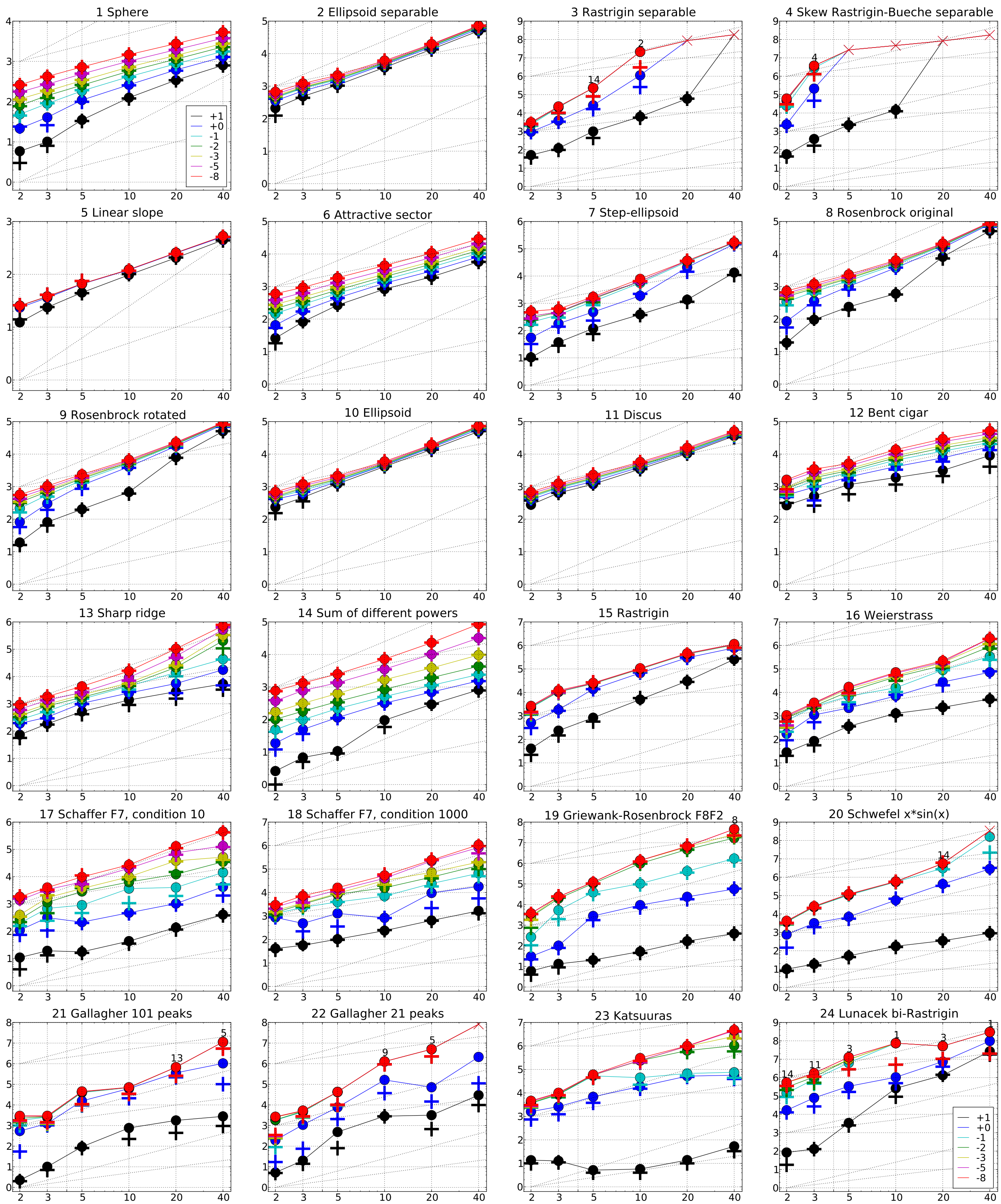

Figure 1: Expected Running Time (ERT, ●) to reach $f_{\rm opt} + \Delta f$ and median number of function evaluations of successful trials (+), shown for $\Delta f = 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-5}, 10^{-8}$ (the exponent is given in the legend of f_1 and f_{24}) versus dimension in log-log presentation. The $\mathrm{ERT}(\Delta f)$ equals to $\#\mathrm{FEs}(\Delta f)$ divided by the number of successful trials, where a trial is successful if $f_{\rm opt} + \Delta f$ was surpassed during the trial. The $\#\mathrm{FEs}(\Delta f)$ are the total number of function evaluations while $f_{\rm opt} + \Delta f$ was not surpassed during the trial from all respective trials (successful and unsuccessful), and $f_{\rm opt}$ denotes the optimal function value. Crosses ($\times$) indicate the total number of function evaluations $\#\mathrm{FEs}(-\infty)$. Numbers above ERT-symbols indicate the number of successful trials. Annotated numbers on the ordinate are decimal logarithms. Additional grid lines show linear and quadratic scaling.

Δf	**f_1 in 5-D**, N=15, mFE=826					**f_1 in 20-D**, N=15, mFE=3062				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	3.5e1	2.8e1	4.2e1	3.5e1	15	3.4e2	3.2e2	3.6e2	3.4e2
1	15	1.1e2	9.6e1	1.2e2	1.1e2	15	6.1e2	5.8e2	6.4e2	6.1e2
1e-1	15	1.8e2	1.7e2	1.9e2	1.8e2	15	8.7e2	8.4e2	8.9e2	8.7e2
1e-3	15	3.3e2	3.2e2	3.5e2	3.3e2	15	1.4e3	1.4e3	1.4e3	1.4e3
1e-5	15	4.9e2	4.8e2	5.0e2	4.9e2	15	1.9e3	1.9e3	2.0e3	1.9e3
1e-8	15	7.3e2	7.1e2	7.4e2	7.3e2	15	2.8e3	2.7e3	2.8e3	2.8e3

Δf	**f_3 in 5-D**, N=15, mFE=1.69e6					**f_3 in 20-D**, N=15, mFE=6.65e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	9.9e2	7.0e2	1.3e3	9.9e2	15	6.0e4	4.9e4	7.0e4	6.0e4
1	15	2.6e4	1.9e4	3.4e4	2.6e4	0	*40e-1*	*20e-1*	*60e-1*	1.0e6
1e-1	14	2.3e5	9.4e4	3.7e5	2.2e5	.	.	.	.	.
1e-3	14	2.3e5	9.8e4	3.7e5	2.2e5	.	.	.	.	.
1e-5	14	2.3e5	9.7e4	3.8e5	2.3e5	.	.	.	.	.
1e-8	14	2.3e5	9.9e4	3.8e5	2.3e5	.	.	.	.	.

Δf	**f_5 in 5-D**, N=15, mFE=90					**f_5 in 20-D**, N=15, mFE=326				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	4.5e1	4.0e1	5.0e1	4.5e1	15	2.1e2	2.0e2	2.1e2	2.1e2
1	15	6.5e1	5.9e1	7.2e1	6.5e1	15	2.5e2	2.4e2	2.6e2	2.5e2
1e-1	15	6.6e1	6.0e1	7.2e1	6.6e1	15	2.6e2	2.4e2	2.7e2	2.6e2
1e-3	15	6.6e1	6.0e1	7.2e1	6.6e1	15	2.6e2	2.4e2	2.7e2	2.6e2
1e-5	15	6.6e1	6.0e1	7.3e1	6.6e1	15	2.6e2	2.4e2	2.7e2	2.6e2
1e-8	15	6.6e1	6.0e1	7.2e1	6.6e1	15	2.6e2	2.4e2	2.7e2	2.6e2

Δf	**f_7 in 5-D**, N=15, mFE=3498					**f_7 in 20-D**, N=15, mFE=41366				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.2e2	8.5e1	1.6e2	1.2e2	15	1.4e3	1.1e3	1.7e3	1.4e3
1	15	4.8e2	3.5e2	6.1e2	4.8e2	15	2.1e4	1.8e4	2.4e4	2.1e4
1e-1	15	1.2e3	8.7e2	1.5e3	1.2e3	15	3.3e4	3.1e4	3.5e4	3.3e4
1e-3	15	1.6e3	1.3e3	1.9e3	1.6e3	15	3.6e4	3.5e4	3.7e4	3.6e4
1e-5	15	1.6e3	1.3e3	1.9e3	1.6e3	15	3.6e4	3.5e4	3.7e4	3.6e4
1e-8	15	1.7e3	1.4e3	2.1e3	1.7e3	15	3.6e4	3.5e4	3.8e4	3.6e4

Δf	**f_9 in 5-D**, N=15, mFE=5564					**f_9 in 20-D**, N=15, mFE=49372				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	2.0e2	1.9e2	2.2e2	2.0e2	15	8.0e3	7.4e3	8.7e3	8.0e3
1	15	1.1e3	8.5e2	1.4e3	1.1e3	15	1.8e4	1.5e4	2.0e4	1.8e4
1e-1	15	1.6e3	1.3e3	1.9e3	1.6e3	15	2.0e4	1.8e4	2.2e4	2.0e4
1e-3	15	1.9e3	1.6e3	2.2e3	1.9e3	15	2.1e4	1.9e4	2.4e4	2.1e4
1e-5	15	2.1e3	1.8e3	2.4e3	2.1e3	15	2.2e4	2.0e4	2.4e4	2.2e4
1e-8	15	2.4e3	2.1e3	2.7e3	2.4e3	15	2.3e4	2.1e4	2.6e4	2.3e4

Δf	**f_{11} in 5-D**, N=15, mFE=2554					**f_{11} in 20-D**, N=15, mFE=16106				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.2e3	1.1e3	1.3e3	1.2e3	15	1.0e4	1.0e4	1.1e4	1.0e4
1	15	1.4e3	1.4e3	1.5e3	1.4e3	15	1.1e4	1.1e4	1.2e4	1.1e4
1e-1	15	1.7e3	1.6e3	1.7e3	1.7e3	15	1.2e4	1.2e4	1.2e4	1.2e4
1e-3	15	1.9e3	1.8e3	1.9e3	1.9e3	15	1.4e4	1.3e4	1.4e4	1.4e4
1e-5	15	2.1e3	2.0e3	2.1e3	2.1e3	15	1.5e4	1.4e4	1.5e4	1.5e4
1e-8	15	2.3e3	2.3e3	2.4e3	2.3e3	15	1.6e4	1.6e4	1.6e4	1.6e4

Δf	**f_{13} in 5-D**, N=15, mFE=7644					**f_{13} in 20-D**, N=15, mFE=165404				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	5.2e2	4.2e2	6.2e2	5.2e2	15	2.8e3	2.0e3	3.7e3	2.8e3
1	15	1.0e3	8.8e2	1.2e3	1.0e3	15	5.5e3	3.8e3	7.4e3	5.5e3
1e-1	15	1.5e3	1.3e3	1.6e3	1.5e3	15	1.4e4	1.0e4	1.8e4	1.4e4
1e-3	15	2.1e3	2.0e3	2.2e3	2.1e3	15	2.8e4	2.4e4	3.1e4	2.8e4
1e-5	15	2.6e3	2.5e3	2.7e3	2.6e3	15	5.6e4	4.4e4	6.7e4	5.6e4
1e-8	15	4.4e3	3.9e3	4.9e3	4.4e3	15	1.0e5	8.9e4	1.2e5	1.0e5

Δf	**f_{15} in 5-D**, N=15, mFE=39838					**f_{15} in 20-D**, N=15, mFE=673663				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	8.4e2	6.1e2	1.1e3	8.4e2	15	3.0e4	2.8e4	3.3e4	3.0e4
1	15	1.4e4	1.1e4	1.7e4	1.4e4	15	3.0e5	2.7e5	3.3e5	3.0e5
1e-1	15	2.3e4	2.0e4	2.7e4	2.3e4	15	4.3e5	3.9e5	4.7e5	4.3e5
1e-3	15	2.4e4	2.1e4	2.7e4	2.4e4	15	4.4e5	4.0e5	4.8e5	4.4e5
1e-5	15	2.5e4	2.1e4	2.8e4	2.5e4	15	4.5e5	4.1e5	4.9e5	4.5e5
1e-8	15	2.5e4	2.2e4	2.9e4	2.5e4	15	4.6e5	4.2e5	5.1e5	4.6e5

Δf	**f_{17} in 5-D**, N=15, mFE=22111					**f_{17} in 20-D**, N=15, mFE=233645				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.7e1	1.3e1	2.2e1	1.7e1	15	1.4e2	1.1e2	1.6e2	1.4e2
1	15	2.1e2	1.9e2	2.4e2	2.1e2	15	1.0e3	9.5e2	1.1e3	1.0e3
1e-1	15	9.0e2	4.8e2	1.4e3	9.0e2	15	4.0e3	2.2e3	5.8e3	4.0e3
1e-3	15	3.7e3	2.9e3	4.4e3	3.7e3	15	3.8e4	3.0e4	4.6e4	3.8e4
1e-5	15	6.4e3	5.4e3	7.3e3	6.4e3	15	7.5e4	6.3e4	8.7e4	7.5e4
1e-8	15	9.8e3	8.3e3	1.1e4	9.8e3	15	1.3e5	1.1e5	1.5e5	1.3e5

Δf	**f_{19} in 5-D**, N=15, mFE=295789					**f_{19} in 20-D**, N=15, mFE=8.33e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	2.0e1	1.7e1	2.4e1	2.0e1	15	1.7e2	1.5e2	1.8e2	1.7e2
1	15	2.8e3	1.8e3	3.9e3	2.8e3	15	2.4e4	1.9e4	2.9e4	2.4e4
1e-1	15	3.9e4	2.9e4	4.9e4	3.9e4	15	4.2e5	3.3e5	5.1e5	4.2e5
1e-3	15	1.2e5	9.6e4	1.4e5	1.2e5	15	6.2e6	5.7e6	6.7e6	6.2e6
1e-5	15	1.2e5	9.7e4	1.5e5	1.2e5	15	6.7e6	6.2e6	7.1e6	6.7e6
1e-8	15	1.2e5	9.9e4	1.5e5	1.2e5	15	6.8e6	6.3e6	7.2e6	6.8e6

Δf	**f_{21} in 5-D**, N=15, mFE=439982					**f_{21} in 20-D**, N=15, mFE=2.72e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	9.3e1	7.4e1	1.1e2	9.3e1	15	1.8e3	5.8e2	3.1e3	1.8e3
1	15	1.6e4	7.5e3	2.4e4	1.6e4	14	3.6e5	1.7e5	5.7e5	3.6e5
1e-1	15	4.0e4	9.6e3	7.2e4	4.0e4	13	6.7e5	3.8e5	9.8e5	6.6e5
1e-3	15	4.3e4	9.9e3	7.8e4	4.3e4	13	6.7e5	3.7e5	9.9e5	6.6e5
1e-5	15	4.4e4	1.0e4	8.0e4	4.4e4	13	6.8e5	3.7e5	1.0e6	6.7e5
1e-8	15	4.5e4	1.1e4	8.2e4	4.5e4	13	6.8e5	3.8e5	1.0e6	6.7e5

Δf	**f_{23} in 5-D**, N=15, mFE=151462					**f_{23} in 20-D**, N=15, mFE=2.13e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	5.1e0	3.7e0	6.5e0	5.1e0	15	1.4e1	1.0e1	1.8e1	1.4e1
1	15	6.8e3	4.6e3	9.1e3	6.8e3	15	5.2e4	3.9e4	6.5e4	5.2e4
1e-1	15	5.3e4	4.0e4	6.7e4	5.3e4	15	6.7e4	5.2e4	8.2e4	6.7e4
1e-3	15	5.8e4	4.4e4	7.1e4	5.8e4	15	9.6e5	8.1e5	1.1e6	9.6e5
1e-5	15	5.9e4	4.5e4	7.3e4	5.9e4	15	9.8e5	8.3e5	1.1e6	9.8e5
1e-8	15	6.1e4	4.6e4	7.5e4	6.1e4	15	1.0e6	8.6e5	1.2e6	1.0e6

Δf	**f_2 in 5-D**, N=15, mFE=2434					**f_2 in 20-D**, N=15, mFE=20690				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.1e3	1.0e3	1.2e3	1.1e3	15	1.4e4	1.3e4	1.4e4	1.4e4
1	15	1.4e3	1.3e3	1.5e3	1.4e3	15	1.6e4	1.5e4	1.6e4	1.6e4
1e-1	15	1.6e3	1.5e3	1.6e3	1.6e3	15	1.7e4	1.7e4	1.7e4	1.7e4
1e-3	15	1.8e3	1.7e3	1.8e3	1.8e3	15	1.8e4	1.8e4	1.8e4	1.8e4
1e-5	15	1.9e3	1.9e3	2.0e3	1.9e3	15	1.9e4	1.9e4	1.9e4	1.9e4
1e-8	15	2.2e3	2.1e3	2.2e3	2.2e3	15	2.0e4	2.0e4	2.0e4	2.0e4

Δf	**f_4 in 5-D**, N=15, mFE=2.05e6					**f_4 in 20-D**, N=15, mFE=6.49e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	2.2e3	1.6e3	2.8e3	2.2e3	0	*12e+0*	*12e+0*	*13e+0*	5.6e5
1	0	*20e-1*	*20e-1*	*27e-1*	5.0e5	.	.	.	.	.
1e-1	.	.	.	.	.	.	.	.	.	.
1e-3	.	.	.	.	.	.	.	.	.	.
1e-5	.	.	.	.	.	.	.	.	.	.
1e-8	.	.	.	.	.	.	.	.	.	.

Δf	**f_6 in 5-D**, N=15, mFE=2194					**f_6 in 20-D**, N=15, mFE=12494				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	2.7e2	2.3e2	3.0e2	2.7e2	15	2.0e3	1.9e3	2.1e3	2.0e3
1	15	4.4e2	4.1e2	4.8e2	4.4e2	15	3.0e3	2.8e3	3.1e3	3.0e3
1e-1	15	6.2e2	5.8e2	6.6e2	6.2e2	15	4.0e3	3.8e3	4.1e3	4.0e3
1e-3	15	9.6e2	9.3e2	9.9e2	9.6e2	15	5.9e3	5.6e3	6.1e3	5.9e3
1e-5	15	1.3e3	1.3e3	1.4e3	1.3e3	15	7.8e3	7.6e3	8.0e3	7.8e3
1e-8	15	1.9e3	1.8e3	1.9e3	1.9e3	15	1.1e4	1.0e4	1.1e4	1.1e4

Δf	**f_8 in 5-D**, N=15, mFE=5092					**f_8 in 20-D**, N=15, mFE=25310				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	2.4e2	1.9e2	2.9e2	2.4e2	15	8.1e3	7.6e3	8.8e3	8.1e3
1	15	1.0e3	7.8e2	1.3e3	1.0e3	15	1.6e4	1.5e4	1.6e4	1.6e4
1e-1	15	1.5e3	1.3e3	1.8e3	1.5e3	15	1.8e4	1.6e4	1.9e4	1.8e4
1e-3	15	1.9e3	1.6e3	2.2e3	1.9e3	15	1.9e4	1.8e4	2.0e4	1.9e4
1e-5	15	2.1e3	1.8e3	2.4e3	2.1e3	15	2.0e4	1.9e4	2.1e4	2.0e4
1e-8	15	2.3e3	2.1e3	2.6e3	2.3e3	15	2.1e4	2.0e4	2.2e4	2.1e4

Δf	**f_{10} in 5-D**, N=15, mFE=2338					**f_{10} in 20-D**, N=15, mFE=20522				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.2e3	1.1e3	1.3e3	1.2e3	15	1.4e4	1.4e4	1.4e4	1.4e4
1	15	1.4e3	1.4e3	1.5e3	1.4e3	15	1.6e4	1.6e4	1.6e4	1.6e4
1e-1	15	1.5e3	1.5e3	1.6e3	1.5e3	15	1.7e4	1.6e4	1.7e4	1.7e4
1e-3	15	1.8e3	1.7e3	1.8e3	1.8e3	15	1.8e4	1.8e4	1.8e4	1.8e4
1e-5	15	1.9e3	1.9e3	2.0e3	1.9e3	15	1.9e4	1.9e4	1.9e4	1.9e4
1e-8	15	2.2e3	2.1e3	2.2e3	2.2e3	15	2.0e4	1.9e4	2.0e4	2.0e4

Δf	**f_{12} in 5-D**, N=15, mFE=9274					**f_{12} in 20-D**, N=15, mFE=47246				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.2e3	8.5e2	1.5e3	1.2e3	15	3.2e3	2.1e3	4.2e3	3.2e3
1	15	2.0e3	1.6e3	2.4e3	2.0e3	15	7.7e3	5.8e3	9.7e3	7.7e3
1e-1	15	2.7e3	2.3e3	3.2e3	2.7e3	15	1.2e4	9.9e3	1.5e4	1.2e4
1e-3	15	3.5e3	3.0e3	4.1e3	3.5e3	15	1.8e4	1.6e4	2.1e4	1.8e4
1e-5	15	4.3e3	3.7e3	5.0e3	4.3e3	15	2.4e4	2.2e4	2.7e4	2.4e4
1e-8	15	5.1e3	4.3e3	5.9e3	5.1e3	15	2.9e4	2.7e4	3.2e4	2.9e4

Δf	**f_{14} in 5-D**, N=15, mFE=2762					**f_{14} in 20-D**, N=15, mFE=26426				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.1e1	8.1e0	1.3e1	1.1e1	15	2.9e2	2.7e2	3.2e2	2.9e2
1	15	1.2e2	1.0e2	1.3e2	1.2e2	15	7.0e2	6.7e2	7.3e2	7.0e2
1e-1	15	2.2e2	2.0e2	2.3e2	2.2e2	15	1.1e3	1.1e3	1.2e3	1.1e3
1e-3	15	6.3e2	6.0e2	6.7e2	6.3e2	15	3.8e3	3.8e3	3.9e3	3.8e3
1e-5	15	1.4e3	1.3e3	1.4e3	1.4e3	15	1.0e4	1.0e4	1.0e4	1.0e4
1e-8	15	2.5e3	2.5e3	2.6e3	2.5e3	15	2.3e4	2.3e4	2.3e4	2.3e4

Δf	**f_{16} in 5-D**, N=15, mFE=50491					**f_{16} in 20-D**, N=15, mFE=479099				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	3.6e2	2.8e2	4.4e2	3.6e2	15	2.3e3	2.2e3	2.5e3	2.3e3
1	15	2.2e3	1.7e3	2.6e3	2.2e3	15	2.8e4	2.1e4	3.5e4	2.8e4
1e-1	15	6.9e3	3.7e3	1.0e4	6.9e3	15	9.0e4	7.5e4	1.0e5	9.0e4
1e-3	15	1.4e4	9.6e3	1.9e4	1.4e4	15	1.9e5	1.5e5	2.3e5	1.9e5
1e-5	15	1.6e4	1.1e4	2.1e4	1.6e4	15	2.0e5	1.6e5	2.4e5	2.0e5
1e-8	15	1.7e4	1.2e4	2.2e4	1.7e4	15	2.2e5	1.8e5	2.6e5	2.2e5

Δf	**f_{18} in 5-D**, N=15, mFE=27718					**f_{18} in 20-D**, N=15, mFE=329369				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	1.0e2	8.6e1	1.2e2	1.0e2	15	6.5e2	6.0e2	7.0e2	6.5e2
1	15	1.3e3	4.2e2	2.2e3	1.3e3	15	9.6e3	5.1e3	1.5e4	9.6e3
1e-1	15	4.0e3	2.8e3	5.2e3	4.0e3	15	2.4e4	1.9e4	2.9e4	2.4e4
1e-3	15	9.3e3	8.5e3	1.0e4	9.3e3	15	7.3e4	6.0e4	8.7e4	7.3e4
1e-5	15	1.3e4	1.1e4	1.4e4	1.3e4	15	2.2e5	2.0e5	2.4e5	2.2e5
1e-8	15	1.6e4	1.4e4	1.8e4	1.6e4	15	2.4e5	2.2e5	2.6e5	2.4e5

Δf	**f_{20} in 5-D**, N=15, mFE=289011					**f_{20} in 20-D**, N=15, mFE=8.60e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	5.3e1	4.3e1	6.4e1	5.3e1	15	3.5e2	3.3e2	3.7e2	3.5e2
1	15	7.0e3	4.9e3	9.0e3	7.0e3	15	4.3e5	3.8e5	4.8e5	4.3e5
1e-1	15	1.1e5	8.9e4	1.2e5	1.1e5	15	3.1e6	2.7e6	3.5e6	3.1e6
1e-3	15	1.2e5	9.8e4	1.4e5	1.2e5	14	5.5e6	4.7e6	6.5e6	5.1e6
1e-5	15	1.2e5	1.0e5	1.4e5	1.2e5	14	5.6e6	4.8e6	6.6e6	5.2e6
1e-8	15	1.2e5	1.0e5	1.4e5	1.2e5	14	5.7e6	4.8e6	6.6e6	5.2e6

Δf	**f_{22} in 5-D**, N=15, mFE=199080					**f_{22} in 20-D**, N=15, mFE=2.70e6				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	4.9e2	2.6e2	7.3e2	4.9e2	15	3.2e3	1.8e3	4.6e3	3.2e3
1	15	7.7e3	2.2e3	1.3e4	7.7e3	15	7.2e4	3.9e4	1.1e5	7.2e4
1e-1	15	4.2e4	2.2e4	6.3e4	4.2e4	5	5.0e6	3.0e6	1.0e7	1.1e6
1e-3	15	4.3e4	2.2e4	6.4e4	4.3e4	5	5.0e6	3.0e6	9.9e6	1.1e6
1e-5	15	4.3e4	2.3e4	6.5e4	4.3e4	5	5.0e6	3.0e6	1.0e7	1.1e6
1e-8	15	4.3e4	2.3e4	6.5e4	4.3e4	5	5.0e6	3.0e6	1.0e7	1.1e6

Δf	**f_{24} in 5-D**, N=15, mFE=3.11e6					**f_{24} in 20-D**, N=15, mFE=1.21e7				
	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$	#	ERT	10%	90%	$\mathrm{RT}_{\mathrm{succ}}$
10	15	3.4e3	2.6e3	4.3e3	3.4e3	15	1.3e6	1.0e6	1.7e6	1.3e6
1	15	3.4e5	2.1e5	4.7e5	3.4e5	12	7.5e6	5.8e6	9.8e6	6.1e6
1e-1	6	6.4e6	4.3e6	1.1e7	2.3e6	3	5.2e7	3.1e7	1.6e8	1.1e7
1e-3	4	9.6e6	5.7e6	2.1e7	2.1e6	3	5.2e7	3.1e7	1.5e8	1.1e7
1e-5	3	1.3e7	7.3e6	4.0e7	2.0e6	3	5.2e7	3.1e7	1.6e8	1.1e7
1e-8	3	1.3e7	7.4e6	4.0e7	2.0e6	3	5.2e7	3.1e7	1.5e8	1.1e7

Table 2: Shown are, for a given target difference to the optimal function value Δf: the number of successful trials (#); the expected running time to surpass $f_{\mathrm{opt}}+\Delta f$ (ERT, see Figure 1); the 10%-tile and 90%-tile of the bootstrap distribution of ERT; the average number of function evaluations in successful trials or, if none was successful, as last entry the median number of function evaluations to reach the best function value ($\mathrm{RT}_{\mathrm{succ}}$). If $f_{\mathrm{opt}}+\Delta f$ was never reached, figures in *italics* denote the best achieved Δf-value of the median trial and the 10% and 90%-tile trial. Furthermore, N denotes the number of trials, and mFE denotes the maximum of number of function evaluations executed in one trial. See Figure 1 for the names of functions.

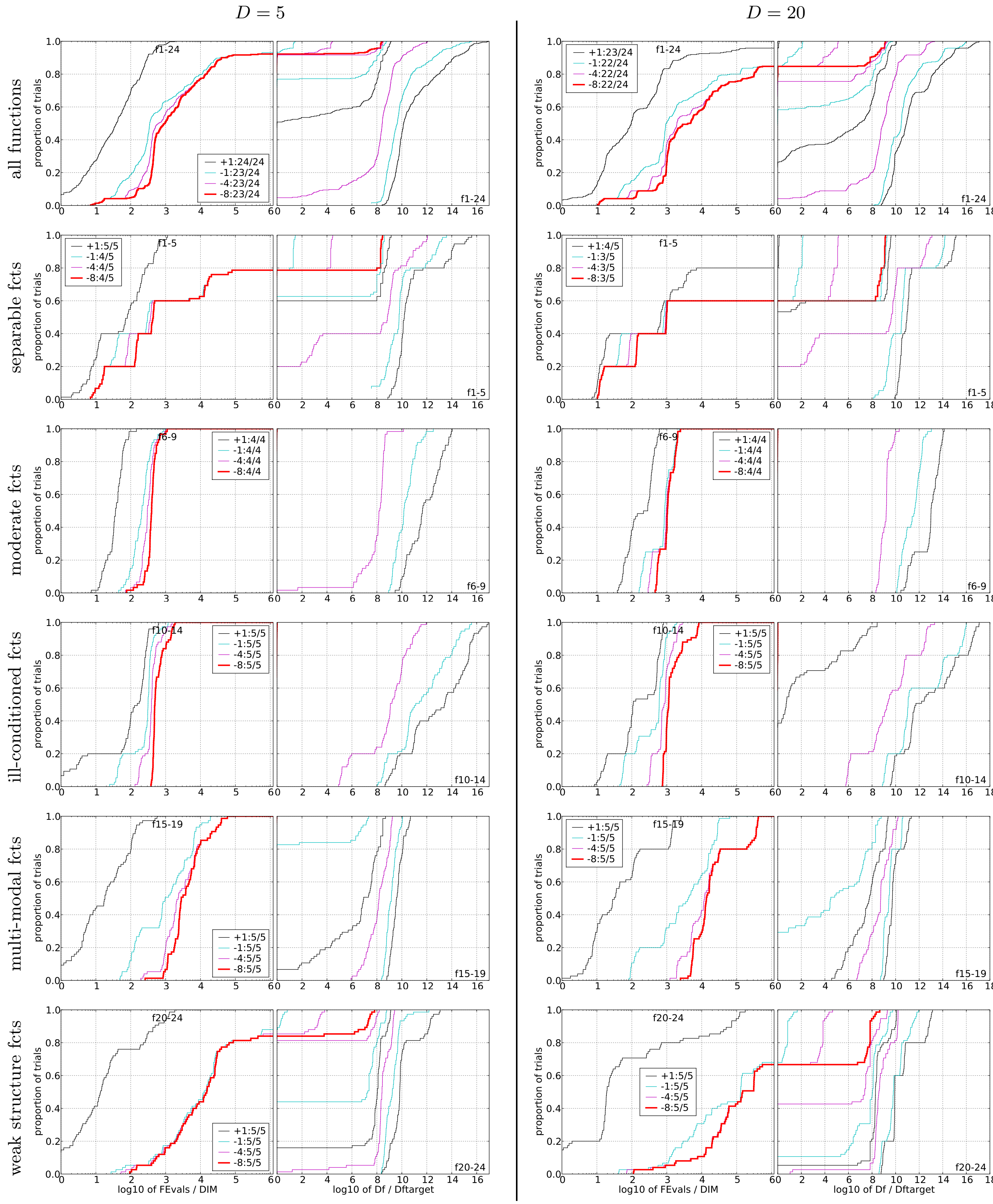

Figure 2: Empirical cumulative distribution functions (ECDFs), plotting the fraction of trials versus running time (left subplots) or versus Δf (right subplots). The thick red line represents the best achieved results. Left subplots: ECDF of the running time (number of function evaluations), divided by search space dimension D, to fall below $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where k is the first value in the legend. Right subplots: ECDF of the best achieved Δf divided by 10^k (upper left lines in continuation of the left subplot), and best achieved Δf divided by 10^{-8} for running times of $D, 10\,D, 100\,D\ldots$ function evaluations (from right to left cycling black-cyan-magenta). Top row: all results from all functions; second row: separable functions; third row: misc. moderate functions; fourth row: ill-conditioned functions; fifth row: multi-modal functions with adequate structure; last row: multi-modal functions with weak structure. The legends indicate the number of functions that were solved in at least one trial. FEvals denotes number of function evaluations, D and DIM denote search space dimension, and Δf and Df denote the difference to the optimal function value.

[2] 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.

[3] N. Hansen. The CMA evolution strategy: a comparing review. In J. Lozano, P. Larranaga, I. Inza, and E. Bengoetxea, editors, *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pages 75–102. Springer, 2006.

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

[5] 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.

[6] N. Hansen and S. Kern. Evaluating the CMA evolution strategy on multimodal test functions. In X. Yao et al., editors, *Parallel Problem Solving from Nature - PPSN VIII, LNCS 3242*, pages 282–291. Springer, 2004.

[7] N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18, 2003.

[8] N. Hansen, A. Niederberger, L. Guzzella, and P. Koumoutsakos. A method for handling uncertainty in evolutionary optimization with an application to feedback control of combustion. *IEEE Transactions on Evolutionary Computation*, 13(1):180–197, 2009.

[9] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.