Comparison Tables: BBOB 2015 Function Testbed with BBOB 2009 as Reference (Expensive Setting)

The BBOBies

July 16, 2015

Abstract

This document provides tabular results of the workshop on Black-Box Optimization Benchmarking held at GECCO 2015 with a focus on benchmarking black-box algorithms for small function evaluation budgets ("expensive setting"), see `http://coco.gforge.inria.fr/doku.php?id=bbob-2015`. Overall, 18 algorithms have been tested on 24 benchmark functions in dimensions between 2 and 20. Only three of them have been tested on the optional instances in dimension 40. A description of the used objective functions can be found in [7, 5]. The experimental set-up is described in [6].

The performance measure provided in the following tables is the expected number of objective function evaluations to reach a given target function value (ERT, expected running time), divided by the respective value for the best algorithm in BBOB-2009 (see [2]) if an algorithm from BBOB-2009 reached the given target function value. The ERT value is given otherwise ($\mathrm{ERT}_{\mathrm{best}}$ is noted as infinite). See [6] for details on how ERT is obtained. Bold entries in the table correspond to values below 3 or the top-three best values. Table 1 gives an overview on all algorithms submitted to the noise-free testbed at GECCO 2015.

Table 1: Names and references of all algorithms submitted for the noise-free testbed

algorithm short name	paper	reference
BSifeg	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
BSif	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
BSqi	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
BSrr	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
CMA-CSA	Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB Noiseless Testbed	[1]
CMA-MSR	Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB Noiseless Testbed	[1]
CMA-TPA	Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB Noiseless Testbed	[1]
GP1-CMAES	SBenchmarking Gaussian Processes and Random Forests Surrogate Models on the BBOB Noiseless Testbed	[3]
GP5-CMAES	Benchmarking Gaussian Processes and Random Forests Surrogate Models on the BBOB Noiseless Testbed	[3]
IPOPCMAv3p61	Benchmarking Gaussian Processes and Random Forests Surrogate Models on the BBOB Noiseless Testbed	[3]
LHD-10xDefault-MATSuMoT	The Impact of Initial Designs on the Performance of MATSuMoTo on the Noiseless BBOB-2015 Testbed: A Preliminary Study	[4]
LHD-2xDefault-MATSuMoTo	The Impact of Initial Designs on the Performance of MATSuMoTo on the Noiseless BBOB-2015 Testbed: A Preliminary Study	[4]
RAND-2xDefault-MATSuMoTo	The Impact of Initial Designs on the Performance of MATSuMoTo on the Noiseless BBOB-2015 Testbed: A Preliminary Study	[4]
RF1-CMAES	Benchmarking Gaussian Processes and Random Forests Surrogate Models on the BBOB Noiseless Testbed	[3]
RF5-CMAES	Benchmarking Gaussian Processes and Random Forests Surrogate Models on the BBOB Noiseless Testbed	[3]
Sifeg	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
Sif	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]
Srr	Dimension Selection in Axis-Parallel Brent-STEP Method for Black-Box Optimization of Separable Continuous Functions	[9]

Table 2: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*1.6e+1*:1.2	*4.0e+0*:2.6	*2.5e-2*:6.2	*1.0e-8*:6.2	*1.0e-8*:6.2	15/15
BSifeg	**1.6**(1)	**2.2**(2)	**1.7**(0.2)	**1.7**(0.2)	**1.7**(0.3)	15/15
BSif	**1.6**(1)	**2.2**(2)	**1.7**(0.2)	**1.7**(0.3)	**1.7**(0.2)	15/15
BSqi	**1.6**(2)	**2.2**(2)	**1.7**(0.3)	**1.7**(0.2)	**1.7**(0.3)	15/15
BSrr	**1.6**(2)	**2.2**(2)	**1.7**(0.3)	**1.7**(0.2)	**1.7**(0.3)	15/15
CMA-CSA	3.1(5)	3.0(3)	11(6)	42(5)	42(6)	15/15
CMA-MSR	**2.7**(1)	4.7(3)	16(11)	69(12)	69(9)	15/15
CMA-TPA	**2.8**(4)	4.5(4)	12(7)	44(8)	44(13)	15/15
GP1-CMAES	**2.4**(5)	**2.5**(2)	6.5(2)	24(4)	24(5)	15/15
GP5-CMAES	3.3(3)	**2.7**(2)	3.4(2)	31(14)	31(16)	15/15
IPOPCMAv3p	4.9(4)	4.6(5)	12(7)	43(9)	43(4)	15/15
LHD-10xDef	**2.6**(2)	3.8(5)	10(0)	∞	∞ *100*	0/15
LHD-2xDefa	**2.4**(1)	**2.8**(2)	3.3(0.8)	∞	∞ *100*	0/15
RAND-2xDef	**2.6**(2)	3.4(2)	3.5(0.9)	∞	∞ *100*	0/15
RF1-CMAES	3.1(4)	4.2(4)	8.0(3)	86(85)	86(111)	10/15
RF5-CMAES	**2.8**(3)	4.7(4)	60(38)	∞	∞ *502*	0/15
Sifeg	**1.6**(1)	**2.2**(2)	**2.3**(0.4)	6.3(0.5)	6.3(0.5)	15/15
Sif	**1.6**(1)	**2.2**(1)	**2.4**(0.3)	6.2(0.6)	6.2(0.6)	15/15
Srr	**1.6**(1)	**2.2**(2)	**2.3**(0.2)	5.9(0.5)	5.9(0.5)	15/15

Table 3: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*1.0e+7*:1.4	*1.6e+6*:2.7	*1.0e+5*:6.1	*6.3e-1*:20	*1.0e-8*:30	15/15
BSifeg	**1.7**(1)	3.1(1)	**1.7**(0.2)	**1.2**(0.3)	**1.2**(0.3)	15/15
BSif	**1.7**(1)	3.1(0.4)	**1.7**(0.2)	**1.1**(0.3)	**1.2**(0.1)	15/15
BSqi	**1.7**(1)	3.1(0.9)	**1.7**(0.2)	**0.93**(0.1)	**1.1**(0.2)	15/15
BSrr	**1.7**(1)	3.1(0)	**1.7**(0.3)	**1.1**(0.3)	**1.3**(0.2)	15/15
CMA-CSA	**1.3**(2)	**2.8**(3)	3.2(2)	15(4)	19(2)	15/15
CMA-MSR	**1.8**(0.5)	**2.6**(2)	**2.8**(2)	16(3)	24(5)	15/15
CMA-TPA	**1.5**(1)	**1.2**(0.5)	**1.7**(1)	12(5)	17(2)	15/15
GP1-CMAES	**2.1**(2)	**2.3**(2)	**2.2**(2)	13(6)	19(2)	12/15
GP5-CMAES	**2.2**(2)	**2.2**(2)	**1.5**(1)	5.1(2)	9.2(6)	14/15
IPOPCMAv3p	**2.3**(2)	**3.0**(2)	3.0(2)	17(6)	∞ *506*	0/15
LHD-10xDef	**1.3**(1)	**1.4**(0.8)	**2.6**(1)	∞	∞ *100*	0/15
LHD-2xDefa	**1.0**(0.7)	**1.2**(1)	**1.7**(1)	∞	∞ *100*	0/15
RAND-2xDef	**1.3**(2)	**1.0**(0.7)	**1.3**(1)	∞	∞ *100*	0/15
RF1-CMAES	**2.5**(3)	**2.2**(2)	**2.8**(3)	168(242)	∞ *506*	0/15
RF5-CMAES	**1.8**(1)	**2.8**(2)	**2.8**(2)	167(194)	∞ *502*	0/15
Sifeg	**1.7**(1)	3.2(2)	**2.0**(0.3)	**1.5**(0.3)	**1.8**(0.2)	15/15
Sif	**1.7**(1)	3.2(1)	**2.0**(0.5)	**1.4**(0.3)	**1.7**(0.3)	15/15
Srr	**1.7**(1)	3.2(0.4)	**2.0**(0.5)	**1.4**(0.2)	**1.9**(0.1)	15/15

Table 4: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*1.0e+2*:1.4	*4.0e+1*:4.1	*2.5e+1*:6.6	*6.3e+0*:26	*2.5e+0*:112	15/15
BSifeg	**2.2**(1)	**1.8**(1)	**1.4**(0.3)	**0.79**(0.6)	**0.30**(0.2)	15/15
BSif	**2.2**(1)	**1.8**(0.7)	**1.4**(0.7)	**0.77**(0.3)	**0.30**(0.2)	15/15
BSqi	**2.2**(1)	**1.8**(1)	**1.4**(0.5)	**0.76**(0.5)	**0.30**(0.2)	15/15
BSrr	**2.2**(0.7)	**1.8**(1)	**1.4**(0.6)	**0.74**(0.6)	**0.28**(0.2)	15/15
CMA-CSA	**1.8**(1)	**2.4**(2)	**2.4**(2)	**2.2**(1)	**2.9**(4)	15/15
CMA-MSR	3.0(3)	**2.9**(1)	**2.6**(1)	3.3(2)	4.6(9)	15/15
CMA-TPA	**2.0**(3)	**2.9**(3)	**2.9**(2)	6.8(8)	5.5(4)	15/15
GP1-CMAES	**2.8**(3)	**1.7**(2)	**2.6**(2)	**2.2**(1)	**2.9**(4)	11/15
GP5-CMAES	**2.0**(1)	**1.9**(2)	4.0(1)	5.8(7)	**2.3**(2)	13/15
IPOPCMAv3p	**1.6**(0.9)	**1.4**(1.0)	**1.6**(0.8)	**2.6**(1)	3.2(2)	10/15
LHD-10xDef	**1.9**(3)	**1.6**(2)	**1.7**(1)	3.0(2)	**1.5**(0.8)	8/15
LHD-2xDefa	**1.6**(0.4)	**1.2**(1)	**1.4**(1)	**1.0**(0.7)	**0.67**(0.8)	11/15
RAND-2xDef	**1.8**(2)	**1.1**(0.7)	**1.9**(1)	**2.4**(1)	**1.4**(2)	7/15
RF1-CMAES	**1.7**(2)	11(1)	7.2(2)	9.0(6)	5.6(6)	7/15
RF5-CMAES	**2.2**(3)	**2.4**(2)	14(2)	15(15)	6.5(8)	7/15
Sifeg	**2.2**(2)	**1.8**(1)	**1.5**(0.9)	**1.1**(0.5)	**0.30**(0.2)	15/15
Sif	**2.2**(2)	**1.8**(1)	**1.5**(0.7)	**1.1**(0.5)	**0.34**(0.2)	15/15
Srr	**2.2**(1)	**1.8**(1)	**1.5**(1)	**1.1**(0.4)	**0.31**(0.1)	15/15

Table 5: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*6.3e+1*:2.4	*4.0e+1*:5.2	*2.5e+1*:8.5	*1.0e+1*:22	*2.5e+0*:120	5/5
BSifeg	**1.8**(1)	**1.1**(0.9)	**1.1**(0.3)	**0.91**(0.4)	**0.42**(0.3)	15/15
BSif	**1.8**(2)	**1.1**(0.8)	**1.1**(0.5)	**0.91**(0.6)	**0.42**(0.3)	15/15
BSqi	**1.8**(2)	**1.1**(0.9)	**1.1**(0.5)	**0.98**(0.4)	**0.45**(0.3)	15/15
BSrr	**1.8**(2)	**1.1**(1.0)	**1.1**(0.6)	**0.93**(0.4)	**0.42**(0.3)	15/15
CMA-CSA	**1.9**(2)	**2.3**(2)	**2.4**(1)	**2.3**(2)	4.3(7)	15/15
CMA-MSR	3.2(2)	**2.6**(2)	**2.3**(2)	6.1(20)	11(13)	15/15
CMA-TPA	3.3(4)	**1.9**(0.7)	**2.5**(2)	**2.9**(4)	6.4(9)	15/15
GP1-CMAES	**2.5**(2)	**1.9**(2)	**2.3**(2)	**2.2**(1)	4.6(3)	8/15
GP5-CMAES	3.2(3)	**2.1**(3)	10(15)	6.4(3)	3.8(3)	9/15
IPOPCMAv3p	**2.9**(3)	**2.2**(2)	**2.8**(3)	**2.9**(3)	5.0(10)	8/15
LHD-10xDef	**1.7**(4)	**1.7**(1)	**2.0**(3)	**2.7**(3)	**1.9**(2)	6/15
LHD-2xDefa	**1.5**(1)	**1.9**(1)	**2.0**(1)	**1.8**(1)	**1.7**(1)	6/15
RAND-2xDef	**2.2**(1)	**1.6**(0.8)	**1.2**(0.8)	**1.7**(2)	**2.1**(1.0)	5/15
RF1-CMAES	**2.1**(2)	**2.3**(3)	**2.1**(2)	3.3(3)	28(27)	2/15
RF5-CMAES	17(2)	8.6(1)	12(16)	16(22)	28(26)	2/15
Sifeg	**1.8**(1)	**1.2**(1.0)	**1.3**(0.2)	**0.96**(0.3)	**0.50**(0.3)	15/15
Sif	**1.8**(2)	**1.2**(1)	**1.3**(0.9)	**0.95**(0.1)	**0.52**(0.3)	15/15
Srr	**1.8**(2)	**1.2**(1)	**1.2**(0.3)	**0.96**(0.6)	**0.51**(0.1)	15/15

Table 6: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*4.0e+1*:1.4	*1.6e+1*:3.5	*1.0e-8*:4.4	*1.0e-8*:4.4	*1.0e-8*:4.4	15/15
BSifeg	**2.9**(1)	**1.8**(0.1)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
BSif	**2.9**(2)	**1.8**(0.1)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
BSqi	**2.9**(2)	**1.8**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.1)	15/15
BSrr	**2.9**(2)	**1.8**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
CMA-CSA	4.2(4)	**3.0**(3)	4.8(3)	4.8(4)	4.8(3)	15/15
CMA-MSR	4.0(5)	**2.9**(2)	6.2(3)	6.2(2)	6.2(2)	15/15
CMA-TPA	3.3(4)	**2.0**(2)	4.0(1)	4.0(1)	4.0(2)	15/15
GP1-CMAES	3.5(4)	**2.2**(1)	9.3(9)	9.3(14)	9.3(12)	15/15
GP5-CMAES	**2.6**(4)	**1.9**(0.9)	4.4(2)	4.4(3)	4.4(2)	15/15
IPOPCMAv3p	5.0(3)	4.1(4)	13(8)	13(9)	13(16)	15/15
LHD-10xDef	**2.6**(2)	**2.2**(2)	14(0.6)	14(0.3)	14(0.6)	15/15
LHD-2xDefa	**1.4**(0.7)	**2.0**(1)	3.3(0.6)	3.3(0.6)	3.3(0.6)	15/15
RAND-2xDef	**1.4**(0.7)	**1.9**(2)	3.6(0.6)	3.6(0.6)	3.6(0.9)	15/15
RF1-CMAES	**2.4**(3)	**1.9**(2)	22(26)	22(32)	22(26)	15/15
RF5-CMAES	3.7(5)	3.2(3)	61(75)	61(34)	61(76)	12/15
Sifeg	**2.9**(2)	**1.8**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
Sif	**2.9**(2)	**1.8**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
Srr	**2.9**(2)	**1.8**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15

Table 7: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*6.3e+4*:1.4	*1.0e+2*:2.8	*1.6e+1*:10	*1.0e+0*:23	*2.5e-6*:103	15/15
BSifeg	**1.3**(1)	**2.0**(3)	245(513)	416(421)	2668(4813)	1/15
BSif	**1.3**(1)	**2.1**(3)	307(617)	442(1192)	∞ *2e4*	0/15
BSqi	**1.3**(1)	**2.0**(1)	355(776)	344(686)	2671(5134)	1/15
BSrr	**1.3**(1)	**2.0**(1)	341(56)	403(547)	1293(1193)	2/15
CMA-CSA	**1.6**(2)	3.6(5)	3.1(2)	**4.3**(4)	**4.2**(0.8)	15/15
CMA-MSR	**1.8**(2)	**2.3**(3)	**2.4**(1)	5.3(3)	5.2(0.8)	15/15
CMA-TPA	**1.3**(0.4)	**1.2**(1)	**1.2**(2)	**3.6**(3)	**3.8**(0.7)	15/15
GP1-CMAES	**2.4**(3)	3.1(5)	**2.8**(2)	4.9(0.7)	∞ *506*	0/15
GP5-CMAES	**1.1**(0.4)	**2.3**(7)	**2.3**(2)	9.2(15)	∞ *506*	0/15
IPOPCMAv3p	**1.5**(2)	**2.7**(2)	**2.6**(5)	**4.6**(2)	**4.8**(4)	13/15
LHD-10xDef	**1.3**(1)	**1.5**(2)	**1.3**(1)	9.0(12)	∞ *100*	0/15
LHD-2xDefa	**1.1**(0)	**1.9**(2)	**1.5**(0.9)	12(16)	∞ *100*	0/15
RAND-2xDef	**1.4**(0.7)	**1.9**(2)	**1.2**(0.7)	6.3(8)	∞ *100*	0/15
RF1-CMAES	**1.8**(0.7)	**2.3**(2)	5.5(14)	67(146)	∞ *506*	0/15
RF5-CMAES	**1.5**(0.9)	3.2(5)	4.8(6)	40(35)	∞ *508*	0/15
Sifeg	**1.3**(1)	**2.3**(4)	333(47)	271(534)	2624(3800)	1/15
Sif	**1.3**(1)	**2.3**(4)	355(1172)	318(750)	2659(2915)	1/15
Srr	**1.3**(1)	**2.1**(3)	319(1229)	242(376)	1251(1158)	2/15

Table 8: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*4.0e+2*:1.6	*1.0e+1*:3.2	*2.5e+0*:14	*1.6e+0*:21	*1.6e-2*:188	15/15
BSifeg	**1.3**(0.5)	**1.7**(1)	135(216)	376(318)	698(696)	2/15
BSif	**1.3**(0.6)	**1.7**(1)	173(395)	376(1149)	252(216)	5/15
BSqi	**1.3**(1)	**1.7**(1)	206(466)	256(513)	355(565)	4/15
BSrr	**1.3**(1)	**1.7**(1)	268(467)	543(2200)	331(424)	4/15
CMA-CSA	**1.3**(0.6)	4.0(3)	3.6(2)	3.5(10)	**1.1**(1)	15/15
CMA-MSR	**1.7**(2)	4.4(5)	**2.0**(3)	**1.6**(2)	**1.0**(0.3)	15/15
CMA-TPA	**1.1**(0.3)	3.7(3)	**2.5**(2)	**1.9**(2)	**0.80**(0.8)	15/15
GP1-CMAES	**1.4**(1)	4.2(1)	**2.1**(0.9)	**1.9**(2)	**1.0**(0.4)	14/15
GP5-CMAES	**0.67**(0.2)	3.3(3)	**2.0**(3)	**2.0**(3)	**0.98**(0.3)	14/15
IPOPCMAv3p	**1.9**(0.9)	3.9(8)	**2.5**(3)	3.6(4)	**1.7**(1)	13/15
LHD-10xDef	**1.2**(1)	7.6(5)	**3.0**(2)	**2.6**(2)	∞ *100*	0/15
LHD-2xDefa	**1**(0.3)	3.0(2)	**1.2**(0.9)	**1.0**(0.8)	**1.3**(1)	5/15
RAND-2xDef	**1.7**(1)	3.8(1)	**1.2**(0.4)	**1.2**(1)	**1.1**(2)	6/15
RF1-CMAES	**1.5**(1)	4.1(1)	**2.0**(2)	**2.7**(2)	**2.2**(2)	11/15
RF5-CMAES	**1.6**(2)	3.8(4)	**2.5**(2)	3.5(2)	13(15)	3/15
Sifeg	**1.3**(0.3)	**1.8**(2)	68(180)	256(954)	444(405)	3/15
Sif	**1.3**(0.9)	**1.8**(2)	78(36)	204(440)	485(501)	3/15
Srr	**1.3**(0.6)	**1.8**(1)	135(144)	287(367)	473(349)	3/15

Table 9: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*2.5e+3*:1.2	*1.0e+2*:3.2	*6.3e+0*:7.0	*1.6e-1*:27	*1.6e-6*:100	15/15
BSifeg	**1.9**(3)	**1.9**(1)	4.0(3)	2208(3277)	∞ *2e4*	0/15
BSif	**1.9**(0.4)	**1.9**(1)	4.2(8)	2303(3150)	∞ *2e4*	0/15
BSqi	**1.9**(2)	**1.9**(1)	4.1(8)	1323(2706)	∞ *2e4*	0/15
BSrr	**1.9**(2)	**1.9**(1)	4.9(5)	2159(3507)	∞ *2e4*	0/15
CMA-CSA	3.5(2)	3.8(3)	8.4(6)	12(11)	**5.9**(3)	15/15
CMA-MSR	**2.1**(0.6)	**1.7**(0.9)	7.3(3)	10(8)	**5.8**(2)	15/15
CMA-TPA	5.2(2)	4.6(3)	4.8(4)	**6.9**(4)	**4.0**(0.9)	15/15
GP1-CMAES	**2.2**(2)	**2.6**(5)	5.0(2)	9.1(6)	10(6)	7/15
GP5-CMAES	**2.6**(2)	**2.6**(2)	**3.8**(6)	11(20)	6.6(6)	9/15
IPOPCMAv3p	3.6(3)	4.0(5)	5.1(4)	**7.8**(12)	8.7(12)	8/15
LHD-10xDef	**2.4**(2)	3.9(8)	6.0(5)	13(12)	∞ *100*	0/15
LHD-2xDefa	**1.6**(2)	**2.4**(2)	**3.0**(2)	**5.1**(2)	∞ *100*	0/15
RAND-2xDef	**2.4**(2)	3.5(1)	**3.2**(1)	10(6)	∞ *100*	0/15
RF1-CMAES	3.2(3)	3.1(2)	9.1(19)	17(10)	∞ *506*	0/15
RF5-CMAES	3.1(3)	22(82)	39(55)	78(92)	∞ *506*	0/15
Sifeg	**1.9**(2)	**2.0**(2)	4.1(6)	933(701)	∞ *2e4*	0/15
Sif	**1.9**(3)	**2.0**(0.6)	4.3(5)	975(1446)	∞ *2e4*	0/15
Srr	**1.9**(2)	**2.0**(1)	3.8(0.4)	1348(1373)	∞ *2e4*	0/15

Table 10: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*6.3e+0*:13	*4.0e+0*:15	*2.5e+0*:15	*2.5e-1*:21	*1.0e-8*:94	15/15
BSifeg	15(24)	14(12)	14(17)	210(120)	∞ *2e4*	0/15
BSif	13(10)	12(19)	12(13)	193(218)	∞ *2e4*	0/15
BSqi	13(14)	12(19)	12(13)	207(294)	∞ *2e4*	0/15
BSrr	12(17)	12(20)	14(20)	241(104)	∞ *2e4*	0/15
CMA-CSA	**2.7**(1)	3.5(1)	4.1(5)	8.7(9)	**6.0**(3)	15/15
CMA-MSR	**2.4**(2)	**2.4**(2)	4.6(2)	10(14)	7.0(3)	15/15
CMA-TPA	**2.5**(2)	3.3(4)	4.6(7)	9.3(7)	**6.0**(2)	15/15
GP1-CMAES	**2.8**(2)	4.1(5)	6.8(17)	15(22)	26(44)	3/15
GP5-CMAES	**1.7**(1)	**1.5**(0.2)	**2.0**(2)	**6.0**(3)	**5.8**(6)	10/15
IPOPCMAv3p	**2.3**(2)	**2.2**(2)	**2.5**(2)	6.1(5)	11(10)	7/15
LHD-10xDef	**2.8**(2)	3.4(2)	3.8(3)	16(15)	∞ *100*	0/15
LHD-2xDefa	**1.3**(0.7)	**1.2**(0.8)	**1.9**(1)	**4.6**(3)	∞ *100*	0/15
RAND-2xDef	**1.6**(1)	**1.6**(1)	**2.0**(0.9)	**6.0**(5)	∞ *100*	0/15
RF1-CMAES	4.8(3)	6.6(35)	11(10)	26(40)	∞ *506*	0/15
RF5-CMAES	21(28)	20(19)	21(31)	57(61)	∞ *504*	0/15
Sifeg	**2.6**(3)	**2.8**(2)	4.0(5)	182(348)	∞ *2e4*	0/15
Sif	**2.5**(3)	5.0(3)	5.8(4)	259(407)	∞ *2e4*	0/15
Srr	**2.5**(3)	**2.8**(2)	3.6(3)	232(223)	∞ *2e4*	0/15

Table 11: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*1.6e+6*:2.0	*4.0e+5*:3.2	*6.3e+2*:8.8	*1.0e+1*:30	*2.5e-8*:101	15/15
BSifeg	**2.7**(2)	**2.5**(0.2)	**1.5**(0.5)	38(59)	∞ *5883*	0/15
BSif	**2.7**(2)	**2.5**(0.7)	**1.5**(0.5)	71(116)	∞ *8444*	0/15
BSqi	**2.7**(2)	**2.5**(2)	**1.5**(0.5)	135(10)	∞ *1e4*	0/15
BSrr	**2.7**(2)	**2.5**(0.2)	**1.5**(0.5)	46(108)	∞ *6681*	0/15
CMA-CSA	3.7(4)	3.3(4)	7.2(3)	7.5(4)	**5.3**(0.9)	15/15
CMA-MSR	**2.1**(2)	**2.3**(2)	7.0(5)	7.4(4)	6.8(0.6)	15/15
CMA-TPA	3.4(4)	**2.8**(3)	5.9(5)	**6.3**(3)	5.3(1)	15/15
GP1-CMAES	**1.5**(2)	**1.5**(0.5)	4.4(3)	**4.6**(3)	**5.2**(2)	12/15
GP5-CMAES	**2.2**(2)	**2.2**(2)	**2.2**(0.9)	**1.6**(0.6)	**2.5**(0.9)	14/15
IPOPCMAv3p	**2.6**(3)	**2.7**(2)	6.8(3)	8.3(7)	∞ *506*	0/15
LHD-10xDef	**1.7**(1)	**1.6**(2)	7.9(1)	16(9)	∞ *100*	0/15
LHD-2xDefa	**1.7**(2)	**2.2**(1)	3.8(2)	6.9(7)	∞ *100*	0/15
RAND-2xDef	**2.7**(3)	**2.7**(2)	3.9(1)	10(9)	∞ *100*	0/15
RF1-CMAES	**2.3**(3)	**2.2**(2)	4.2(1)	19(25)	∞ *506*	0/15
RF5-CMAES	3.4(2)	**2.9**(2)	8.7(6)	14(18)	∞ *502*	0/15
Sifeg	**2.7**(2)	**2.8**(1)	**1.9**(0.3)	8.6(12)	∞ *2159*	0/15
Sif	**2.7**(2)	**2.8**(2)	**1.9**(0.4)	12(17)	∞ *2178*	0/15
Srr	**2.7**(2)	**2.8**(1)	**1.9**(0.3)	7.9(0.1)	∞ *2193*	0/15

Table 12: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*1.0e+7*:1.1	*1.6e+6*:3.2	*1.0e+4*:6.6	*4.0e+1*:23	*4.0e-8*:100	15/15
BSifeg	**1.2**(0)	**1.6**(1)	**1.4**(0.3)	47(49)	∞ *6926*	0/15
BSif	**1.2**(0.5)	**1.6**(1)	**1.4**(0.3)	39(23)	∞ *5430*	0/15
BSqi	**1.2**(0)	**1.6**(1)	**1.4**(0.3)	71(92)	∞ *5998*	0/15
BSrr	**1.2**(0.5)	**1.6**(1)	**1.4**(0.6)	40(109)	∞ *6000*	0/15
CMA-CSA	**1.6**(0.2)	**1.3**(1)	3.5(3)	4.7(5)	**5.2**(0.8)	15/15
CMA-MSR	**1.1**(0.9)	**0.96**(1)	3.8(5)	5.3(0.8)	6.6(0.6)	15/15
CMA-TPA	**1.9**(2)	**1.5**(1)	4.2(3)	6.0(3)	5.3(0.4)	15/15
GP1-CMAES	**2.2**(4)	**1.6**(2)	3.4(3)	**3.6**(3)	**4.7**(0.8)	14/15
GP5-CMAES	3.6(4)	**1.8**(2)	**2.2**(0.9)	**1.3**(0.4)	**2.4**(0.7)	15/15
IPOPCMAv3p	3.1(2)	**1.9**(1)	3.7(1)	4.3(2)	∞ *506*	0/15
LHD-10xDef	**1.5**(0.7)	**1.3**(0.8)	4.9(5)	11(7)	∞ *100*	0/15
LHD-2xDefa	**1.7**(1)	**1.3**(0.9)	**2.5**(1)	5.6(8)	∞ *100*	0/15
RAND-2xDef	**1.8**(1)	**1.2**(1)	**2.7**(2)	**3.0**(0.5)	∞ *100*	0/15
RF1-CMAES	**2.7**(4)	**2.9**(5)	4.9(4)	5.5(7)	∞ *506*	0/15
RF5-CMAES	**1.7**(0.9)	**2.2**(3)	3.8(3)	6.4(12)	∞ *508*	0/15
Sifeg	**1.2**(1)	**1.7**(1)	**1.7**(1.0)	10(6)	∞ *2391*	0/15
Sif	**1.2**(1)	**1.7**(1)	**1.7**(0.6)	10(29)	∞ *2405*	0/15
Srr	**1.2**(0.5)	**1.7**(1)	**1.7**(0.8)	11(11)	∞ *2357*	0/15

Table 13: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*2.5e+8*:1.3	*6.3e+6*:2.7	*6.3e+5*:6.3	*4.0e+1*:21	*1.6e-3*:101	15/15
BSifeg	**0.85**(0.4)	**1.2**(1)	**1.3**(0.8)	12(28)	295(517)	5/15
BSif	**0.85**(0.4)	**1.2**(1)	**1.3**(1.0)	29(3)	368(224)	4/15
BSqi	**0.85**(0.2)	**1.2**(1)	**1.3**(0.5)	23(5)	320(351)	4/15
BSrr	**0.85**(0.2)	**1.2**(1)	**1.3**(0.9)	19(0.4)	326(335)	4/15
CMA-CSA	**1.4**(2)	**1.9**(3)	**1.4**(2)	6.5(5)	**8.5**(17)	15/15
CMA-MSR	**1.3**(0.8)	**1.3**(0.7)	**0.87**(0.8)	9.1(7)	9.3(5)	15/15
CMA-TPA	**1.4**(2)	**1.8**(2)	**1.8**(2)	7.0(7)	**6.8**(9)	15/15
GP1-CMAES	**2.0**(6)	**1.5**(2)	**1.2**(1)	3.6(1.0)	10(9)	7/15
GP5-CMAES	**1.2**(2)	**1.6**(1)	**1.3**(2)	**3.5**(6)	**7.8**(5)	7/15
IPOPCMAv3p	**1.7**(3)	**1.8**(1)	**1.7**(2)	4.0(1)	14(26)	5/15
LHD-10xDef	**0.90**(0.4)	**1.2**(0.7)	**0.97**(1)	24(19)	∞ *100*	0/15
LHD-2xDefa	**1**(0.2)	**0.78**(0.6)	**0.97**(0.9)	5.0(4)	∞ *100*	0/15
RAND-2xDef	**1.1**(0.4)	**0.95**(0.7)	**1.1**(0.8)	6.3(11)	∞ *100*	0/15
RF1-CMAES	**1.6**(4)	**1.5**(2)	**1.4**(0.8)	4.8(3)	∞ *506*	0/15
RF5-CMAES	**1.2**(0.2)	**1.9**(3)	**1.9**(2)	11(6)	∞ *504*	0/15
Sifeg	**0.85**(0)	**1.2**(1)	**1.3**(1)	**2.8**(6)	76(138)	5/15
Sif	**0.85**(0)	**1.2**(1)	**1.4**(0.6)	4.9(14)	83(92)	5/15
Srr	**0.85**(0.4)	**1.2**(1)	**1.4**(1)	**3.4**(1)	124(113)	4/15

Table 14: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*4.0e+2*:1.6	*2.5e+2*:3.1	*6.3e+1*:8.7	*1.0e+1*:23	*4.0e-6*:100	15/15
BSifeg	**2.3**(2)	**1.6**(1)	184(1215)	541(602)	∞ *2e4*	0/15
BSif	**2.3**(2)	**1.6**(1)	224(780)	789(712)	∞ *2e4*	0/15
BSqi	**2.3**(2)	**1.6**(1)	252(879)	430(542)	∞ *2e4*	0/15
BSrr	**2.3**(2)	**1.6**(1)	184(0.7)	491(498)	∞ *2e4*	0/15
CMA-CSA	**2.5**(2)	**2.1**(1)	**2.8**(3)	**2.9**(3)	**5.0**(0.5)	15/15
CMA-MSR	**2.8**(5)	**2.1**(2)	4.1(4)	3.7(1)	**6.2**(1)	15/15
CMA-TPA	**2.8**(3)	**2.5**(2)	3.5(5)	4.5(3)	**5.3**(0.6)	15/15
GP1-CMAES	**1.6**(2)	**1.4**(1)	**1.9**(2)	**2.7**(2)	24(20)	3/15
GP5-CMAES	**2.4**(3)	**1.9**(3)	5.5(1)	3.8(2)	11(4)	6/15
IPOPCMAv3p	**2.2**(3)	**1.8**(2)	**2.6**(2)	3.9(2)	∞ *506*	0/15
LHD-10xDef	**2.0**(1)	**2.2**(2)	**2.8**(2)	**2.9**(1)	∞ *100*	0/15
LHD-2xDefa	**1.7**(2)	**1.5**(0.6)	**1.4**(0.7)	**1.4**(0.8)	∞ *100*	0/15
RAND-2xDef	**1.6**(1)	**1.8**(2)	**1.7**(0.8)	**1.4**(0.8)	∞ *100*	0/15
RF1-CMAES	**2.1**(0.9)	**1.8**(3)	**2.2**(2)	7.9(7)	∞ *506*	0/15
RF5-CMAES	**2.6**(2)	**2.1**(1)	**1.9**(3)	11(33)	∞ *508*	0/15
Sifeg	**2.3**(2)	**1.6**(1)	178(899)	563(826)	∞ *2e4*	0/15
Sif	**2.3**(2)	**1.6**(1)	258(1006)	410(561)	∞ *2e4*	0/15
Srr	**2.3**(2)	**1.6**(1.0)	278(884)	428(499)	∞ *2e4*	0/15

Table 15: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*1.6e+1*:1.4	*2.5e+0*:4.2	*1.0e+0*:7.4	*2.5e-2*:21	*1.0e-8*:101	15/15
BSifeg	**2.0**(3)	12(11)	8.5(20)	7.3(22)	∞ *2e4*	0/15
BSif	**2.0**(3)	12(20)	8.7(11)	6.8(12)	∞ *2e4*	0/15
BSqi	**2.0**(3)	7.4(15)	5.7(7)	4.2(4)	∞ *2e4*	0/15
BSrr	**2.0**(2)	11(24)	8.4(13)	6.2(6)	∞ *2e4*	0/15
CMA-CSA	**1.3**(0.7)	**1.7**(1)	**1.9**(2)	4.1(2)	**5.5**(0.9)	15/15
CMA-MSR	**2.1**(2)	**2.4**(1.0)	**2.8**(2)	6.0(2)	**6.5**(0.7)	15/15
CMA-TPA	**2.4**(1)	3.5(4)	3.5(2)	5.8(2)	**5.8**(0.8)	15/15
GP1-CMAES	**0.95**(1)	**2.2**(0.6)	**1.9**(2)	**2.5**(0.7)	∞ *506*	0/15
GP5-CMAES	**1.9**(3)	**2.9**(2)	**2.3**(1)	**1.7**(0.8)	∞ *506*	0/15
IPOPCMAv3p	**1.2**(0.7)	3.3(2)	3.1(2)	3.8(1)	∞ *506*	0/15
LHD-10xDef	**1.4**(0.7)	**1.2**(1)	3.1(3)	8.2(9)	∞ *100*	0/15
LHD-2xDefa	**1.4**(0.9)	**1.6**(2)	**1.4**(1)	**2.0**(1)	∞ *100*	0/15
RAND-2xDef	**1.1**(1)	**2.2**(1)	**1.9**(1)	**2.1**(0.9)	∞ *100*	0/15
RF1-CMAES	**1.4**(2)	5.9(6)	6.3(9)	8.8(9)	∞ *506*	0/15
RF5-CMAES	**1.2**(0.9)	44(46)	48(59)	79(74)	∞ *506*	0/15
Sifeg	**2.0**(1)	**2.8**(1)	**2.3**(1)	3.2(5)	∞ *2e4*	0/15
Sif	**2.0**(2)	**2.8**(3)	**2.4**(0.7)	4.1(4)	∞ *2e4*	0/15
Srr	**2.0**(1)	**2.6**(2)	**2.1**(0.7)	**2.1**(3)	∞ *2e4*	0/15

Table 16: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*1.6e+2*:1.2	*4.0e+1*:4.7	*2.5e+1*:10	*1.0e+1*:37	*2.5e+0*:118	5/5
BSifeg	**1.4**(0.4)	**1.8**(1.0)	**2.2**(0.8)	27(5)	111(252)	10/15
BSif	**1.4**(1)	151(561)	70(257)	22(140)	118(120)	10/15
BSqi	**1.4**(2)	3.3(0.9)	4.0(9)	3.3(4)	62(119)	13/15
BSrr	**1.4**(0.4)	6.5(39)	6.4(0.1)	4.2(11)	59(91)	13/15
CMA-CSA	**1.3**(0)	**1.0**(1)	**1.1**(1)	**1.1**(0.5)	**1.6**(3)	15/15
CMA-MSR	**2.1**(0.8)	**1.3**(2)	**1.3**(2)	**0.86**(0.5)	**2.2**(0.6)	15/15
CMA-TPA	**1.8**(2)	**2.2**(5)	**1.5**(0.7)	**1.5**(1)	3.6(5)	15/15
GP1-CMAES	**2.3**(5)	**1.6**(1)	**1.3**(1.0)	**1.3**(1)	3.1(3)	10/15
GP5-CMAES	**2.4**(2)	**2.1**(2)	**1.3**(0.4)	**0.75**(0.3)	**1.6**(3)	14/15
IPOPCMAv3p	**1.4**(0.4)	**2.0**(2)	**1.4**(0.5)	**1.3**(1)	4.8(5)	8/15
LHD-10xDef	**2.5**(2)	**1.8**(2)	**1.6**(2)	**2.2**(1)	**3.0**(2)	4/15
LHD-2xDefa	**1.4**(1)	**0.81**(1)	**0.94**(0.7)	**0.90**(0.8)	**1.4**(2)	7/15
RAND-2xDef	**2.1**(0.8)	**1.6**(2)	**1.1**(0.8)	**1.4**(1.0)	**1.4**(2)	7/15
RF1-CMAES	3.2(6)	**1.8**(0.8)	**1.5**(1)	**1.9**(0.7)	**1.8**(2)	12/15
RF5-CMAES	**2.0**(1)	**1.8**(1)	**1.9**(3)	6.1(1)	8.0(6)	6/15
Sifeg	**1.4**(2)	**1.5**(2)	**1.8**(0.7)	**1.7**(3)	50(103)	12/15
Sif	**1.4**(1)	**1.5**(2)	**1.9**(3)	**1.7**(0.4)	33(42)	13/15
Srr	**1.4**(1)	**1.5**(1)	**1.5**(0.5)	**1.6**(3)	35(46)	15/15

Table 17: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*1.0e+2*:1.1	*2.5e+1*:3.9	*1.6e+1*:6.5	*4.0e+0*:31	*2.5e-1*:127	5/5
BSifeg	**1.8**(1)	**2.9**(5)	**2.7**(3)	**1.0**(0.8)	12(29)	15/15
BSif	**1.8**(0.9)	**2.9**(2)	**2.9**(2)	**1.1**(1)	7.0(12)	15/15
BSqi	**1.8**(3)	3.9(1)	**2.9**(5)	**1.1**(0.9)	17(14)	15/15
BSrr	**1.8**(3)	3.4(1)	3.8(4)	**1.2**(1)	20(54)	15/15
CMA-CSA	**1.8**(2)	3.3(3)	4.1(3)	6.1(9)	3.6(4)	15/15
CMA-MSR	**1.5**(0.9)	**1.9**(3)	**2.3**(2)	6.7(0.9)	8.1(7)	15/15
CMA-TPA	**1.8**(2)	**1.6**(2)	**2.1**(2)	3.7(7)	**3.0**(4)	15/15
GP1-CMAES	**1.5**(0.7)	**1.8**(0.6)	**1.5**(1)	**2.3**(1)	4.6(7)	8/15
GP5-CMAES	**1.6**(2)	**2.3**(2)	**2.4**(2)	11(15)	6.6(10)	7/15
IPOPCMAv3p	**1.2**(0.4)	**1.6**(1)	**1.7**(2)	**1.2**(0.9)	**2.5**(4)	11/15
LHD-10xDef	**1.5**(0.9)	**1.8**(0.8)	**1.5**(1)	**1.5**(1)	5.8(6)	2/15
LHD-2xDefa	**1.5**(0.9)	**1.7**(3)	**2.0**(2)	**1.4**(1)	**2.6**(3)	4/15
RAND-2xDef	**1.2**(0.4)	**1.7**(2)	**1.5**(2)	**1.0**(1)	**1.1**(1.0)	8/15
RF1-CMAES	**1.8**(2)	**1.7**(1)	**1.7**(3)	3.6(5)	5.3(4)	7/15
RF5-CMAES	**1.1**(0)	**1.3**(1.0)	**1.7**(2)	6.1(8)	6.4(9)	6/15
Sifeg	**1.8**(3)	**2.4**(2)	**2.1**(0.7)	**0.94**(0.6)	7.4(0.8)	15/15
Sif	**1.8**(3)	**2.4**(2)	**2.0**(1)	**0.89**(0.2)	14(42)	15/15
Srr	**1.8**(3)	**2.3**(2)	**2.0**(1)	**0.91**(0.4)	12(29)	15/15

Table 18: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*4.0e+1*:1.2	*1.0e+1*:2.7	*4.0e+0*:10	*2.5e+0*:28	*1.6e-1*:119	5/5
BSifeg	**1.2**(1)	**2.2**(2)	**1.9**(2)	6.6(22)	41(41)	14/15
BSif	**1.2**(1)	**2.2**(2)	**2.1**(3)	11(21)	40(48)	15/15
BSqi	**1.2**(0.8)	**2.2**(2)	**1.9**(2)	32(0.8)	21(28)	15/15
BSrr	**1.2**(0.6)	**2.2**(2)	**1.9**(2)	21(0.7)	40(57)	15/15
CMA-CSA	**1.3**(0.6)	3.1(10)	9.3(14)	3.5(11)	**2.1**(2)	15/15
CMA-MSR	**1.7**(1)	20(59)	12(20)	5.8(1)	3.5(3)	15/15
CMA-TPA	**1.4**(0.8)	3.6(6)	3.3(6)	**1.4**(1)	**2.2**(0.3)	15/15
GP1-CMAES	**2.2**(3)	3.8(9)	6.6(26)	4.1(10)	4.7(8)	8/15
GP5-CMAES	**1.4**(0.6)	7.6(7)	17(29)	8.8(12)	5.7(4)	7/15
IPOPCMAv3p	**1.2**(0.4)	**2.5**(2)	**1.8**(2)	**2.2**(0.4)	**2.3**(3)	12/15
LHD-10xDef	**1.4**(0.6)	**1.8**(2)	**1.8**(2)	**1.6**(1.0)	4.1(5)	3/15
LHD-2xDefa	**1.2**(0.8)	**1.5**(2)	**1.5**(1)	**0.77**(0.6)	3.9(6)	3/15
RAND-2xDef	**1.5**(0.6)	**1.9**(2)	**1.8**(1)	**0.79**(0.3)	**2.6**(4)	4/15
RF1-CMAES	**1.3**(1)	**2.6**(4)	5.6(1)	**2.3**(0.7)	7.2(10)	6/15
RF5-CMAES	**1.4**(0)	50(50)	29(35)	16(15)	61(86)	1/15
Sifeg	**1.2**(0.6)	**2.2**(2)	**1.5**(2)	**1.4**(2)	6.8(5)	15/15
Sif	**1.2**(0)	**2.2**(1)	**1.6**(2)	6.4(1)	8.5(11)	15/15
Srr	**1.2**(0.2)	**2.2**(2)	**1.4**(2)	**1.9**(4)	3.6(5)	15/15

Table 19: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*4.0e+2*:1.2	*1.0e+2*:3.2	*4.0e+1*:7.2	*6.3e+0*:32	*1.6e+0*:104	5/5
BSifeg	**1.2**(0)	**1.2**(1)	**1.0**(0.6)	10(2)	133(112)	10/15
BSif	**1.2**(1)	**1.2**(1)	**1.0**(0.7)	12(45)	126(193)	10/15
BSqi	**1.2**(0.6)	**1.2**(0.9)	**1.0**(0.6)	16(37)	119(162)	10/15
BSrr	**1.2**(0.6)	**1.2**(1)	**1.0**(0.5)	13(0.4)	98(172)	12/15
CMA-CSA	**1**(0.4)	**2.0**(4)	**1.8**(2)	**2.7**(6)	3.5(3)	15/15
CMA-MSR	**1.3**(1)	**1.8**(1)	**1.6**(2)	6.1(6)	6.2(6)	15/15
CMA-TPA	**0.94**(0)	**1.5**(0.5)	**1.2**(0.3)	**1.2**(1)	**3.0**(3)	15/15
GP1-CMAES	**2.7**(3)	**2.2**(2)	4.2(9)	**2.0**(5)	4.3(4)	9/15
GP5-CMAES	14(0.6)	26(1)	17(53)	9.5(16)	6.5(10)	8/15
IPOPCMAv3p	**2.4**(8)	**1.9**(0.6)	**1.7**(0.8)	**1.1**(0.9)	**2.7**(10)	11/15
LHD-10xDef	**1.2**(1)	**1.1**(1)	**1.2**(0.8)	**1.8**(1)	**2.3**(2)	6/15
LHD-2xDefa	**1.3**(1)	**1.3**(0.7)	**0.80**(0.5)	**1.0**(0.9)	**1.1**(0.9)	9/15
RAND-2xDef	**1.2**(0.8)	**1.1**(1)	**0.79**(0.8)	**0.96**(0.6)	**1.7**(2)	7/15
RF1-CMAES	**1.2**(0.8)	**1.9**(2)	**1.1**(1)	**1.3**(0.6)	8.0(2)	6/15
RF5-CMAES	**1.6**(2)	3.9(8)	6.4(2)	5.0(12)	9.0(15)	6/15
Sifeg	**1.2**(0.6)	**1.1**(0.8)	**1.0**(0.7)	24(2)	52(98)	14/15
Sif	**1.2**(1)	**1.1**(1)	**1.0**(0.8)	19(1)	119(97)	11/15
Srr	**1.2**(1)	**1.1**(0.8)	**1.0**(0.7)	40(255)	129(97)	10/15

Table 20: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:23	*1.0e-1*:26	*6.3e-2*:38	*4.0e-2*:40	*1.0e-2*:216	15/15
BSifeg	**4.0**(3)	**4.0**(3)	**2.8**(2)	**3.7**(4)	15(10)	15/15
BSif	4.2(3)	4.2(4)	**3.0**(3)	**3.7**(3)	33(77)	12/15
BSqi	**3.6**(2)	**3.7**(2)	**2.6**(2)	4.2(5)	35(31)	14/15
BSrr	**3.5**(2)	**3.6**(2)	**2.6**(3)	**3.7**(4)	28(58)	13/15
CMA-CSA	13(4)	21(45)	18(20)	26(22)	13(8)	15/15
CMA-MSR	8.0(11)	12(10)	16(23)	26(43)	13(14)	15/15
CMA-TPA	7.0(3)	6.2(6)	14(15)	18(23)	**6.2**(8)	15/15
GP1-CMAES	8.8(9)	11(15)	12(15)	12(17)	16(21)	2/15
GP5-CMAES	12(11)	17(20)	13(13)	15(14)	16(9)	2/15
IPOPCMAv3p	7.9(8)	14(24)	11(16)	12(8)	34(37)	1/15
LHD-10xDef	8.9(12)	18(15)	38(99)	36(28)	**6.9**(11)	1/15
LHD-2xDefa	7.5(7)	8.3(18)	12(18)	37(45)	∞ *100*	0/15
RAND-2xDef	5.6(6)	16(17)	11(13)	11(15)	**6.6**(13)	1/15
RF1-CMAES	17(34)	19(11)	17(17)	28(27)	10(4)	3/15
RF5-CMAES	12(31)	12(10)	15(17)	20(40)	∞ *504*	0/15
Sifeg	7.3(2)	6.8(6)	6.5(6)	7.2(5)	31(22)	14/15
Sif	6.5(5)	6.2(4)	6.7(5)	7.6(4)	25(22)	15/15
Srr	5.0(3)	5.1(4)	5.1(3)	6.3(4)	28(57)	14/15

Table 21: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*4.0e+3*:1.9	*2.5e+2*:2.8	*4.0e+0*:6.3	*2.5e+0*:21	*6.3e-1*:139	15/15
BSifeg	**2.1**(2)	14(4)	17(32)	10(22)	28(47)	15/15
BSif	**2.1**(2)	13(3)	19(32)	8.0(3)	62(71)	12/15
BSqi	**2.1**(2)	12(1)	19(24)	8.1(14)	16(40)	15/15
BSrr	**2.1**(2)	14(5)	16(1)	8.3(9)	18(17)	15/15
CMA-CSA	**2.4**(1)	**3.0**(3)	**2.4**(2)	**1.6**(2)	**6.0**(7)	15/15
CMA-MSR	**1.7**(2)	**1.9**(2)	**2.6**(4)	**2.2**(0.9)	13(19)	15/15
CMA-TPA	3.2(3)	4.2(2)	4.0(5)	**2.1**(2)	11(13)	15/15
GP1-CMAES	**1.9**(3)	**2.3**(2)	**1.6**(0.9)	**1.6**(0.8)	**4.3**(4)	8/15
GP5-CMAES	**2.0**(3)	**2.0**(1)	**2.2**(2)	**2.0**(0.3)	8.1(8)	5/15
IPOPCMAv3p	**2.3**(2)	**2.5**(3)	**2.3**(3)	**1.5**(2)	16(21)	3/15
LHD-10xDef	**1.8**(2)	4.0(3)	3.3(2)	**2.1**(2)	**5.1**(4)	2/15
LHD-2xDefa	**1.8**(2)	**2.0**(1)	**2.1**(0.8)	**1.7**(2)	11(14)	1/15
RAND-2xDef	**1.3**(0.7)	**2.8**(2)	**2.5**(1)	**1.9**(0.4)	11(10)	1/15
RF1-CMAES	**2.1**(2)	**2.7**(3)	3.2(4)	3.5(2)	16(11)	3/15
RF5-CMAES	**2.6**(3)	21(8)	27(23)	12(18)	26(33)	2/15
Sifeg	**2.1**(1)	3.6(3)	3.6(0.9)	**1.8**(2)	17(12)	15/15
Sif	**2.1**(2)	3.7(2)	3.7(4)	**1.9**(1)	20(28)	15/15
Srr	**2.1**(2)	3.5(3)	3.0(2)	**1.8**(2)	16(71)	15/15

Table 22: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*1.0e+1*:1.7	*6.3e+0*:2.6	*2.5e+0*:7.9	*1.6e+0*:30	*4.0e-1*:105	15/15
BSifeg	**2.1**(1)	**3.0**(2)	3.0(6)	**2.2**(4)	42(163)	13/15
BSif	**2.1**(2)	**2.9**(2)	**3.0**(2)	**2.9**(5)	62(180)	12/15
BSqi	**2.1**(1)	**2.5**(2)	3.4(8)	**2.4**(4)	53(124)	13/15
BSrr	**2.1**(2)	**2.9**(2)	3.4(4)	**2.3**(4)	62(61)	12/15
CMA-CSA	**1.5**(1)	**2.4**(4)	5.2(13)	4.0(14)	4.5(8)	15/15
CMA-MSR	**1.9**(1)	**2.6**(1)	**1.7**(2)	12(35)	156(985)	14/15
CMA-TPA	**1.5**(2)	**2.7**(4)	**2.5**(2)	5.2(12)	37(165)	15/15
GP1-CMAES	**1.8**(2)	**2.1**(1)	**1.6**(2)	10(12)	6.2(14)	7/15
GP5-CMAES	**1.5**(0.6)	**2.5**(2)	3.8(1)	6.4(6)	8.9(8)	6/15
IPOPCMAv3p	**1.6**(0.9)	**2.0**(3)	**1.9**(0.9)	3.4(6)	3.7(9)	9/15
LHD-10xDef	**1.4**(2)	**1.8**(2)	**1.2**(2)	**1.2**(1)	**0.74**(0.7)	12/15
LHD-2xDefa	**1.4**(2)	**1.7**(2)	**1.3**(0.8)	**0.75**(0.5)	**0.45**(0.5)	14/15
RAND-2xDef	**1.5**(0.9)	**2.8**(3)	**1.5**(0.9)	**0.90**(0.6)	**0.49**(0.2)	14/15
RF1-CMAES	**1.2**(0.6)	**3.0**(4)	7.5(32)	13(26)	20(37)	3/15
RF5-CMAES	**1.7**(2)	**2.1**(3)	**1.2**(0.9)	4.4(4)	11(6)	5/15
Sifeg	**2.1**(1)	**2.3**(2)	**1.3**(0.7)	32(2)	88(96)	12/15
Sif	**2.1**(1)	**2.3**(2)	**1.3**(0.7)	49(4)	98(79)	11/15
Srr	**2.1**(1)	**2.3**(2)	**1.3**(0.8)	21(1)	107(79)	11/15

Table 23: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*4.0e+1*:1.3	*1.6e+1*:3.2	*6.3e+0*:9.3	*1.6e+0*:25	*1.0e-1*:168	15/15
BSifeg	**1.9**(3)	**1.9**(2)	3.9(2)	18(12)	94(77)	9/15
BSif	**1.9**(3)	**1.9**(2)	3.6(7)	37(107)	80(115)	10/15
BSqi	**1.9**(2)	**1.9**(2)	4.8(5)	14(5)	91(161)	9/15
BSrr	**1.9**(2)	**1.9**(2)	5.5(9)	26(38)	85(90)	10/15
CMA-CSA	**1.8**(2)	**1.9**(3)	5.6(2)	17(34)	14(5)	15/15
CMA-MSR	**1.9**(2)	**2.2**(3)	7.3(22)	22(11)	18(52)	15/15
CMA-TPA	**1.7**(1)	**1.4**(1)	**1.9**(2)	14(24)	7.6(5)	15/15
GP1-CMAES	**1.3**(0.4)	**1.1**(1)	10(14)	6.1(15)	6.8(8)	5/15
GP5-CMAES	**1.6**(2)	**1.4**(1)	3.6(14)	4.1(6)	**2.1**(2)	11/15
IPOPCMAv3p	**1.7**(0.8)	**1.3**(1)	**2.3**(4)	4.8(16)	4.2(6)	7/15
LHD-10xDef	**1.3**(1.0)	**1.2**(0.9)	**1.3**(1)	**1.8**(1)	**0.81**(0.7)	9/15
LHD-2xDefa	**1.3**(1)	**1.9**(2)	**1.3**(2)	**1.5**(1)	**1.9**(2)	4/15
RAND-2xDef	**1.3**(0.6)	**1.3**(1)	**1.2**(0.7)	**1.3**(1)	**0.58**(0.8)	10/15
RF1-CMAES	**1.6**(0.8)	**1.3**(2)	6.3(4)	12(15)	12(8)	3/15
RF5-CMAES	**1.6**(1.0)	**1.6**(0.9)	8.5(26)	13(20)	12(11)	3/15
Sifeg	**2.0**(4)	**2.1**(3)	**2.3**(0.5)	8.9(22)	84(96)	10/15
Sif	**2.0**(0)	**2.3**(1)	**2.3**(2)	10(41)	103(119)	9/15
Srr	**2.0**(2)	**2.0**(2)	**2.4**(4)	27(9)	89(149)	11/15

Table 24: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*4.0e+1*:1.5	*2.5e+1*:2.6	*1.0e+1*:7.8	*4.0e+0*:55	*2.5e+0*:103	5/5
BSifeg	**2.0**(3)	**1.6**(2)	**1.9**(1)	**1.3**(2)	**0.91**(0.7)	15/15
BSif	**2.0**(3)	**1.6**(2)	**1.9**(2)	**1.3**(2)	**1.0**(0.9)	15/15
BSqi	**2.0**(2)	**1.6**(2)	**1.8**(2)	**1.6**(1)	**1.3**(1)	15/15
BSrr	**2.0**(3)	**1.6**(1)	**2.1**(1)	**1.5**(1.0)	**1.1**(0.8)	15/15
CMA-CSA	**1.3**(2)	**1.1**(0.6)	3.2(2)	**3.0**(2)	**2.8**(6)	15/15
CMA-MSR	**1.7**(1.0)	**2.3**(3)	**2.2**(3)	6.3(6)	4.8(6)	15/15
CMA-TPA	**1.1**(0.2)	**1.2**(1)	**1.9**(3)	3.2(4)	4.8(2)	15/15
GP1-CMAES	**1.2**(0.8)	**1.6**(2)	**2.2**(3)	3.0(3)	4.4(8)	10/15
GP5-CMAES	**1.7**(2)	**1.7**(1.0)	**2.0**(2)	**2.9**(5)	**2.2**(2)	13/15
IPOPCMAv3p	**1.3**(0.7)	**1.0**(0.6)	**0.99**(2)	**1.8**(1)	**2.8**(3)	13/15
LHD-10xDef	**1.6**(1.0)	**1.1**(0.7)	**1.5**(1.0)	**1.7**(2)	**1.8**(1)	6/15
LHD-2xDefa	**1.8**(1)	**2.1**(3)	**2.4**(3)	**1.6**(2)	4.5(7)	3/15
RAND-2xDef	**1.0**(1.0)	**1.4**(1)	**2.0**(2)	**1.5**(0.9)	4.2(5)	3/15
RF1-CMAES	**1.7**(1)	**1.4**(1)	**1.8**(3)	3.5(4)	**2.4**(1)	14/15
RF5-CMAES	**1.8**(1)	**1.4**(1)	**1.5**(2)	**1.8**(3)	**1.9**(2)	14/15
Sifeg	**1.9**(3)	**1.5**(1)	**1.9**(1)	**1.9**(2)	**2.2**(2)	15/15
Sif	**1.9**(3)	**1.5**(2)	**1.8**(2)	**1.8**(2)	**2.2**(2)	15/15
Srr	**1.9**(2)	**1.5**(1)	**1.9**(1)	**1.9**(2)	**2.3**(3)	15/15

Table 25: 02-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*4.0e+1*:1.1	*2.5e+1*:2.7	*1.6e+1*:7.7	*6.3e+0*:44	*2.5e+0*:275	5/5
BSifeg	**2.1**(3)	**2.3**(2)	**1.8**(3)	**2.4**(1)	6.3(1.0)	14/15
BSif	**2.1**(4)	**2.3**(2)	**1.6**(2)	**2.3**(2)	8.3(28)	15/15
BSqi	**2.1**(0.9)	**2.3**(0.8)	**2.9**(0.5)	**2.9**(3)	11(29)	14/15
BSrr	**2.1**(4)	**2.3**(2)	**1.9**(6)	**2.6**(2)	6.4(18)	14/15
CMA-CSA	**2.3**(3)	**2.2**(2)	**1.8**(1)	**1.2**(1)	3.3(6)	15/15
CMA-MSR	3.1(4)	**2.4**(2)	**1.6**(2)	3.8(0.6)	5.6(7)	15/15
CMA-TPA	**1.6**(0.5)	**2.2**(3)	**1.3**(1)	**2.3**(5)	4.1(3)	15/15
GP1-CMAES	**1.6**(1)	**1.7**(2)	**1.9**(3)	**1.8**(6)	**3.3**(2)	7/15
GP5-CMAES	**1.8**(0)	**1.9**(2)	**1.4**(3)	3.4(6)	**2.9**(4)	7/15
IPOPCMAv3p	**1.6**(0.9)	**2.1**(4)	**1.5**(1)	**1.2**(1)	3.9(6)	6/15
LHD-10xDef	**1.9**(1)	**1.4**(1)	**1.5**(2)	**1.5**(1)	∞ *100*	0/15
LHD-2xDefa	**1.8**(1)	**1.6**(0.8)	**2.0**(3)	**2.6**(4)	∞ *100*	0/15
RAND-2xDef	**1.2**(0.9)	**2.4**(2)	**1.8**(1)	**1.9**(3)	**2.5**(4)	2/15
RF1-CMAES	**1.4**(0.2)	**1.8**(3)	**1.3**(1)	**0.80**(0.6)	5.3(7)	4/15
RF5-CMAES	**1.8**(2)	**1.2**(0.5)	**1.2**(1)	**2.5**(3)	13(20)	2/15
Sifeg	**2.1**(3)	**2.4**(2)	**1.4**(0.6)	**2.3**(0.5)	4.9(2)	15/15
Sif	**2.1**(3)	**2.4**(3)	**1.4**(0.6)	**2.1**(2)	5.1(13)	15/15
Srr	**2.1**(2)	**2.4**(2)	**1.4**(0.8)	**2.1**(1)	7.4(4)	15/15

Table 26: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*1.6e+1*:3.0	*1.0e+1*:3.6	*1.0e-8*:8.0	*1.0e-8*:8.0	*1.0e-8*:8.0	15/15
BSifeg	**1.7**(2)	**1.5**(1)	**2.1**(0.3)	**2.1**(0.3)	**2.1**(0.3)	15/15
BSif	**1.7**(2)	**1.5**(2)	**2.1**(0.3)	**2.1**(0.3)	**2.1**(0.2)	15/15
BSqi	**1.7**(1)	**1.5**(1)	**2.1**(0.3)	**2.1**(0.2)	**2.1**(0.3)	15/15
BSrr	**1.7**(1)	**1.5**(1)	**2.1**(0.2)	**2.1**(0.2)	**2.1**(0.3)	15/15
CMA-CSA	5.4(2)	5.6(5)	52(3)	52(4)	52(4)	15/15
CMA-MSR	**2.1**(2)	**2.4**(3)	83(8)	83(9)	83(12)	15/15
CMA-TPA	**2.9**(2)	3.4(3)	55(20)	55(14)	55(18)	15/15
GP1-CMAES	**2.4**(2)	3.2(2)	31(3)	31(5)	31(5)	15/15
GP5-CMAES	**2.2**(2)	**2.8**(2)	46(36)	46(29)	46(36)	15/15
IPOPCMAv3p	**2.7**(4)	**2.8**(2)	52(7)	52(8)	52(6)	15/15
LHD-10xDef	**2.1**(2)	3.6(4)	∞	∞	∞ *150*	0/15
LHD-2xDefa	**2.1**(2)	**2.2**(2)	∞	∞	∞ *150*	0/15
RAND-2xDef	**1.8**(2)	**2.3**(2)	∞	∞	∞ *150*	0/15
RF1-CMAES	**2.4**(2)	**2.3**(1)	133(108)	133(106)	133(74)	9/15
RF5-CMAES	12(41)	11(2)	∞	∞	∞ *753*	0/15
Sifeg	**1.7**(2)	**1.5**(1)	7.5(0.7)	7.5(1)	7.5(1)	15/15
Sif	**1.7**(2)	**1.5**(1.0)	7.5(1)	7.5(0.7)	7.5(0.8)	15/15
Srr	**1.7**(1)	**1.5**(1)	7.2(0.6)	7.2(0.8)	7.2(0.9)	15/15

Table 27: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*6.3e+6*:1.5	*6.3e+5*:4.3	*4.0e+4*:10	*1.0e+2*:32	*1.0e-8*:49	15/15
BSifeg	3.7(4)	**2.8**(0.4)	**1.4**(0.2)	**0.66**(0.3)	**1.2**(0.3)	15/15
BSif	3.7(4)	**2.8**(0.9)	**1.4**(0.5)	**0.68**(0.2)	**1.2**(0.1)	15/15
BSqi	3.7(4)	**2.8**(0.4)	**1.4**(0.2)	**0.62**(0.1)	**1.1**(0.3)	15/15
BSrr	3.7(4)	**2.8**(1)	**1.4**(0.5)	**0.66**(0.0)	**1.3**(0.2)	15/15
CMA-CSA	**1.5**(0.3)	**1.4**(0.7)	3.5(4)	7.3(2)	18(2)	15/15
CMA-MSR	**2.3**(1.0)	**1.3**(0.5)	3.1(5)	6.9(2)	22(3)	15/15
CMA-TPA	**2.2**(1.0)	**2.0**(3)	3.3(5)	7.2(2)	19(3)	15/15
GP1-CMAES	3.1(3)	**2.5**(2)	**2.4**(1)	5.4(1)	45(38)	5/15
GP5-CMAES	3.1(3)	**1.9**(0.9)	**2.3**(1)	**2.5**(1.0)	21(28)	8/15
IPOPCMAv3p	**3.0**(4)	**2.3**(2)	**2.4**(3)	8.9(6)	∞ *751*	0/15
LHD-10xDef	**1.0**(0.3)	**1.2**(1)	**2.7**(3)	70(92)	∞ *150*	0/15
LHD-2xDefa	**1.2**(1.0)	**1.2**(1)	**1.3**(1)	70(89)	∞ *150*	0/15
RAND-2xDef	**1.2**(1.0)	**1.0**(1)	**2.3**(1)	16(16)	∞ *150*	0/15
RF1-CMAES	**2.2**(2)	**1.3**(1)	15(2)	60(69)	∞ *751*	0/15
RF5-CMAES	**1.9**(3)	**1.4**(1)	7.0(1)	167(184)	∞ *760*	0/15
Sifeg	3.7(4)	**2.9**(2)	**1.9**(0.8)	**1.1**(0.1)	**1.7**(0.3)	15/15
Sif	3.7(4)	**2.9**(0.4)	**1.9**(0.8)	**1.2**(0.4)	**1.6**(0.3)	15/15
Srr	3.7(4)	**2.9**(0.2)	**1.9**(1)	**1.0**(0.1)	**1.7**(0.1)	15/15

Table 28: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*1.0e+2*:2.2	*6.3e+1*:6.1	*4.0e+1*:10	*1.6e+1*:32	*4.0e+0*:319	15/15
BSifeg	3.1(3)	**1.6**(1.0)	**1.3**(0.2)	**0.64**(0.3)	**0.18**(0.1)	15/15
BSif	3.1(3)	**1.6**(1.0)	**1.3**(0.2)	**0.64**(0.4)	**0.19**(0.1)	15/15
BSqi	3.1(2)	**1.6**(0.9)	**1.3**(0.2)	**0.62**(0.2)	**0.19**(0.1)	15/15
BSrr	3.1(2)	**1.6**(0.9)	**1.3**(0.2)	**0.60**(0.2)	**0.19**(0.1)	15/15
CMA-CSA	3.5(4)	**2.1**(2)	**2.5**(2)	3.2(2)	3.3(3)	15/15
CMA-MSR	3.0(3)	**1.8**(2)	**2.5**(2)	5.2(2)	**2.7**(3)	15/15
CMA-TPA	**3.0**(0.9)	**1.8**(2)	**2.3**(2)	**2.3**(2)	**2.6**(2)	15/15
GP1-CMAES	**2.0**(3)	**1.6**(2)	**1.9**(2)	3.7(2)	**1.1**(0.9)	13/15
GP5-CMAES	**1.9**(2)	**1.4**(1)	**1.5**(0.9)	**1.4**(0.6)	**1.6**(2)	11/15
IPOPCMAv3p	**2.8**(3)	**2.3**(3)	3.5(3)	**2.9**(3)	**1.9**(3)	10/15
LHD-10xDef	**2.9**(2)	**1.7**(1)	4.1(2)	4.0(0.3)	3.5(4)	2/15
LHD-2xDefa	**1.6**(0.9)	**1.3**(1)	**1.4**(1)	**1.7**(0.5)	**1.2**(1)	5/15
RAND-2xDef	**1.7**(1)	**1.1**(1)	**1.6**(1.0)	**1.6**(1.0)	**1.2**(1)	5/15
RF1-CMAES	**2.7**(2)	**2.6**(2)	**2.3**(0.6)	5.9(12)	6.8(13)	4/15
RF5-CMAES	**2.4**(1)	8.2(1)	9.2(28)	14(26)	11(18)	3/15
Sifeg	3.1(3)	**1.6**(0.6)	**1.4**(0.3)	**0.72**(0.1)	**0.21**(0.0)	15/15
Sif	3.1(2)	**1.6**(0.2)	**1.4**(0.3)	**0.72**(0.3)	**0.22**(0.1)	15/15
Srr	3.1(3)	**1.6**(1)	**1.4**(0.3)	**0.69**(0.5)	**0.20**(0.0)	15/15

Table 29: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*1.0e+2*:5.4	*6.3e+1*:10	*6.3e+1*:10	*2.5e+1*:36	*4.0e+0*:617	15/15
BSifeg	**2.3**(1)	**2.0**(2)	**2.0**(2)	**1.1**(0.6)	**0.16**(0.1)	15/15
BSif	**2.2**(3)	**2.1**(2)	**2.1**(2)	**1.1**(0.6)	**0.16**(0.1)	15/15
BSqi	**2.2**(3)	**2.0**(2)	**2.0**(1)	**1.1**(0.5)	**0.17**(0.0)	15/15
BSrr	**2.3**(3)	**2.0**(1)	**2.0**(2)	**1.1**(0.6)	**0.16**(0.1)	15/15
CMA-CSA	**1.9**(3)	**1.7**(2)	**1.7**(0.8)	**1.7**(1)	3.9(5)	15/15
CMA-MSR	**3.0**(0.6)	**2.1**(2)	**2.1**(1)	**2.4**(0.9)	7.0(7)	15/15
CMA-TPA	**1.9**(2)	**2.1**(2)	**2.1**(2)	**2.7**(2)	4.0(4)	15/15
GP1-CMAES	**2.5**(3)	**2.3**(1)	**2.3**(2)	**2.0**(0.6)	8.3(10)	2/15
GP5-CMAES	3.1(0.9)	6.0(2)	6.0(5)	4.6(11)	5.5(5)	3/15
IPOPCMAv3p	**1.9**(1)	**2.5**(2)	**2.5**(2)	3.5(1.0)	8.3(13)	2/15
LHD-10xDef	**1.0**(0.9)	**2.1**(1)	**2.1**(4)	**2.9**(2)	∞ *150*	0/15
LHD-2xDefa	**1.2**(2)	**1.4**(2)	**1.4**(2)	**2.5**(3)	3.6(5)	1/15
RAND-2xDef	**1.2**(0.7)	**2.1**(2)	**2.1**(1)	**2.1**(3)	∞ *150*	0/15
RF1-CMAES	**1.4**(2)	3.2(3)	3.2(3)	33(16)	∞ *751*	0/15
RF5-CMAES	19(0.3)	16(15)	16(28)	37(26)	∞ *753*	0/15
Sifeg	**2.1**(2)	**1.7**(0.9)	**1.7**(0.5)	**0.81**(0.5)	**0.14**(0.0)	15/15
Sif	**2.1**(2)	**1.8**(0.5)	**1.8**(1)	**0.85**(0.2)	**0.14**(0.1)	15/15
Srr	**2.1**(2)	**1.7**(1)	**1.7**(1)	**0.80**(0.3)	**0.14**(0.1)	15/15

Table 30: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*4.0e+1*:2.2	*2.5e+1*:4.8	*1.0e-8*:6.6	*1.0e-8*:6.6	*1.0e-8*:6.6	15/15
BSifeg	3.5(1)	**1.9**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
BSif	3.5(0.5)	**1.9**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
BSqi	3.5(0.9)	**1.9**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
BSrr	3.5(0.7)	**1.9**(0.2)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
CMA-CSA	3.1(3)	**2.4**(5)	5.5(2)	5.5(4)	5.5(3)	15/15
CMA-MSR	**3.0**(2)	**2.0**(1)	4.9(2)	4.9(3)	4.9(3)	15/15
CMA-TPA	**1.8**(2)	**1.9**(1)	3.9(2)	3.9(2)	3.9(2)	15/15
GP1-CMAES	3.7(3)	**2.2**(1)	24(56)	24(6)	24(28)	14/15
GP5-CMAES	3.2(2)	**2.4**(1)	4.7(3)	4.7(3)	4.7(4)	15/15
IPOPCMAv3p	4.3(4)	3.1(4)	10(13)	10(16)	10(13)	15/15
LHD-10xDef	**1.5**(2)	**2.2**(2)	13(0.4)	13(0.4)	13(0.4)	15/15
LHD-2xDefa	**1.8**(1)	**2.0**(2)	3.1(0.8)	3.1(0.4)	3.1(0.4)	15/15
RAND-2xDef	**2.2**(0.9)	**2.2**(1)	3.1(0.4)	3.1(0.4)	3.1(0.4)	15/15
RF1-CMAES	3.7(5)	**2.2**(2)	19(41)	19(22)	19(40)	15/15
RF5-CMAES	3.8(2)	11(32)	150(114)	150(263)	150(97)	7/15
Sifeg	3.5(0.6)	**1.9**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.1)	15/15
Sif	3.5(0.7)	**1.9**(0.2)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15
Srr	3.5(0.5)	**1.9**(0.3)	**1.4**(0.2)	**1.4**(0.2)	**1.4**(0.2)	15/15

Table 31: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*6.3e+4*:1.8	*6.3e+3*:3.7	*4.0e+1*:13	*1.0e+1*:34	*6.3e-4*:159	15/15
BSifeg	**1.9**(3)	**1.2**(1)	79(273)	146(132)	1254(864)	2/15
BSif	**1.9**(2)	**1.2**(2)	99(184)	209(699)	∞ *3e4*	0/15
BSqi	**1.9**(2)	**1.2**(0.7)	30(19)	231(9)	853(961)	3/15
BSrr	**1.9**(2)	**1.2**(1)	119(211)	114(178)	1324(805)	2/15
CMA-CSA	**1.7**(3)	**1.3**(0.3)	**1.6**(0.5)	**1.5**(2)	**2.6**(0.7)	15/15
CMA-MSR	**2.3**(6)	**2.1**(10)	**2.1**(2)	**2.8**(1)	3.6(0.5)	15/15
CMA-TPA	**2.4**(4)	**2.2**(2)	3.0(2)	3.1(0.7)	**3.0**(0.9)	15/15
GP1-CMAES	3.7(6)	**2.9**(3)	**1.6**(1)	**2.7**(3)	∞ *751*	0/15
GP5-CMAES	3.7(6)	3.3(4)	3.4(4)	**2.5**(2)	∞ *760*	0/15
IPOPCMAv3p	4.0(2)	3.8(3)	**2.8**(4)	**2.8**(2)	**3.2**(0.7)	15/15
LHD-10xDef	**0.81**(0.8)	**1.6**(0.5)	**2.0**(1)	4.2(6)	∞ *150*	0/15
LHD-2xDefa	**0.85**(0.3)	**0.85**(0.5)	**1.4**(2)	4.8(6)	∞ *150*	0/15
RAND-2xDef	**0.96**(0.8)	**1.1**(0.5)	**2.6**(3)	3.1(3)	∞ *150*	0/15
RF1-CMAES	**2.7**(4)	**2.3**(2)	**2.7**(1)	13(33)	∞ *751*	0/15
RF5-CMAES	3.2(2)	34(103)	23(48)	42(54)	∞ *760*	0/15
Sifeg	**1.9**(2)	**1.2**(1)	9.4(11)	81(244)	798(891)	3/15
Sif	**1.9**(3)	**1.2**(1)	25(75)	106(430)	∞ *3e4*	0/15
Srr	**1.9**(2)	**1.2**(1)	35(25)	75(104)	543(577)	4/15

Table 32: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*2.5e+2*:1.5	*6.3e+1*:4.2	*1.0e+1*:11	*2.5e+0*:38	*4.0e-1*:174	15/15
BSifeg	**1.7**(1.0)	34(241)	49(133)	142(494)	119(208)	10/15
BSif	**1.7**(2)	34(241)	61(45)	235(400)	233(252)	7/15
BSqi	**1.7**(3)	34(1)	167(178)	215(402)	131(196)	10/15
BSrr	**1.7**(2)	34(121)	131(443)	325(814)	158(233)	9/15
CMA-CSA	**1.9**(1)	**2.4**(2)	**2.8**(2)	**1.9**(1)	**0.89**(1)	15/15
CMA-MSR	**2.2**(2)	**1.8**(2)	3.5(3)	**2.1**(2)	**1.3**(1)	15/15
CMA-TPA	**3.0**(2)	**1.8**(2)	3.8(6)	**2.7**(1)	**1.4**(1)	15/15
GP1-CMAES	**1.4**(0.2)	**1.3**(0.7)	**2.0**(2)	**1.5**(2)	**0.56**(0.6)	15/15
GP5-CMAES	**2.1**(3)	**2.3**(2)	**2.0**(1)	**1.1**(0.8)	**0.63**(1)	15/15
IPOPCMAv3p	**2.8**(2)	3.9(3)	4.9(5)	**2.8**(3)	**1.2**(0.7)	15/15
LHD-10xDef	**1.2**(0.7)	**1.8**(2)	3.7(4)	**2.1**(1)	**1.9**(2)	6/15
LHD-2xDefa	**1.8**(1.0)	**1.8**(2)	**1.9**(1)	**1.5**(2)	**0.98**(0.8)	9/15
RAND-2xDef	**1.7**(1)	**1.2**(1)	**2.4**(2)	**1.6**(2)	**0.90**(0.9)	10/15
RF1-CMAES	**2.0**(2)	3.3(1.0)	7.5(7)	3.8(6)	**2.1**(5)	12/15
RF5-CMAES	**2.5**(3)	**2.7**(2)	10(24)	20(11)	31(36)	2/15
Sifeg	**1.7**(0.5)	**1.9**(2)	37(45)	205(380)	122(76)	11/15
Sif	**1.7**(2)	**2.0**(2)	61(0.6)	234(188)	131(277)	10/15
Srr	**1.7**(3)	**1.9**(2)	97(0.7)	297(543)	154(280)	9/15

Table 33: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.0e+4*:1.8	*1.6e+3*:4.0	*1.0e+2*:15	*6.3e+0*:31	*1.0e-1*:152	15/15
BSifeg	**1.7**(2)	**2.4**(1)	6.6(14)	7.3(14)	398(343)	6/15
BSif	**1.7**(1)	**2.4**(1)	6.0(10)	6.1(11)	266(223)	7/15
BSqi	**1.7**(1)	**2.4**(1)	4.8(10)	6.7(11)	789(1355)	3/15
BSrr	**1.7**(3)	**2.4**(2)	4.2(12)	4.5(5)	595(363)	4/15
CMA-CSA	4.8(7)	4.1(5)	3.2(2)	3.1(1)	**3.3**(1)	15/15
CMA-MSR	3.4(2)	3.2(4)	3.4(2)	4.0(3)	4.5(1)	15/15
CMA-TPA	**2.6**(3)	**2.5**(4)	**2.9**(2)	3.5(2)	**3.7**(1.0)	15/15
GP1-CMAES	3.2(4)	**2.6**(3)	**2.7**(2)	**2.9**(2)	13(8)	5/15
GP5-CMAES	**2.9**(2)	**2.7**(3)	**2.1**(1.0)	3.0(7)	**4.4**(4)	10/15
IPOPCMAv3p	3.9(3)	3.3(2)	3.2(1)	3.2(1)	5.5(3)	10/15
LHD-10xDef	**2.3**(1)	**2.9**(2)	5.0(3)	10(8)	∞ *150*	0/15
LHD-2xDefa	**1.5**(0.8)	**2.0**(2)	**1.8**(0.5)	3.3(3)	∞ *150*	0/15
RAND-2xDef	**2.0**(1)	**2.2**(2)	**1.9**(1)	3.0(2)	∞ *150*	0/15
RF1-CMAES	**2.0**(0.8)	**2.7**(3)	4.0(9)	13(20)	∞ *751*	0/15
RF5-CMAES	**2.6**(4)	**1.9**(4)	19(22)	53(52)	∞ *753*	0/15
Sifeg	**1.7**(2)	**2.4**(0.8)	**1.4**(1)	**1.8**(2)	123(135)	11/15
Sif	**1.7**(1)	**2.4**(0.8)	**1.3**(0.2)	**2.1**(3)	261(107)	7/15
Srr	**1.7**(1)	**2.4**(1)	**1.3**(2)	**1.7**(2)	211(359)	8/15

Table 34: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*1.0e+1*:21	*6.3e+0*:25	*4.0e+0*:32	*2.5e+0*:48	*6.3e-3*:152	15/15
BSifeg	14(30)	12(25)	48(152)	107(222)	∞ *3e4*	0/15
BSif	12(21)	11(6)	39(15)	179(220)	∞ *3e4*	0/15
BSqi	8.5(14)	7.7(12)	50(13)	88(15)	1267(1585)	2/15
BSrr	8.0(15)	8.2(15)	41(9)	75(17)	∞ *3e4*	0/15
CMA-CSA	3.8(0.9)	3.7(1)	3.2(2)	**3.2**(4)	**4.3**(3)	15/15
CMA-MSR	5.8(3)	6.9(13)	7.0(7)	7.3(7)	**5.9**(3)	15/15
CMA-TPA	4.3(2)	4.6(2)	4.9(1)	6.0(6)	**5.4**(2)	15/15
GP1-CMAES	3.5(1)	4.0(6)	5.5(1)	7.7(8)	35(40)	2/15
GP5-CMAES	**2.4**(1.0)	**2.5**(1)	**3.2**(1)	4.0(0.9)	12(22)	5/15
IPOPCMAv3p	3.5(3)	3.5(2)	**3.2**(2)	**3.3**(4)	8.2(6)	8/15
LHD-10xDef	10(12)	16(14)	16(11)	22(28)	∞ *150*	0/15
LHD-2xDefa	**2.5**(2)	**2.9**(2)	3.3(2)	5.8(3)	∞ *150*	0/15
RAND-2xDef	**2.6**(1)	3.3(3)	**3.0**(1)	**4.0**(4)	∞ *150*	0/15
RF1-CMAES	8.1(20)	7.8(16)	9.1(24)	12(17)	∞ *751*	0/15
RF5-CMAES	37(38)	34(52)	33(34)	32(31)	∞ *753*	0/15
Sifeg	**1.9**(0.8)	**1.8**(2)	63(76)	73(209)	∞ *3e4*	0/15
Sif	**1.8**(1)	**1.7**(2)	40(0.8)	87(211)	∞ *3e4*	0/15
Srr	**1.5**(0.4)	**1.4**(1)	101(184)	135(249)	∞ *2e4*	0/15

Table 35: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*6.3e+6*:1.7	*1.6e+5*:4.4	*4.0e+4*:12	*4.0e+2*:37	*1.0e+0*:152	15/15
BSifeg	**1.0**(0.3)	**2.5**(1)	4.7(13)	411(417)	1211(992)	1/15
BSif	**1.0**(0.3)	**2.5**(1)	5.7(0.2)	578(334)	1253(1696)	1/15
BSqi	**1.0**(0.3)	**2.5**(1)	6.9(0.2)	304(351)	725(915)	2/15
BSrr	**1.0**(0.3)	**2.5**(1)	5.4(0.3)	691(454)	1155(1570)	1/15
CMA-CSA	**1.4**(2)	**2.3**(2)	**2.8**(2)	4.8(2)	**3.0**(1.0)	15/15
CMA-MSR	**1.4**(2)	**3.0**(1)	**2.0**(2)	**4.0**(2)	3.4(1)	15/15
CMA-TPA	**1.2**(0.9)	3.2(3)	**1.9**(1)	**3.2**(2)	3.1(0.9)	15/15
GP1-CMAES	**2.6**(0.6)	**2.7**(3)	**2.4**(2)	4.3(2)	**2.8**(1)	15/15
GP5-CMAES	**1.3**(0.4)	**1.8**(3)	**1.6**(1)	**2.3**(1)	**1.1**(0.4)	15/15
IPOPCMAv3p	**2.2**(3)	**2.7**(4)	**2.0**(1)	5.8(5)	4.4(6)	13/15
LHD-10xDef	**1.4**(1)	**2.4**(2)	**2.9**(3)	14(16)	∞ *150*	0/15
LHD-2xDefa	**1.2**(0.6)	3.0(2)	**1.8**(1)	11(11)	∞ *150*	0/15
RAND-2xDef	**1.3**(2)	**3.0**(3)	**2.1**(1)	8.4(12)	∞ *150*	0/15
RF1-CMAES	**1.5**(2)	**2.4**(2)	7.8(32)	34(55)	∞ *751*	0/15
RF5-CMAES	**2.0**(2)	3.2(3)	**1.5**(1)	49(79)	∞ *753*	0/15
Sifeg	**1.0**(1)	**2.7**(1)	**1.5**(0.3)	173(211)	467(452)	1/15
Sif	**1.0**(1)	**2.7**(1)	**1.5**(0.4)	109(121)	482(552)	1/15
Srr	**1.0**(0.3)	**2.7**(0.8)	**1.5**(0.4)	105(181)	194(326)	2/15

Table 36: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*2.5e+6*:1.9	*4.0e+5*:4.5	*6.3e+4*:9.4	*2.5e+1*:36	*2.5e-1*:174	15/15
BSifeg	**1.6**(1)	**2.0**(1)	**1.2**(0.7)	138(224)	1140(884)	1/15
BSif	**1.6**(1)	**2.0**(1.0)	**1.2**(0.7)	168(241)	365(243)	3/15
BSqi	**1.6**(3)	**2.0**(1)	**1.2**(0.2)	141(245)	∞ *2e4*	0/15
BSrr	**1.6**(1)	**2.0**(1)	**1.2**(0.2)	126(274)	1061(1784)	1/15
CMA-CSA	**1.0**(0.8)	**1.1**(1)	**1.3**(2)	**6.2**(4)	3.0(0.8)	15/15
CMA-MSR	**2.0**(2)	**2.6**(2)	**1.8**(2)	10(5)	**3.0**(0.5)	15/15
CMA-TPA	**1.5**(2)	**1.5**(1)	**1.5**(1)	7.8(5)	3.1(0.9)	15/15
GP1-CMAES	**2.5**(3)	**2.1**(2)	**1.7**(1)	**6.7**(6)	**2.8**(0.6)	15/15
GP5-CMAES	**2.3**(2)	**1.9**(1)	**1.9**(0.6)	**2.7**(1)	**1.2**(0.2)	15/15
IPOPCMAv3p	**1.8**(2)	**2.3**(2)	**1.9**(2)	11(8)	13(9)	5/15
LHD-10xDef	**1.9**(1)	**2.4**(4)	**2.4**(3)	30(44)	∞ *150*	0/15
LHD-2xDefa	**1.8**(2)	**2.1**(2)	**1.5**(1)	11(16)	∞ *150*	0/15
RAND-2xDef	**1.4**(1)	**1.4**(1)	**1.4**(0.9)	11(12)	∞ *150*	0/15
RF1-CMAES	**1.6**(2)	**1.3**(1)	**1.3**(0.9)	34(48)	61(90)	1/15
RF5-CMAES	**2.3**(2)	**1.9**(1)	**1.9**(1)	15(25)	∞ *753*	0/15
Sifeg	**1.7**(2)	**2.0**(1)	**1.4**(0.9)	50(104)	∞ *7533*	0/15
Sif	**1.7**(1)	**2.0**(1)	**1.4**(0.3)	66(123)	∞ *7579*	0/15
Srr	**1.7**(4)	**2.0**(1)	**1.4**(0.5)	60(177)	189(350)	3/15

Table 37: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*1.0e+8*:1.5	*1.0e+7*:3.6	*6.3e+5*:13	*6.3e+2*:31	*1.0e+0*:168	15/15
BSifeg	**1.1**(0.7)	**1.3**(1.0)	**1.1**(0.2)	10(7)	48(52)	11/15
BSif	**1.1**(1.0)	**1.3**(1)	**1.1**(0.3)	8.6(17)	80(72)	9/15
BSqi	**1.1**(2)	**1.3**(2)	**1.1**(0.4)	43(0.1)	71(83)	11/15
BSrr	**1.1**(0.7)	**1.3**(0.6)	**1.1**(0.4)	27(8)	46(128)	10/15
CMA-CSA	**1.2**(2)	**2.2**(1)	**1.6**(1)	5.1(2)	5.1(3)	15/15
CMA-MSR	**1.7**(2)	**1.9**(2)	3.2(2)	7.1(2)	6.7(14)	15/15
CMA-TPA	**1.7**(2)	**2.5**(2)	3.1(2)	5.6(2)	**5.0**(7)	15/15
GP1-CMAES	**1.0**(1)	**1.7**(2)	**2.3**(2)	**4.2**(3)	**4.3**(3)	10/15
GP5-CMAES	**1.2**(0.3)	**2.5**(2)	**2.3**(1)	5.6(1)	**4.9**(10)	8/15
IPOPCMAv3p	**1.4**(1)	**2.6**(3)	**2.9**(3)	5.3(2)	5.9(6)	8/15
LHD-10xDef	**0.96**(1.0)	**1.6**(2)	4.2(3)	∞	∞ *150*	0/15
LHD-2xDefa	**1.1**(0.7)	**1.8**(2)	**1.6**(0.8)	4.7(3)	13(12)	1/15
RAND-2xDef	**1**(0.3)	**1.5**(0.9)	**1.6**(0.9)	4.2(5)	∞ *150*	0/15
RF1-CMAES	**1.0**(3)	**1.6**(2)	**2.4**(2)	5.1(2)	20(20)	3/15
RF5-CMAES	**1.2**(0.5)	10(34)	18(26)	65(104)	∞ *753*	0/15
Sifeg	**1.1**(0)	**1.3**(2)	**1.2**(0.4)	**2.9**(0.6)	20(18)	9/15
Sif	**1.1**(1)	**1.3**(0.9)	**1.4**(0.3)	5.8(13)	20(26)	9/15
Srr	**1.1**(0)	**1.3**(1.0)	**1.3**(0.3)	**3.4**(11)	19(33)	10/15

Table 38: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*1.0e+3*:1.6	*4.0e+2*:6.8	*2.5e+2*:11	*4.0e+1*:30	*2.5e-3*:182	15/15
BSifeg	**1.7**(1)	**1.6**(0.7)	**1.4**(0.9)	89(513)	∞ *3e4*	0/15
BSif	**1.7**(2)	**1.6**(1)	**1.4**(0.5)	59(19)	∞ *3e4*	0/15
BSqi	**1.7**(2)	**1.6**(0.7)	**1.3**(0.3)	26(68)	∞ *3e4*	0/15
BSrr	**1.7**(3)	**1.6**(0.8)	**1.4**(0.7)	50(165)	∞ *2e4*	0/15
CMA-CSA	**3.0**(4)	**2.2**(3)	**1.8**(2)	3.7(2)	**3.4**(0.5)	15/15
CMA-MSR	**2.4**(3)	**1.8**(2)	**2.0**(1)	4.2(2)	3.9(0.6)	15/15
CMA-TPA	**1.8**(0.9)	**1.8**(2)	**2.6**(5)	3.5(2)	**3.9**(0.7)	15/15
GP1-CMAES	**0.75**(0.3)	**1.7**(1)	**1.5**(0.8)	**2.0**(0.9)	19(16)	3/15
GP5-CMAES	**2.0**(2)	**1.4**(2)	**1.2**(0.9)	**1.4**(0.4)	**3.2**(3)	11/15
IPOPCMAv3p	**1.6**(3)	**1.7**(2)	**1.9**(0.8)	3.4(2)	12(7)	5/15
LHD-10xDef	**1.6**(4)	**1.5**(0.6)	**2.9**(3)	3.1(0.7)	∞ *150*	0/15
LHD-2xDefa	**1.2**(0.8)	**1.1**(0.8)	**1.3**(0.8)	**1.3**(0.6)	∞ *150*	0/15
RAND-2xDef	**1.4**(0.9)	**0.99**(1)	**1.1**(0.9)	**1.5**(2)	∞ *150*	0/15
RF1-CMAES	**1.9**(2)	**1.6**(1)	**1.8**(2)	4.2(0.7)	∞ *751*	0/15
RF5-CMAES	**1.9**(2)	**2.0**(2)	**2.5**(5)	21(11)	∞ *760*	0/15
Sifeg	**1.7**(3)	**1.5**(1)	**1.3**(0.3)	169(270)	∞ *2e4*	0/15
Sif	**1.7**(1)	**1.5**(1)	**1.3**(0.4)	76(0.9)	∞ *2e4*	0/15
Srr	**1.7**(1)	**1.5**(0.9)	**1.3**(0.2)	58(45)	∞ *2e4*	0/15

Table 39: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*1.0e+1*:2.2	*6.3e+0*:4.2	*2.5e+0*:10	*6.3e-2*:31	*2.5e-6*:160	15/15
BSifeg	**1.8**(1)	**2.1**(3)	**2.7**(1)	5.0(8)	∞ *3e4*	0/15
BSif	**1.8**(2)	**2.2**(1)	4.3(10)	5.5(7)	∞ *3e4*	0/15
BSqi	**1.8**(2)	**2.3**(2)	**2.4**(0.9)	4.2(6)	∞ *3e4*	0/15
BSrr	**1.8**(2)	**1.9**(2)	**2.5**(1)	5.6(9)	∞ *3e4*	0/15
CMA-CSA	3.8(3)	3.1(1)	**2.3**(2)	4.1(0.7)	**4.0**(0.5)	15/15
CMA-MSR	**2.5**(2)	**1.8**(2)	**2.4**(2)	5.3(2)	**4.4**(0.9)	15/15
CMA-TPA	4.4(5)	3.7(3)	3.4(2)	4.5(2)	**3.8**(0.5)	15/15
GP1-CMAES	3.9(5)	**2.7**(3)	**2.2**(1)	**2.7**(0.8)	∞ *751*	0/15
GP5-CMAES	3.3(4)	**2.3**(2)	**2.0**(1)	**1.9**(0.6)	68(40)	1/15
IPOPCMAv3p	**2.2**(3)	**2.3**(4)	**2.7**(2)	4.0(1)	23(38)	3/15
LHD-10xDef	**1.5**(2)	**1.4**(3)	**2.0**(2)	4.0(0.4)	∞ *150*	0/15
LHD-2xDefa	**1.9**(1)	**1.6**(1)	**1.6**(0.9)	**1.8**(0.9)	∞ *150*	0/15
RAND-2xDef	**2.3**(2)	**1.5**(0.7)	**1.5**(1)	**2.3**(1.0)	∞ *150*	0/15
RF1-CMAES	**2.9**(5)	**2.7**(4)	4.7(8)	16(30)	∞ *751*	0/15
RF5-CMAES	**1.9**(2)	**1.1**(1.0)	20(57)	47(43)	∞ *753*	0/15
Sifeg	**1.8**(2)	**1.7**(1)	**2.0**(1)	**2.6**(2)	∞ *3e4*	0/15
Sif	**1.8**(2)	**1.7**(2)	**2.0**(2)	**3.0**(3)	∞ *3e4*	0/15
Srr	**1.8**(2)	**1.7**(1)	**1.9**(1)	**1.9**(1)	∞ *3e4*	0/15

Table 40: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*1.6e+2*:1.6	*6.3e+1*:5.6	*4.0e+1*:12	*1.6e+1*:68	*6.3e+0*:221	15/15
BSifeg	**1.7**(2)	**1.7**(0.9)	**2.6**(0.4)	43(44)	72(50)	11/15
BSif	**1.7**(1)	**1.7**(2)	**1.9**(4)	44(53)	59(104)	13/15
BSqi	**1.7**(2)	**1.7**(1)	3.4(9)	51(119)	64(117)	12/15
BSrr	**1.7**(2)	**1.7**(1)	3.3(10)	31(5)	56(96)	13/15
CMA-CSA	**3.0**(4)	**2.7**(3)	**2.1**(1)	**1.3**(0.5)	**0.85**(0.5)	15/15
CMA-MSR	**1.7**(1)	**1.6**(0.8)	**1.4**(1)	**1.4**(1)	3.7(3)	15/15
CMA-TPA	**1.9**(1)	**1.5**(2)	**2.2**(2)	**1.4**(0.8)	**0.91**(0.5)	15/15
GP1-CMAES	4.4(3)	**2.9**(2)	**2.3**(1)	**1.0**(0.2)	**1.9**(2)	11/15
GP5-CMAES	**2.8**(3)	**2.0**(1)	**1.7**(0.7)	**0.67**(0.8)	**1.6**(2)	13/15
IPOPCMAv3p	4.2(11)	**2.5**(3)	**1.8**(2)	**1.4**(0.4)	**1.1**(0.5)	14/15
LHD-10xDef	**1.9**(1)	**2.2**(2)	3.3(2)	**2.0**(1)	3.3(4)	3/15
LHD-2xDefa	**2.2**(2)	**1.3**(1)	**1.3**(1)	**0.68**(0.4)	**1.1**(0.7)	7/15
RAND-2xDef	**1.4**(0.9)	**0.99**(0.8)	**0.98**(0.5)	**0.74**(0.5)	**0.57**(0.3)	12/15
RF1-CMAES	**2.4**(3)	**2.3**(3)	**1.7**(0.7)	**0.79**(0.7)	**2.0**(6)	11/15
RF5-CMAES	**2.4**(2)	18(3)	13(2)	5.5(9)	11(14)	4/15
Sifeg	**1.7**(2)	**1.6**(1)	**0.97**(0.6)	17(0.3)	36(78)	13/15
Sif	**1.7**(1)	**1.6**(1)	**0.97**(0.4)	23(83)	45(26)	13/15
Srr	**1.7**(2)	**1.6**(1.0)	**0.95**(0.6)	11(0.3)	32(50)	15/15

Table 41: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*6.3e+1*:1.5	*2.5e+1*:8.2	*1.6e+1*:10	*1.0e+1*:41	*2.5e+0*:208	15/15
BSifeg	**1.6**(1)	**1.5**(1)	**1.9**(1.0)	**1.6**(2)	24(7)	14/15
BSif	**1.6**(1)	**1.5**(1)	**2.7**(7)	**2.0**(1)	20(41)	15/15
BSqi	**1.6**(1)	**1.5**(1)	**1.8**(1)	**2.0**(2)	13(23)	15/15
BSrr	**1.6**(1)	**1.5**(1)	**1.9**(1)	**1.5**(1)	10(2)	15/15
CMA-CSA	**2.8**(3)	**1.5**(1)	4.4(5)	**1.7**(1)	**3.0**(5)	15/15
CMA-MSR	**1.7**(1)	3.3(3)	7.7(4)	6.7(11)	4.1(5)	15/15
CMA-TPA	**1.7**(2)	**1.8**(2)	3.4(7)	3.2(4)	**2.7**(4)	15/15
GP1-CMAES	**2.0**(2)	**1.6**(2)	**2.9**(2)	**1.3**(1)	**2.7**(6)	10/15
GP5-CMAES	**2.1**(2)	**1.2**(2)	**1.4**(2)	**0.78**(0.9)	**2.1**(4)	12/15
IPOPCMAv3p	**1.5**(1)	**1.7**(1)	**2.2**(1)	**1.6**(2)	**1.9**(2)	12/15
LHD-10xDef	**1.7**(1)	**0.76**(0.8)	**2.1**(4)	**0.99**(0.9)	**1.0**(1.0)	8/15
LHD-2xDefa	**1.3**(0)	**1.1**(0.7)	**2.6**(2)	**1.0**(1)	**0.53**(0.5)	11/15
RAND-2xDef	**1.9**(2)	**1.6**(1)	**1.8**(2)	**1.4**(0.7)	**1.0**(0.4)	8/15
RF1-CMAES	**2.5**(4)	**1.8**(2)	**2.8**(3)	**1.1**(1)	**2.6**(3)	10/15
RF5-CMAES	**2.6**(2)	**1.4**(2)	6.4(35)	**2.8**(0.4)	**2.6**(4)	10/15
Sifeg	**1.6**(1)	**1.9**(2)	**2.7**(2)	**1.1**(0.6)	6.3(21)	15/15
Sif	**1.6**(1)	**1.8**(2)	**2.6**(1)	**1.0**(0.7)	3.6(4)	15/15
Srr	**1.6**(1)	**1.9**(2)	3.0(1)	**1.2**(0.7)	4.2(8)	15/15

Table 42: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.6e+1*:1.8	*1.0e+1*:3.6	*6.3e+0*:14	*2.5e+0*:34	*2.5e-1*:189	5/5
BSifeg	4.3(11)	3.6(1)	**1.4**(3)	30(105)	107(81)	10/15
BSif	3.3(11)	**2.9**(2)	**1.2**(0.2)	18(65)	114(177)	10/15
BSqi	3.1(2)	**3.0**(2)	**1.2**(1)	13(42)	82(80)	11/15
BSrr	4.1(2)	5.2(24)	**2.9**(0.6)	**2.7**(5)	112(70)	10/15
CMA-CSA	3.3(2)	**2.3**(0.4)	**1.3**(0.4)	**1.6**(1.0)	**0.88**(0.3)	15/15
CMA-MSR	3.7(3)	3.1(5)	**1.5**(1)	**1.9**(1)	**2.4**(3)	15/15
CMA-TPA	3.8(5)	4.5(2)	**1.8**(2)	**2.1**(1)	**1.0**(0.3)	15/15
GP1-CMAES	**1.8**(2)	**2.3**(2)	**1.2**(1)	**2.9**(6)	**1.3**(2)	13/15
GP5-CMAES	**2.0**(2)	3.2(5)	**2.7**(7)	3.9(3)	3.1(5)	11/15
IPOPCMAv3p	4.9(4)	5.4(4)	**2.8**(4)	**2.5**(1)	**1.1**(0.5)	15/15
LHD-10xDef	**2.9**(2)	**2.8**(4)	**1.6**(2)	**2.4**(1)	**2.2**(2)	5/15
LHD-2xDefa	**2.1**(2)	**2.4**(1)	**1.1**(0.7)	**1.2**(0.4)	5.7(13)	2/15
RAND-2xDef	3.3(3)	**2.6**(1)	**1.1**(0.8)	**1.0**(0.4)	3.8(4)	3/15
RF1-CMAES	3.3(4)	4.1(4)	**1.7**(2)	7.5(16)	5.7(9)	7/15
RF5-CMAES	22(0.6)	22(37)	11(18)	13(11)	13(24)	4/15
Sifeg	4.5(3)	3.8(7)	**1.9**(3)	23(3)	33(27)	14/15
Sif	4.4(19)	4.7(11)	**1.9**(3)	**2.5**(1)	55(76)	13/15
Srr	4.0(2)	3.7(2)	**1.6**(2)	**1.6**(2)	31(26)	15/15

Table 43: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*6.3e+1*:1.8	*4.0e+1*:4.8	*2.5e+1*:13	*1.0e+1*:40	*6.3e-1*:184	15/15
BSifeg	**2.7**(2)	**1.9**(2)	**1.5**(2)	**1.6**(0.3)	334(420)	5/15
BSif	**2.7**(3)	**1.9**(1)	**1.5**(3)	**1.3**(2)	218(273)	7/15
BSqi	**2.7**(2)	**1.9**(2)	**1.5**(2)	**1.1**(3)	158(197)	9/15
BSrr	**2.7**(2)	**1.9**(2)	**1.3**(0.8)	17(60)	187(356)	7/15
CMA-CSA	3.9(4)	**2.7**(2)	**1.7**(2)	**1.4**(1)	**3.6**(1)	15/15
CMA-MSR	4.6(5)	**2.7**(4)	**1.8**(2)	**1.4**(1.0)	4.7(17)	15/15
CMA-TPA	6.1(9)	3.7(5)	**2.4**(1)	**1.7**(1)	**3.8**(2)	15/15
GP1-CMAES	4.2(4)	**2.2**(0.8)	**1.7**(2)	**1.3**(0.5)	3.8(7)	9/15
GP5-CMAES	5.8(4)	14(78)	14(21)	5.6(12)	**2.7**(3)	11/15
IPOPCMAv3p	3.3(3)	4.2(4)	**2.6**(1)	**1.7**(3)	5.4(4)	8/15
LHD-10xDef	**2.9**(4)	**1.6**(2)	**1.4**(2)	**2.1**(0.2)	12(15)	1/15
LHD-2xDefa	**2.6**(0.8)	**1.8**(1)	**1.2**(0.7)	**0.93**(0.6)	3.9(5)	3/15
RAND-2xDef	3.4(3)	**2.0**(2)	**1.3**(1)	**0.97**(1)	6.0(7)	2/15
RF1-CMAES	4.4(9)	3.3(4)	5.6(2)	7.7(19)	17(16)	3/15
RF5-CMAES	3.5(3)	**2.1**(0.8)	**1.1**(1)	5.1(5)	58(77)	1/15
Sifeg	**2.7**(3)	**2.1**(1)	**1.8**(3)	**1.5**(1)	139(166)	10/15
Sif	**2.7**(3)	**2.1**(1)	4.8(15)	3.7(0.9)	153(143)	9/15
Srr	**2.7**(3)	**2.1**(1)	**1.7**(1)	**1.2**(1)	94(119)	11/15

Table 44: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:81	*1.0e-1*:109	*6.3e-2*:109	*4.0e-2*:119	*1.6e-2*:1230	15/15
BSifeg	**12**(18)	**17**(46)	**35**(50)	**62**(61)	54(36)	5/15
BSif	**19**(18)	36(58)	66(179)	139(103)	42(51)	6/15
BSqi	19(10)	45(15)	62(56)	99(193)	**37**(64)	7/15
BSrr	23(24)	33(51)	78(127)	149(256)	115(55)	3/15
CMA-CSA	30(31)	39(48)	59(54)	**59**(45)	**12**(14)	15/15
CMA-MSR	69(83)	96(148)	274(546)	677(1164)	131(193)	12/15
CMA-TPA	28(15)	41(49)	**46**(68)	**66**(55)	**7.5**(4)	15/15
GP1-CMAES	43(31)	48(31)	**48**(21)	92(129)	∞ *753*	0/15
GP5-CMAES	68(48)	104(86)	104(117)	∞	∞ *762*	0/15
IPOPCMAv3p	28(47)	∞	∞	∞	∞ *751*	0/15
LHD-10xDef	∞	∞	∞	∞	∞ *150*	0/15
LHD-2xDefa	∞	∞	∞	∞	∞ *150*	0/15
RAND-2xDef	∞	∞	∞	∞	∞ *150*	0/15
RF1-CMAES	23(22)	24(13)	100(105)	93(87)	∞ *751*	0/15
RF5-CMAES	44(61)	∞	∞	∞	∞ *755*	0/15
Sifeg	22(25)	**19**(33)	70(135)	176(205)	170(346)	2/15
Sif	20(51)	36(61)	54(96)	169(173)	48(76)	6/15
Srr	**15**(23)	**21**(35)	62(95)	133(119)	48(14)	6/15

Table 45: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*4.0e+3*:3.5	*2.5e+3*:4.3	*4.0e+0*:13	*1.6e+0*:41	*1.0e+0*:385	5/5
BSifeg	**2.4**(2)	**2.2**(2)	4.4(0.2)	5.3(6)	13(24)	15/15
BSif	**2.4**(2)	**2.2**(2)	4.0(9)	6.4(9)	17(13)	15/15
BSqi	**2.4**(2)	**2.2**(2)	**2.8**(5)	5.4(6)	11(16)	15/15
BSrr	**2.4**(2)	**2.2**(2)	**2.0**(1)	4.7(6)	14(27)	14/15
CMA-CSA	**2.0**(2)	**1.6**(2)	**2.2**(1.0)	11(13)	**4.5**(2)	15/15
CMA-MSR	**1.7**(0.5)	**1.7**(2)	**2.6**(2)	20(35)	13(16)	15/15
CMA-TPA	**1.7**(1)	**2.4**(2)	**2.9**(1)	10(10)	7.6(10)	15/15
GP1-CMAES	**2.0**(3)	**1.8**(0.6)	**2.4**(2)	8.6(15)	**3.7**(2)	6/15
GP5-CMAES	**1.2**(0.6)	**1.2**(2)	**2.0**(1)	5.4(5)	**2.0**(3)	9/15
IPOPCMAv3p	**1.4**(1)	**1.5**(1)	**2.8**(2)	12(18)	4.5(9)	5/15
LHD-10xDef	**0.79**(0.5)	**1.1**(2)	4.3(3)	∞	∞ *150*	0/15
LHD-2xDefa	**1.4**(1)	**1.3**(0.9)	**1.8**(1)	17(17)	∞ *150*	0/15
RAND-2xDef	**1.2**(2)	**1.4**(2)	**1.9**(1)	13(15)	∞ *150*	0/15
RF1-CMAES	**1.6**(2)	**1.7**(0.9)	6.4(13)	17(9)	5.9(6)	4/15
RF5-CMAES	**1.8**(2)	**2.0**(2)	33(24)	63(42)	∞ *760*	0/15
Sifeg	**2.5**(2)	**2.3**(2)	**2.0**(1.0)	**3.3**(4)	7.3(9)	15/15
Sif	**2.5**(2)	**2.3**(2)	**2.2**(3)	**3.8**(3)	10(25)	15/15
Srr	**2.5**(2)	**2.3**(1.0)	**1.9**(0.7)	**2.7**(0.9)	9.0(26)	15/15

Table 46: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*1.6e+1*:2.5	*1.0e+1*:5.9	*6.3e+0*:14	*2.5e+0*:41	*1.6e+0*:167	15/15
BSifeg	**2.8**(3)	**1.6**(1)	152(524)	117(70)	84(90)	12/15
BSif	**2.8**(3)	**1.6**(1)	152(7)	249(439)	114(154)	10/15
BSqi	**2.8**(2)	**1.6**(1)	152(521)	167(259)	150(139)	9/15
BSrr	**2.8**(3)	**1.6**(1)	150(0.9)	113(46)	130(135)	10/15
CMA-CSA	**2.2**(3)	**1.3**(1)	5.8(17)	5.1(1)	6.9(3)	15/15
CMA-MSR	**2.9**(5)	**2.1**(2)	**2.8**(5)	6.7(1.0)	5.9(8)	15/15
CMA-TPA	**2.7**(3)	**1.6**(2)	**2.3**(4)	**1.6**(2)	**1.9**(3)	15/15
GP1-CMAES	**1.1**(2)	**0.93**(0.8)	4.9(28)	5.3(19)	7.1(8)	6/15
GP5-CMAES	**2.2**(2)	**1.4**(2)	**1.3**(1)	**1.2**(0.5)	**2.4**(2)	11/15
IPOPCMAv3p	**2.5**(3)	**1.9**(1)	**1.4**(2)	**1.6**(1)	**2.5**(2)	11/15
LHD-10xDef	**1.2**(0.8)	**1.7**(2)	**1.3**(2)	**1.4**(0.9)	**0.66**(0.2)	14/15
LHD-2xDefa	**2.2**(2)	**1.3**(0.8)	**1.6**(2)	**1.1**(0.7)	**0.89**(1)	10/15
RAND-2xDef	**1.5**(0.6)	**1.2**(2)	**1.1**(0.8)	**1.1**(1)	**0.80**(0.8)	10/15
RF1-CMAES	**2.3**(2)	**2.1**(2)	**1.9**(1)	**2.5**(9)	4.4(8)	8/15
RF5-CMAES	**1.9**(2)	**1.9**(1)	5.8(0.7)	5.2(6)	7.3(11)	7/15
Sifeg	**2.8**(3)	**1.9**(2)	151(522)	92(561)	95(145)	11/15
Sif	**2.8**(3)	**2.1**(2)	152(523)	101(1)	104(204)	10/15
Srr	**2.8**(3)	**1.9**(2)	151(522)	62(183)	142(147)	9/15

Table 47: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*4.0e+1*:2.9	*2.5e+1*:5.2	*1.0e+1*:18	*6.3e+0*:33	*1.0e+0*:170	5/5
BSifeg	**3.0**(2)	**2.1**(3)	7.6(2)	79(214)	87(138)	11/15
BSif	**2.6**(2)	**2.0**(2)	10(30)	171(237)	219(410)	7/15
BSqi	3.1(3)	**2.2**(3)	4.8(6)	85(265)	130(261)	10/15
BSrr	3.2(6)	**2.2**(5)	3.9(9)	87(180)	146(210)	9/15
CMA-CSA	**1.8**(2)	**1.7**(2)	**1.4**(1)	3.0(7)	11(20)	15/15
CMA-MSR	**2.0**(1)	**2.3**(3)	**2.0**(2)	6.8(1)	5.9(25)	15/15
CMA-TPA	**1.8**(2)	**2.0**(4)	**1.8**(3)	**1.5**(2)	19(50)	15/15
GP1-CMAES	**1.4**(0.9)	**2.0**(1)	**1.7**(2)	5.2(12)	3.8(9)	9/15
GP5-CMAES	**1.5**(1)	**1.5**(3)	4.2(3)	6.4(6)	10(15)	5/15
IPOPCMAv3p	**2.7**(3)	**2.6**(3)	**2.1**(1)	**1.8**(0.9)	10(24)	5/15
LHD-10xDef	**2.8**(2)	**2.1**(1)	**1.7**(1)	**1.9**(1)	**0.88**(0.9)	11/15
LHD-2xDefa	**1.4**(1)	**1.7**(0.9)	**1.4**(0.7)	**1.4**(1.0)	**2.2**(3)	5/15
RAND-2xDef	**1.6**(0.6)	**2.2**(1)	**0.97**(0.6)	**1.3**(1)	**1.1**(1.0)	9/15
RF1-CMAES	**1.1**(1)	**0.96**(0.5)	**1.7**(1)	7.1(24)	5.5(8)	7/15
RF5-CMAES	12(1)	14(0.8)	6.9(19)	10(27)	11(11)	5/15
Sifeg	**2.3**(3)	**1.8**(2)	**1.8**(3)	31(3)	61(62)	13/15
Sif	**2.3**(3)	**1.8**(3)	3.2(8)	45(12)	67(162)	12/15
Srr	**2.3**(2)	**1.8**(1)	**1.5**(4)	**1.6**(2)	66(172)	12/15

Table 48: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*1.0e+1*:2.6	*6.3e+0*:16	*4.0e+0*:44	*2.5e+0*:79	*1.6e+0*:198	15/15
BSifeg	3.8(2)	**1.4**(1)	**1.2**(1)	**2.2**(4)	**1.5**(2)	15/15
BSif	3.8(2)	**1.4**(1)	**1.6**(2)	**2.7**(3)	**2.0**(2)	15/15
BSqi	3.7(3)	**1.2**(1)	**1.1**(2)	**1.8**(3)	**1.4**(1)	15/15
BSrr	3.9(3)	**1.4**(1)	**1.4**(1)	**1.6**(1)	**1.3**(1)	15/15
CMA-CSA	**3.3**(6)	**1.5**(2)	**1.4**(2)	**2.8**(2)	8.5(9)	15/15
CMA-MSR	**2.6**(2)	**2.2**(1)	3.4(2)	5.6(4)	4.8(7)	15/15
CMA-TPA	4.2(6)	3.3(6)	7.6(3)	6.5(3)	4.2(2)	15/15
GP1-CMAES	**3.3**(4)	**1.7**(0.7)	**2.3**(2)	3.5(4)	4.3(6)	10/15
GP5-CMAES	5.7(7)	**1.7**(2)	**2.1**(4)	**2.1**(2)	**1.3**(1)	14/15
IPOPCMAv3p	4.3(7)	**1.5**(2)	**1.7**(2)	3.1(4)	4.7(4)	9/15
LHD-10xDef	6.4(3)	**2.0**(2)	**2.3**(1)	8.5(9)	11(20)	1/15
LHD-2xDefa	4.0(4)	**1.3**(0.8)	3.6(5)	3.5(6)	∞ *150*	0/15
RAND-2xDef	4.5(4)	**2.7**(4)	**2.9**(3)	7.7(8)	11(14)	1/15
RF1-CMAES	5.1(3)	**1.7**(1)	**1.9**(1)	**2.5**(4)	6.9(5)	6/15
RF5-CMAES	3.5(2)	**1.4**(1)	**1.6**(1)	5.4(7)	7.9(6)	6/15
Sifeg	3.7(4)	**1.9**(2)	**2.3**(2)	3.9(2)	**2.9**(2)	15/15
Sif	3.7(2)	**1.7**(3)	**2.3**(2)	3.8(3)	**2.8**(1)	15/15
Srr	3.7(4)	**1.7**(3)	**2.2**(2)	3.9(2)	**2.3**(1)	15/15

Table 49: 03-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*4.0e+1*:4.6	*2.5e+1*:13	*1.6e+1*:47	*1.6e+1*:47	*6.3e+0*:382	15/15
BSifeg	**1.5**(1)	**2.2**(0.5)	**2.5**(4)	**2.5**(4)	17(35)	14/15
BSif	**1.6**(0.5)	**2.4**(1)	**2.3**(1)	**2.3**(2)	24(49)	12/15
BSqi	**1.6**(1)	**1.7**(1)	**2.6**(3)	**2.6**(1)	13(29)	14/15
BSrr	**1.5**(1)	**1.7**(0.7)	**2.5**(5)	**2.5**(4)	8.1(21)	15/15
CMA-CSA	**1.8**(2)	**2.1**(2)	**1.5**(0.5)	**1.5**(0.9)	**2.0**(3)	15/15
CMA-MSR	**1.7**(0.8)	**2.6**(2)	**1.6**(0.9)	**1.6**(2)	3.4(4)	15/15
CMA-TPA	**1.4**(1)	**2.0**(1)	**1.3**(2)	**1.3**(0.3)	**2.0**(0.7)	15/15
GP1-CMAES	**2.2**(2)	**2.2**(1)	**1.7**(0.8)	**1.7**(0.9)	**2.0**(3)	9/15
GP5-CMAES	**1.5**(2)	**2.4**(2)	**1.2**(1)	**1.2**(2)	**2.3**(4)	8/15
IPOPCMAv3p	**2.0**(2)	**2.3**(3)	**1.8**(1)	**1.8**(1)	**1.5**(0.9)	12/15
LHD-10xDef	**2.1**(2)	3.3(3)	3.0(4)	3.0(4)	5.8(5)	1/15
LHD-2xDefa	**1.2**(2)	**2.2**(3)	**1.7**(0.7)	**1.7**(2)	5.6(9)	1/15
RAND-2xDef	**0.93**(1)	**2.3**(1)	**2.2**(3)	**2.2**(0.9)	5.7(8)	1/15
RF1-CMAES	**1.9**(4)	**2.4**(2)	**1.3**(0.8)	**1.3**(0.7)	**1.6**(1)	10/15
RF5-CMAES	**1.2**(1)	**1.3**(0.8)	5.2(5)	5.2(11)	**2.3**(3)	9/15
Sifeg	**1.4**(2)	**1.7**(0.6)	**1.3**(0.9)	**1.3**(1)	6.1(34)	14/15
Sif	**1.4**(1)	**1.6**(0.6)	**1.4**(1)	**1.4**(1)	6.2(3)	15/15
Srr	**1.4**(2)	**1.8**(1.0)	**1.3**(0.6)	**1.3**(0.6)	11(34)	14/15

Table 50: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\mathrm{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*2.5e+1*:4.8	*1.6e+1*:7.6	*1.0e-8*:12	*1.0e-8*:12	*1.0e-8*:12	15/15
BSifeg	**1.8**(2)	**1.7**(0.8)	**2.2**(0.1)	**2.2**(0.1)	**2.2**(0.1)	15/15
BSif	**1.8**(2)	**1.7**(1)	**2.2**(0.2)	**2.2**(0.1)	**2.2**(0.2)	15/15
BSqi	**1.8**(2)	**1.7**(1)	**2.2**(0.1)	**2.2**(0.1)	**2.2**(0.2)	15/15
BSrr	**1.8**(2)	**1.7**(0.9)	**2.2**(0.0)	**2.2**(0.2)	**2.2**(0.2)	15/15
CMA-CSA	4.3(3)	3.6(2)	58(5)	58(2)	58(5)	15/15
CMA-MSR	3.0(2)	3.1(3)	90(7)	90(9)	90(9)	15/15
CMA-TPA	**2.7**(4)	**2.5**(3)	52(9)	52(4)	52(10)	15/15
GP1-CMAES	**2.5**(2)	**2.2**(2)	36(6)	36(5)	36(5)	15/15
GP5-CMAES	**2.3**(2)	**2.1**(1)	92(54)	92(86)	92(122)	11/15
IPOPCMAv3p	3.0(4)	**2.9**(4)	59(5)	59(9)	59(7)	15/15
LHD-10xDef	**2.8**(1)	3.8(6)	∞	∞	∞ *250*	0/15
LHD-2xDefa	**2.6**(2)	**2.8**(1)	∞	∞	∞ *250*	0/15
RAND-2xDef	**2.0**(1)	**2.2**(1)	∞	∞	∞ *250*	0/15
RF1-CMAES	3.7(3)	**3.0**(1)	1520(1031)	1520(2191)	1520(799)	1/15
RF5-CMAES	**2.2**(3)	**2.4**(1)	∞	∞	∞ *1252*	0/15
Sifeg	**1.8**(2)	**1.7**(1)	8.9(0.9)	8.9(1.0)	8.9(1)	15/15
Sif	**1.8**(2)	**1.7**(0.7)	8.7(1)	8.7(1)	8.7(0.9)	15/15
Srr	**1.8**(1)	**1.7**(1)	8.4(1)	8.4(1)	8.4(1)	15/15

Table 51: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*1.6e+6*:2.9	*4.0e+5*:11	*4.0e+4*:15	*6.3e+2*:58	*1.0e-8*:95	15/15
BSifeg	6.0(3)	**1.8**(1)	**1.5**(0.2)	**0.52**(0.2)	**1.1**(0.1)	15/15
BSif	6.0(4)	**1.8**(1)	**1.5**(0.2)	**0.53**(0.3)	**1.1**(0.2)	15/15
BSqi	6.0(4)	**1.8**(1)	**1.5**(0.2)	**0.49**(0.1)	**0.92**(0.2)	15/15
BSrr	6.0(2)	**1.8**(1)	**1.5**(0.2)	**0.52**(0.2)	**1.1**(0.2)	15/15
CMA-CSA	**2.3**(2)	**1.8**(2)	6.1(4)	7.5(2)	18(0.9)	15/15
CMA-MSR	**2.5**(2)	**1.9**(2)	5.9(4)	7.4(3)	21(1)	15/15
CMA-TPA	3.7(6)	**2.1**(2)	3.9(3)	7.7(3)	18(3)	15/15
GP1-CMAES	**2.1**(1.0)	**1.9**(2)	4.0(3)	5.8(4)	∞ *1258*	0/15
GP5-CMAES	**2.5**(3)	**1.7**(2)	3.2(1)	**2.7**(1)	94(92)	2/15
IPOPCMAv3p	3.1(4)	**1.5**(2)	5.6(3)	10(9)	∞ *1258*	0/15
LHD-10xDef	**1.1**(0.7)	**1.2**(2)	8.4(6)	65(37)	∞ *250*	0/15
LHD-2xDefa	**1.3**(1)	**0.79**(2)	4.3(2)	32(47)	∞ *250*	0/15
RAND-2xDef	**1.6**(1)	**1.0**(0.6)	3.3(2)	62(33)	∞ *250*	0/15
RF1-CMAES	**2.9**(4)	**2.7**(2)	6.8(5)	51(50)	∞ *1258*	0/15
RF5-CMAES	**2.1**(1)	7.1(23)	25(29)	∞	∞ *1260*	0/15
Sifeg	6.4(4)	**2.0**(1)	**2.1**(0.6)	**0.84**(0.2)	**1.4**(0.1)	15/15
Sif	6.4(2)	**2.0**(0.2)	**2.1**(0.6)	**0.82**(0.1)	**1.4**(0.3)	15/15
Srr	6.4(3)	**2.0**(0.6)	**2.1**(0.5)	**0.79**(0.1)	**1.5**(0.2)	15/15

Table 52: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*1.6e+2*:4.1	*1.0e+2*:15	*6.3e+1*:23	*2.5e+1*:73	*1.0e+1*:716	15/15
BSifeg	**2.6**(2)	**1.2**(0.2)	**0.98**(0.1)	**0.46**(0.1)	**0.11**(0.1)	15/15
BSif	**2.6**(2)	**1.2**(0.2)	**0.98**(0.1)	**0.46**(0.1)	**0.11**(0.0)	15/15
BSqi	**2.6**(2)	**1.2**(0.5)	**0.98**(0.1)	**0.46**(0.1)	**0.10**(0.1)	15/15
BSrr	**2.6**(2)	**1.2**(0.4)	**0.98**(0.1)	**0.46**(0.1)	**0.09**(0.1)	15/15
CMA-CSA	4.2(5)	**2.4**(2)	**2.6**(2)	**2.5**(1)	**1.4**(0.3)	15/15
CMA-MSR	3.3(3)	**2.4**(1)	3.4(2)	6.7(14)	**1.7**(2)	15/15
CMA-TPA	3.5(7)	**2.2**(1)	**2.6**(2)	**2.5**(1)	**0.81**(0.7)	15/15
GP1-CMAES	**2.8**(3)	**1.5**(2)	**2.3**(0.7)	**2.5**(5)	**1.6**(1)	11/15
GP5-CMAES	3.0(2)	**1.3**(0.9)	**1.6**(1)	**2.2**(0.5)	**2.6**(3)	8/15
IPOPCMAv3p	**2.7**(2)	**1.4**(2)	**2.5**(0.8)	3.1(1)	**1.1**(0.9)	12/15
LHD-10xDef	**2.1**(3)	**2.8**(2)	3.7(2)	3.6(0.5)	**1.0**(0.8)	5/15
LHD-2xDefa	**2.1**(2)	**1.5**(1)	**1.9**(1)	**2.1**(3)	**2.5**(4)	2/15
RAND-2xDef	**2.3**(2)	**1.4**(1)	**1.6**(0.4)	**2.2**(0.5)	**0.58**(0.3)	7/15
RF1-CMAES	**1.9**(3)	**1.6**(1)	**2.2**(0.8)	**2.4**(1)	3.0(3)	6/15
RF5-CMAES	3.4(3)	**1.6**(1)	4.6(19)	14(15)	6.1(7)	4/15
Sifeg	**2.6**(2)	**1.2**(0.5)	**0.99**(0.1)	**0.57**(0.4)	**0.13**(0.1)	15/15
Sif	**2.6**(2)	**1.2**(0.3)	**0.99**(0.1)	**0.57**(0.2)	**0.13**(0.1)	15/15
Srr	**2.6**(2)	**1.2**(0.3)	**0.99**(0.2)	**0.56**(0.1)	**0.12**(0.0)	15/15

Table 53: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*2.5e+2*:2.6	*1.6e+2*:10	*1.0e+2*:19	*4.0e+1*:65	*1.6e+1*:434	15/15
BSifeg	**2.1**(2)	**0.90**(0.8)	**1.4**(0.8)	**1.0**(0.6)	**0.25**(0.1)	15/15
BSif	**2.1**(3)	**0.90**(0.7)	**1.4**(0.5)	**1.1**(0.5)	**0.26**(0.1)	15/15
BSqi	**2.1**(2)	**0.90**(1)	**1.3**(0.4)	**1.3**(0.8)	**0.28**(0.2)	15/15
BSrr	**2.1**(2)	**0.90**(0.9)	**1.3**(1.0)	**1.1**(0.4)	**0.23**(0.1)	15/15
CMA-CSA	4.6(9)	**2.9**(4)	3.0(2)	4.3(1)	**2.3**(3)	15/15
CMA-MSR	4.2(6)	**2.7**(4)	**2.3**(1)	**2.8**(2)	**2.0**(3)	15/15
CMA-TPA	4.7(5)	**2.5**(1)	3.4(2)	**3.0**(2)	**2.4**(2)	15/15
GP1-CMAES	**2.7**(2)	**2.4**(2)	**2.6**(3)	5.1(8)	3.2(2)	9/15
GP5-CMAES	**1.6**(1)	**1.6**(1)	**3.0**(2)	5.2(6)	7.5(7)	5/15
IPOPCMAv3p	**2.3**(3)	**1.5**(1)	**2.4**(1)	**2.4**(1)	**1.2**(0.9)	14/15
LHD-10xDef	3.9(2)	3.2(3)	5.1(3)	8.3(6)	∞ *250*	0/15
LHD-2xDefa	**2.2**(3)	**1.5**(1)	3.0(2)	6.8(11)	∞ *250*	0/15
RAND-2xDef	**2.8**(3)	**2.2**(2)	3.1(2)	4.2(3)	8.4(13)	1/15
RF1-CMAES	**2.4**(3)	**1.9**(0.8)	**2.5**(2)	7.9(10)	12(20)	3/15
RF5-CMAES	**2.7**(4)	**1.9**(1)	4.9(11)	31(26)	∞ *1252*	0/15
Sifeg	**2.1**(2)	**0.92**(1.0)	**1.3**(0.4)	**0.85**(0.3)	**0.21**(0.1)	15/15
Sif	**2.1**(3)	**0.92**(0.7)	**1.3**(0.5)	**0.85**(0.3)	**0.21**(0.1)	15/15
Srr	**2.1**(2)	**0.92**(1)	**1.3**(0.8)	**0.83**(0.2)	**0.20**(0.0)	15/15

Table 54: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*6.3e+1*:4.0	*4.0e+1*:10	*1.0e-8*:10	*1.0e-8*:10	*1.0e-8*:10	15/15
BSifeg	**2.9**(0.7)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
BSif	**2.9**(0.4)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
BSqi	**2.9**(0.9)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
BSrr	**2.9**(0.4)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
CMA-CSA	**1.9**(2)	**1.8**(0.7)	5.2(2)	5.2(2)	5.2(2)	15/15
CMA-MSR	**2.2**(2)	**1.8**(1)	5.9(3)	5.9(2)	5.9(1)	15/15
CMA-TPA	**2.6**(3)	**1.9**(2)	5.1(2)	5.1(2)	5.1(2)	15/15
GP1-CMAES	**2.0**(2)	**1.5**(1.0)	26(9)	26(20)	26(37)	15/15
GP5-CMAES	3.0(1)	**1.7**(0.1)	6.4(4)	6.4(4)	6.4(3)	15/15
IPOPCMAv3p	3.0(3)	**2.2**(1)	21(12)	21(20)	21(14)	15/15
LHD-10xDef	3.2(5)	4.6(5)	13(0.2)	13(0.2)	13(0.2)	15/15
LHD-2xDefa	**1.9**(1)	**1.8**(1)	3.5(2)	3.5(2)	3.5(3)	15/15
RAND-2xDef	**2.0**(2)	**2.0**(1.0)	3.1(0.2)	3.1(0.2)	3.1(0.2)	15/15
RF1-CMAES	**2.3**(2)	**1.7**(0.9)	45(26)	45(35)	45(26)	15/15
RF5-CMAES	**2.8**(3)	**1.9**(2)	137(220)	137(56)	137(117)	10/15
Sifeg	**2.9**(0.8)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
Sif	**2.9**(0.3)	**1.4**(0.1)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15
Srr	**2.9**(0.8)	**1.4**(0.2)	**1.5**(0.1)	**1.5**(0.1)	**1.5**(0.1)	15/15

Table 55: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*1.0e+5*:3.0	*2.5e+4*:8.4	*1.0e+2*:16	*2.5e+1*:54	*2.5e-1*:254	15/15
BSifeg	**1.2**(0.9)	**1.2**(0.8)	**2.3**(3)	66(371)	210(215)	9/15
BSif	**1.2**(1)	**1.2**(1.0)	**2.9**(6)	79(83)	1315(1887)	2/15
BSqi	**1.2**(2)	**1.2**(0.8)	3.6(0.8)	71(214)	228(417)	8/15
BSrr	**1.2**(2)	**1.2**(1)	**2.3**(1)	66(403)	226(149)	8/15
CMA-CSA	**2.0**(2)	**1.6**(0.9)	**2.9**(1)	**2.7**(2)	**1.9**(0.5)	15/15
CMA-MSR	3.2(3)	**1.9**(1)	3.5(4)	3.3(1)	**2.1**(0.5)	15/15
CMA-TPA	**2.9**(2)	**2.0**(2)	5.9(3)	3.5(2)	**2.0**(0.2)	15/15
GP1-CMAES	**2.2**(3)	**1.4**(1)	**2.3**(2)	**2.2**(2)	74(77)	1/15
GP5-CMAES	**3.0**(3)	**1.8**(1)	**2.8**(7)	3.7(7)	∞ *1260*	0/15
IPOPCMAv3p	3.1(4)	**1.8**(1)	3.3(3)	**3.1**(3)	**2.2**(0.6)	15/15
LHD-10xDef	**1.6**(2)	**2.2**(4)	5.5(4)	3.6(3)	∞ *250*	0/15
LHD-2xDefa	**1.7**(2)	**1.3**(1)	3.2(5)	4.9(5)	∞ *250*	0/15
RAND-2xDef	**2.3**(2)	**1.8**(1)	**2.8**(4)	9.0(8)	∞ *250*	0/15
RF1-CMAES	**3.0**(3)	**2.0**(2)	4.5(5)	8.5(4)	∞ *1258*	0/15
RF5-CMAES	3.6(4)	**2.4**(3)	19(54)	154(117)	∞ *1260*	0/15
Sifeg	**1.2**(1)	**1.2**(0.9)	5.3(0.8)	36(10)	120(143)	12/15
Sif	**1.2**(2)	**1.2**(1.0)	8.2(26)	64(388)	458(237)	5/15
Srr	**1.2**(1)	**1.2**(1)	6.7(21)	34(112)	101(59)	11/15

Table 56: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*1.6e+2*:4.2	*1.0e+2*:6.2	*2.5e+1*:20	*4.0e+0*:54	*1.0e+0*:324	15/15
BSifeg	34(249)	24(2)	55(128)	620(1442)	754(1633)	3/15
BSif	34(2)	24(85)	64(137)	600(462)	1037(1119)	2/15
BSqi	34(1)	24(2)	68(76)	488(365)	726(522)	3/15
BSrr	35(127)	46(168)	96(0.6)	669(912)	1050(649)	2/15
CMA-CSA	**2.0**(2)	**2.8**(1)	3.6(3)	3.9(1.0)	**1.3**(1.0)	15/15
CMA-MSR	3.5(4)	3.7(3)	3.2(3)	4.5(8)	**1.1**(1.0)	15/15
CMA-TPA	**2.5**(3)	**2.3**(2)	**2.5**(2)	**2.7**(1)	**0.98**(0.1)	15/15
GP1-CMAES	**1.7**(1)	**2.0**(1)	**2.1**(1)	4.7(2)	**1.4**(1.0)	15/15
GP5-CMAES	**2.2**(2)	**2.5**(2)	**1.9**(0.9)	**1.2**(0.6)*	**0.82**(0.9)	15/15
IPOPCMAv3p	3.3(2)	3.4(4)	3.3(2)	**3.5**(0.9)	**1.5**(0.7)	14/15
LHD-10xDef	**1.5**(1)	**2.2**(3)	5.2(2)	4.8(3)	5.5(7)	2/15
LHD-2xDefa	**1.4**(1)	**2.0**(3)	**1.8**(2)	8.0(8)	11(11)	1/15
RAND-2xDef	**1.3**(1.0)	**1.7**(2)	**2.5**(3)	9.3(7)	11(24)	1/15
RF1-CMAES	**1.9**(2)	**2.6**(1)	**2.7**(2)	15(15)	10(9)	5/15
RF5-CMAES	**2.5**(3)	**2.7**(2)	5.3(13)	33(23)	17(16)	3/15
Sifeg	**1.6**(3)	**2.0**(3)	**1.5**(2)	396(707)	276(160)	6/15
Sif	**1.6**(2)	**2.0**(2)	8.1(26)	296(420)	204(167)	8/15
Srr	**1.6**(3)	**2.1**(2)	8.3(27)	317(429)	306(449)	6/15

Table 57: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.0e+4*:4.6	*6.3e+3*:6.8	*1.0e+3*:18	*6.3e+1*:54	*1.6e+0*:258	15/15
BSifeg	**1.6**(0.3)	**1.8**(1)	**1.0**(0.1)	13(30)	38(40)	15/15
BSif	**1.6**(2)	**1.8**(1)	**1.0**(0.2)	34(73)	64(66)	14/15
BSqi	**1.6**(2)	**1.8**(1)	**1.0**(0.4)	8.2(11)	35(49)	14/15
BSrr	**1.6**(2)	**1.8**(1)	**1.0**(0.1)	12(12)	37(37)	14/15
CMA-CSA	**2.6**(1)	**2.8**(1)	**2.8**(1)	**2.7**(1)	**4.9**(1)	15/15
CMA-MSR	**2.2**(3)	**2.0**(4)	**2.6**(3)	4.2(1)	**3.4**(2)	15/15
CMA-TPA	3.4(6)	**2.6**(6)	**2.7**(2)	4.0(1)	5.8(5)	15/15
GP1-CMAES	**2.8**(3)	**2.2**(3)	**2.3**(1)	**2.1**(0.9)	7.2(11)	8/15
GP5-CMAES	**2.1**(1)	**1.9**(2)	**1.9**(1)	**2.5**(2)	70(35)	1/15
IPOPCMAv3p	3.1(3)	**2.8**(2)	**2.1**(1.0)	3.8(2)	**4.9**(4)	11/15
LHD-10xDef	**2.7**(3)	**2.8**(2)	5.2(3)	12(8)	∞ *250*	0/15
LHD-2xDefa	**2.4**(1)	**2.3**(2)	**1.5**(0.9)	3.6(9)	∞ *250*	0/15
RAND-2xDef	3.1(2)	**2.3**(2)	**1.5**(0.6)	3.1(2)	∞ *250*	0/15
RF1-CMAES	**2.6**(2)	**2.3**(2)	3.1(2)	7.8(12)	36(38)	2/15
RF5-CMAES	**2.3**(2)	**2.0**(1)	**2.8**(3)	56(29)	∞ *1252*	0/15
Sifeg	**1.6**(2)	**1.8**(1)	**1.1**(0.1)	**2.2**(3)	17(22)	15/15
Sif	**1.6**(2)	**1.8**(1)	**1.1**(0.5)	3.4(1)	48(43)	14/15
Srr	**1.6**(1)	**1.8**(1)	**1.1**(0.1)	**1.7**(1)	21(34)	15/15

Table 58: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*2.5e+1*:20	*1.6e+1*:26	*1.0e+1*:35	*4.0e+0*:62	*1.6e-2*:256	15/15
BSifeg	19(58)	16(75)	14(16)	223(484)	∞ *5e4*	0/15
BSif	57(363)	45(279)	36(15)	416(392)	∞ *5e4*	0/15
BSqi	15(6)	13(0.9)	11(28)	181(58)	∞ *5e4*	0/15
BSrr	17(13)	14(20)	15(14)	248(306)	∞ *4e4*	0/15
CMA-CSA	**7.5**(2)	**6.7**(2)	**5.7**(0.6)	**6.6**(7)	**7.1**(4)	15/15
CMA-MSR	10(4)	8.5(2)	7.2(1)	7.9(3)	**6.9**(7)	15/15
CMA-TPA	7.8(3)	**6.8**(3)	**5.4**(2)	**4.6**(2)	**5.0**(1)	15/15
GP1-CMAES	12(10)	10(15)	8.2(7)	14(24)	∞ *1258*	0/15
GP5-CMAES	15(5)	14(12)	13(19)	27(17)	70(55)	1/15
IPOPCMAv3p	10(3)	8.7(3)	7.5(2)	**7.2**(10)	36(18)	2/15
LHD-10xDef	185(180)	144(128)	∞	∞	∞ *250*	0/15
LHD-2xDefa	12(3)	12(7)	25(40)	∞	∞ *250*	0/15
RAND-2xDef	11(5)	11(9)	20(3)	60(85)	∞ *250*	0/15
RF1-CMAES	38(82)	36(51)	30(15)	48(42)	∞ *1258*	0/15
RF5-CMAES	106(204)	126(50)	257(205)	∞	∞ *1252*	0/15
Sifeg	**7.5**(6)	6.9(14)	5.8(4)	102(220)	∞ *5e4*	0/15
Sif	36(105)	28(93)	24(4)	221(505)	∞ *5e4*	0/15
Srr	**6.2**(2)	**5.3**(6)	**4.5**(7)	95(237)	∞ *4e4*	0/15

Table 59: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*2.5e+6*:2.9	*6.3e+5*:7.0	*2.5e+5*:17	*6.3e+3*:54	*2.5e+1*:297	15/15
BSifeg	**1.8**(2)	**2.1**(0.9)	12(41)	463(716)	∞ *3e4*	0/15
BSif	**1.8**(2)	**2.1**(1)	9.2(0.7)	481(550)	∞ *3e4*	0/15
BSqi	**1.8**(2)	**2.1**(1)	7.1(0.4)	325(448)	∞ *3e4*	0/15
BSrr	**1.8**(1)	**2.1**(1)	10(0.6)	529(749)	∞ *2e4*	0/15
CMA-CSA	**2.5**(2)	**1.8**(1)	**1.5**(0.8)	4.0(1)	**2.7**(0.6)	15/15
CMA-MSR	**2.3**(1)	**2.0**(3)	**1.6**(1)	4.0(3)	**2.9**(0.8)	15/15
CMA-TPA	**1.9**(3)	**2.3**(2)	**1.4**(2)	**3.3**(2)	**2.3**(1)	15/15
GP1-CMAES	**1.2**(2)	**0.94**(1.0)	**0.85**(0.6)	**2.9**(2)	**2.0**(0.4)	15/15
GP5-CMAES	**2.4**(3)	**1.6**(1)	**1.0**(1)	**1.6**(0.6)	**0.95**(0.2)	15/15
IPOPCMAv3p	**1.8**(1)	**2.2**(2)	**1.6**(2)	4.7(2)	4.4(5)	11/15
LHD-10xDef	**1.5**(2)	**1.7**(0.9)	**1.4**(2)	12(9)	∞ *250*	0/15
LHD-2xDefa	**1.9**(2)	**1.4**(1)	**1.6**(1)	12(13)	∞ *250*	0/15
RAND-2xDef	**2.3**(2)	**2.1**(2)	**1.5**(0.9)	3.5(1)	∞ *250*	0/15
RF1-CMAES	3.2(4)	**2.3**(2)	**1.6**(0.8)	10(18)	63(90)	1/15
RF5-CMAES	**2.7**(2)	**1.7**(2)	4.5(0.4)	43(28)	∞ *1260*	0/15
Sifeg	**1.8**(2)	**1.9**(1)	**1.5**(1)	76(134)	∞ *1e4*	0/15
Sif	**1.8**(2)	**1.9**(2)	**1.4**(1)	97(142)	∞ *1e4*	0/15
Srr	**1.8**(2)	**1.9**(1)	**1.4**(1)	93(70)	∞ *1e4*	0/15

Table 60: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*1.0e+6*:3.0	*6.3e+4*:6.2	*6.3e+2*:16	*6.3e+1*:74	*6.3e-1*:298	15/15
BSifeg	**2.4**(3)	**2.2**(1)	**1.5**(0.5)	116(123)	∞ *3e4*	0/15
BSif	**2.4**(1)	**2.2**(1)	**1.5**(0.5)	93(86)	∞ *3e4*	0/15
BSqi	**2.4**(2)	**2.2**(1)	**1.5**(0.6)	64(103)	∞ *4e4*	0/15
BSrr	**2.4**(1)	**2.2**(1)	**1.5**(0.6)	85(79)	∞ *3e4*	0/15
CMA-CSA	**1.5**(1)	**2.0**(2)	3.8(2)	6.0(4)	**3.0**(0.5)	15/15
CMA-MSR	3.2(4)	3.1(2)	6.6(2)	7.3(3)	3.5(0.5)	15/15
CMA-TPA	**2.8**(7)	3.6(0.8)	4.4(2)	5.7(4)	**3.2**(0.4)	15/15
GP1-CMAES	**1.4**(0.9)	**2.4**(3)	4.5(3)	**2.3**(2)	5.1(4)	11/15
GP5-CMAES	**1.5**(1)	**2.5**(2)	**2.9**(2)	**1.9**(1)	**2.3**(5)	14/15
IPOPCMAv3p	**2.5**(4)	**2.9**(2)	5.1(3)	5.3(6)	∞ *1258*	0/15
LHD-10xDef	**1.7**(2)	3.1(3)	6.1(6)	5.2(9)	∞ *250*	0/15
LHD-2xDefa	**1.5**(1)	**3.0**(2)	4.1(3)	**3.5**(4)	∞ *250*	0/15
RAND-2xDef	**1.6**(0.6)	3.6(3)	5.8(4)	8.7(7)	∞ *250*	0/15
RF1-CMAES	**2.2**(1)	**2.6**(2)	3.9(2)	5.1(12)	∞ *1258*	0/15
RF5-CMAES	**2.2**(3)	**2.3**(3)	**3.0**(2)	11(11)	∞ *1260*	0/15
Sifeg	**2.4**(2)	**2.3**(2)	**1.8**(0.7)	11(10)	∞ *2e4*	0/15
Sif	**2.4**(2)	**2.3**(2)	**1.8**(0.5)	24(30)	∞ *2e4*	0/15
Srr	**2.4**(2)	**2.3**(1)	**1.7**(0.8)	23(0.2)	∞ *2e4*	0/15

Table 61: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*4.0e+7*:3.6	*1.6e+7*:7.6	*4.0e+6*:19	*1.6e+4*:52	*1.0e+0*:268	15/15
BSifeg	**1.3**(1)	**2.0**(1)	**1.2**(1)	8.1(5)	144(213)	6/15
BSif	**1.3**(1)	**2.1**(1)	**1.4**(0.2)	11(9)	157(334)	6/15
BSqi	**1.3**(2)	**2.0**(0.8)	**1.1**(0.4)	6.7(15)	42(34)	12/15
BSrr	**1.3**(2)	**2.0**(1)	**1.2**(0.9)	14(8)	51(63)	10/15
CMA-CSA	3.5(3)	**2.9**(3)	**2.4**(2)	4.7(1)	7.1(3)	15/15
CMA-MSR	**2.1**(6)	3.1(4)	**2.9**(2)	5.7(1)	**5.4**(6)	15/15
CMA-TPA	**2.8**(2)	3.3(2)	4.0(4)	4.6(1)	**6.1**(1)	15/15
GP1-CMAES	**2.0**(2)	**2.5**(2)	**1.9**(0.6)	**3.7**(2)	**6.2**(6)	8/15
GP5-CMAES	3.1(2)	**2.9**(2)	**3.0**(2)	21(17)	8.5(15)	6/15
IPOPCMAv3p	**1.9**(3)	**2.9**(2)	3.0(3)	4.6(1.0)	10(12)	6/15
LHD-10xDef	**1.1**(1)	**2.8**(3)	4.5(2)	8.2(4)	∞ *250*	0/15
LHD-2xDefa	**1.5**(1)	**1.9**(1)	**1.4**(0.9)	**3.0**(1)	∞ *250*	0/15
RAND-2xDef	**1.4**(3)	**2.0**(2)	**1.5**(0.4)	**2.9**(0.8)	∞ *250*	0/15
RF1-CMAES	**2.1**(1)	3.2(3)	3.2(1)	5.1(4)	22(12)	3/15
RF5-CMAES	**1.9**(2)	**1.7**(2)	5.6(2)	117(171)	∞ *1260*	0/15
Sifeg	**1.3**(1)	8.2(48)	5.1(0.4)	10(17)	50(69)	5/15
Sif	**1.3**(0.9)	4.4(1)	**2.0**(0.4)	7.5(8)	56(58)	5/15
Srr	**1.3**(1)	6.9(20)	4.6(13)	11(22)	21(17)	9/15

Table 62: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*1.0e+3*:2.8	*6.3e+2*:8.4	*4.0e+2*:17	*6.3e+1*:52	*6.3e-2*:264	15/15
BSifeg	**1.5**(2)	**1.5**(0.9)	**1.1**(0.6)	223(318)	2385(2219)	1/15
BSif	**1.5**(3)	**1.5**(0.9)	**1.1**(0.7)	227(323)	∞ *4e4*	0/15
BSqi	**1.5**(0.7)	**1.5**(0.9)	**1.1**(0.7)	208(285)	∞ *4e4*	0/15
BSrr	**1.5**(3)	**1.5**(0.5)	**1.1**(1.0)	221(422)	1122(1443)	2/15
CMA-CSA	3.2(4)	3.1(3)	**3.0**(2)	3.9(3)	**4.0**(1)	15/15
CMA-MSR	3.8(3)	**2.5**(2)	3.5(1)	4.5(2)	**3.8**(0.5)	15/15
CMA-TPA	4.6(3)	**2.9**(3)	3.5(2)	3.8(2)	**4.3**(2)	15/15
GP1-CMAES	**2.3**(2)	**1.6**(1)	**2.0**(1)	**2.3**(0.7)	70(92)	1/15
GP5-CMAES	3.2(3)	**2.1**(2)	**1.6**(0.5)	**1.4**(0.3)	11(14)	5/15
IPOPCMAv3p	3.0(4)	**2.2**(3)	3.0(2)	3.9(0.6)	10(15)	7/15
LHD-10xDef	**2.6**(1)	**3.0**(3)	4.8(3)	**2.9**(0.5)	∞ *250*	0/15
LHD-2xDefa	**2.1**(1)	**2.0**(1)	**1.9**(0.3)	**1.8**(0.4)	∞ *250*	0/15
RAND-2xDef	**2.0**(1)	**1.8**(1)	**1.8**(0.9)	**1.7**(0.5)	∞ *250*	0/15
RF1-CMAES	3.2(2)	**2.0**(1)	**2.4**(1)	8.1(19)	69(82)	1/15
RF5-CMAES	3.5(2)	**2.2**(2)	**2.1**(1)	35(21)	∞ *1252*	0/15
Sifeg	**1.5**(3)	**1.5**(0.8)	**1.1**(0.8)	83(183)	2257(3322)	1/15
Sif	**1.5**(3)	**1.5**(0.9)	**1.1**(0.8)	118(282)	∞ *4e4*	0/15
Srr	**1.5**(1)	**1.5**(1.0)	**1.1**(0.7)	178(277)	1104(1218)	2/15

Table 63: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*1.6e+1*:3.0	*1.0e+1*:10	*6.3e+0*:15	*2.5e-1*:53	*1.0e-5*:251	15/15
BSifeg	**3.0**(2)	**1.5**(0.7)	**1.4**(0.6)	8.7(7)	∞ *5e4*	0/15
BSif	**3.0**(3)	**1.5**(1.0)	**1.4**(1)	10(11)	∞ *5e4*	0/15
BSqi	**3.0**(3)	**1.5**(0.9)	**1.4**(0.8)	5.6(4)	∞ *5e4*	0/15
BSrr	**3.0**(3)	**1.5**(0.8)	**1.4**(1)	7.7(9)	∞ *5e4*	0/15
CMA-CSA	4.1(4)	**1.7**(3)	**2.6**(2)	3.2(1)	**3.9**(0.6)	15/15
CMA-MSR	4.3(4)	**2.5**(1)	**2.5**(1)	4.2(0.9)	**4.1**(0.6)	15/15
CMA-TPA	3.5(10)	**2.1**(1)	**2.7**(4)	3.6(1)	**4.0**(0.7)	15/15
GP1-CMAES	3.2(2)	**1.6**(1.0)	**1.9**(2)	**2.2**(0.9)	∞ *1258*	0/15
GP5-CMAES	4.0(4)	**1.8**(1)	**1.5**(2)	**1.6**(0.6)	∞ *1260*	0/15
IPOPCMAv3p	4.8(3)	**2.4**(3)	**2.4**(2)	3.8(0.5)	24(23)	3/15
LHD-10xDef	**2.0**(1)	**1.2**(1)	**2.2**(2)	3.2(0.4)	∞ *250*	0/15
LHD-2xDefa	**2.2**(3)	**1.5**(1)	**1.4**(0.6)	**2.0**(0.9)	∞ *250*	0/15
RAND-2xDef	3.0(2)	**1.4**(2)	**1.7**(0.9)	4.8(10)	∞ *250*	0/15
RF1-CMAES	3.2(5)	**2.1**(2)	**2.8**(4)	4.2(1)	∞ *1258*	0/15
RF5-CMAES	**2.8**(2)	**1.2**(1)	**2.0**(3)	81(163)	∞ *1260*	0/15
Sifeg	**3.0**(3)	**1.5**(1)	**1.4**(0.7)	**2.7**(3)	∞ *5e4*	0/15
Sif	**3.0**(3)	**1.5**(1)	**1.4**(0.7)	**2.8**(1)	∞ *5e4*	0/15
Srr	**3.0**(3)	**1.5**(1)	**1.4**(0.9)	**1.9**(1)	∞ *5e4*	0/15

Table 64: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*1.6e+2*:3.0	*1.0e+2*:13	*6.3e+1*:24	*4.0e+1*:55	*1.6e+1*:289	5/5
BSifeg	**2.9**(2)	3.1(8)	10(34)	70(255)	160(181)	9/15
BSif	**2.9**(3)	**1.8**(1)	27(7)	50(220)	228(268)	7/15
BSqi	**2.9**(2)	**2.7**(12)	26(80)	52(43)	184(139)	8/15
BSrr	**2.9**(3)	**2.6**(6)	10(14)	104(142)	237(276)	7/15
CMA-CSA	4.5(5)	**2.3**(1.0)	**2.6**(1.0)	**2.0**(1.0)	**1.4**(0.9)	15/15
CMA-MSR	5.1(6)	**2.3**(3)	**2.4**(2)	**2.0**(1)	**1.0**(0.5)	15/15
CMA-TPA	7.8(14)	**2.8**(3)	**2.5**(1)	**2.2**(0.8)	**1.2**(0.7)	15/15
GP1-CMAES	4.6(3)	**1.9**(1)	**1.9**(1)	**1.4**(0.5)	**1.2**(1)	15/15
GP5-CMAES	3.5(3)	**1.6**(1)	**1.6**(1)	**1.1**(0.4)	3.0(4)	11/15
IPOPCMAv3p	**2.3**(2)	**1.5**(0.6)	**1.6**(1.0)	**1.8**(1)	**1.3**(0.8)	15/15
LHD-10xDef	4.4(5)	**2.9**(3)	3.8(2)	**2.7**(0.5)	**1.3**(2)	9/15
LHD-2xDefa	**2.1**(1)	**1.5**(0.8)	**1.5**(0.5)	**1.5**(0.4)	**1.0**(1)	9/15
RAND-2xDef	3.7(3)	**1.6**(1)	**1.3**(0.6)	**1.4**(1.0)	**1.2**(2)	8/15
RF1-CMAES	5.4(4)	**2.2**(1)	**2.4**(1)	**1.7**(1)	**1.3**(2)	14/15
RF5-CMAES	3.5(5)	**1.6**(0.8)	**2.4**(3)	7.0(9)	5.7(8)	8/15
Sifeg	**2.9**(3)	**1.2**(1.0)	**1.0**(0.4)	**0.90**(0.4)	52(63)	14/15
Sif	**2.9**(3)	**1.2**(0.7)	**1.0**(0.3)	**0.87**(0.7)	51(78)	13/15
Srr	**2.9**(2)	**1.2**(0.9)	**1.0**(0.2)	**0.85**(0.6)	39(56)	14/15

Table 65: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:4.8	*2.5e+1*:16	*1.6e+1*:46	*1.0e+1*:120	*4.0e+0*:334	15/15
BSifeg	**1.7**(2)	**1.4**(0.9)	**0.85**(0.8)	**1.3**(2)	40(46)	15/15
BSif	**1.7**(1)	**1.4**(1.0)	**1.0**(1)	10(35)	32(25)	15/15
BSqi	**1.7**(2)	**1.4**(0.7)	**1.1**(0.4)	**2.5**(6)	22(37)	15/15
BSrr	**1.7**(2)	**1.4**(0.7)	**1.1**(1)	**1.4**(0.8)	30(21)	15/15
CMA-CSA	3.7(4)	**2.9**(5)	3.1(3)	**2.2**(1)	**1.5**(0.7)	15/15
CMA-MSR	**2.0**(1)	**2.2**(2)	**1.9**(2)	5.9(11)	4.6(9)	15/15
CMA-TPA	**2.3**(1)	**2.9**(3)	**3.0**(2)	**1.7**(2)	**1.8**(0.6)	15/15
GP1-CMAES	**1.6**(1)	**1.3**(1)	**0.90**(0.8)	**1.2**(0.9)	**1.4**(2)	13/15
GP5-CMAES	**2.0**(2)	**1.5**(0.4)	**2.7**(4)	**1.3**(2)	**1.9**(2)	13/15
IPOPCMAv3p	**1.4**(0.6)	**1.2**(0.9)	**1.5**(1)	**2.4**(2)	**2.9**(3)	11/15
LHD-10xDef	**1.4**(0.7)	**1.6**(1)	**1.9**(2)	**1.5**(0.5)	**1.4**(0.8)	7/15
LHD-2xDefa	**2.5**(2)	**1.4**(0.6)	**1.6**(1)	**2.2**(2)	11(19)	1/15
RAND-2xDef	**1.3**(0.6)	**1.2**(1)	**1.8**(2)	**1.7**(3)	**1.3**(1)	7/15
RF1-CMAES	**1.3**(1)	**0.90**(1)	**1.3**(2)	**1.8**(2)	**2.1**(2)	11/15
RF5-CMAES	**2.5**(3)	**1.3**(1)	**1.1**(1)	**1.7**(5)	3.2(5)	9/15
Sifeg	**1.7**(1)	**1.6**(2)	**0.89**(0.7)	**0.62**(0.4)	7.2(6)	15/15
Sif	**1.7**(1)	**1.6**(1)	**0.89**(0.4)	**0.69**(0.5)	7.7(6)	15/15
Srr	**1.7**(1)	**1.6**(1)	**0.88**(0.5)	**0.68**(0.6)	3.4(10)	15/15

Table 66: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.0e+1*:5.2	*6.3e+0*:26	*4.0e+0*:57	*2.5e+0*:110	*6.3e-1*:412	15/15
BSifeg	6.3(21)	5.5(14)	174(222)	189(123)	268(547)	5/15
BSif	7.0(3)	298(478)	322(447)	188(116)	157(92)	7/15
BSqi	4.1(4)	139(944)	141(444)	160(229)	150(162)	7/15
BSrr	4.3(3)	6.4(7)	182(447)	137(183)	229(260)	6/15
CMA-CSA	4.2(5)	**1.8**(1)	**1.5**(1.0)	**1.0**(0.4)	**0.61**(0.3)	15/15
CMA-MSR	4.2(2)	**2.0**(2)	**1.5**(0.6)	**1.1**(0.5)	**0.60**(0.1)	15/15
CMA-TPA	24(72)	6.0(3)	3.5(1)	**2.1**(0.7)	**1.5**(3)	15/15
GP1-CMAES	4.5(4)	**1.6**(1)	**1.1**(0.6)	**0.78**(0.1)	**0.45**(0.3)	15/15
GP5-CMAES	3.6(6)	**1.6**(2)	**1.6**(4)	**1.8**(0.1)	**2.5**(3)	10/15
IPOPCMAv3p	4.1(4)	**1.8**(0.9)	**1.3**(0.9)	**1.2**(0.5)	**0.81**(0.9)	14/15
LHD-10xDef	**2.1**(1)	**2.4**(2)	**1.8**(1)	**1.4**(0.4)	4.4(5)	2/15
LHD-2xDefa	**2.4**(2)	**1.4**(0.6)	**1.0**(0.7)	**1.1**(2)	8.9(13)	1/15
RAND-2xDef	**2.3**(3)	**1.2**(1)	**1.4**(0.8)	**1.8**(2)	4.3(4)	2/15
RF1-CMAES	3.0(3)	**2.2**(2)	3.3(7)	3.9(11)	4.1(5)	7/15
RF5-CMAES	4.8(7)	8.6(2)	9.0(11)	10(8)	22(8)	2/15
Sifeg	3.9(2)	10(68)	76(84)	61(206)	109(171)	9/15
Sif	3.9(3)	**1.4**(1)	132(428)	112(142)	112(80)	9/15
Srr	3.9(3)	**1.5**(1)	99(343)	121(360)	191(303)	6/15

Table 67: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*6.3e+1*:3.4	*4.0e+1*:7.2	*2.5e+1*:20	*1.6e+1*:58	*1.6e+0*:318	15/15
BSifeg	**1.5**(3)	**2.2**(2)	75(529)	114(282)	184(219)	8/15
BSif	**1.5**(0.9)	**2.2**(3)	115(414)	120(119)	228(210)	7/15
BSqi	**1.5**(2)	**2.3**(3)	91(3)	90(299)	302(205)	6/15
BSrr	**1.5**(3)	**2.3**(2)	8.4(26)	106(168)	204(290)	7/15
CMA-CSA	4.4(6)	3.9(1)	**2.1**(1)	**1.4**(0.7)	**2.1**(2)	15/15
CMA-MSR	**1.6**(1.0)	**2.0**(2)	**2.0**(1)	**1.3**(0.7)	4.1(12)	15/15
CMA-TPA	**2.5**(1)	**2.7**(3)	**1.8**(2)	**1.1**(0.8)	**1.4**(2)	15/15
GP1-CMAES	**2.4**(4)	3.6(2)	**2.1**(2)	**1.2**(0.5)	**1.2**(1)	14/15
GP5-CMAES	**1.7**(0.9)	**3.0**(3)	5.2(11)	**2.1**(3)	5.8(6)	7/15
IPOPCMAv3p	**1.2**(0.7)	**1.2**(2)	**1.8**(2)	**1.3**(1)	**1.4**(0.3)	14/15
LHD-10xDef	**2.6**(5)	3.4(4)	3.3(3)	**2.0**(0.8)	12(12)	1/15
LHD-2xDefa	**1.3**(2)	**1.6**(2)	**2.2**(2)	**1.3**(0.6)	5.5(7)	2/15
RAND-2xDef	**1.4**(1)	**2.2**(2)	**1.7**(0.9)	**1.5**(1)	12(14)	1/15
RF1-CMAES	**1.5**(3)	**1.7**(2)	**1.3**(1)	**0.89**(0.9)	**2.2**(3)	11/15
RF5-CMAES	**2.2**(2)	**3.0**(4)	9.5(22)	7.6(8)	29(28)	2/15
Sifeg	**1.5**(3)	**2.1**(1)	6.1(0.4)	29(19)	145(192)	9/15
Sif	**1.5**(2)	**2.1**(2)	63(232)	40(145)	118(135)	10/15
Srr	**1.5**(3)	**2.2**(1)	3.6(10)	38(50)	67(191)	11/15

Table 68: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\text{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:172	*1.0e-1*:242	*6.3e-2*:675	*4.0e-2*:3078	*2.5e-2*:4946	15/15
BSifeg	917(1787)	909(830)	1077(864)	∞	∞ *5e4*	0/15
BSif	732(937)	694(668)	∞	∞	∞ *5e4*	0/15
BSqi	969(780)	1440(975)	∞	∞	∞ *5e4*	0/15
BSrr	737(913)	925(869)	∞	∞	∞ *5e4*	0/15
CMA-CSA	**154**(83)	**153**(146)	**70**(55)	**19**(26)	**15**(16)	15/15
CMA-MSR	**222**(95)	**306**(76)	**229**(559)	**139**(167)	**137**(100)	7/15
CMA-TPA	**91**(67)	**84**(68)	**39**(44)	**18**(18)	**14**(13)	15/15
GP1-CMAES	∞	∞	∞	∞	∞ *1260*	0/15
GP5-CMAES	∞	∞	∞	∞	∞ *1262*	0/15
IPOPCMAv3p	∞	∞	∞	∞	∞ *1258*	0/15
LHD-10xDef	∞	∞	∞	∞	∞ *250*	0/15
LHD-2xDefa	∞	∞	∞	∞	∞ *250*	0/15
RAND-2xDef	∞	∞	∞	∞	∞ *250*	0/15
RF1-CMAES	∞	∞	∞	∞	∞ *1258*	0/15
RF5-CMAES	∞	∞	∞	∞	∞ *1262*	0/15
Sifeg	437(515)	477(723)	1039(1787)	∞	∞ *5e4*	0/15
Sif	521(333)	1385(2297)	1021(1558)	∞	∞ *5e4*	0/15
Srr	670(770)	671(1338)	1042(946)	∞	∞ *5e4*	0/15

Table 69: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*6.3e+3*:5.1	*4.0e+3*:8.4	*4.0e+1*:15	*2.5e+0*:69	*1.0e+0*:851	15/15
BSifeg	3.0(2)	**2.0**(1)	**1.9**(1)	**2.2**(3)	9.3(20)	14/15
BSif	3.0(2)	**2.0**(1)	**2.0**(0.5)	**2.1**(2)	23(35)	12/15
BSqi	3.0(1)	**2.0**(0.7)	**1.8**(1)	**2.3**(1)	8.7(27)	15/15
BSrr	3.0(2)	**2.0**(0.9)	**1.8**(2)	**1.5**(2)	11(30)	14/15
CMA-CSA	**2.3**(3)	**2.0**(1)	3.7(0.9)	**2.5**(0.9)	9.2(5)	15/15
CMA-MSR	3.3(5)	**2.4**(3)	4.9(2)	**2.8**(0.8)	1666(1444)	4/15
CMA-TPA	**2.8**(3)	**2.3**(2)	3.8(1)	3.1(2)	17(19)	15/15
GP1-CMAES	**1.9**(2)	**1.8**(1)	**3.0**(2)	4.2(2)	11(4)	2/15
GP5-CMAES	**2.5**(2)	**1.8**(2)	**2.1**(0.9)	**2.1**(1.0)	∞ *1260*	0/15
IPOPCMAv3p	**2.1**(3)	**1.8**(1)	4.1(2)	3.3(2)	21(21)	1/15
LHD-10xDef	**1.8**(2)	**1.4**(0.8)	6.3(4)	11(8)	∞ *250*	0/15
LHD-2xDefa	**2.8**(2)	**2.2**(1)	**2.2**(0.8)	7.6(8)	∞ *250*	0/15
RAND-2xDef	**2.0**(2)	**2.1**(2)	**2.6**(1.0)	3.3(2)	∞ *250*	0/15
RF1-CMAES	**2.9**(5)	**2.4**(3)	3.9(2)	5.6(9)	∞ *1258*	0/15
RF5-CMAES	**2.5**(2)	**1.8**(1)	25(26)	20(24)	∞ *1260*	0/15
Sifeg	3.0(2)	**2.1**(0.3)	**1.8**(0.6)	**1.1**(0.4)	**3.1**(0.6)	15/15
Sif	3.0(2)	**2.1**(1)	**1.8**(0.4)	**1.2**(0.3)	**6.6**(21)	14/15
Srr	3.0(2)	**2.1**(0.9)	**1.8**(1)	**0.83**(0.5)	**2.8**(6)	15/15

Table 70: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*4.0e+1*:3.9	*2.5e+1*:11	*1.6e+1*:31	*6.3e+0*:73	*1.6e+0*:347	5/5
BSifeg	**2.9**(3)	**1.6**(2)	11(0.4)	90(245)	141(232)	8/15
BSif	**2.9**(2)	**1.7**(1)	101(4)	183(253)	290(217)	5/15
BSqi	**2.9**(2)	**2.4**(0.9)	16(58)	85(226)	221(325)	6/15
BSrr	**2.9**(3)	**1.6**(2)	14(50)	73(23)	137(142)	9/15
CMA-CSA	**2.5**(2)	**1.9**(1)	**1.7**(2)	**1.8**(1)	92(2)	14/15
CMA-MSR	3.2(6)	**2.0**(0.8)	**1.7**(1)	3.8(2)	249(589)	13/15
CMA-TPA	**2.4**(2)	**2.1**(2)	**2.0**(0.9)	**2.0**(1.0)	45(148)	15/15
GP1-CMAES	**1.1**(1)	**1.4**(2)	**0.99**(0.5)	4.7(22)	4.9(6)	7/15
GP5-CMAES	**2.5**(2)	**1.7**(1)	**0.89**(0.4)	**1.4**(3)	4.7(10)	8/15
IPOPCMAv3p	**1.8**(2)	**1.6**(1)	**1.7**(2)	10(18)	15(20)	3/15
LHD-10xDef	**1.5**(1)	**1.7**(2)	**1.5**(2)	**2.2**(0.6)	**2.0**(3)	5/15
LHD-2xDefa	**1.7**(2)	**2.1**(2)	**1.3**(0.7)	**1.5**(0.9)	**1.7**(1)	5/15
RAND-2xDef	**1.4**(2)	**1.1**(1)	**1.3**(1.0)	**1.4**(0.9)	**3.3**(3)	3/15
RF1-CMAES	**2.0**(2)	**1.8**(0.9)	**1.4**(2)	4.1(5)	4.8(3)	7/15
RF5-CMAES	**1.8**(2)	**2.6**(2)	**2.9**(7)	7.5(8)	10(5)	5/15
Sifeg	**2.9**(2)	**1.4**(1)	**0.82**(0.6)	110(0.4)	195(505)	7/15
Sif	**2.9**(2)	**1.6**(1)	**0.93**(0.5)	158(341)	150(91)	8/15
Srr	**2.9**(2)	**1.4**(1)	**0.84**(0.5)	73(195)	179(411)	7/15

Table 71: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:3.6	*4.0e+1*:15	*2.5e+1*:32	*1.0e+1*:71	*1.6e+0*:341	5/5
BSifeg	3.5(2)	7.4(46)	15(80)	34(14)	88(126)	10/15
BSif	3.5(4)	5.0(0.6)	114(1)	102(180)	217(183)	7/15
BSqi	3.5(4)	3.8(10)	12(9)	31(43)	59(178)	13/15
BSrr	3.5(2)	5.5(16)	13(0.4)	37(178)	100(155)	10/15
CMA-CSA	**2.7**(2)	**1.6**(1)	**1.4**(0.9)	4.1(6)	151(274)	14/15
CMA-MSR	4.3(6)	**1.8**(2)	5.2(15)	14(16)	113(136)	15/15
CMA-TPA	3.7(3)	**1.9**(1)	**1.2**(0.6)	**2.5**(0.7)	253(236)	12/15
GP1-CMAES	**2.6**(3)	**1.4**(1)	**1.2**(0.9)	3.6(0.1)	10(10)	4/15
GP5-CMAES	**2.9**(2)	7.5(7)	4.6(3)	4.3(6)	11(17)	4/15
IPOPCMAv3p	**2.8**(1)	**2.2**(3)	**1.7**(1)	5.8(9)	6.3(9)	6/15
LHD-10xDef	**2.2**(2)	**2.0**(0.8)	**2.3**(2)	**1.9**(0.5)	**2.5**(2)	4/15
LHD-2xDefa	3.3(3)	**1.8**(2)	**1.3**(0.8)	**1.4**(2)	**1.8**(3)	5/15
RAND-2xDef	**1.8**(2)	**1.5**(0.9)	**0.98**(0.6)	**0.79**(0.6)	**2.4**(6)	4/15
RF1-CMAES	3.1(5)	**2.0**(2)	**1.3**(1)	5.5(13)	3.2(5)	9/15
RF5-CMAES	**2.1**(4)	4.7(27)	**2.9**(2)	7.3(3)	24(16)	2/15
Sifeg	3.4(3)	**1.8**(2)	115(0.7)	123(212)	116(262)	9/15
Sif	3.4(3)	**1.7**(1)	5.8(18)	117(379)	125(83)	9/15
Srr	3.4(2)	**1.6**(1)	**2.5**(0.8)	62(177)	75(98)	11/15

Table 72: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*1.0e+1*:3.0	*6.3e+0*:9.0	*4.0e+0*:33	*2.5e+0*:84	*1.0e+0*:518	15/15
BSifeg	**2.6**(2)	**2.5**(2)	**1.5**(0.8)	**1.7**(1.0)	4.3(5)	15/15
BSif	**2.6**(3)	**2.5**(1)	**1.6**(2)	**1.9**(3)	3.3(5)	15/15
BSqi	**2.6**(2)	**2.4**(2)	**1.9**(3)	**2.0**(2)	6.6(6)	15/15
BSrr	**2.6**(2)	**2.5**(2)	**1.6**(1)	**2.0**(2)	3.7(6)	15/15
CMA-CSA	**2.3**(2)	3.2(3)	5.9(5)	8.9(8)	13(18)	15/15
CMA-MSR	**2.5**(4)	3.7(10)	3.1(3)	6.0(3)	3.2(4)	15/15
CMA-TPA	3.2(2)	3.8(5)	3.4(3)	12(5)	16(12)	15/15
GP1-CMAES	**1.9**(3)	**2.7**(2)	3.2(3)	6.5(4)	4.9(2)	6/15
GP5-CMAES	**2.4**(4)	**2.0**(2)	**2.4**(2)	3.0(4)	**2.2**(4)	11/15
IPOPCMAv3p	**2.3**(4)	**2.2**(3)	3.4(1)	3.8(3)	12(29)	3/15
LHD-10xDef	3.9(5)	3.4(3)	3.3(2)	4.3(8)	6.8(5)	1/15
LHD-2xDefa	3.1(4)	3.8(2)	3.8(3)	10(11)	∞ *250*	0/15
RAND-2xDef	**2.5**(1)	**2.7**(2)	**2.1**(2)	4.6(4)	7.1(7)	1/15
RF1-CMAES	**1.8**(2)	**2.0**(2)	**3.0**(2)	8.7(13)	∞ *1260*	0/15
RF5-CMAES	**2.4**(2)	**1.5**(3)	3.3(2)	4.0(3)	∞ *1288*	0/15
Sifeg	3.4(9)	**2.9**(3)	**1.7**(1)	**2.8**(2)	**2.7**(2)	15/15
Sif	3.4(2)	3.2(3)	**1.8**(1)	3.0(3)	**2.8**(3)	15/15
Srr	3.4(2)	3.2(3)	**2.3**(3)	3.4(3)	**2.5**(3)	15/15

Table 73: 05-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*6.3e+1*:15	*4.0e+1*:37	*4.0e+1*:37	*2.5e+1*:118	*1.6e+1*:692	15/15
BSifeg	**1.6**(0.4)	3.8(2)	3.8(6)	32(14)	20(19)	14/15
BSif	**1.6**(0.9)	**2.5**(7)	**2.5**(4)	37(52)	17(11)	14/15
BSqi	**2.8**(0.3)	**2.9**(3)	**2.9**(7)	31(3)	18(18)	14/15
BSrr	**1.5**(0.6)	3.3(2)	3.3(0.6)	41(100)	32(77)	11/15
CMA-CSA	**2.0**(1)	**2.1**(2)	**2.1**(0.9)	**1.5**(2)	**1.6**(2)	15/15
CMA-MSR	**2.2**(2)	**2.6**(0.6)	**2.6**(1)	**1.8**(0.8)	**0.93**(2)	15/15
CMA-TPA	**1.9**(1)	**1.9**(0.8)	**1.9**(1)	**1.9**(1.0)	**1.4**(1)	15/15
GP1-CMAES	**1.7**(1)	**1.7**(0.5)	**1.7**(0.9)	**1.3**(1.0)	**1.1**(1.0)	13/15
GP5-CMAES	**1.5**(1.0)	**1.1**(0.4)	**1.1**(0.4)	**2.5**(2)	**1.6**(2)	9/15
IPOPCMAv3p	**2.0**(1)	**2.3**(2)	**2.3**(1)	**1.6**(1)	**1.1**(0.7)	12/15
LHD-10xDef	**2.7**(2)	6.5(2)	6.5(9)	4.8(3)	∞ *250*	0/15
LHD-2xDefa	**1.9**(2)	3.4(4)	3.4(5)	∞	∞ *250*	0/15
RAND-2xDef	**1.7**(2)	7.3(9)	7.3(6)	15(16)	∞ *250*	0/15
RF1-CMAES	**1.9**(1)	**2.4**(3)	**2.4**(3)	**2.4**(4)	**1.9**(2)	9/15
RF5-CMAES	**1.8**(2)	3.6(6)	3.6(6)	4.1(4)	**2.9**(2)	7/15
Sifeg	**1.8**(2)	**1.7**(2)	**1.7**(2)	13(2)	4.8(2)	15/15
Sif	**1.7**(1)	**1.7**(3)	**1.7**(2)	**2.9**(2)	4.8(11)	15/15
Srr	**1.8**(0.6)	**1.6**(1)	**1.6**(1)	7.3(3)	5.8(1.0)	15/15

Table 74: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*4.0e+1*:8.0	*2.5e+1*:16	*1.0e-8*:23	*1.0e-8*:23	*1.0e-8*:23	15/15
BSifeg	**3.2**(1.0)	**2.0**(0.2)	**2.4**(0.3)	**2.4**(0.3)	**2.4**(0.2)	15/15
BSif	**3.2**(0.9)	**2.0**(0.2)	**2.4**(0.2)	**2.4**(0.2)	**2.4**(0.2)	15/15
BSqi	**3.2**(0.8)	**2.0**(0.2)	**2.4**(0.2)	**2.4**(0.3)	**2.4**(0.2)	15/15
BSrr	**3.2**(0.7)	**2.0**(0.1)	**2.4**(0.2)	**2.4**(0.2)	**2.4**(0.2)	15/15
CMA-CSA	5.2(4)	3.7(1)	63(5)	63(7)	63(5)	15/15
CMA-MSR	4.4(2)	5.2(2)	87(5)	87(10)	87(5)	15/15
CMA-TPA	7.5(6)	5.5(3)	51(5)	51(4)	51(6)	15/15
GP1-CMAES	3.6(3)	**2.6**(0.7)	43(7)	43(8)	43(8)	15/15
GP5-CMAES	**3.0**(2)	**1.9**(1.0)	404(275)	404(741)	404(248)	4/15
IPOPCMAv3p	5.0(4)	3.9(2)	65(4)	65(3)	65(2)	15/15
LHD-10xDef	9.4(9)	12(5)	∞	∞	∞ *500*	0/15
LHD-2xDefa	5.0(2)	**2.9**(1)	∞	∞	∞ *500*	0/15
RAND-2xDef	4.2(2)	**2.9**(1.0)	∞	∞	∞ *500*	0/15
RF1-CMAES	4.3(2)	3.2(2)	522(508)	522(673)	522(761)	3/15
RF5-CMAES	3.9(2)	**2.7**(1)	∞	∞	∞ *2514*	0/15
Sifeg	**3.2**(1)	**2.0**(0.2)	12(2)	12(2)	12(1)	15/15
Sif	**3.2**(1)	**2.0**(0.2)	11(2)	11(2)	11(2)	15/15
Srr	**3.2**(0.9)	**2.0**(0.5)	11(2)	11(2)	11(2)	15/15

Table 75: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*2.5e+6*:5.6	*1.0e+6*:17	*1.0e+5*:33	*2.5e+3*:118	*1.0e-8*:196	15/15
BSifeg	5.4(4)	**2.4**(0.9)	**1.5**(0.3)	**0.54**(0.2)	**1.2**(0.2)	15/15
BSif	5.4(2)	**2.4**(1)	**1.5**(0.1)	**0.56**(0.1)	**1.2**(0.3)	15/15
BSqi	5.4(4)	**2.4**(0.9)	**1.5**(0.3)	**0.51**(0.1)	**1.2**(0.2)	15/15
BSrr	5.4(4)	**2.4**(1)	**1.5**(0.2)	**0.54**(0.2)	**1.2**(0.2)	15/15
CMA-CSA	**2.3**(2)	**2.1**(3)	7.4(3)	8.5(4)	23(1.0)	15/15
CMA-MSR	3.2(2)	**3.0**(2)	7.3(2)	8.7(4)	26(2)	15/15
CMA-TPA	**1.8**(2)	**1.8**(1)	5.8(3)	7.4(2)	23(1)	15/15
GP1-CMAES	**1.4**(1)	**1.1**(0.6)	4.8(3)	7.5(2)	∞ *2502*	0/15
GP5-CMAES	**2.8**(3)	**1.9**(2)	3.5(1)	3.2(0.8)	∞ *2502*	0/15
IPOPCMAv3p	**2.5**(3)	**2.0**(3)	6.0(3)	10(3)	∞ *2502*	0/15
LHD-10xDef	**1.5**(0.8)	**1.2**(0.8)	9.0(3)	∞	∞ *500*	0/15
LHD-2xDefa	**1.5**(2)	**1.5**(1)	5.5(2)	∞	∞ *500*	0/15
RAND-2xDef	**1.8**(2)	**1.8**(1)	5.1(4)	∞	∞ *500*	0/15
RF1-CMAES	**2.3**(2)	**2.4**(2)	7.7(7)	58(30)	∞ *2502*	0/15
RF5-CMAES	**2.4**(3)	**2.3**(1)	34(35)	317(661)	∞ *2504*	0/15
Sifeg	5.5(4)	**2.6**(0.5)	**1.8**(0.3)	**0.69**(0.1)	**1.7**(0.2)	15/15
Sif	5.5(4)	**2.6**(0.9)	**1.8**(0.3)	**0.73**(0.1)	**1.7**(0.2)	15/15
Srr	5.5(4)	**2.6**(0.9)	**1.8**(0.2)	**0.68**(0.0)	**1.7**(0.2)	15/15

Table 76: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*4.0e+2*:8.2	*1.6e+2*:37	*1.0e+2*:69	*6.3e+1*:147	*2.5e+1*:1129	15/15
BSifeg	**2.9**(0.5)	**1.1**(0.1)	**0.69**(0.1)	**0.41**(0.1)	**0.11**(0.0)	15/15
BSif	**2.9**(0.9)	**1.1**(0.1)	**0.69**(0.1)	**0.43**(0.2)	**0.12**(0.0)	15/15
BSqi	**2.9**(0.9)	**1.1**(0.1)	**0.69**(0.1)	**0.41**(0.1)	**0.10**(0.0)	15/15
BSrr	**2.9**(0.8)	**1.1**(0.0)	**0.69**(0.1)	**0.42**(0.1)	**0.09**(0.0)	15/15
CMA-CSA	**2.6**(2)	**2.8**(2)	3.1(1)	**2.8**(0.7)	**1.4**(1)	15/15
CMA-MSR	**2.3**(2)	3.3(1)	3.1(1)	**2.8**(0.6)	**1.4**(2)	15/15
CMA-TPA	3.2(4)	3.3(2)	**2.8**(0.7)	**2.4**(1)	**1.1**(1)	15/15
GP1-CMAES	**1.8**(2)	**2.2**(0.8)	**2.2**(0.7)	**2.5**(1)	**1.8**(3)	11/15
GP5-CMAES	**1.9**(2)	**1.7**(0.6)	**1.6**(0.6)	6.6(7)	32(68)	1/15
IPOPCMAv3p	**2.8**(3)	**2.7**(1)	3.1(2)	3.4(1)	**1.3**(0.3)	13/15
LHD-10xDef	**2.3**(3)	5.7(0.1)	3.7(0.4)	**2.4**(0.2)	**2.1**(2)	3/15
LHD-2xDefa	**1.8**(2)	**2.2**(1)	**2.5**(1)	3.4(4)	∞ *500*	0/15
RAND-2xDef	**1.1**(0.7)	**1.8**(0.4)	**1.7**(0.3)	**1.8**(0.7)	**2.0**(2)	3/15
RF1-CMAES	**1.7**(2)	**1.9**(1)	**2.1**(0.7)	**2.1**(0.7)	**1.6**(2)	10/15
RF5-CMAES	**2.5**(1)	4.6(0.7)	5.4(6)	8.5(8)	31(34)	1/15
Sifeg	**2.9**(0.9)	**1.1**(0.1)	**0.74**(0.1)	**0.46**(0.1)	**0.14**(0.0)	15/15
Sif	**2.9**(0.5)	**1.1**(0.1)	**0.74**(0.1)	**0.48**(0.1)	**0.14**(0.0)	15/15
Srr	**2.9**(0.6)	**1.1**(0.1)	**0.74**(0.1)	**0.46**(0.1)	**0.14**(0.0)	15/15

Table 77: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*2.5e+2*:21	*1.6e+2*:59	*1.6e+2*:59	*6.3e+1*:139	*4.0e+1*:854	15/15
BSifeg	**1.6**(1)	**0.86**(0.4)	**0.86**(0.4)	**0.89**(0.3)	**0.20**(0.1)	15/15
BSif	**1.6**(1)	**0.86**(0.2)	**0.86**(0.2)	**0.97**(0.2)	**0.20**(0.0)	15/15
BSqi	**1.6**(0.9)	**0.86**(0.3)	**0.86**(0.1)	**0.92**(0.5)	**0.23**(0.1)	15/15
BSrr	**1.6**(0.6)	**0.86**(0.4)	**0.86**(0.5)	**0.85**(0.2)	**0.20**(0.1)	15/15
CMA-CSA	4.6(2)	**2.9**(1)	**2.9**(1)	4.3(2)	**1.3**(3)	15/15
CMA-MSR	5.3(3)	3.4(1)	3.4(0.9)	5.2(0.4)	**2.2**(1)	15/15
CMA-TPA	5.3(4)	**3.0**(1)	**3.0**(0.9)	5.4(0.7)	**2.3**(2)	15/15
GP1-CMAES	4.3(2)	3.5(1)	3.5(1)	11(6)	4.2(4)	8/15
GP5-CMAES	3.8(3)	3.1(3)	3.1(3)	∞	∞ *2516*	0/15
IPOPCMAv3p	4.4(4)	3.2(1)	3.2(1)	5.2(3)	**1.6**(0.3)	13/15
LHD-10xDef	12(4)	8.2(7)	8.2(6)	∞	∞ *500*	0/15
LHD-2xDefa	4.4(3)	3.8(2)	3.8(1)	53(72)	∞ *500*	0/15
RAND-2xDef	4.2(2)	3.2(2)	3.2(1)	12(29)	8.7(11)	1/15
RF1-CMAES	4.8(2)	3.2(2)	3.2(2)	13(14)	42(111)	1/15
RF5-CMAES	17(46)	24(17)	24(19)	267(238)	∞ *2504*	0/15
Sifeg	**1.7**(0.8)	**0.94**(0.3)	**0.94**(0.3)	**0.82**(0.2)	**0.17**(0.0)	15/15
Sif	**1.7**(1)	**0.94**(0.2)	**0.94**(0.4)	**0.82**(0.3)	**0.16**(0.0)	15/15
Srr	**1.7**(0.4)	**0.95**(0.2)	**0.95**(0.2)	**0.79**(0.1)	**0.16**(0.0)	15/15

Table 78: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*1.0e+2*:16	*6.3e+1*:19	*1.0e-8*:20	*1.0e-8*:20	*1.0e-8*:20	15/15
BSifeg	**1.5**(0.2)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSif	**1.5**(0.2)	**1.4**(0.2)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSqi	**1.5**(0.2)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSrr	**1.5**(0.2)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
CMA-CSA	**1.8**(2)	**2.4**(1)	6.1(1)	6.1(2)	6.1(1)	15/15
CMA-MSR	**1.6**(1)	**2.0**(1)	5.4(2)	5.4(1)	5.4(2)	15/15
CMA-TPA	**1.3**(0.8)	**2.1**(1)	5.0(2)	5.0(2)	5.0(1)	15/15
GP1-CMAES	**1.4**(0.8)	**1.9**(0.3)	42(26)	42(39)	42(30)	15/15
GP5-CMAES	**1.3**(0.9)	**1.6**(0.5)	5.4(4)	5.4(3)	5.4(2)	15/15
IPOPCMAv3p	**1.6**(1)	**2.8**(0.6)	29(11)	29(9)	29(6)	15/15
LHD-10xDef	4.5(6)	11(1.0)	12(0.3)	12(0.4)	12(0.2)	15/15
LHD-2xDefa	**2.4**(0.8)	**2.3**(0.3)	**3.0**(0.1)	**3.0**(0.2)	**3.0**(0.2)	15/15
RAND-2xDef	**2.6**(0.2)	**2.4**(0.1)	3.1(0.3)	3.1(0.2)	3.1(0.2)	15/15
RF1-CMAES	**1.7**(0.9)	**2.4**(1)	35(28)	35(19)	35(17)	15/15
RF5-CMAES	**1.4**(1)	**2.2**(0.9)	120(198)	120(86)	120(81)	10/15
Sifeg	**1.5**(0.3)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
Sif	**1.5**(0.2)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
Srr	**1.5**(0.2)	**1.4**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15

Table 79: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*1.6e+5*:7.0	*6.3e+4*:16	*4.0e+2*:36	*1.0e+2*:102	*4.0e+0*:504	15/15
BSifeg	**1.6**(2)	**1.4**(1)	**1.2**(0.8)	38(23)	1260(1610)	2/15
BSif	**1.6**(2)	**1.4**(1.0)	**1.4**(0.4)	64(103)	862(1268)	3/15
BSqi	**1.6**(1)	**1.4**(0.6)	**1.2**(0.8)	30(86)	255(280)	7/15
BSrr	**1.6**(1)	**1.4**(0.9)	**1.2**(0.9)	42(24)	694(177)	3/15
CMA-CSA	**1.6**(1)	**1.5**(1)	**1.8**(2)	3.6(1)	**1.8**(0.3)	15/15
CMA-MSR	**2.7**(4)	**1.8**(3)	**2.2**(1)	**2.8**(0.7)	**1.6**(0.4)	15/15
CMA-TPA	**2.4**(3)	**2.2**(1)	4.8(3)	3.8(1)	**1.8**(0.3)	15/15
GP1-CMAES	**2.6**(2)	**1.8**(1)	**2.2**(1)	**2.1**(0.5)	4.8(4)	11/15
GP5-CMAES	**2.4**(3)	**1.5**(1)	**1.3**(0.3)	5.7(8)	∞ *2516*	0/15
IPOPCMAv3p	**2.7**(2)	**1.9**(4)	**2.8**(3)	3.0(2)	**1.7**(0.6)	15/15
LHD-10xDef	**1.9**(3)	4.1(2)	5.8(0.3)	7.2(8)	∞ *500*	0/15
LHD-2xDefa	**2.1**(2)	**1.8**(1)	**1.5**(0.6)	6.3(9)	∞ *500*	0/15
RAND-2xDef	**2.0**(3)	**1.8**(1)	**1.5**(0.2)	6.2(5)	∞ *500*	0/15
RF1-CMAES	**1.7**(2)	**1.4**(1)	**1.9**(0.9)	**2.7**(2)	71(106)	1/15
RF5-CMAES	**2.4**(2)	**1.7**(3)	5.9(3)	39(47)	∞ *2518*	0/15
Sifeg	**1.6**(2)	**1.4**(1)	**1.4**(0.3)	5.9(7)	118(49)	11/15
Sif	**1.6**(2)	**1.4**(0.9)	**1.4**(0.2)	8.6(11)	300(388)	7/15
Srr	**1.6**(1)	**1.4**(1)	**1.3**(0.3)	8.1(5)	110(53)	11/15

Table 80: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*2.5e+2*:9.2	*1.6e+2*:18	*1.0e+2*:33	*1.0e+1*:172	*4.0e+0*:678	15/15
BSifeg	**2.2**(2)	**1.5**(0.9)	14(32)	1100(1642)	2095(2465)	1/15
BSif	**2.2**(2)	**1.5**(0.9)	14(16)	882(1047)	2193(1494)	1/15
BSqi	**2.2**(2)	**1.5**(1)	14(16)	931(1437)	1037(898)	2/15
BSrr	**2.2**(2)	**1.5**(0.8)	14(31)	894(1247)	∞ *1e5*	0/15
CMA-CSA	4.2(5)	3.3(2)	**2.9**(1)	**2.3**(0.8)	**1.1**(0.8)	15/15
CMA-MSR	3.5(2)	3.1(2)	3.2(1)	**1.9**(0.5)	**0.83**(0.9)	15/15
CMA-TPA	**2.8**(2)	**2.8**(2)	**2.7**(2)	**1.7**(0.6)	**1.2**(1.0)	15/15
GP1-CMAES	3.2(2)	**2.3**(1)	**2.1**(0.9)	**1.6**(0.7)	**1.1**(1)	15/15
GP5-CMAES	**2.3**(2)	**1.7**(1)	**1.4**(0.6)	**1.0**(0.3)⋆	**0.79**(0.4)	15/15
IPOPCMAv3p	**2.8**(3)	**3.0**(3)	**2.5**(1)	**2.6**(3)	**0.85**(1)	15/15
LHD-10xDef	**2.4**(3)	4.5(4)	5.4(3)	10(5)	∞ *500*	0/15
LHD-2xDefa	**2.3**(3)	**2.4**(2)	**2.3**(0.3)	43(20)	∞ *500*	0/15
RAND-2xDef	3.2(3)	3.0(1)	**2.6**(1)	13(9)	∞ *500*	0/15
RF1-CMAES	4.2(3)	3.6(3)	**3.0**(2)	13(26)	12(12)	4/15
RF5-CMAES	3.3(2)	**2.9**(2)	3.2(5)	31(47)	28(21)	2/15
Sifeg	**2.2**(2)	**1.6**(0.9)	5.5(0.3)	281(354)	363(508)	5/15
Sif	**2.2**(2)	**1.6**(0.9)	5.5(17)	166(150)	444(486)	4/15
Srr	**2.2**(2)	**1.6**(1)	5.5(0.7)	217(540)	220(211)	7/15

Table 81: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.6e+4*:15	*1.0e+4*:22	*1.6e+3*:34	*2.5e+2*:103	*4.0e+0*:727	15/15
BSifeg	**1.9**(0.1)	**1.5**(0.1)	**1.2**(0.1)	**1.3**(0.4)	51(29)	13/15
BSif	**1.9**(0.9)	**1.5**(0.9)	**1.2**(0.1)	**0.93**(0.1)	123(268)	10/15
BSqi	**1.9**(0.9)	**1.5**(0.5)	**1.2**(0.1)	**0.99**(2)	30(40)	15/15
BSrr	**1.9**(1)	**1.5**(0.4)	**1.2**(0.1)	**0.68**(0.4)	35(44)	13/15
CMA-CSA	4.2(5)	3.6(4)	5.0(3)	3.6(2)	**4.9**(4)	15/15
CMA-MSR	4.4(3)	3.9(3)	6.0(2)	3.3(1)	**5.3**(12)	15/15
CMA-TPA	3.8(4)	3.3(2)	4.0(1)	**2.2**(0.8)	5.4(5)	15/15
GP1-CMAES	**2.9**(2)	**2.5**(1)	3.1(0.4)	**2.3**(1)	12(16)	4/15
GP5-CMAES	**2.3**(0.9)	**1.7**(0.7)	**2.2**(0.7)	**1.8**(0.5)	∞ *2516*	0/15
IPOPCMAv3p	3.2(3)	**3.0**(3)	4.4(2)	**2.5**(0.6)	**4.1**(2)	11/15
LHD-10xDef	8.1(7)	7.1(4)	6.8(0.4)	3.2(0.5)	∞ *500*	0/15
LHD-2xDefa	3.3(0.3)	**2.3**(0.4)	**2.5**(0.4)	**2.5**(0.6)	∞ *500*	0/15
RAND-2xDef	3.3(0.2)	**2.5**(0.5)	**2.6**(0.9)	**2.4**(0.6)	∞ *500*	0/15
RF1-CMAES	4.1(2)	3.4(1)	4.3(2)	3.3(5)	∞ *2502*	0/15
RF5-CMAES	**2.9**(2)	**2.4**(1)	10(1)	15(18)	∞ *2514*	0/15
Sifeg	**1.9**(1)	**1.5**(0.1)	**1.3**(0.1)	**0.66**(0.2)	26(43)	15/15
Sif	**1.9**(1)	**1.5**(0.5)	**1.3**(0.1)	**0.68**(0.3)	31(66)	14/15
Srr	**1.9**(1)	**1.5**(0.4)	**1.3**(0.1)	**0.65**(0.1)	39(125)	13/15

Table 82: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*4.0e+1*:125	*2.5e+1*:148	*1.6e+1*:180	*1.0e+1*:200	*1.6e+0*:563	15/15
BSifeg	59(140)	53(131)	46(17)	47(169)	1100(1053)	2/15
BSif	239(371)	267(723)	239(393)	247(362)	2383(4152)	1/15
BSqi	47(120)	43(24)	38(81)	37(140)	1041(1197)	2/15
BSrr	47(38)	43(10)	37(24)	37(19)	1006(1142)	2/15
CMA-CSA	**3.5**(2)	**3.4**(2)	**3.1**(0.4)	**3.3**(0.5)	**5.2**(1)	15/15
CMA-MSR	4.5(2)	4.2(1.0)	3.8(1)	4.0(0.5)	**5.8**(4)	15/15
CMA-TPA	4.6(7)	**4.1**(4)	**3.6**(8)	**3.7**(2)	**5.0**(3)	15/15
GP1-CMAES	**4.3**(4)	4.2(0.6)	3.9(2)	4.3(1.0)	66(29)	1/15
GP5-CMAES	28(11)	71(76)	93(137)	84(139)	∞ *2526*	0/15
IPOPCMAv3p	**3.8**(2)	**3.5**(2)	**3.2**(0.4)	**3.5**(2)	33(101)	2/15
LHD-10xDef	19(16)	50(54)	42(31)	∞	∞ *500*	0/15
LHD-2xDefa	11(9)	12(7)	∞	∞	∞ *500*	0/15
RAND-2xDef	14(15)	25(54)	∞	∞	∞ *500*	0/15
RF1-CMAES	14(13)	15(19)	13(14)	16(16)	∞ *2502*	0/15
RF5-CMAES	∞	∞	∞	∞	∞ *2504*	0/15
Sifeg	69(179)	59(151)	49(126)	46(114)	2334(2673)	1/15
Sif	52(180)	45(151)	38(125)	37(3)	1159(1546)	2/15
Srr	33(53)	28(55)	24(47)	22(4)	2211(2228)	1/15

Table 83: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*2.5e+6*:6.0	*1.0e+6*:21	*4.0e+5*:38	*2.5e+4*:104	*6.3e+2*:512	15/15
BSifeg	**1.9**(3)	**1.2**(1)	7.7(50)	481(363)	∞ *7e4*	0/15
BSif	**1.9**(2)	**1.2**(0.9)	7.5(25)	413(1203)	∞ *7e4*	0/15
BSqi	**1.9**(2)	**1.2**(1)	4.5(26)	438(558)	∞ *8e4*	0/15
BSrr	**1.9**(2)	**1.2**(1.0)	3.6(0.7)	334(281)	∞ *5e4*	0/15
CMA-CSA	**2.4**(2)	**1.1**(0.7)	**2.5**(3)	5.0(1)	**2.8**(0.6)	15/15
CMA-MSR	**2.8**(2)	**1.6**(1)	**2.2**(1)	3.5(0.9)	**2.7**(0.3)	15/15
CMA-TPA	**1.2**(2)	**1.4**(1)	**1.9**(1)	**3.5**(1.0)	**2.4**(0.7)	15/15
GP1-CMAES	**2.0**(2)	**1.2**(2)	**1.2**(0.9)	**2.9**(2)	**2.2**(0.6)	15/15
GP5-CMAES	3.2(3)	**1.5**(0.9)	**1.2**(0.9)	**1.8**(1)	**1.0**(0.5)	15/15
IPOPCMAv3p	**1.9**(1)	**1.2**(2)	**2.3**(2)	4.1(2)	3.1(0.7)	15/15
LHD-10xDef	**2.4**(4)	**1.6**(2)	3.7(2)	11(7)	∞ *500*	0/15
LHD-2xDefa	**2.1**(2)	**1.5**(1)	**1.9**(0.9)	6.6(7)	∞ *500*	0/15
RAND-2xDef	**1.5**(1)	**1.1**(1)	**1.4**(1)	12(22)	∞ *500*	0/15
RF1-CMAES	**1.4**(0.9)	**1.3**(1)	**1.7**(2)	7.6(8)	∞ *2502*	0/15
RF5-CMAES	**2.2**(2)	**1.2**(0.6)	**2.3**(2)	61(79)	∞ *2504*	0/15
Sifeg	**1.9**(1)	**1.1**(0.9)	**1.1**(0.4)	117(244)	∞ *4e4*	0/15
Sif	**1.9**(3)	**1.1**(0.8)	**1.1**(0.4)	139(186)	∞ *4e4*	0/15
Srr	**1.9**(3)	**1.1**(0.6)	**1.1**(0.4)	50(13)	∞ *3e4*	0/15

Table 84: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*4.0e+4*:6.4	*2.5e+3*:15	*6.3e+1*:217	*4.0e+1*:244	*2.5e+0*:675	15/15
BSifeg	**2.4**(1)	**2.2**(0.9)	1105(1450)	∞	∞ *7e4*	0/15
BSif	**2.4**(2)	**2.2**(1)	339(361)	2153(1833)	∞ *8e4*	0/15
BSqi	**2.4**(1)	**2.2**(1)	363(529)	1107(417)	∞ *8e4*	0/15
BSrr	**2.4**(1)	**2.2**(0.9)	503(409)	1283(718)	∞ *4e4*	0/15
CMA-CSA	**1.5**(1.0)	4.0(3)	7.1(1)	6.6(0.6)	**2.9**(0.2)	15/15
CMA-MSR	3.8(4)	4.2(5)	6.5(0.9)	**6.3**(1)	**3.3**(0.2)	15/15
CMA-TPA	**2.1**(2)	3.1(2)	**5.8**(2)	**5.9**(1)	**3.0**(0.4)	15/15
GP1-CMAES	3.1(3)	3.3(2)	**6.4**(3)	7.8(1)	55(65)	1/15
GP5-CMAES	**2.9**(2)	**2.9**(2)	**3.5**(2)	**5.1**(1)	4.3(1)	11/15
IPOPCMAv3p	3.6(6)	4.1(3)	11(13)	18(10)	∞ *2502*	0/15
LHD-10xDef	5.3(7)	5.3(4)	∞	∞	∞ *500*	0/15
LHD-2xDefa	4.3(3)	3.7(3)	16(13)	∞	∞ *500*	0/15
RAND-2xDef	3.6(4)	3.9(3)	33(49)	∞	∞ *500*	0/15
RF1-CMAES	3.2(4)	4.3(3)	19(12)	153(192)	∞ *2502*	0/15
RF5-CMAES	**2.4**(3)	3.2(2)	55(32)	154(201)	∞ *2514*	0/15
Sifeg	**2.4**(2)	**2.3**(1)	225(495)	1601(2467)	∞ *6e4*	0/15
Sif	**2.4**(2)	**2.3**(2)	244(228)	3570(4531)	∞ *6e4*	0/15
Srr	**2.4**(1)	**2.3**(1)	208(235)	577(785)	∞ *3e4*	0/15

Table 85: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*4.0e+7*:15	*2.5e+7*:24	*1.6e+7*:34	*1.0e+6*:103	*1.0e+1*:515	15/15
BSifeg	**2.2**(0.7)	**1.6**(0.3)	**1.3**(0.2)	**3.0**(11)	262(255)	5/15
BSif	**2.2**(0.1)	**1.6**(0.3)	**1.3**(0.3)	3.3(7)	353(240)	4/15
BSqi	**2.2**(0.6)	**1.6**(0.3)	**1.3**(0.1)	6.1(0.2)	42(61)	14/15
BSrr	**2.2**(0.3)	**1.6**(0.3)	**1.3**(0.3)	**2.8**(0.4)	62(108)	11/15
CMA-CSA	3.4(3)	3.3(2)	**2.7**(1)	**2.9**(1)	4.2(5)	15/15
CMA-MSR	3.9(3)	3.8(2)	3.7(1)	3.3(0.6)	4.9(4)	15/15
CMA-TPA	3.7(4)	4.2(3)	3.6(2)	**2.5**(0.5)	**3.6**(3)	15/15
GP1-CMAES	**2.3**(2)	**2.6**(1)	**3.0**(2)	**2.9**(3)	**2.9**(3)	13/15
GP5-CMAES	6.8(4)	13(9)	17(8)	28(34)	21(13)	3/15
IPOPCMAv3p	**3.0**(2)	3.2(2)	3.3(2)	**2.8**(0.7)	**2.8**(4)	13/15
LHD-10xDef	6.7(4)	8.6(2)	6.9(1)	6.2(8)	∞ *500*	0/15
LHD-2xDefa	3.6(2)	**2.8**(0.6)	**2.4**(0.8)	**2.6**(1)	∞ *500*	0/15
RAND-2xDef	**2.4**(2)	**2.3**(0.7)	**2.1**(1)	**2.2**(1)	∞ *500*	0/15
RF1-CMAES	4.4(4)	3.9(2)	3.1(1)	**2.5**(1)	4.2(3)	11/15
RF5-CMAES	3.5(4)	3.1(2)	3.2(3)	21(19)	∞ *2504*	0/15
Sifeg	**2.2**(0.4)	**1.6**(0.1)	**1.2**(0.2)	7.3(17)	74(104)	5/15
Sif	**2.2**(0.3)	**1.6**(0.2)	**1.2**(0.2)	4.1(6)	99(113)	4/15
Srr	**2.2**(0.8)	**1.6**(0.3)	**1.2**(0.0)	3.6(4)	20(33)	10/15

Table 86: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*1.0e+3*:12	*6.3e+2*:32	*4.0e+2*:40	*6.3e+1*:154	*2.5e+0*:521	15/15
BSifeg	**2.9**(0.8)	**1.4**(0.4)	3.4(0.1)	45(68)	163(104)	11/15
BSif	**2.9**(0.5)	**1.4**(0.2)	**1.6**(0.2)	88(7)	∞ *9e4*	0/15
BSqi	**2.9**(0.7)	**1.4**(0.3)	**1.6**(0.2)	34(96)	207(193)	8/15
BSrr	**2.9**(1.0)	**1.4**(0.1)	**1.5**(0.2)	35(52)	208(202)	8/15
CMA-CSA	4.5(2)	3.6(0.5)	4.6(2)	**3.0**(0.3)	**3.3**(2)	15/15
CMA-MSR	5.7(3)	4.3(2)	5.9(1)	3.4(0.8)	**2.7**(0.4)	15/15
CMA-TPA	6.5(3)	4.0(1)	4.7(0.9)	**2.7**(0.5)	3.6(2)	15/15
GP1-CMAES	3.7(1)	**2.5**(0.6)	3.1(0.5)	**1.9**(0.5)	5.9(4)	8/15
GP5-CMAES	**2.6**(1)	**1.6**(0.3)	**1.7**(0.4)	**0.91**(0.1)	3.9(3)	10/15
IPOPCMAv3p	4.5(3)	3.5(3)	5.0(1)	3.2(0.6)	5.3(4)	9/15
LHD-10xDef	14(7)	7.0(0.2)	5.8(0.2)	**2.2**(0.2)	∞ *500*	0/15
LHD-2xDefa	3.7(1)	**2.2**(0.7)	**2.5**(0.8)	**1.6**(0.4)	**3.4**(3)	4/15
RAND-2xDef	3.6(2)	**2.1**(0.5)	**2.2**(0.6)	**1.6**(0.4)	7.2(13)	2/15
RF1-CMAES	3.4(1.0)	3.0(1)	3.5(0.9)	3.1(1)	16(20)	4/15
RF5-CMAES	3.8(2)	3.4(2)	11(16)	245(241)	∞ *2514*	0/15
Sifeg	**2.9**(0.5)	**1.4**(0.2)	**1.4**(0.3)	16(56)	75(153)	13/15
Sif	**2.9**(1)	**1.4**(0.2)	**1.4**(0.1)	17(56)	97(132)	11/15
Srr	**2.9**(1)	**1.4**(0.2)	**1.4**(0.1)	21(0.6)	76(103)	13/15

Table 87: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*4.0e+1*:7.7	*1.6e+1*:27	*1.0e+1*:37	*6.3e-1*:107	*1.0e-4*:505	15/15
BSifeg	**1.3**(1)	**1.2**(0.2)	**1.4**(1)	10(12)	∞ *1e5*	0/15
BSif	**1.3**(1)	**1.2**(0.4)	**1.4**(0.2)	33(91)	∞ *1e5*	0/15
BSqi	**1.3**(1)	**1.2**(0.4)	**1.3**(0.9)	5.9(3)	∞ *1e5*	0/15
BSrr	**1.3**(2)	**1.2**(0.4)	**1.3**(0.2)	9.0(9)	∞ *1e5*	0/15
CMA-CSA	**2.6**(1)	**2.3**(1)	**2.7**(1)	3.3(0.6)	**3.6**(0.3)	15/15
CMA-MSR	**2.1**(2)	**2.1**(2)	3.1(1)	3.6(0.8)	**3.6**(0.4)	15/15
CMA-TPA	**2.2**(2)	**2.4**(2)	3.0(1)	3.1(0.7)	**3.5**(0.2)	15/15
GP1-CMAES	**1.9**(2)	**1.6**(1.0)	**2.0**(1)	**2.3**(0.6)	25(36)	3/15
GP5-CMAES	**1.7**(1)	**1.5**(0.4)	**1.6**(0.4)	**1.7**(0.3)	∞ *2526*	0/15
IPOPCMAv3p	**1.2**(0.8)	**2.2**(1.0)	**2.4**(3)	3.2(1)	7.1(4)	10/15
LHD-10xDef	**1.8**(2)	3.8(4)	5.8(0.8)	4.0(0.5)	∞ *500*	0/15
LHD-2xDefa	**1.0**(0.5)	**1.8**(0.8)	**2.1**(1.0)	6.4(12)	∞ *500*	0/15
RAND-2xDef	**1.4**(2)	**1.8**(0.5)	**1.8**(0.4)	4.1(4)	∞ *500*	0/15
RF1-CMAES	**1.6**(1)	**1.7**(2)	**2.2**(1)	3.6(2)	∞ *2502*	0/15
RF5-CMAES	**1.5**(1)	**1.6**(1)	5.8(4)	42(41)	∞ *2504*	0/15
Sifeg	**1.3**(2)	**1.2**(0.5)	**1.3**(0.2)	**1.7**(0.7)	∞ *1e5*	0/15
Sif	**1.3**(2)	**1.2**(0.6)	**1.3**(0.3)	**2.1**(3)	∞ *1e5*	0/15
Srr	**1.3**(1)	**1.2**(0.7)	**1.2**(0.5)	**1.5**(0.6)	∞ *1e5*	0/15

Table 88: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*2.5e+2*:9.0	*1.6e+2*:72	*1.0e+2*:186	*6.3e+1*:450	*4.0e+1*:872	15/15
BSifeg	3.5(1)	88(161)	78(251)	267(366)	1489(1436)	1/15
BSif	**3.5**(3)	72(0.8)	49(112)	198(356)	722(809)	2/15
BSqi	3.5(0.8)	83(300)	68(179)	277(480)	702(460)	2/15
BSrr	3.5(2)	46(12)	94(122)	163(100)	236(176)	5/15
CMA-CSA	4.9(2)	**1.4**(0.7)	**0.95**(0.5)	**1.1**(0.5)	**1.2**(0.4)	15/15
CMA-MSR	5.6(2)	**1.5**(0.7)	**1.2**(0.2)	**0.81**(0.2)	**0.57**(0.1)	15/15
CMA-TPA	6.9(4)	**1.8**(0.6)	**1.1**(0.3)	**0.94**(0.3)	**1.2**(0.7)	15/15
GP1-CMAES	4.3(5)	**1.1**(0.2)	**0.78**(0.3)	**0.79**(0.5)	**2.2**(2)	11/15
GP5-CMAES	**3.4**(2)	**0.75**(0.3)	**0.70**(0.4)	**3.0**(3)	41(65)	1/15
IPOPCMAv3p	3.9(3)	**1.3**(0.8)	**0.94**(0.4)	**0.92**(0.5)	**1.3**(2)	14/15
LHD-10xDef	10(10)	3.1(0.2)	**1.5**(0.2)	**0.83**(0.2)	**0.98**(0.6)	8/15
LHD-2xDefa	4.0(3)	**0.97**(0.4)	**0.71**(0.2)	**0.75**(1)	**0.97**(1)	7/15
RAND-2xDef	4.7(2)	**0.98**(0.4)	**0.71**(0.3)	**0.91**(1)	**1.9**(3)	4/15
RF1-CMAES	**3.2**(2)	**1.0**(0.5)	**0.83**(0.1)	**0.89**(0.4)	**0.99**(0.3)	14/15
RF5-CMAES	5.6(4)	3.3(0.5)	3.2(6)	4.4(5)	20(22)	2/15
Sifeg	3.8(2)	61(233)	56(43)	51(37)	138(134)	8/15
Sif	3.8(1)	61(114)	59(92)	76(154)	195(181)	6/15
Srr	3.7(3)	59(221)	65(130)	102(134)	153(115)	7/15

Table 89: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:12	*2.5e+1*:47	*1.6e+1*:88	*1.0e+1*:425	*4.0e+0*:989	15/15
BSifeg	**1.9**(2)	**1.0**(0.4)	13(33)	36(39)	105(92)	9/15
BSif	**1.9**(2)	**1.1**(0.7)	17(32)	38(45)	90(126)	9/15
BSqi	**1.9**(2)	**1.1**(0.6)	42(60)	60(96)	122(95)	8/15
BSrr	**1.9**(2)	**1.1**(0.6)	23(77)	18(27)	109(80)	8/15
CMA-CSA	**1.8**(1)	4.6(3)	6.9(4)	**1.8**(0.6)	**1.5**(3)	15/15
CMA-MSR	**2.5**(3)	**2.7**(2)	6.3(14)	**1.5**(0.4)	**2.5**(5)	15/15
CMA-TPA	4.5(5)	4.8(5)	10(8)	3.1(2)	**2.6**(3)	15/15
GP1-CMAES	**1.1**(0.8)	**2.4**(2)	3.2(3)	**1.1**(0.3)	**1.5**(2)	12/15
GP5-CMAES	**1.4**(2)	**1.6**(1.0)	**1.4**(0.4)	**0.39**(0.1)	**0.89**(1)	13/15
IPOPCMAv3p	**1.1**(1)	3.3(3)	7.7(3)	**2.4**(0.9)	**1.6**(1)	14/15
LHD-10xDef	**1.7**(2)	3.6(2)	3.7(3)	**1.6**(1)	∞ *500*	0/15
LHD-2xDefa	**1.4**(1)	**1.8**(0.7)	6.9(7)	5.4(4)	∞ *500*	0/15
RAND-2xDef	**1.3**(0.9)	**2.6**(3)	3.6(4)	**1.8**(2)	7.3(6)	1/15
RF1-CMAES	**1.6**(1)	3.1(2)	4.5(3)	**1.3**(0.8)	4.4(3)	6/15
RF5-CMAES	**1.5**(2)	**2.0**(2)	**2.4**(1)	3.6(6)	17(21)	2/15
Sifeg	**2.0**(3)	**1.5**(0.8)	**1.4**(0.8)	5.9(18)	28(12)	15/15
Sif	**2.0**(2)	**1.5**(1)	8.9(30)	4.2(4)	39(29)	14/15
Srr	**2.0**(3)	**1.5**(0.9)	12(54)	4.5(6)	21(27)	15/15

Table 90: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.0e+1*:26	*6.3e+0*:85	*4.0e+0*:155	*2.5e+0*:238	*6.3e-1*:585	15/15
BSifeg	**1.4**(1)	301(1177)	982(1814)	2734(4204)	∞ *1e5*	0/15
BSif	**1.4**(0.9)	237(464)	965(1440)	1182(1806)	∞ *1e5*	0/15
BSqi	**1.4**(0.6)	205(661)	740(647)	1684(1788)	∞ *1e5*	0/15
BSrr	**1.4**(0.7)	296(592)	761(1122)	1812(1492)	∞ *1e5*	0/15
CMA-CSA	3.4(2)	**2.1**(1.0)	**1.7**(0.7)	**1.6**(0.7)	**1.7**(3)	15/15
CMA-MSR	**2.0**(0.9)	**1.4**(0.6)	**1.2**(0.7)	**1.1**(0.3)	**2.8**(2)	15/15
CMA-TPA	**2.4**(2)	**1.4**(0.3)	**1.1**(0.3)	**1.0**(0.4)	**1.2**(0.2)	15/15
GP1-CMAES	**1.7**(0.8)	**1.0**(0.6)	**0.88**(0.3)	**0.80**(0.3)	**1.0**(0.2)	14/15
GP5-CMAES	**1.7**(1)	**1.1**(0.5)	**0.99**(0.8)	**2.4**(6)	18(26)	3/15
IPOPCMAv3p	**2.4**(2)	**1.4**(1)	**1.4**(0.4)	**1.3**(0.6)	**1.0**(0.2)	15/15
LHD-10xDef	3.5(4)	**2.7**(0.4)	**2.1**(0.6)	4.4(6)	∞ *500*	0/15
LHD-2xDefa	**1.6**(0.8)	**0.99**(0.6)	**2.4**(2)	6.9(7)	∞ *500*	0/15
RAND-2xDef	**1.8**(0.7)	**1.2**(0.8)	**1.7**(0.7)	3.7(3)	∞ *500*	0/15
RF1-CMAES	**1.7**(2)	**1.3**(0.6)	**1.2**(0.4)	**3.0**(0.7)	13(17)	4/15
RF5-CMAES	**1.6**(0.2)	**2.6**(0.2)	5.0(6)	10(10)	64(29)	1/15
Sifeg	**1.3**(0.7)	208(202)	402(557)	440(736)	1223(2308)	2/15
Sif	**1.3**(0.8)	194(337)	490(1441)	834(1353)	∞ *1e5*	0/15
Srr	**1.3**(0.4)	182(4)	429(1444)	487(629)	∞ *1e5*	0/15

Table 91: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*4.0e+1*:11	*2.5e+1*:56	*1.6e+1*:172	*1.6e+1*:172	*2.5e+0*:561	15/15
BSifeg	3.2(5)	382(511)	354(420)	354(544)	2408(1070)	1/15
BSif	3.5(7)	362(748)	317(415)	317(529)	∞ *1e5*	0/15
BSqi	81(589)	223(397)	281(731)	281(287)	2341(1144)	1/15
BSrr	**3.0**(2)	305(691)	528(814)	528(532)	∞ *9e4*	0/15
CMA-CSA	6.7(6)	**2.5**(0.5)	**1.2**(0.7)	**1.2**(0.4)	**1.0**(0.2)	15/15
CMA-MSR	4.9(4)	**1.8**(0.3)	**0.99**(0.4)	**0.99**(0.3)	**1.9**(4)	15/15
CMA-TPA	5.1(4)	**2.1**(0.9)	**1.1**(0.3)	**1.1**(0.2)	**1.5**(0.3)	15/15
GP1-CMAES	4.4(5)	**1.7**(0.8)	**0.85**(0.5)	**0.85**(0.5)	**1.6**(1)	13/15
GP5-CMAES	3.7(6)	**1.3**(0.5)	**0.71**(0.2)	**0.71**(0.4)	6.0(10)	7/15
IPOPCMAv3p	4.7(3)	**2.1**(1)	**1.2**(0.6)	**1.2**(0.4)	**1.5**(0.3)	14/15
LHD-10xDef	8.2(9)	4.0(0.2)	**2.0**(2)	**2.0**(0.4)	∞ *500*	0/15
LHD-2xDefa	4.2(3)	**1.8**(1)	**1.8**(2)	**1.8**(2)	∞ *500*	0/15
RAND-2xDef	**3.0**(2)	**1.5**(0.6)	**1.6**(2)	**1.6**(3)	∞ *500*	0/15
RF1-CMAES	**2.8**(2)	**1.4**(0.8)	**0.95**(0.4)	**0.95**(0.5)	3.2(2)	10/15
RF5-CMAES	4.0(2)	4.5(22)	3.8(10)	3.8(10)	66(54)	1/15
Sifeg	8.3(2)	83(7)	171(286)	171(236)	1179(1377)	2/15
Sif	320(3)	194(671)	223(324)	223(553)	∞ *1e5*	0/15
Srr	6.7(2)	137(290)	184(432)	184(327)	∞ *9e4*	0/15

Table 92: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:618	*1.0e-1*:10609	*6.3e-2*:10623	*4.0e-2*:10625	*2.5e-2*:10644	15/15
BSifeg	∞	∞	∞	∞	∞ *1e5*	0/15
BSif	∞	∞	∞	∞	∞ *1e5*	0/15
BSqi	∞	∞	∞	∞	∞ *1e5*	0/15
BSrr	∞	∞	∞	∞	∞ *1e5*	0/15
CMA-CSA	**171**(101)	**12**(5)	**19**(12)	**23**(10)	**26**(25)	15/15
CMA-MSR	**237**(437)	**21**(31)	**35**(55)	**62**(128)	**290**(257)	4/15
CMA-TPA	**156**(83)	**12**(5)	**18**(11)	**24**(13)	**34**(18)	15/15
GP1-CMAES	∞	∞	∞	∞	∞ *2504*	0/15
GP5-CMAES	∞	∞	∞	∞	∞ *2528*	0/15
IPOPCMAv3p	∞	∞	∞	∞	∞ *2504*	0/15
LHD-10xDef	∞	∞	∞	∞	∞ *500*	0/15
LHD-2xDefa	∞	∞	∞	∞	∞ *500*	0/15
RAND-2xDef	∞	∞	∞	∞	∞ *500*	0/15
RF1-CMAES	∞	∞	∞	∞	∞ *2502*	0/15
RF5-CMAES	∞	∞	∞	∞	∞ *2516*	0/15
Sifeg	∞	∞	∞	∞	∞ *1e5*	0/15
Sif	∞	∞	∞	∞	∞ *1e5*	0/15
Srr	∞	∞	∞	∞	∞ *1e5*	0/15

Table 93: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*1.0e+4*:17	*6.3e+3*:21	*6.3e+1*:30	*2.5e+0*:122	*1.0e+0*:15426	13/15
BSifeg	**1.9**(0.6)	**1.9**(0.1)	**1.7**(0.8)	**1.9**(2)	3.2(4)	12/15
BSif	**1.9**(0.1)	**1.9**(0.1)	**1.8**(1)	**1.8**(2)	3.5(3)	12/15
BSqi	**1.9**(0.1)	**1.9**(0.1)	**1.7**(0.5)	**1.1**(1)	3.4(7)	12/15
BSrr	**1.9**(1)	**1.9**(0.1)	**1.7**(0.4)	**1.4**(0.3)	**2.8**(3)	13/15
CMA-CSA	**2.3**(1)	**2.4**(2)	5.2(1)	5.5(2)	**1.8**(0.9)	15/15
CMA-MSR	3.1(3)	3.5(2)	5.7(2)	23(2)	∞ *1e6*	0/15
CMA-TPA	3.1(2)	3.2(2)	5.1(0.9)	3.0(1)	18(0.3)	14/15
GP1-CMAES	**1.9**(2)	**2.0**(1)	3.8(1.0)	8.3(9)	∞ *2502*	0/15
GP5-CMAES	**2.1**(1.0)	**1.9**(0.8)	**2.4**(0.6)	25(23)	∞ *2526*	0/15
IPOPCMAv3p	3.2(3)	3.3(2)	5.4(2)	5.1(3)	**2.3**(5)	1/15
LHD-10xDef	3.6(6)	6.0(5)	7.7(0.3)	14(14)	∞ *500*	0/15
LHD-2xDefa	**2.9**(0.7)	**2.9**(0.6)	**2.9**(0.7)	6.2(5)	∞ *500*	0/15
RAND-2xDef	**1.8**(2)	**2.2**(0.5)	**2.9**(2)	4.9(3)	∞ *500*	0/15
RF1-CMAES	**2.6**(1)	3.0(2)	5.4(2)	4.4(2)	∞ *2502*	0/15
RF5-CMAES	**2.4**(2)	**2.5**(2)	21(32)	35(65)	∞ *2514*	0/15
Sifeg	**1.9**(1)	**1.9**(0.2)	**2.0**(0.6)	**0.72**(0.4)	**1.4**(2)	15/15
Sif	**1.9**(0.7)	**1.9**(0.1)	**1.9**(0.7)	**0.76**(0.4)	**0.87**(1)	15/15
Srr	**1.9**(0.7)	**1.9**(0.2)	**1.9**(0.4)	**0.76**(0.2)	**1.9**(3)	14/15

Table 94: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*4.0e+1*:30	*2.5e+1*:46	*1.6e+1*:56	*1.0e+1*:130	*6.3e+0*:639	15/15
BSifeg	**2.3**(0.5)	263(547)	374(0.5)	171(317)	116(123)	10/15
BSif	5.7(0.7)	354(54)	520(656)	331(768)	226(279)	7/15
BSqi	**2.5**(8)	217(557)	333(33)	151(29)	92(121)	11/15
BSrr	4.1(0.7)	222(762)	338(25)	159(282)	95(147)	11/15
CMA-CSA	4.7(3)	7.6(16)	11(25)	7.7(11)	5.0(10)	15/15
CMA-MSR	3.5(2)	3.7(3)	15(30)	10(13)	8.4(3)	15/15
CMA-TPA	3.3(1)	3.0(0.8)	8.4(12)	4.8(5)	**1.4**(2)	15/15
GP1-CMAES	**2.1**(1)	**2.1**(0.9)	**2.0**(0.6)	**2.4**(10)	**1.2**(3)	12/15
GP5-CMAES	**1.8**(0.5)	**2.6**(4)	7.1(8)	3.7(13)	**1.3**(2)	13/15
IPOPCMAv3p	3.5(3)	8.1(15)	17(37)	10(10)	3.2(5)	9/15
LHD-10xDef	6.5(3)	4.7(0.2)	4.6(0.3)	**2.7**(2)	**1.0**(2)	9/15
LHD-2xDefa	**2.1**(2)	**1.8**(0.3)	**2.1**(0.8)	**1.3**(1)	**0.52**(0.2)	11/15
RAND-2xDef	**2.0**(0.3)	**1.7**(0.4)	**2.2**(0.3)	**1.2**(0.2)	**0.39**(0.3)	12/15
RF1-CMAES	9.1(2)	7.5(15)	11(2)	9.1(11)	4.0(9)	8/15
RF5-CMAES	**2.9**(2)	6.2(0.8)	6.9(12)	13(20)	12(14)	4/15
Sifeg	**1.6**(0.9)	216(1297)	177(612)	79(385)	88(100)	11/15
Sif	**1.6**(0.6)	256(1092)	210(284)	110(384)	109(150)	10/15
Srr	**1.6**(0.7)	197(0.8)	162(4)	78(78)	65(59)	12/15

Table 95: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:18	*4.0e+1*:30	*4.0e+1*:30	*6.3e+0*:155	*4.0e+0*:631	14/15
BSifeg	404(1)	737(829)	737(1656)	530(487)	135(475)	9/15
BSif	404(1405)	1042(2485)	1042(2927)	813(798)	251(673)	6/15
BSqi	404(2810)	713(1339)	713(828)	434(647)	109(119)	10/15
BSrr	404(0.8)	802(3787)	802(1302)	469(416)	116(116)	10/15
CMA-CSA	3.3(4)	3.9(1)	3.9(1)	303(900)	183(249)	13/15
CMA-MSR	5.1(3)	11(25)	11(2)	428(2713)	349(1667)	12/15
CMA-TPA	4.7(3)	4.1(3)	4.1(3)	604(1968)	148(2)	13/15
GP1-CMAES	**3.2**(2)	**3.1**(2)	**3.1**(0.7)	**3.9**(13)	**1.3**(2)	12/15
GP5-CMAES	**2.9**(1)	4.0(1)	4.0(12)	12(14)	3.4(8)	9/15
IPOPCMAv3p	4.5(3)	4.7(2)	4.7(5)	17(21)	5.6(3)	7/15
LHD-10xDef	6.2(6)	7.7(0.1)	7.7(0.3)	5.7(4)	**1.4**(2)	7/15
LHD-2xDefa	3.4(1.0)	**3.1**(2)	**3.1**(2)	**4.4**(6)	**1.3**(1)	7/15
RAND-2xDef	3.4(2)	**2.8**(2)	**2.8**(0.7)	**4.6**(3)	**1.4**(4)	7/15
RF1-CMAES	**3.2**(3)	10(23)	10(65)	16(24)	8.3(8)	5/15
RF5-CMAES	3.9(3)	4.4(2)	4.4(2)	36(52)	9.0(12)	5/15
Sifeg	8.0(22)	442(1261)	442(699)	280(380)	83(125)	11/15
Sif	404(0.5)	772(1364)	772(1968)	611(835)	151(88)	8/15
Srr	404(1406)	672(1655)	672(1655)	389(317)	109(195)	10/15

Table 96: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*6.3e+0*:10	*4.0e+0*:62	*2.5e+0*:162	*2.5e+0*:162	*1.0e+0*:915	15/15
BSifeg	**2.9**(4)	**1.5**(0.7)	**1.7**(2)	**1.7**(2)	13(11)	15/15
BSif	**2.9**(2)	**1.2**(0.9)	**1.6**(2)	**1.6**(2)	12(24)	15/15
BSqi	**2.9**(4)	**1.2**(0.5)	**1.1**(2)	**1.1**(1)	13(10)	15/15
BSrr	**2.9**(4)	**1.2**(1)	**1.5**(2)	**1.5**(2)	11(16)	15/15
CMA-CSA	6.3(5)	6.0(6)	18(14)	18(14)	23(16)	15/15
CMA-MSR	4.8(3)	3.2(3)	**2.8**(2)	**2.8**(0.8)	**2.9**(4)	15/15
CMA-TPA	3.4(4)	3.9(2)	16(23)	16(19)	12(5)	15/15
GP1-CMAES	**2.0**(3)	**1.7**(1)	5.3(4)	5.3(4)	**2.7**(2)	10/15
GP5-CMAES	3.6(4)	**1.7**(2)	**1.6**(0.3)	**1.6**(4)	**0.92**(0.4)	13/15
IPOPCMAv3p	**2.3**(0.8)	**2.8**(2)	4.4(3)	4.4(2)	∞ *2514*	0/15
LHD-10xDef	**2.1**(2)	**2.3**(2)	5.9(4)	5.9(2)	∞ *500*	0/15
LHD-2xDefa	**2.0**(1)	**1.6**(2)	8.3(11)	8.3(17)	∞ *500*	0/15
RAND-2xDef	**2.0**(2)	3.8(9)	10(13)	10(5)	∞ *500*	0/15
RF1-CMAES	**1.7**(1)	3.5(4)	7.1(5)	7.1(10)	∞ *2506*	0/15
RF5-CMAES	**1.6**(3)	3.1(3)	10(8)	10(7)	∞ *2548*	0/15
Sifeg	**2.5**(3)	**1.9**(2)	3.5(3)	3.5(4)	4.6(9)	15/15
Sif	**2.5**(3)	**1.9**(2)	3.7(4)	3.7(4)	6.4(11)	15/15
Srr	**2.5**(3)	**1.9**(2)	3.7(2)	3.7(2)	5.7(4)	15/15

Table 97: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*1.0e+2*:66	*6.3e+1*:596	*4.0e+1*:3181	*2.5e+1*:7668	*1.6e+1*:14353	15/15
BSifeg	22(16)	50(73)	34(31)	87(130)	∞ *9e4*	0/15
BSif	37(13)	37(13)	72(62)	∞	∞ *9e4*	0/15
BSqi	18(10)	71(104)	55(63)	174(352)	∞ *9e4*	0/15
BSrr	19(18)	39(56)	41(40)	162(215)	∞ *8e4*	0/15
CMA-CSA	**2.6**(0.9)	**1.1**(1)	**1.2**(1)	**0.62**(0.5)	**1.2**(0.9)	15/15
CMA-MSR	**2.6**(1)	**0.84**(0.3)	**0.45**(0.6)	**0.50**(0.4)	**0.83**(0.6)	15/15
CMA-TPA	**2.6**(1)	**1.0**(0.5)	**0.67**(0.2)	**0.70**(0.8)	**0.94**(0.6)	15/15
GP1-CMAES	**1.7**(0.8)	**0.90**(0.5)	**2.6**(3)	**1.1**(0.9)	∞ *2514*	0/15
GP5-CMAES	**1.2**(0.3)	**0.81**(0.9)	∞	∞	∞ *2528*	0/15
IPOPCMAv3p	**2.7**(1)	**1.4**(0.9)	**2.7**(1)	4.8(4)	∞ *2504*	0/15
LHD-10xDef	4.8(1)	**2.9**(3)	**2.4**(2)	∞	∞ *500*	0/15
LHD-2xDefa	8.3(8)	∞	∞	∞	∞ *500*	0/15
RAND-2xDef	4.8(2)	∞	∞	∞	∞ *500*	0/15
RF1-CMAES	**2.2**(0.9)	**1.1**(0.9)	**1.2**(2)	**2.3**(2)	∞ *2502*	0/15
RF5-CMAES	**2.4**(1)	3.4(3)	∞	∞	∞ *2514*	0/15
Sifeg	3.0(3)	5.9(13)	9.0(7)	26(33)	∞ *9e4*	0/15
Sif	3.0(2)	10(16)	13(19)	35(33)	44(21)	2/15
Srr	**2.8**(2)	**1.5**(2)	8.8(5)	22(27)	46(55)	2/15

Table 98: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*6.3e+1*:24	*4.0e+1*:42	*1.0e-8*:43	*1.0e-8*:43	*1.0e-8*:43	15/15
BSifeg	**2.2**(1.0)	**1.7**(0.1)	**2.5**(0.2)	**2.5**(0.2)	**2.5**(0.3)	15/15
BSif	**2.2**(0.9)	**1.7**(0.1)	**2.5**(0.3)	**2.5**(0.2)	**2.5**(0.3)	15/15
BSqi	**2.2**(0.9)	**1.7**(0.1)	**2.5**(0.2)	**2.5**(0.3)	**2.5**(0.2)	15/15
BSrr	**2.2**(0.4)	**1.7**(0.1)	**2.5**(0.2)	**2.5**(0.2)	**2.5**(0.2)	15/15
CMA-CSA	4.7(2)	4.4(1)	64(4)	64(3)	64(5)	15/15
CMA-MSR	6.0(2)	5.1(2)	75(4)	75(3)	75(3)	15/15
CMA-TPA	5.4(2)	4.0(0.9)	46(3)	46(4)	46(1)	15/15
GP1-CMAES	3.9(2)	3.1(0.6)	58(6)	58(10)	58(7)	15/15
GP5-CMAES	**2.9**(0.9)	**2.0**(0.2)	∞	∞	∞ *5034*	0/15
IPOPCMAv3p	4.8(2)	4.6(1)	64(2)	64(2)	64(2)	15/15
LHD-10xDef	17(0.1)	10(0.1)	∞	∞	∞ *1000*	0/15
LHD-2xDefa	3.9(0.2)	**2.5**(0.2)	∞	∞	∞ *1000*	0/15
RAND-2xDef	4.0(0.5)	**2.8**(0.4)	∞	∞	∞ *1000*	0/15
RF1-CMAES	3.9(1)	3.5(0.8)	73(20)	73(18)	73(18)	15/15
RF5-CMAES	3.7(2)	3.0(0.9)	∞	∞	∞ *5006*	0/15
Sifeg	**2.2**(0.6)	**1.7**(0.2)	16(0.9)	16(0.9)	16(1)	15/15
Sif	**2.2**(0.7)	**1.7**(0.1)	16(1)	16(1)	16(2)	15/15
Srr	**2.2**(0.8)	**1.7**(0.1)	16(1)	16(1)	16(0.7)	15/15

Table 99: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*4.0e+6*:29	*2.5e+6*:42	*1.0e+5*:65	*1.0e+4*:207	*1.0e-8*:412	15/15
BSifeg	**2.3**(1.0)	**1.8**(0.2)	**1.5**(0.2)	**0.54**(0.1)	**1.2**(0.1)	15/15
BSif	**2.3**(1.0)	**1.8**(0.2)	**1.5**(0.0)	**0.55**(0.1)	**1.2**(0.1)	15/15
BSqi	**2.3**(0.0)	**1.8**(0.2)	**1.5**(0.1)	**0.53**(0.1)	**1.3**(0.2)	15/15
BSrr	**2.3**(1)	**1.8**(0.3)	**1.5**(0.1)	**0.54**(0.1)	**1.3**(0.2)	15/15
CMA-CSA	**1.2**(0.6)	**1.1**(0.8)	14(3)	11(3)	33(1)	15/15
CMA-MSR	**1.0**(0.4)	**1.6**(2)	11(5)	7.9(3)	38(3)	15/15
CMA-TPA	**1.4**(1)	**1.5**(0.7)	11(4)	9.5(3)	36(2)	15/15
GP1-CMAES	**1.5**(0.9)	**1.6**(1)	8.9(5)	8.0(3)	∞ *5006*	0/15
GP5-CMAES	**0.90**(0.9)	**1.3**(1)	5.4(2)	3.9(0.8)	∞ *5006*	0/15
IPOPCMAv3p	**0.93**(0.6)	**1.4**(1)	14(4)	11(2)	∞ *5006*	0/15
LHD-10xDef	**1.6**(1)	4.1(4)	30(19)	∞	∞ *1000*	0/15
LHD-2xDefa	**1.1**(0.8)	**1.9**(0.8)	8.5(5)	72(67)	∞ *1000*	0/15
RAND-2xDef	**1.3**(2)	**1.6**(2)	7.8(4)	70(116)	∞ *1000*	0/15
RF1-CMAES	**1.1**(0.8)	**1.2**(1)	11(3)	29(27)	∞ *5006*	0/15
RF5-CMAES	**1.2**(1)	**1.3**(0.9)	184(149)	∞	∞ *5006*	0/15
Sifeg	**2.3**(0.9)	**1.9**(0.3)	**1.9**(0.2)	**0.77**(0.1)	**2.2**(0.2)	15/15
Sif	**2.3**(0.5)	**1.9**(0.5)	**1.9**(0.2)	**0.82**(0.2)	**2.2**(0.3)	15/15
Srr	**2.3**(1)	**1.9**(0.5)	**1.9**(0.3)	**0.76**(0.1)	**2.2**(0.3)	15/15

Table 100: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*6.3e+2*:33	*4.0e+2*:44	*1.6e+2*:109	*1.0e+2*:255	*2.5e+1*:3277	15/15
BSifeg	**1.7**(0.2)	**1.6**(0.2)	**0.92**(0.1)	**0.52**(0.0)	**0.12**(0.0)	15/15
BSif	**1.7**(0.2)	**1.6**(0.2)	**0.92**(0.0)	**0.52**(0.0)	**0.12**(0.0)	15/15
BSqi	**1.7**(0.2)	**1.6**(0.2)	**0.92**(0.1)	**0.52**(0.1)	**0.11**(0.0)	15/15
BSrr	**1.7**(0.2)	**1.6**(0.1)	**0.92**(0.1)	**0.52**(0.1)	**0.11**(0.0)	15/15
CMA-CSA	**1.9**(2)	4.0(0.5)	7.4(2)	7.0(3)	3.5(1.0)	15/15
CMA-MSR	**2.9**(2)	4.5(0.8)	5.2(1.0)	**3.0**(0.3)	3.5(1)	15/15
CMA-TPA	3.1(1)	4.0(1.0)	6.0(4)	4.4(1)	**2.5**(1)	15/15
GP1-CMAES	**2.3**(1)	3.2(1.0)	5.9(1)	4.2(2)	22(34)	1/15
GP5-CMAES	**1.9**(0.9)	**2.9**(1)	15(23)	43(50)	∞ *5034*	0/15
IPOPCMAv3p	**1.9**(1)	3.6(2)	7.6(4)	7.0(2)	22(30)	1/15
LHD-10xDef	9.1(4)	10(0.3)	8.7(0.7)	9.0(6)	∞ *1000*	0/15
LHD-2xDefa	**2.5**(0.5)	3.0(0.5)	8.6(7)	18(40)	∞ *1000*	0/15
RAND-2xDef	**2.7**(0.3)	**2.8**(0.4)	5.7(5)	10(24)	∞ *1000*	0/15
RF1-CMAES	**2.0**(1)	3.7(0.9)	6.4(2)	4.0(0.8)	22(41)	1/15
RF5-CMAES	**1.7**(0.8)	**2.6**(0.5)	18(17)	82(128)	∞ *5006*	0/15
Sifeg	**1.7**(0.2)	**1.6**(0.1)	**0.97**(0.1)	**0.54**(0.1)	**0.14**(0.0)	15/15
Sif	**1.7**(0.4)	**1.6**(0.2)	**0.97**(0.1)	**0.55**(0.1)	**0.15**(0.0)	15/15
Srr	**1.7**(0.2)	**1.6**(0.2)	**0.97**(0.2)	**0.54**(0.1)	**0.13**(0.0)	15/15

Table 101: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*6.3e+2*:22	*4.0e+2*:91	*2.5e+2*:250	*1.6e+2*:332	*6.3e+1*:1927	15/15
BSifeg	**2.8**(1)	**0.98**(0.1)	**0.58**(0.3)	**0.66**(0.3)	**0.20**(0.0)	15/15
BSif	**2.8**(0.9)	**0.98**(0.1)	**0.59**(0.3)	**0.68**(0.4)	**0.21**(0.0)	15/15
BSqi	**2.8**(1)	**0.98**(0.2)	**0.62**(0.5)	**0.70**(0.3)	**0.24**(0.0)	15/15
BSrr	**2.8**(0.5)	**0.98**(0.2)	**0.60**(0.3)	**0.69**(0.3)	**0.21**(0.0)	15/15
CMA-CSA	7.1(3)	3.1(0.4)	**2.1**(0.4)	3.7(1)	**2.2**(1)	15/15
CMA-MSR	8.5(3)	3.3(0.8)	**2.0**(0.8)	3.6(5)	5.4(6)	15/15
CMA-TPA	8.2(2)	**2.9**(0.8)	**1.9**(0.5)	**2.9**(0.6)	3.5(2)	15/15
GP1-CMAES	9.2(8)	4.5(2)	11(4)	42(88)	39(22)	1/15
GP5-CMAES	7.5(2)	5.5(9)	13(45)	215(159)	∞ *5022*	0/15
IPOPCMAv3p	7.4(3)	3.3(1)	**2.3**(0.6)	4.1(1)	**2.4**(2)	10/15
LHD-10xDef	22(10)	18(12)	∞	∞	∞ *1000*	0/15
LHD-2xDefa	9.1(3)	6.8(1)	58(71)	∞	∞ *1000*	0/15
RAND-2xDef	10(8)	8.4(8)	13(11)	∞	∞ *1000*	0/15
RF1-CMAES	8.1(2)	3.4(1)	**2.5**(0.6)	7.7(13)	∞ *5006*	0/15
RF5-CMAES	8.9(3)	13(2)	57(110)	∞	∞ *5006*	0/15
Sifeg	**2.8**(0.8)	**1.1**(0.3)	**0.56**(0.2)	**0.60**(0.1)	**0.17**(0.0)	15/15
Sif	**2.8**(1)	**1.1**(0.2)	**0.58**(0.1)	**0.61**(0.1)	**0.17**(0.0)	15/15
Srr	**2.8**(1)	**1.1**(0.2)	**0.56**(0.1)	**0.59**(0.1)	**0.16**(9e-3)	15/15

Table 102: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*2.5e+2*:19	*1.6e+2*:34	*1.0e-8*:41	*1.0e-8*:41	*1.0e-8*:41	15/15
BSifeg	**2.1**(0.3)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSif	**2.1**(0.3)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSqi	**2.1**(0.2)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
BSrr	**2.1**(0.2)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
CMA-CSA	**1.5**(0.7)	**1.8**(0.5)	6.0(0.8)	6.0(1)	6.0(1)	15/15
CMA-MSR	**1.9**(0.6)	**1.9**(0.7)	5.6(1)	5.6(1)	5.6(1)	15/15
CMA-TPA	**1.3**(1)	**1.4**(0.4)	4.9(2)	4.9(1)	4.9(1)	15/15
GP1-CMAES	**1.6**(0.8)	**1.9**(0.5)	92(115)	92(62)	92(85)	11/15
GP5-CMAES	**1.7**(0.9)	**1.7**(0.4)	4.8(0.7)	4.8(2)	4.8(1)	15/15
IPOPCMAv3p	**2.0**(2)	**2.0**(0.9)	36(13)	36(15)	36(19)	15/15
LHD-10xDef	8.1(9)	12(0.0)	11(0.4)	11(0.2)	11(0.3)	15/15
LHD-2xDefa	3.7(2)	**2.5**(0.0)	**3.0**(0.1)	**3.0**(0.2)	**3.0**(0.2)	15/15
RAND-2xDef	3.4(2)	**2.6**(0.1)	3.4(2)	3.4(3)	3.4(3)	15/15
RF1-CMAES	**1.8**(1)	**2.3**(0.8)	50(22)	50(26)	50(24)	15/15
RF5-CMAES	**2.0**(0.9)	**2.0**(0.6)	265(451)	265(361)	265(330)	6/15
Sifeg	**2.1**(0.2)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
Sif	**2.1**(0.3)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15
Srr	**2.1**(0.2)	**1.5**(0.1)	**1.5**(0.0)	**1.5**(0.0)	**1.5**(0.0)	15/15

Table 103: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*2.5e+5*:16	*6.3e+4*:43	*1.6e+4*:62	*1.6e+2*:353	*1.6e+1*:1078	15/15
BSifeg	**2.3**(1)	**1.3**(0.8)	**1.2**(0.4)	39(42)	713(1132)	3/15
BSif	**2.3**(2)	**1.3**(0.3)	**1.2**(0.5)	159(199)	2558(3325)	1/15
BSqi	**1.9**(2)	**1.1**(0.7)	**1.1**(0.3)	34(59)	419(453)	2/7
BSrr	**2.3**(1)	**1.3**(0.7)	**1.2**(0.4)	36(47)	210(81)	8/15
CMA-CSA	3.6(2)	**2.5**(1)	**2.4**(1)	**2.8**(0.8)	**1.8**(0.3)	15/15
CMA-MSR	3.6(2)	**2.4**(0.6)	**2.4**(0.9)	**1.9**(0.8)	**1.5**(0.6)	15/15
CMA-TPA	3.1(2)	**2.2**(2)	**2.0**(0.7)	**2.3**(0.7)	**1.6**(0.5)	15/15
GP1-CMAES	**2.8**(2)	**2.0**(0.9)	**2.0**(0.8)	**1.7**(1)	4.2(4)	11/15
GP5-CMAES	**2.6**(2)	**1.6**(0.5)	**1.5**(0.4)	19(24)	∞ *5024*	0/15
IPOPCMAv3p	**2.9**(2)	**2.3**(1)	**2.3**(2)	**2.0**(0.5)	**1.5**(0.2)	15/15
LHD-10xDef	17(12)	10(1)	7.1(0.2)	4.3(4)	∞ *1000*	0/15
LHD-2xDefa	4.6(2)	**2.3**(0.3)	**1.9**(0.4)	6.8(6)	∞ *1000*	0/15
RAND-2xDef	4.8(0.2)	**2.3**(0.3)	**1.8**(0.5)	5.7(9)	∞ *1000*	0/15
RF1-CMAES	**2.7**(2)	**2.1**(0.7)	**2.3**(1)	**2.9**(1)	66(118)	1/15
RF5-CMAES	**2.1**(2)	**1.6**(0.6)	**1.8**(0.4)	32(28)	∞ *5006*	0/15
Sifeg	**2.3**(2)	**1.2**(0.5)	**1.2**(0.4)	7.4(9)	151(146)	10/15
Sif	**2.3**(2)	**1.2**(0.7)	**1.2**(0.5)	37(71)	294(301)	7/15
Srr	**2.3**(2)	**1.2**(0.7)	**1.2**(0.3)	10(5)	94(67)	12/15

Table 104: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*1.0e+3*:11	*4.0e+2*:39	*2.5e+2*:74	*6.3e+1*:319	*1.0e+1*:1351	15/15
BSifeg	**1.4**(0.7)	**1.5**(0.5)	**1.0**(0.5)	187(162)	∞ *2e5*	0/15
BSif	**1.4**(2)	**1.5**(0.5)	7.0(0.4)	258(349)	∞ *2e5*	0/15
BSqi	**1.4**(2)	**1.5**(0.5)	**1.0**(0.4)	288(335)	∞ *2e5*	0/15
BSrr	**1.4**(1)	**1.5**(0.6)	**1.0**(0.5)	174(69)	∞ *2e5*	0/15
CMA-CSA	**2.7**(2)	3.5(2)	3.1(1)	**1.7**(0.4)	**1.7**(2)	15/15
CMA-MSR	3.5(2)	3.1(0.6)	**2.7**(1.0)	**1.4**(0.4)	**2.1**(1)	15/15
CMA-TPA	3.2(3)	3.7(1)	**2.6**(0.7)	**1.3**(0.3)	**2.1**(1)	15/15
GP1-CMAES	**2.1**(2)	**2.2**(1)	**1.7**(0.7)	**0.92**(0.4)	3.0(3)	10/15
GP5-CMAES	**2.4**(2)	**1.8**(0.6)	**1.3**(0.1)	**0.58**(0.0)*	**1.6**(0.8)	14/15
IPOPCMAv3p	**1.3**(1)	**2.2**(2)	**2.5**(0.8)	**1.6**(0.2)	**1.3**(0.7)	15/15
LHD-10xDef	**2.4**(2)	10(3)	5.7(0.8)	8.1(13)	∞ *1000*	0/15
LHD-2xDefa	**1.5**(1)	**2.6**(0.3)	**1.9**(0.5)	8.1(4)	∞ *1000*	0/15
RAND-2xDef	**2.2**(3)	**2.8**(2)	**2.1**(0.7)	6.4(6)	∞ *1000*	0/15
RF1-CMAES	**2.0**(1)	**2.6**(2)	**2.3**(0.9)	**1.5**(0.7)	54(73)	1/15
RF5-CMAES	**1.8**(2)	**1.9**(0.9)	**1.8**(0.4)	7.0(7)	∞ *5034*	0/15
Sifeg	**1.4**(2)	**1.5**(0.6)	**3.0**(0.4)	48(6)	∞ *2e5*	0/15
Sif	**1.4**(1)	**1.5**(0.5)	**1.1**(0.6)	21(40)	∞ *2e5*	0/15
Srr	**1.4**(2)	**1.5**(0.7)	**1.1**(0.3)	12(4)	∞ *2e5*	0/15

Table 105: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*4.0e+4*:19	*2.5e+4*:35	*4.0e+3*:67	*2.5e+2*:231	*1.6e+1*:1470	15/15
BSifeg	**3.3**(0.4)	**2.0**(0.3)	**1.2**(0.1)	4.2(6)	65(107)	12/15
BSif	**3.3**(0.7)	**2.0**(0.4)	**1.2**(0.1)	4.9(14)	518(422)	3/15
BSqi	**3.1**(0.5)	**1.9**(0.4)	**1.2**(0.1)	**1.6**(2)	32(16)	7/8
BSrr	**3.3**(1)	**2.0**(0.3)	**1.2**(0.1)	3.7(0.2)	75(77)	10/15
CMA-CSA	7.3(5)	5.8(2)	4.9(1)	**2.9**(0.5)	**2.4**(0.7)	15/15
CMA-MSR	8.7(3)	5.9(1)	5.0(1.0)	**2.9**(0.7)	**2.5**(0.9)	15/15
CMA-TPA	7.2(2)	4.3(1)	3.8(0.8)	**2.2**(0.5)	**1.9**(0.5)	15/15
GP1-CMAES	5.5(2)	3.4(0.7)	3.1(0.8)	**2.3**(0.3)	3.0(2)	12/15
GP5-CMAES	4.4(0.6)	**2.6**(0.2)	**2.2**(0.4)	11(22)	8.2(9)	5/15
IPOPCMAv3p	6.3(3)	4.8(1.0)	4.4(0.3)	**2.8**(1.0)	**2.9**(2)	12/15
LHD-10xDef	23(0.2)	12(0.1)	6.7(0.4)	3.8(2)	∞ *1000*	0/15
LHD-2xDefa	5.5(0.7)	3.2(0.6)	3.1(0.7)	**2.8**(0.7)	∞ *1000*	0/15
RAND-2xDef	5.4(0.5)	3.2(0.4)	**2.8**(0.4)	**2.5**(0.7)	∞ *1000*	0/15
RF1-CMAES	5.8(1)	4.0(0.6)	4.0(1)	**2.7**(0.8)	25(19)	2/15
RF5-CMAES	5.8(2)	3.7(2)	4.2(2)	96(108)	∞ *5006*	0/15
Sifeg	3.3(1)	**2.0**(0.3)	**1.2**(0.1)	**0.98**(0.6)	49(53)	14/15
Sif	3.3(0.6)	**2.0**(0.1)	**1.2**(0.1)	**1.1**(0.8)	75(113)	12/15
Srr	3.3(0.7)	**2.0**(0.2)	**1.2**(0.1)	**0.95**(0.2)	77(55)	10/15

Table 106: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*1.0e+2*:357	*6.3e+1*:560	*4.0e+1*:684	*2.5e+1*:756	*1.0e+1*:1716	15/15
BSifeg	19(39)	66(100)	59(81)	57(78)	∞ *2e5*	0/15
BSif	73(20)	115(91)	147(295)	181(258)	∞ *2e5*	0/15
BSqi	7.4(3)	26(26)	25(45)	27(24)	∞ *2e5*	0/5
BSrr	16(11)	65(100)	57(4)	55(107)	∞ *2e5*	0/15
CMA-CSA	**2.2**(1)	**2.0**(4)	**1.8**(2)	**1.9**(0.1)	**3.8**(0.6)	15/15
CMA-MSR	**2.0**(0.2)	**2.4**(1)	**2.1**(4)	**2.1**(1)	**3.8**(0.9)	15/15
CMA-TPA	**1.5**(0.3)	**1.0**(0.3)	**0.97**(0.2)	**1.0**(0.4)	**3.8**(0.7)	15/15
GP1-CMAES	**2.1**(0.8)	**1.6**(0.5)	**1.4**(0.4)	**1.5**(0.4)	∞ *5006*	0/15
GP5-CMAES	12(7)	12(14)	13(9)	12(24)	∞ *5020*	0/15
IPOPCMAv3p	3.6(4)	3.6(0.9)	3.1(5)	**3.0**(3)	∞ *5006*	0/15
LHD-10xDef	8.0(4)	27(35)	∞	∞	∞ *1000*	0/15
LHD-2xDefa	5.3(6)	13(8)	22(21)	∞	∞ *1000*	0/15
RAND-2xDef	5.3(6)	13(23)	22(24)	∞	∞ *1000*	0/15
RF1-CMAES	3.8(1)	5.8(8)	5.3(4)	5.8(4)	∞ *5006*	0/15
RF5-CMAES	∞	∞	∞	∞	∞ *5006*	0/15
Sifeg	3.4(5)	19(28)	16(45)	15(6)	∞ *2e5*	0/15
Sif	16(6)	29(76)	26(29)	25(45)	∞ *2e5*	0/15
Srr	**1.9**(3)	16(3)	14(27)	13(19)	∞ *2e5*	0/15

Table 107: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*1.6e+6*:15	*1.0e+6*:27	*4.0e+5*:70	*6.3e+4*:231	*4.0e+3*:1015	15/15
BSifeg	**2.4**(2)	**2.0**(0.9)	**1.6**(1)	479(636)	∞ *2e5*	0/15
BSif	**2.4**(2)	**2.0**(1)	**1.6**(0.4)	494(497)	∞ *2e5*	0/15
BSqi	**2.4**(2)	**2.0**(1)	**1.5**(0.6)	221(501)	∞ *2e5*	0/15
BSrr	**2.4**(2)	**2.0**(1)	**1.5**(0.6)	204(583)	∞ *9e4*	0/15
CMA-CSA	7.8(4)	8.3(8)	7.5(4)	5.8(1)	**3.0**(0.5)	15/15
CMA-MSR	8.1(4)	5.8(3)	4.0(2)	**3.4**(1)	**2.7**(0.7)	15/15
CMA-TPA	7.5(6)	6.5(2)	4.7(1)	**3.3**(1)	**2.6**(0.6)	15/15
GP1-CMAES	5.4(4)	5.4(3)	4.2(0.8)	3.5(0.7)	**2.3**(0.9)	15/15
GP5-CMAES	4.8(1)	3.6(2)	3.1(0.7)	**2.0**(0.4)*	**1.1**(0.4)	15/15
IPOPCMAv3p	4.8(5)	4.3(4)	5.1(2)	5.0(1.0)	3.2(0.7)	15/15
LHD-10xDef	15(10)	12(8)	7.5(0.9)	21(19)	∞ *1000*	0/15
LHD-2xDefa	6.7(4)	4.8(3)	3.3(2)	5.8(4)	∞ *1000*	0/15
RAND-2xDef	6.4(2)	4.6(3)	3.6(1)	6.9(9)	∞ *1000*	0/15
RF1-CMAES	5.9(3)	5.3(3)	4.1(0.9)	3.9(1)	74(110)	1/15
RF5-CMAES	5.1(4)	5.2(3)	10(19)	145(184)	∞ *5006*	0/15
Sifeg	**2.5**(2)	**2.0**(2)	**1.9**(1)	21(30)	∞ *1e5*	0/15
Sif	**2.5**(1.0)	**2.0**(1)	**1.7**(1)	32(15)	∞ *1e5*	0/15
Srr	**2.5**(2)	**2.0**(2)	**1.6**(1.0)	14(18)	∞ *7e4*	0/15

Table 108: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*4.0e+4*:11	*2.5e+3*:27	*1.6e+2*:313	*1.0e+2*:481	*1.0e+1*:1002	15/15
BSifeg	**1.2**(1.0)	**1.3**(0.5)	796(1249)	4369(6014)	∞ *1e5*	0/15
BSif	**1.2**(1)	**1.3**(0.6)	1191(1314)	∞	∞ *2e5*	0/15
BSqi	**1.2**(0.5)	**1.3**(1)	462(666)	2574(3002)	∞ *2e5*	0/15
BSrr	**1.2**(0.8)	**1.3**(1)	1822(2282)	2506(3240)	∞ *8e4*	0/15
CMA-CSA	**2.2**(2)	3.3(4)	12(1)	**8.1**(1.0)	**4.6**(0.3)	15/15
CMA-MSR	**2.0**(2)	**2.9**(2)	**9.2**(2)	**6.9**(1)	**4.7**(0.5)	15/15
CMA-TPA	**2.2**(2)	**2.5**(2)	10(1)	**7.2**(0.8)	**4.5**(0.3)	15/15
GP1-CMAES	**1.3**(1.0)	**1.8**(3)	14(14)	15(5)	∞ *5006*	0/15
GP5-CMAES	**1.8**(1)	**2.3**(2)	**5.2**(3)	17(12)	∞ *5008*	0/15
IPOPCMAv3p	**1.5**(0.9)	**2.5**(2)	52(91)	74(55)	∞ *5006*	0/15
LHD-10xDef	**2.6**(2)	4.0(4)	22(30)	∞	∞ *1000*	0/15
LHD-2xDefa	**1.9**(1)	**2.6**(3)	22(10)	∞	∞ *1000*	0/15
RAND-2xDef	3.0(1)	**2.5**(3)	∞	∞	∞ *1000*	0/15
RF1-CMAES	**1.5**(0.6)	3.2(3)	**8.6**(5)	12(7)	∞ *5006*	0/15
RF5-CMAES	**1.9**(2)	**2.8**(2)	105(152)	146(174)	∞ *5008*	0/15
Sifeg	**1.2**(1)	**1.3**(1)	478(717)	3784(2685)	∞ *1e5*	0/15
Sif	**1.2**(1)	**1.3**(0.5)	668(1268)	∞	∞ *1e5*	0/15
Srr	**1.2**(1)	**1.3**(1)	529(1045)	2003(2431)	∞ *7e4*	0/15

Table 109: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*1.0e+8*:23	*6.3e+7*:39	*2.5e+7*:76	*4.0e+6*:209	*1.0e+1*:1042	15/15
BSifeg	3.5(0.9)	11(0.6)	16(38)	25(62)	∞ *1e5*	0/15
BSif	**3.4**(2)	9.3(28)	29(130)	39(27)	∞ *1e5*	0/15
BSqi	**3.2**(0.7)	12(36)	47(64)	63(231)	173(106)	9/15
BSrr	**3.3**(2)	11(33)	89(48)	69(211)	677(608)	2/15
CMA-CSA	5.4(2)	4.5(2)	**3.7**(0.7)	**2.3**(0.4)	**3.6**(3)	15/15
CMA-MSR	6.2(2)	4.9(2)	3.8(0.8)	**2.5**(0.2)	3.7(3)	15/15
CMA-TPA	7.2(3)	5.2(1)	3.8(0.6)	**2.0**(0.5)	3.8(4)	15/15
GP1-CMAES	4.7(3)	3.9(3)	3.7(0.9)	**2.8**(0.9)	**2.4**(4)	13/15
GP5-CMAES	18(82)	22(40)	38(31)	74(142)	21(12)	3/15
IPOPCMAv3p	5.1(3)	4.7(0.4)	3.8(1)	**2.5**(0.3)	3.8(2)	11/15
LHD-10xDef	17(7)	12(0.7)	8.9(4)	17(14)	∞ *1000*	0/15
LHD-2xDefa	4.6(2)	4.1(0.7)	3.8(0.8)	3.2(1)	∞ *1000*	0/15
RAND-2xDef	5.0(1)	4.0(2)	**3.7**(1)	3.8(3)	∞ *1000*	0/15
RF1-CMAES	3.9(2)	3.9(1)	**3.1**(0.5)	**1.8**(0.1)	**3.0**(2)	12/15
RF5-CMAES	5.2(2)	4.6(1)	9.4(2)	39(108)	∞ *5006*	0/15
Sifeg	3.7(0.6)	**3.0**(0.3)	8.1(20)	26(19)	∞ *4e4*	0/15
Sif	3.7(3)	**3.2**(4)	12(37)	23(12)	∞ *5e4*	0/15
Srr	4.6(0.6)	**3.8**(3)	28(19)	20(28)	∞ *4e4*	0/15

Table 110: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*1.6e+3*:28	*1.0e+3*:64	*6.3e+2*:79	*4.0e+1*:211	*2.5e+0*:1724	15/15
BSifeg	**2.2**(0.3)	**1.3**(0.1)	**1.5**(0.1)	125(442)	99(70)	11/15
BSif	**2.2**(0.7)	**1.3**(0.2)	**1.5**(0.1)	1842(1760)	∞ *2e5*	0/15
BSqi	**2.2**(0.6)	**1.3**(0.2)	**1.4**(0.2)	71(9)	106(67)	9/15
BSrr	**2.2**(0.5)	**1.3**(0.1)	**1.4**(0.2)	67(143)	88(62)	11/15
CMA-CSA	4.3(2)	3.6(0.7)	4.7(0.8)	**5.4**(0.4)	**2.8**(2)	15/15
CMA-MSR	5.6(1)	4.3(1)	4.9(0.4)	6.1(0.3)	**2.8**(1)	15/15
CMA-TPA	3.9(1.0)	3.0(0.9)	3.7(0.6)	**5.3**(0.3)	**4.0**(3)	15/15
GP1-CMAES	**2.5**(0.9)	**2.4**(0.4)	**2.9**(0.5)	49(26)	42(52)	1/15
GP5-CMAES	**2.4**(0.4)	**1.5**(0.3)	**1.6**(0.3)	**2.9**(0.8)$^{\star 2}$	4.5(14)	7/15
IPOPCMAv3p	3.3(2)	3.9(1)	5.5(0.8)	8.0(3)	9.1(7)	4/15
LHD-10xDef	15(0.1)	6.9(0.1)	6.3(0.2)	8.5(7)	∞ *1000*	0/15
LHD-2xDefa	3.3(1)	**2.2**(0.2)	**2.8**(2)	9.4(9)	∞ *1000*	0/15
RAND-2xDef	3.5(0.5)	**2.2**(0.3)	**2.9**(0.5)	7.1(6)	∞ *1000*	0/15
RF1-CMAES	3.2(1)	**3.0**(1)	3.9(1)	7.1(1)	7.3(5)	5/15
RF5-CMAES	3.4(2)	**3.0**(0.7)	4.2(1)	343(659)	∞ *5006*	0/15
Sifeg	**2.2**(0.7)	**1.3**(0.1)	**1.4**(0.1)	19(41)	43(29)	15/15
Sif	**2.2**(0.8)	**1.3**(0.1)	**1.4**(0.1)	23(84)	81(59)	11/15
Srr	**2.2**(0.6)	**1.3**(0.1)	**1.4**(0.1)	16(77)	64(111)	13/15

Table 111: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*2.5e+1*:15	*1.6e+1*:42	*1.0e+1*:75	*1.6e+0*:219	*6.3e-4*:1106	15/15
BSifeg	3.3(2)	**1.7**(0.4)	**1.5**(0.6)	10(9)	∞ *2e5*	0/15
BSif	3.3(2)	**1.7**(0.7)	**1.5**(0.8)	9.1(10)	∞ *2e5*	0/15
BSqi	3.9(1)	**2.0**(0.3)	**1.5**(0.3)	4.7(3)	∞ *2e5*	0/4
BSrr	3.3(2)	**1.7**(1.0)	**1.4**(0.6)	6.4(5)	∞ *2e5*	0/15
CMA-CSA	9.1(7)	4.8(2)	4.2(1)	**2.9**(0.3)	**3.1**(0.4)	15/15
CMA-MSR	8.7(2)	4.7(2)	4.2(0.8)	**2.7**(0.4)	**2.7**(0.2)	15/15
CMA-TPA	8.9(8)	4.8(3)	3.5(2)	**2.4**(0.3)	**2.6**(0.4)	15/15
GP1-CMAES	7.9(5)	3.9(1)	**3.0**(0.7)	**2.3**(0.6)	4.6(2)	13/15
GP5-CMAES	5.7(2)	**2.6**(0.6)	**2.1**(0.3)	**1.7**(0.5)	67(23)	1/15
IPOPCMAv3p	10(3)	4.6(1)	3.7(1)	**2.9**(0.3)	3.9(0.4)	15/15
LHD-10xDef	25(7)	11(0.8)	6.9(0.4)	7.7(6)	∞ *1000*	0/15
LHD-2xDefa	8.1(2)	3.7(1)	3.1(1)	3.8(3)	∞ *1000*	0/15
RAND-2xDef	8.8(2)	4.1(0.8)	3.3(0.9)	7.3(6)	∞ *1000*	0/15
RF1-CMAES	7.1(3)	4.4(2)	3.5(0.9)	3.1(0.9)	33(32)	2/15
RF5-CMAES	6.5(7)	3.9(2)	3.7(1)	153(104)	∞ *5006*	0/15
Sifeg	**3.3**(2)	**1.8**(0.5)	**1.3**(0.4)	**1.4**(0.7)	∞ *2e5*	0/15
Sif	**3.3**(2)	**1.8**(0.8)	**1.3**(0.4)	**1.5**(0.9)	∞ *2e5*	0/15
Srr	**3.3**(2)	**1.8**(0.7)	**1.3**(0.5)	**1.1**(0.7)	∞ *2e5*	0/15

Table 112: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*6.3e+2*:15	*4.0e+2*:67	*2.5e+2*:292	*1.6e+2*:846	*1.0e+2*:1671	15/15
BSifeg	**2.2**(2)	18(0.6)	274(399)	525(570)	1563(1672)	1/15
BSif	**2.2**(2)	76(563)	296(322)	943(1318)	1581(3030)	1/15
BSqi	**2.1**(1)	15(0.3)	224(291)	346(493)	∞ *2e5*	0/11
BSrr	**2.2**(1)	5.7(33)	272(279)	450(534)	1391(3523)	1/15
CMA-CSA	4.9(2)	**2.5**(0.6)	**1.1**(0.2)	**0.86**(0.2)	**1.5**(0.6)	15/15
CMA-MSR	6.2(4)	**2.8**(1)	**1.1**(0.3)	**0.68**(0.1)	**0.44**(0.1)	15/15
CMA-TPA	6.1(3)	**2.3**(0.4)	**0.87**(0.3)	**0.70**(0.2)	**0.72**(0.1)	15/15
GP1-CMAES	3.6(2)	**1.9**(0.8)	**0.81**(0.2)	**0.78**(0.4)	**0.64**(0.2)	15/15
GP5-CMAES	3.7(3)	**1.5**(0.3)	**0.60**(0.2)	**2.2**(8)	3.9(3)	7/15
IPOPCMAv3p	4.5(3)	**2.6**(1)	**1.2**(0.3)	**1.2**(0.4)	**1.4**(2)	14/15
LHD-10xDef	16(14)	6.4(0.1)	**1.9**(0.4)	**1.1**(0.2)	**1.2**(0.5)	7/15
LHD-2xDefa	5.5(2)	**2.0**(0.9)	**1.1**(0.9)	**1.5**(1)	8.7(6)	1/15
RAND-2xDef	5.1(2)	**1.9**(0.5)	**1.1**(0.5)	**1.6**(0.4)	8.8(6)	1/15
RF1-CMAES	4.5(2)	**2.2**(1)	**0.93**(0.2)	**0.92**(0.2)	**0.70**(0.1)	15/15
RF5-CMAES	3.6(2)	**1.9**(0.4)	**0.95**(0.5)	**1.7**(2)	6.5(6)	5/15
Sifeg	**2.2**(1)	46(168)	49(91)	126(362)	774(620)	2/15
Sif	**2.2**(2)	52(0.5)	74(144)	170(170)	703(1812)	2/15
Srr	**2.2**(2)	52(190)	72(135)	100(124)	352(371)	4/15

Table 113: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:26	*2.5e+1*:127	*1.6e+1*:540	*1.6e+1*:540	*1.0e+1*:1384	15/15
BSifeg	**1.7**(1)	**0.92**(0.2)	44(119)	44(89)	123(354)	9/15
BSif	**1.7**(1)	**0.92**(0.2)	42(89)	42(88)	178(188)	8/15
BSqi	**1.7**(1)	**0.93**(0.2)	21(42)	21(0.7)	98(223)	10/13
BSrr	**1.7**(0.9)	**0.93**(0.2)	40(46)	40(134)	119(89)	10/15
CMA-CSA	7.8(10)	16(6)	4.3(2)	4.3(1)	**1.9**(0.6)	15/15
CMA-MSR	5.2(2)	**2.1**(0.4)	**0.81**(0.3)	**0.81**(0.4)	**0.80**(1)	15/15
CMA-TPA	4.5(8)	8.4(6)	**2.2**(1)	**2.2**(1)	**1.2**(0.7)	15/15
GP1-CMAES	3.3(2)	4.3(2)	**1.4**(0.3)	**1.4**(0.2)	**0.90**(0.1)	14/15
GP5-CMAES	3.4(3)	**1.6**(0.6)	**0.54**(0.2)	**0.54**(0.2)	**0.57**(0.2)	15/15
IPOPCMAv3p	3.4(4)	10(4)	3.2(1)	3.2(1)	**1.4**(0.4)	15/15
LHD-10xDef	6.5(7)	5.4(2)	4.8(3)	4.8(3)	3.4(4)	3/15
LHD-2xDefa	3.2(3)	3.6(2)	**2.6**(3)	**2.6**(3)	5.1(8)	2/15
RAND-2xDef	3.8(3)	6.4(8)	5.0(8)	5.0(8)	11(16)	1/15
RF1-CMAES	**2.8**(4)	4.6(3)	**1.6**(0.5)	**1.6**(0.4)	**0.79**(0.2)	15/15
RF5-CMAES	4.2(4)	**2.2**(2)	**0.83**(0.2)	**0.83**(0.3)	**1.0**(3)	13/15
Sifeg	**1.9**(2)	**1.5**(0.6)	10(9)	10(7)	15(22)	15/15
Sif	**1.9**(2)	**1.6**(1)	6.1(6)	6.1(6)	12(12)	15/15
Srr	**1.9**(1)	**1.8**(3)	3.3(12)	3.3(0.9)	14(10)	15/15

Table 114: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.6e+1*:11	*1.0e+1*:63	*6.3e+0*:305	*4.0e+0*:468	*1.0e+0*:1030	15/15
BSifeg	22(2)	493(0.8)	1042(852)	∞	∞ *2e5*	0/15
BSif	35(3)	488(791)	2637(4581)	∞	∞ *2e5*	0/15
BSqi	**2.6**(1)	**1.7**(0.6)	346(328)	∞	∞ *2e5*	0/3
BSrr	27(3)	476(1)	1359(2006)	∞	∞ *2e5*	0/15
CMA-CSA	7.5(5)	**3.0**(2)	**1.1**(0.2)	**1.0**(0.2)	**1.0**(0.2)	15/15
CMA-MSR	7.9(3)	**2.7**(0.6)	**0.90**(0.2)	**1.0**(0.2)	6.5(3)	15/15
CMA-TPA	8.8(7)	**2.7**(0.7)	**0.93**(0.1)	**0.95**(0.3)	**1.4**(0.4)	15/15
GP1-CMAES	**2.4**(3)	**1.4**(0.8)	**0.73**(0.3)	**0.79**(0.3)	3.4(10)	10/15
GP5-CMAES	3.3(2)	**1.6**(0.8)	**0.79**(0.4)	**0.87**(0.6)	11(27)	5/15
IPOPCMAv3p	**2.8**(2)	**2.0**(2)	**0.89**(0.3)	**0.95**(0.3)	**0.99**(0.2)	15/15
LHD-10xDef	12(9)	7.3(2)	**2.2**(0.4)	5.7(2)	∞ *1000*	0/15
LHD-2xDefa	3.3(4)	**2.6**(0.8)	**1.4**(2)	10(8)	∞ *1000*	0/15
RAND-2xDef	5.0(4)	**2.7**(2)	**2.1**(2)	31(38)	∞ *1000*	0/15
RF1-CMAES	3.1(3)	**1.9**(0.7)	**0.78**(0.3)	**0.83**(0.3)	4.2(10)	9/15
RF5-CMAES	4.0(2)	**2.7**(0.8)	3.6(5)	17(13)	∞ *5006*	0/15
Sifeg	3.8(6)	230(797)	797(1392)	5983(5449)	∞ *2e5*	0/15
Sif	3.6(5)	7.1(18)	750(822)	5964(6604)	∞ *2e5*	0/15
Srr	3.6(3)	3.8(6)	979(1795)	5963(5002)	∞ *2e5*	0/15

Table 115: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*4.0e+1*:116	*2.5e+1*:252	*1.6e+1*:430	*1.0e+1*:621	*4.0e+0*:1090	15/15
BSifeg	5.9(13)	468(752)	5872(6811)	∞	∞ *2e5*	0/14
BSif	123(7)	758(760)	∞	∞	∞ *2e5*	0/15
BSqi	**0.80**(0.3)	368(889)	∞	∞	∞ *2e5*	0/4
BSrr	26(0.2)	649(989)	∞	∞	∞ *2e5*	0/15
CMA-CSA	**1.3**(0.3)	**1.0**(0.3)	**0.96**(0.1)	**0.96**(0.3)	**0.96**(0.3)	15/15
CMA-MSR	**1.3**(0.5)	**1.0**(0.3)	**0.97**(0.6)	**2.8**(0.2)	4.8(13)	15/15
CMA-TPA	**1.3**(0.7)	**1.0**(0.7)	**0.92**(0.3)	**1.6**(3)	**1.3**(0.3)	15/15
GP1-CMAES	**0.81**(0.3)	**0.84**(0.3)	**0.85**(0.3)	**0.93**(0.5)	5.2(6)	8/15
GP5-CMAES	**0.95**(0.9)	**0.85**(0.6)	**1.7**(3)	**2.8**(4)	19(22)	3/15
IPOPCMAv3p	**0.94**(0.7)	**1.1**(0.2)	**1.1**(0.4)	**1.1**(0.4)	**1.1**(0.2)	15/15
LHD-10xDef	3.5(0.5)	**2.7**(0.8)	5.2(3)	∞	∞ *1000*	0/15
LHD-2xDefa	**1.00**(0.6)	**1.6**(0.7)	11(6)	∞	∞ *1000*	0/15
RAND-2xDef	**0.99**(0.4)	**1.6**(1)	16(11)	∞	∞ *1000*	0/15
RF1-CMAES	**0.87**(0.4)	**0.82**(0.2)	**0.87**(0.2)	**1.0**(0.2)	10(9)	5/15
RF5-CMAES	**1.3**(0.4)	**2.7**(1)	5.4(3)	53(99)	∞ *5006*	0/15
Sifeg	48(9)	217(303)	1821(1148)	∞	∞ *2e5*	0/15
Sif	91(59)	367(370)	1063(2314)	4564(4318)	∞ *2e5*	0/15
Srr	30(8)	132(7)	801(1299)	∞	∞ *2e5*	0/15

Table 116: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\text{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:2.5e5	*1.0e-1*:3.4e5	*6.3e-2*:3.4e5	*4.0e-2*:3.4e5	*2.5e-2*:3.4e5	3/15
BSifeg	∞	∞	∞	∞	∞ *2e5*	0/15
BSif	∞	∞	∞	∞	∞ *2e5*	0/15
BSqi	∞	∞	∞	∞	∞ *2e5*	0/8
BSrr	∞	∞	∞	∞	∞ *2e5*	0/15
CMA-CSA	**0.89**(0.5)	**0.82**(0.5)	**0.87**(0.5)	**2.9**(2)	**5.6**(4)	11/15
CMA-MSR	**1.2**(0.5)	**1.2**(0.4)	**1.7**(0.9)	**3.0**(5)	**15**(28)	5/15
CMA-TPA	**1.3**(0.6)	**1.6**(0.6)	**2.3**(0.9)	**3.4**(2)	**7.9**(8)	9/15
GP1-CMAES	∞	∞	∞	∞	∞ *5006*	0/15
GP5-CMAES	∞	∞	∞	∞	∞ *5020*	0/15
IPOPCMAv3p	∞	∞	∞	∞	∞ *5008*	0/15
LHD-10xDef	∞	∞	∞	∞	∞ *1000*	0/15
LHD-2xDefa	∞	∞	∞	∞	∞ *1000*	0/15
RAND-2xDef	∞	∞	∞	∞	∞ *1000*	0/15
RF1-CMAES	∞	∞	∞	∞	∞ *5008*	0/15
RF5-CMAES	∞	∞	∞	∞	∞ *5034*	0/15
Sifeg	∞	∞	∞	∞	∞ *2e5*	0/15
Sif	∞	∞	∞	∞	∞ *2e5*	0/15
Srr	∞	∞	∞	∞	∞ *2e5*	0/15

Table 117: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*1.6e+4*:38	*1.0e+4*:42	*2.5e+2*:62	*2.5e+0*:250	*1.6e+0*:2536	15/15
BSifeg	**1.8**(0.6)	**1.8**(0.6)	**1.6**(0.6)	**2.1**(4)	3.2(0.8)	15/15
BSif	**1.8**(1)	**1.8**(0.3)	**1.6**(0.8)	**1.6**(0.8)	**1.4**(1)	15/15
BSqi	**1.1**(1)	**1.5**(1)	**1.5**(0.5)	**1.0**(0.3)	**1.3**(0.3)	4/4
BSrr	**1.8**(0.5)	**1.8**(1)	**1.6**(0.5)	**1.1**(0.5)	4.8(14)	15/15
CMA-CSA	3.2(1)	3.7(0.6)	5.5(2)	7.6(9)	18(13)	15/15
CMA-MSR	3.8(0.8)	4.2(1)	5.9(1)	3.2(0.5)	4477(3623)	2/15
CMA-TPA	3.4(0.7)	3.8(0.4)	4.3(0.9)	5.8(0.6)	432(576)	10/15
GP1-CMAES	**2.8**(1)	3.2(1)	3.6(0.2)	4.0(2)	∞ *5006*	0/15
GP5-CMAES	**2.3**(0.5)	**2.2**(0.2)	**2.5**(0.4)	284(618)	∞ *5022*	0/15
IPOPCMAv3p	3.5(1)	4.3(2)	6.1(0.9)	5.9(0.5)	6.2(6)	4/15
LHD-10xDef	11(2)	10(0.2)	8.7(0.5)	7.9(10)	∞ *1000*	0/15
LHD-2xDefa	3.3(0.6)	3.3(0.6)	4.4(2)	**2.8**(2)	5.7(4)	1/15
RAND-2xDef	3.1(0.6)	3.1(0.7)	4.8(3)	4.9(2)	∞ *1000*	0/15
RF1-CMAES	3.5(0.7)	3.9(0.6)	5.2(1)	3.4(1)	6.1(7)	4/15
RF5-CMAES	**2.6**(0.8)	**2.9**(1)	6.6(4)	134(150)	∞ *5006*	0/15
Sifeg	**1.8**(0.6)	**1.8**(0.3)	**1.9**(0.3)	**0.75**(0.1)	**0.55**(0.6)	15/15
Sif	**1.8**(1)	**1.8**(0.6)	**1.9**(0.7)	**0.82**(0.4)	**0.64**(0.5)	15/15
Srr	**1.8**(0.6)	**1.8**(0.7)	**1.9**(0.4)	**0.71**(0.1)	**0.51**(0.3)	15/15

Table 118: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*6.3e+1*:36	*4.0e+1*:77	*4.0e+1*:77	*1.6e+1*:456	*4.0e+0*:1094	15/15
BSifeg	**2.0**(0.1)	**1.2**(0.2)	**1.2**(0.2)	87(219)	98(149)	11/15
BSif	**2.0**(0.3)	**1.2**(0.2)	**1.2**(0.2)	93(158)	146(109)	9/15
BSqi	**2.1**(0.1)	**1.4**(0.5)	**1.4**(0.5)	187(401)	79(117)	3/4
BSrr	**2.0**(0.1)	**1.2**(0.1)	**1.2**(0.6)	80(44)	162(230)	8/15
CMA-CSA	5.6(2)	3.4(0.9)	3.4(0.5)	**2.5**(3)	99(520)	14/15
CMA-MSR	5.0(1)	3.1(0.8)	3.1(0.9)	**1.6**(0.2)	234(2)	13/15
CMA-TPA	4.9(2)	3.3(1)	3.3(1)	17(4)	198(577)	13/15
GP1-CMAES	3.8(2)	6.3(0.4)	6.3(0.6)	**1.8**(5)	3.4(1)	10/15
GP5-CMAES	3.2(1.0)	**2.9**(0.4)	**2.9**(8)	**1.8**(2)	7.7(10)	6/15
IPOPCMAv3p	5.8(5)	5.5(4)	5.5(2)	3.2(4)	7.8(7)	6/15
LHD-10xDef	13(0.8)	6.6(0.6)	6.6(0.9)	**2.0**(2)	**2.4**(4)	5/15
LHD-2xDefa	3.6(0.6)	**2.6**(1)	**2.6**(3)	**0.88**(0.3)	**1.2**(0.8)	9/15
RAND-2xDef	3.4(0.6)	**1.9**(0.4)	**1.9**(0.5)	**0.46**(0.1)	**0.57**(0.5)	12/15
RF1-CMAES	5.2(2)	3.7(1)	3.7(1)	5.5(6)	8.0(10)	6/15
RF5-CMAES	7.6(2)	6.1(2)	6.1(1)	8.7(6)	14(7)	4/15
Sifeg	**2.0**(0.3)	**1.3**(0.3)	**1.3**(0.5)	65(164)	87(267)	12/15
Sif	**2.0**(0.5)	**1.3**(0.5)	**1.3**(0.1)	106(141)	136(259)	9/15
Srr	**2.0**(0.3)	**1.3**(0.6)	**1.3**(0.3)	67(103)	160(92)	8/15

Table 119: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:45	*4.0e+1*:68	*4.0e+1*:68	*1.6e+1*:231	*6.3e+0*:1219	15/15
BSifeg	**1.6**(0.2)	4.4(24)	4.4(12)	166(86)	156(297)	8/15
BSif	**1.6**(0.3)	6.1(0.3)	6.1(0.1)	196(144)	225(246)	7/15
BSqi	**1.6**(0.0)	6.6(14)	6.6(0.0)	228(445)	230(456)	2/5
BSrr	**1.6**(0.3)	3.5(9)	3.5(9)	167(263)	117(123)	10/15
CMA-CSA	4.6(1)	9.1(2)	9.1(17)	38(98)	230(458)	12/15
CMA-MSR	5.8(1)	14(18)	14(35)	7.9(11)	206(329)	13/15
CMA-TPA	4.0(1)	6.8(1)	6.8(1)	327(0.8)	428(431)	10/15
GP1-CMAES	12(29)	17(17)	17(20)	12(11)	**2.7**(1)	10/15
GP5-CMAES	**2.5**(0.3)	3.6(6)	3.6(12)	4.8(8)	**2.0**(7)	11/15
IPOPCMAv3p	4.6(2)	10(2)	10(2)	6.7(6)	4.6(5)	8/15
LHD-10xDef	10(0.4)	7.8(2)	7.8(2)	5.4(5)	**2.1**(1)	5/15
LHD-2xDefa	3.4(0.6)	3.3(3)	3.3(2)	**2.0**(3)	**1.2**(2)	8/15
RAND-2xDef	3.3(0.6)	3.2(2)	3.2(0.3)	**1.9**(2)	**1.1**(1)	8/15
RF1-CMAES	4.6(3)	5.2(3)	5.2(1)	**4.6**(6)	**3.0**(0.4)	10/15
RF5-CMAES	5.8(3)	11(38)	11(22)	11(17)	5.7(4)	7/15
Sifeg	**1.6**(0.4)	**1.7**(4)	**1.7**(0.2)	230(351)	159(287)	9/15
Sif	**1.6**(0.1)	**1.9**(6)	**1.9**(0.2)	118(156)	125(125)	10/15
Srr	**1.6**(0.4)	**1.5**(0.2)	**1.5**(3)	170(497)	133(101)	9/15

Table 120: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*6.3e+0*:29	*4.0e+0*:118	*2.5e+0*:306	*2.5e+0*:306	*1.0e+0*:1614	15/15
BSifeg	**2.3**(2)	**1.4**(0.4)	**1.9**(2)	**1.9**(2)	60(52)	14/15
BSif	**2.3**(2)	**1.4**(0.4)	**2.8**(10)	**2.8**(5)	37(39)	15/15
BSrr	**2.3**(2)	**1.4**(0.5)	**2.5**(2)	**2.5**(5)	54(25)	14/15
CMA-CSA	6.7(4)	8.1(5)	37(35)	37(45)	93(10)	15/15
CMA-MSR	4.1(3)	3.9(2)	4.3(0.6)	4.3(10)	**2.0**(6)	15/15
CMA-TPA	5.2(3)	12(10)	34(47)	34(50)	23(38)	15/15
GP1-CMAES	**2.1**(2)	6.4(4)	5.8(5)	5.8(2)	**1.6**(0.9)	14/15
GP5-CMAES	**2.2**(3)	**2.9**(2)	**1.9**(4)	**1.9**(0.2)	**0.84**(0.8)	15/15
IPOPCMAv3p	**1.5**(2)	6.7(4)	75(126)	75(72)	∞ *5020*	0/15
LHD-10xDef	**2.3**(3)	8.2(16)	48(43)	48(25)	∞ *1000*	0/15
LHD-2xDefa	**1.6**(2)	5.6(7)	23(20)	23(22)	∞ *1000*	0/15
RAND-2xDef	**2.1**(0.8)	6.2(5)	∞	∞	∞ *1000*	0/15
RF1-CMAES	**1.3**(0.9)	3.4(6)	244(139)	244(275)	∞ *5010*	0/15
RF5-CMAES	**1.8**(3)	4.5(6)	113(313)	113(144)	∞ *5086*	0/15
Sifeg	**2.3**(2)	4.7(3)	4.2(3)	4.2(2)	6.5(5)	15/15
Sif	**2.3**(2)	4.7(2)	4.4(2)	4.4(2)	11(9)	15/15
Srr	**2.3**(1)	4.8(3)	3.9(1)	3.9(2)	5.2(8)	15/15

Table 121: 20-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*2.5e+2*:208	*1.6e+2*:918	*1.0e+2*:6628	*6.3e+1*:9885	*4.0e+1*:31629	15/15
BSifeg	61(21)	73(139)	56(58)	∞	∞ *2e5*	0/15
BSif	76(70)	187(223)	196(301)	∞	∞ *2e5*	0/15
BSrr	63(152)	78(66)	75(81)	∞	∞ *2e5*	0/15
CMA-CSA	**1.5**(0.2)	**1.4**(0.2)	**1.4**(0.5)	**0.99**(1)	**0.87**(0.8)	15/15
CMA-MSR	**1.3**(0.3)	**0.73**(0.1)	**0.21**(0.3)	**0.61**(0.4)	**0.34**(0.2)	15/15
CMA-TPA	**1.2**(0.3)	**1.7**(1.0)	**0.71**(0.3)	**0.67**(0.3)	**0.92**(0.3)	15/15
GP1-CMAES	**0.83**(0.3)	**0.84**(0.3)	**0.20**(0.0)	**0.24**(0.1)	**0.38**(0.6)	5/15
GP5-CMAES	**0.51**(0.1)$^{\star}_{\downarrow 4}$	**0.86**(0.2)	**0.37**(0.2)	**0.32**(0.3)	**1.1**(0.9)	2/15
IPOPCMAv3p	**1.4**(0.4)	**1.7**(1)	5.5(3)	7.5(9)	∞ *5008*	0/15
LHD-10xDef	3.0(3)	**1.6**(0.8)	**0.74**(0.8)	∞	∞ *1000*	0/15
LHD-2xDefa	7.5(3)	7.8(10)	∞	∞	∞ *1000*	0/15
RAND-2xDef	9.3(12)	∞	∞	∞	∞ *1000*	0/15
RF1-CMAES	**1.4**(0.3)	**2.5**(2)	**1.0**(1)	**2.4**(4)	**2.3**(3)	1/15
RF5-CMAES	**2.2**(1)	4.1(4)	11(21)	∞	∞ *5034*	0/15
Sifeg	16(45)	19(27)	23(9)	258(219)	∞ *2e5*	0/15
Sif	19(16)	20(30)	28(44)	∞	∞ *2e5*	0/15
Srr	17(12)	12(11)	13(17)	∞	∞ *2e5*	0/15

Table 122: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_1 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*2.5e+2*:48	*1.6e+2*:82	*1.0e-8*:83	*1.0e-8*:83	*1.0e-8*:83	15/15
CMA-CSA	**2.0**(0.5)	2.3(0.9)	64(4)	64(2)	64(4)	15/15
CMA-MSR	2.6(0.9)	2.8(0.7)	66(3)	66(3)	66(2)	15/15
CMA-TPA	2.5(0.7)	**2.3**(0.4)	**43**(1)$^{\star 4}$	**43**(2)$^{\star 4}$	**43**(2)$^{\star 4}$	15/15

Table 123: 40-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\mathrm{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*1.0e+7*:39	*6.3e+6*:71	*4.0e+5*:121	*2.5e+4*:499	*1.0e-8*:1188	15/15
CMA-CSA	**1.3**(0.9)	1.9(0.9)	17(8)	12(1)	**40**(1)$^{\star 4}$	15/15
CMA-MSR	2.0(1)	**1.8**(1)	**8.0**(2)	**8.2**(2)	47(0.9)	15/15
CMA-TPA	2.1(2)	2.2(0.8)	8.5(2)	8.9(3)	46(0.9)	15/15

Table 124: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*1.6e+3*:68	*1.0e+3*:222	*6.3e+2*:471	*4.0e+2*:662	*6.3e+1*:6332	15/15
CMA-CSA	**1.8**(1)	1.4(0.4)	1.1(0.2)	1.8(0.3)	3.6(1)	15/15
CMA-MSR	2.3(1)	1.3(0.5)	1.1(0.2)	**1.3**(0.2)	4.0(1)	15/15
CMA-TPA	2.3(1)	**1.1**(0.3)	**0.83**(0.1)*	1.4(0.5)	**2.7**(2)	15/15

Table 125: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*1.0e+3*:439	*6.3e+2*:670	*4.0e+2*:707	*2.5e+2*:735	*1.0e+2*:5369	15/15
CMA-CSA	1.2(0.4)	1.5(0.2)	3.3(0.8)	4.5(0.8)	**3.6**(3)	15/15
CMA-MSR	1.1(0.2)	1.2(0.3)	**1.8**(0.4)	12(8)	8.1(3)	15/15
CMA-TPA	**0.94**(0.2)	**1.1**(0.1)	2.0(0.4)	**2.5**(0.4)	4.5(2)	15/15

Table 126: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*4.0e+2*:51	*2.5e+2*:81	*1.0e-1*:120	*1.0e-8*:121	*1.0e-8*:121	15/15
CMA-CSA	1.7(0.5)	1.8(0.4)	4.4(0.5)	4.4(0.3)	4.4(0.4)	15/15
CMA-MSR	1.7(0.9)	1.8(0.3)	3.6(0.3)	**3.6**(0.7)	**3.6**(0.5)	15/15
CMA-TPA	**1.2**(0.5)*	**1.2**(0.4)$^{\star 2}$	**3.5**(0.5)	3.6(0.4)	3.6(0.3)	15/15

Table 127: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*6.3e+5*:50	*4.0e+5*:82	*4.0e+4*:127	*4.0e+2*:734	*6.3e+1*:2121	15/15
CMA-CSA	**1.6**(0.5)	**1.4**(0.7)	2.1(0.5)	2.9(0.2)	1.8(0.2)	15/15
CMA-MSR	1.8(0.8)	1.6(0.7)	**2.0**(0.5)	**1.7**(0.2)	1.7(0.9)	15/15
CMA-TPA	1.6(0.4)	1.9(0.5)	2.3(0.7)	1.9(0.3)	**1.5**(0.5)	15/15

Table 128: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*1.6e+3*:35	*1.0e+3*:106	*6.3e+2*:165	*2.5e+2*:489	*2.5e+1*:2987	15/15
CMA-CSA	**2.9**(1)	2.4(0.8)	2.7(1)	1.7(0.3)	**1.1**(0.8)	15/15
CMA-MSR	3.5(2)	2.2(0.2)	2.1(0.5)	1.2(0.3)	6.7(4)	15/15
CMA-TPA	3.4(1)	**2.0**(0.2)	**1.8**(0.4)	**1.1**(0.2)	10(30)	15/15

Table 129: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.0e+5*:85	*6.3e+4*:111	*4.0e+4*:125	*2.5e+3*:430	*6.3e+1*:2106	15/15
CMA-CSA	3.4(1)	3.2(0.3)	3.4(0.7)	1.9(0.2)	3.3(8)	15/15
CMA-MSR	3.4(0.9)	3.1(0.8)	3.1(0.5)	1.7(0.2)	3.3(4)	15/15
CMA-TPA	**2.6**(0.6)	**2.2**(0.5)$^{\star 2}$	**2.3**(0.3)$^{\star 2}$	**1.3**(0.1)$^{\star 3}$	**1.4**(3)	15/15

Table 130: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*2.5e+2*:676	*1.6e+2*:865	*1.0e+2*:1397	*6.3e+1*:1896	*4.0e+1*:2180	15/15
CMA-CSA	2.1(0.3)	1.9(0.3)	1.6(0.4)	2.2(8)	2.2(4)	15/15
CMA-MSR	2.1(0.5)	1.9(0.5)	1.7(1)	2.5(5)	2.4(8)	15/15
CMA-TPA	**1.5**(0.6)*	**1.5**(1)	**1.1**(0.6)	**0.91**(0.5)*	**0.98**(0.3)*	15/15

Table 131: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*1.0e+7*:44	*6.3e+6*:80	*2.5e+6*:126	*2.5e+5*:408	*6.3e+3*:2376	15/15
CMA-CSA	1.5(0.3)	1.7(0.9)	5.3(2)	6.1(1)	4.2(0.7)	15/15
CMA-MSR	1.8(2)	**1.6**(0.7)	**2.3**(0.9)	**3.2**(1)	**3.2**(0.6)	15/15
CMA-TPA	**1.4**(1)	2.0(1)	2.7(0.4)	3.3(0.9)	3.3(0.7)	15/15

Table 132: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*1.0e+4*:22	*2.5e+3*:52	*2.5e+2*:432	*1.6e+2*:887	*1.6e+1*:2204	15/15
CMA-CSA	1.8(2)	1.8(3)	23(1)	12(0.8)	5.3(0.2)	15/15
CMA-MSR	**1.0**(1)	1.5(1.0)	**15**(2)$^{\star 2}$	**8.5**(0.4)$^{\star 2}$	5.0(0.3)	15/15
CMA-TPA	1.6(1)	**1.3**(0.9)	18(2)	9.5(0.7)	**4.8**(0.2)*	15/15

Table 133: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*2.5e+8*:54	*1.6e+8*:218	*1.0e+8*:284	*1.0e+7*:424	*4.0e+1*:2479	15/15
CMA-CSA	**4.7**(2)	1.7(0.4)	1.7(0.2)	2.2(0.1)	**1.5**(0.1)	15/15
CMA-MSR	5.6(2)	1.8(0.5)	1.7(0.5)	2.2(0.3)	2.2(2)	15/15
CMA-TPA	4.8(2)	**1.5**(0.5)	**1.4**(0.3)	**1.7**(0.4)$^{\star 2}$	1.9(2)	15/15

Table 134: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*2.5e+3*:85	*1.6e+3*:121	*1.6e+3*:121	*6.3e+1*:429	*1.0e+1*:2029	15/15
CMA-CSA	2.3(0.6)	3.7(0.6)	3.7(0.3)	5.0(0.6)	2.5(2)	15/15
CMA-MSR	2.4(0.8)	3.4(0.5)	3.4(0.5)	4.5(0.3)	2.8(3)	15/15
CMA-TPA	**2.2**(0.5)	**2.9**(0.5)	**2.9**(0.6)	**3.8**(0.4)$^{\star 3}$	**2.3**(3)	15/15

Table 135: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*6.3e+1*:34	*4.0e+1*:137	*2.5e+1*:176	*4.0e+0*:438	*1.0e-3*:2207	15/15
CMA-CSA	**4.1**(1)	2.2(0.6)	3.0(0.4)	2.7(0.4)	3.6(0.2)	15/15
CMA-MSR	5.7(2)	2.3(0.5)	2.5(0.7)	2.2(0.4)	**2.5**(0.3)	15/15
CMA-TPA	6.0(1)	**2.1**(0.6)	**2.3**(0.6)	**2.1**(0.2)	2.6(0.1)	15/15

Table 136: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*1.0e+3*:192	*6.3e+2*:458	*4.0e+2*:1170	*2.5e+2*:3875	*2.5e+2*:3875	15/15
CMA-CSA	1.5(0.4)	1.1(0.3)	1.1(0.2)	1.0(0.2)	1.0(0.2)	15/15
CMA-MSR	1.4(0.6)	0.98(0.3)	**0.73**(0.1)	**0.30**(0.0)$^{\star 4}$	**0.30**(0.0)$^{\star 4}$	15/15
CMA-TPA	**1.2**(0.3)	**0.88**(0.2)	0.92(0.7)	0.57(0.1)	0.57(0.2)	15/15

Table 137: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:117	*2.5e+1*:297	*1.6e+1*:4010	*1.6e+1*:4010	*1.0e+1*:5244	15/15
CMA-CSA	25(17)	14(3)	1.2(0.3)	1.2(0.1)	1.0(0.5)	15/15
CMA-MSR	**2.4**(0.4)$^{\star 4}$	**1.6**(0.4)$^{\star 4}$	**0.17**(0.0)$^{\star 4}$	**0.17**(0.0)$^{\star 4}$	**0.43**(0.1)$^{\star 2}$	15/15
CMA-TPA	11(4)	5.3(1)	0.48(0.1)	0.48(0.1)	1.3(1)	15/15

Table 138: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.6e+1*:54	*1.0e+1*:399	*6.3e+0*:688	*4.0e+0*:1115	*1.0e+0*:4220	15/15
CMA-CSA	5.4(3)	1.3(0.6)	1.1(0.3)	1.0(0.3)	**0.56**(0.1)	15/15
CMA-MSR	3.9(2)	0.96(0.3)	1.1(0.5)	9.1(4)	7.3(7)	15/15
CMA-TPA	**3.8**(1)	**0.93**(0.3)	**0.99**(0.3)	**1.0**(0.4)	5.0(9)	15/15

Table 139: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*6.3e+1*:55	*4.0e+1*:329	*4.0e+1*:329	*2.5e+1*:579	*6.3e+0*:2006	15/15
CMA-CSA	3.4(2)	1.2(0.4)	1.2(0.2)	1.2(0.3)	**0.97**(0.4)	15/15
CMA-MSR	3.2(0.9)	0.99(0.4)	0.99(0.2)	1.00(0.6)	10(15)	15/15
CMA-TPA	**2.7**(1)	**0.89**(0.2)	**0.89**(0.2)	**0.86**(0.4)	1.6(3)	15/15

Table 140: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:8.6e5	*1.0e-1*:1.4e6	*6.3e-2*:3.1e6	*4.0e-2*:5.2e6	*2.5e-2*:8.7e6	15/15
CMA-CSA	**1.2**(0.7)	**1.0**(0.6)	**0.56**(0.4)	**0.54**(0.2)	**0.66**(0.7)	9/15
CMA-MSR	1.8(0.9)	1.4(0.7)	0.76(0.3)	0.75(0.9)	1.3(2)	5/15
CMA-TPA	1.2(0.6)	1.0(0.2)	0.60(0.2)	0.62(0.4)	0.74(0.4)	8/15

Table 141: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*2.5e+4*:83	*1.6e+4*:86	*1.0e+3*:125	*2.5e+0*:515	*1.6e+0*:5582	15/15
CMA-CSA	4.4(0.8)	4.9(0.7)	5.6(0.9)	5.2(0.7)	**57**(24)$^{\star 2}$	15/15
CMA-MSR	3.8(0.7)	4.2(0.7)	4.6(0.8)	2.6(0.2)	∞ *3e6*	0/15
CMA-TPA	**3.1**(0.4)*	**3.4**(0.5)$^{\star 2}$	**3.7**(0.5)$^{\star 2}$	**2.3**(0.2)	9480(1e4)	1/15

Table 142: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*6.3e+1*:160	*4.0e+1*:305	*2.5e+1*:380	*1.6e+1*:784	*6.3e+0*:2510	30/30
CMA-CSA	2.8(0.5)	1.9(0.4)	1.8(0.5)	**1.1**(0.6)	2.4(3)	15/15
CMA-MSR	2.2(0.7)	1.5(0.4)	**1.6**(0.5)	2.9(0.1)	96(4)	14/15
CMA-TPA	**2.0**(0.2)	**1.4**(0.2)	2.9(13)	2.2(3)	**2.0**(4)	15/15

Table 143: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:160	*4.0e+1*:231	*2.5e+1*:687	*1.6e+1*:1392	*1.0e+1*:3090	15/15
CMA-CSA	2.9(1)	9.1(25)	4.2(9)	57(205)	87(211)	14/15
CMA-MSR	5.8(12)	**4.5**(0.9)	347(2415)	172(2)	77(0.8)	14/15
CMA-TPA	**2.4**(0.3)	7.2(0.3)	**3.2**(3)	**3.0**(7)	**1.6**(1)	15/15

Table 144: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*6.3e+0*:68	*4.0e+0*:292	*2.5e+0*:603	*2.5e+0*:603	*1.6e+0*:2487	15/15
CMA-CSA	12(6)	72(51)	108(79)	108(121)	27(11)	15/15
CMA-MSR	**7.0**(3)	**2.5**(0.2)$^{\star 4}$	**1.4**(0.2)$^{\star 4}$	**1.4**(0.2)$^{\star 4}$	**0.41**(0.1)$^{\star 4}$	15/15
CMA-TPA	12(6)	85(213)	86(205)	86(114)	27(51)	15/15

Table 145: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*4.0e+2*:1404	*2.5e+2*:17825	*1.6e+2*:18980	*1.0e+2*:38677	*6.3e+1*:1.6e5	15/15
CMA-CSA	0.98(0.4)	1.4(0.9)	1.4(2)	1.2(0.5)	0.80(0.6)	15/15
CMA-MSR	**0.63**(0.1)	**0.07**(4e-3)$^{\star 4}$	**0.43**(0.4)	**0.47**(0.2)	**0.31**(0.1)	15/15
CMA-TPA	0.76(0.4)	0.48(0.3)	0.48(0.3)	0.68(0.4)	0.60(0.5)	15/15

References

[1] Asma Atamna. Benchmarking IPOP-CMA-ES-TPA and IPOP-CMA-ES-MSR on the BBOB noiseless testbed. In Laredo et al. [8], pages 1135–1142.

[2] Anne Auger, Steffen Finck, Nikolaus Hansen, and Raymond Ros. BBOB 2009: Comparison tables of all algorithms on all noiseless functions. Technical Report RT-0383, INRIA, April 2010.

[3] Lukás Bajer, Zbynek Pitra, and Martin Holena. Benchmarking gaussian processes and random forests surrogate models on the BBOB noiseless testbed. In Laredo et al. [8], pages 1143–1150.

[4] Dimo Brockhoff, Bernd Bischl, and Tobias Wagner. The impact of initial designs on the performance of matsumoto on the noiseless BBOB-2015 testbed: A preliminary study. In Laredo et al. [8], pages 1159–1166.

[5] 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.

[6] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2012: Experimental setup. Technical report, INRIA, 2012.

[7] 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.

[8] Juan Luis Jiménez Laredo, Sara Silva, and Anna Isabel Esparcia-Alcázar, editors. *Genetic and Evolutionary Computation Conference, GECCO 2015, Madrid, Spain, July 11-15, 2015, Companion Material Proceedings*. ACM, 2015.

[9] Petr Posík and Petr Baudis. Dimension selection in axis-parallel brent-step method for black-box optimization of separable continuous functions. In Laredo et al. [8], pages 1151–1158.