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

The BBOBies

May 27, 2015

Abstract

This document provides tabular results of the special session on Black-Box Optimization Benchmarking at CEC 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=cec-bbob-2015`. Overall, eight algorithms have been tested on 24 benchmark functions in dimensions between 2 and 20. A description of the used objective functions can be found in [6, 4]. The experimental set-up is described in [5].

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 [1]) 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 [5] 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 CEC 2015.

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

algorithm short name	paper	reference
MATSuMoTo	Comparison of the MATSuMoTo Library for Expensive Optimization on the Noiseless Black-Box Optimization Benchmarking Testbed	[2]
R-DE-10e2	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
R-DE-10e5	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
R-SHADE-10e2	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
R-SHADE-10e5	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
RL-SHADE-10e2	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
RL-SHADE-10e5	Parameter Tuning for Differential Evolution for Cheap, Medium, and Expensive Computational Budgets	[7]
SOO	Simultaneous Optimistic Optimization on the Noiseless BBOB Testbed	[3]

Table 2: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.7**(2)	**1.8**(1)	**2.8**(0.8)$^{\star 2}$	∞	∞ *100*	0/15
R-DE-10e2-	**2.7**(1)	**2.3**(3)	10(5)	∞	∞ *200*	0/15
R-DE-10e5-	**2.3**(1)	**3.0**(4)	66(222)	102(6)	102(150)	15/15
RL-SHADE-1	**1.8**(4)	**2.1**(0.8)	11(4)	240(210)	240(330)	2/15
RL-SHADE-1	**2.1**(2)	3.3(4)	30(17)	225(20)	225(26)	15/15
R-SHADE-10	**2.7**(2)	**2.8**(2)	13(5)	∞	∞ *200*	0/15
R-SHADE-10	**1.9**(1)	**2.9**(1)	11(4)	**45**(6)	**45**(7)	15/15
SOO-Derbel	**1.1**(0.2)	**1.2**(0.8)	**5.6**(2)	**52**(7)	**52**(3)	15/15

Table 3: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(0.7)	**1.6**(0.5)	**1.3**(1)	∞	∞ *100*	0/15
R-DE-10e2-	**1.6**(1)	**1.9**(2)	**2.2**(2)	**7.3**(3)	∞ *200*	0/15
R-DE-10e5-	**1**(0.5)	**1.2**(2)	**2.1**(2)	13(46)	**24**(3)	15/15
RL-SHADE-1	**1.1**(1)	**1.0**(0.5)	3.2(2)	**6.4**(2)	∞ *200*	0/15
RL-SHADE-1	**0.90**(0.4)	**1.5**(0.9)	**2.8**(3)	27(13)	73(4)	15/15
R-SHADE-10	**1.1**(0.9)	**1.1**(0.7)	5.0(3)	12(6)	∞ *200*	0/15
R-SHADE-10	**1.1**(1)	**2.3**(2)	**2.2**(3)	8.7(1)	**15**(2)	15/15
SOO-Derbel	**1.1**(0.4)	**1.2**(1)	**2.6**(1)	8.9(2)	27(4)	15/15

Table 4: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(0.7)	**1.4**(1)	**1.5**(0.9)	**2.7**(2)	**1.4**(0.5)	8/15
R-DE-10e2-	**1.3**(1)	**1.3**(1)	**1.7**(1)	**2.2**(2)	**0.87**(0.6)	15/15
R-DE-10e5-	**1.0**(0.7)	**2.1**(2)	**2.3**(1)	**2.1**(2)	**2.2**(2)	15/15
RL-SHADE-1	**1.4**(1)	**1.5**(1)	**2.3**(2)	**1.8**(0.7)	**1.0**(0.2)	13/15
RL-SHADE-1	**1.8**(2)	**2.3**(2)	3.8(3)	4.9(4)	**2.8**(1.0)	15/15
R-SHADE-10	**1.2**(0.5)	**2.0**(2)	**2.0**(2)	**1.9**(1)	**0.80**(0.6)	15/15
R-SHADE-10	**1.5**(1)	**2.3**(2)	**2.8**(3)	5.4(9)	**2.0**(2)	15/15
SOO-Derbel	**0.86**(0.2)	**1.3**(0.9)	**1.7**(2)	**1.9**(1)	**0.90**(0.5)	15/15

Table 5: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.9**(1)	**1.5**(0.9)	**1.8**(1)	**1.9**(1)	**2.2**(2)	5/15
R-DE-10e2-	**1.6**(0.9)	**1.7**(1)	**2.0**(2)	**1.9**(2)	**1.2**(0.4)	13/15
R-DE-10e5-	**2.5**(2)	**2.8**(2)	5.5(2)	5.1(11)	**1.7**(0.3)	15/15
RL-SHADE-1	**1.4**(1)	**1.6**(1)	**1.5**(2)	**1.8**(0.9)	**0.81**(0.2)	14/15
RL-SHADE-1	**2.1**(2)	**1.5**(0.9)	4.1(3)	5.2(1)	**2.9**(1)	15/15
R-SHADE-10	**1.8**(1)	**2.4**(2)	**1.9**(1)	**2.0**(2)	**1.1**(0.5)	13/15
R-SHADE-10	**1.9**(2)	**1.6**(2)	**1.8**(1)	**2.4**(0.6)	**3.0**(6)	15/15
SOO-Derbel	**0.83**(0.4)	**1.2**(0.6)	**1.5**(1)	**1.5**(0.8)	**1.2**(0.4)	15/15

Table 6: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(1)	**1.2**(0.9)	**2.5**(0.9)$^{\star 4}$	**2.5**(1)$^{\star 4}$	**2.5**(0.6)$^{\star 4}$	15/15
R-DE-10e2-	**1.9**(0.7)	3.1(3)	∞	∞	∞ *200*	0/15
R-DE-10e5-	**1.5**(1)	**2.7**(2)	319(272)	319(367)	319(295)	15/15
RL-SHADE-1	**2.0**(1)	3.0(2)	320(474)	320(253)	320(276)	2/15
RL-SHADE-1	**1.3**(0.4)	**1.7**(3)	349(24)	349(33)	349(14)	15/15
R-SHADE-10	**2.0**(1)	**2.9**(3)	∞	∞	∞ *200*	0/15
R-SHADE-10	**1.9**(1)	**2.0**(0.9)	**133**(49)	**133**(52)	**133**(22)	15/15
SOO-Derbel	**1.8**(0.4)	**2.5**(0.1)	338(0.1)	338(0.1)	338(0.1)	15/15

Table 7: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(0.9)	**1.7**(1)	**1.3**(0.9)	14(9)	∞ *100*	0/15
R-DE-10e2-	**1.4**(0.9)	**2.1**(2)	**1.4**(2)	**2.3**(0.8)	∞ *200*	0/15
R-DE-10e5-	**1.5**(0.9)	**1.8**(2)	12(4)	20(66)	18(21)	15/15
RL-SHADE-1	**1.2**(0.4)	**2.0**(2)	**2.0**(2)	4.2(2)	∞ *200*	0/15
RL-SHADE-1	**1.4**(2)	**2.0**(2)	**2.4**(2)	7.1(5)	**16**(4)	15/15
R-SHADE-10	**1.0**(0.5)	**1.3**(1)	**1.2**(0.8)	4.2(2)	∞ *200*	0/15
R-SHADE-10	**1.4**(2)	**2.3**(3)	**1.7**(3)	**3.2**(2)	**4.9**(3)	15/15
SOO-Derbel	**1**(1)	**0.95**(0.4)	**1.2**(2)	7.6(7)	1.4e4(2e4)	2/15

Table 8: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.4**(1)	**2.8**(2)	**1.2**(0.9)	**1.1**(0.3)	**0.55**(0.4)	10/15
R-DE-10e2-	**1.5**(0.8)	3.1(2)	**1.9**(1)	**2.0**(2)	**1.0**(1)	11/15
R-DE-10e5-	**0.87**(0.6)	**2.2**(2)	**1.5**(2)	**2.6**(5)	**1.6**(0.8)	15/15
RL-SHADE-1	**1**(1)	5.9(4)	**2.5**(2)	**2.8**(4)	**2.4**(2)	6/15
RL-SHADE-1	**0.96**(0.5)	4.9(7)	**2.9**(4)	**3.0**(4)	**2.3**(1)	15/15
R-SHADE-10	**1.0**(0.3)	**2.9**(1)	**1.4**(0.8)	**2.0**(2)	4.9(4)	3/15
R-SHADE-10	**0.75**(0.3)	**2.6**(3)	**1.4**(2)	**1.4**(1)	**0.81**(0.6)	15/15
SOO-Derbel	**1.2**(1)	**1.8**(2)	**1.0**(1.0)	**1.3**(0.7)	**0.76**(0.4)	15/15

Table 9: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.7**(1)	**2.1**(1)	**2.6**(1)	7.7(11)	∞ *100*	0/15
R-DE-10e2-	**1.5**(0.8)	4.2(4)	4.1(4)	16(5)	∞ *200*	0/15
R-DE-10e5-	**2.0**(0.6)	**2.5**(0.8)	**2.8**(3)	16(8)	**12**(6)	15/15
RL-SHADE-1	**1.8**(1)	**2.5**(3)	3.8(4)	6.3(2)	∞ *200*	0/15
RL-SHADE-1	**2.1**(0.4)	**2.6**(1)	7.8(8)	12(5)	23(3)	15/15
R-SHADE-10	**2.3**(1)	**3.0**(3)	7.6(8)	15(34)	∞ *200*	0/15
R-SHADE-10	**2.0**(2)	**2.7**(2)	3.4(2)	**4.6**(5)	**5.3**(5)	15/15
SOO-Derbel	**1.7**(0.4)	**1.4**(1)	**1.9**(1)	**4.4**(3)	21(13)	15/15

Table 10: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f9	*6.3e+0*:13	*4.0e+0*:15	*2.5e+0*:15	*2.5e-1*:21	*1.0e-8*:94	15/15
MATSUMOTO-	**1.3**(1)	**1.4**(0.9)	**1.6**(1)	**4.3**(2)	∞ *100*	0/15
R-DE-10e2-	**3.0**(6)	3.5(5)	4.9(1)	8.3(9)	∞ *200*	0/15
R-DE-10e5-	3.8(3)	4.6(2)	6.0(5)	27(86)	**20**(13)	15/15
RL-SHADE-1	3.9(6)	4.0(3)	4.4(7)	10(8)	∞ *200*	0/15
RL-SHADE-1	3.2(2)	3.4(4)	3.6(4)	15(8)	28(2)	15/15
R-SHADE-10	3.4(2)	3.7(3)	5.0(4)	16(15)	∞ *200*	0/15
R-SHADE-10	**2.2**(2)	**2.0**(2)	**2.3**(3)	7.3(3)	**8.0**(12)	15/15
SOO-Derbel	**1.4**(0.8)	**1.6**(0.9)	**1.9**(0.8)	**2.8**(2)	25(16)	15/15

Table 11: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**2.3**(2)	**2.4**(2)	**2.9**(3)	9.1(11)	∞ *100*	0/15
R-DE-10e2-	**2.2**(1)	**2.0**(1)	6.8(6)	7.3(5)	∞ *200*	0/15
R-DE-10e5-	**2.8**(2)	**2.8**(2)	6.9(6)	27(30)	74(51)	15/15
RL-SHADE-1	**1.7**(2)	**1.7**(1)	7.1(6)	18(24)	∞ *200*	0/15
RL-SHADE-1	**2.0**(2)	**2.2**(1)	14(8)	15(8)	**26**(2)	15/15
R-SHADE-10	**2.2**(2)	**1.6**(1)	11(10)	13(11)	∞ *200*	0/15
R-SHADE-10	**2.2**(2)	**2.4**(2)	5.0(3)	**3.2**(0.9)	**4.4**(1)	15/15
SOO-Derbel	**1.3**(0.8)	**1.1**(0.9)	**3.9**(3)	**3.3**(2)	494(196)	15/15

Table 12: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(0.9)	**1.2**(1)	**2.4**(1)	**2.8**(2)	∞ *100*	0/15
R-DE-10e2-	**1.8**(0.5)	**1.1**(0.8)	4.2(4)	5.1(3)	∞ *200*	0/15
R-DE-10e5-	**1.9**(2)	**1.8**(1)	**3.0**(2)	10(20)	94(103)	15/15
RL-SHADE-1	**2.0**(0.7)	**1.6**(2)	4.4(4)	5.7(5)	∞ *200*	0/15
RL-SHADE-1	**2.1**(3)	**1.9**(2)	8.6(12)	12(5)	**25**(3)	15/15
R-SHADE-10	**1.8**(0.9)	**1.5**(1)	4.1(2)	9.0(7)	∞ *200*	0/15
R-SHADE-10	**1.8**(0.5)	**2.2**(1)	4.4(4)	4.1(5)	**5.0**(2)	15/15
SOO-Derbel	**1.2**(0.5)	**0.94**(0.8)	**2.7**(1)	**2.9**(0.9)	481(197)	15/15

Table 13: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.1**(0.8)	**0.93**(0.7)	**0.77**(0.4)	4.6(3)	∞ *100*	0/15
R-DE-10e2-	**0.85**(0.2)	**0.80**(0.5)	**1.0**(1.0)	5.3(5)	29(22)	1/15
R-DE-10e5-	**0.95**(0)	**0.98**(0.5)	**0.98**(0.5)	8.0(1.0)	37(55)	15/15
RL-SHADE-1	**1.6**(2)	**1.5**(2)	**1.1**(0.6)	**3.9**(2)	29(27)	1/15
RL-SHADE-1	**1**(1)	**0.76**(0.7)	**1.4**(1)	13(5)	**26**(9)	15/15
R-SHADE-10	**1.1**(0.6)	**1.7**(1)	**1.4**(1.0)	7.8(3)	∞ *200*	0/15
R-SHADE-10	**0.80**(0.2)	**0.90**(0.7)	**1.6**(0.8)	4.7(4)	**15**(15)	15/15
SOO-Derbel	**0.95**(0.8)	**0.73**(0.7)	**0.59**(0.5)	**4.3**(2)	33(20)	15/15

Table 14: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(0.9)	**1.5**(0.6)	**1.4**(0.5)	**1.6**(0.4)	∞ *100*	0/15
R-DE-10e2-	**1.5**(2)	**1.6**(4)	**2.1**(2)	4.0(0.7)	∞ *200*	0/15
R-DE-10e5-	**1.8**(1)	**1.5**(1)	**2.4**(5)	4.5(9)	100(94)	15/15
RL-SHADE-1	**1.4**(2)	**1.5**(0.6)	**2.4**(2)	3.9(3)	∞ *200*	0/15
RL-SHADE-1	**1.2**(0.3)	**0.68**(0.5)	**2.5**(4)	4.7(4)	**25**(2)	15/15
R-SHADE-10	**2.3**(1)	**1.6**(1)	**3.0**(2)	5.8(6)	∞ *200*	0/15
R-SHADE-10	**1.5**(1)	**1.3**(1)	**1.8**(1)	**2.6**(0.8)	**4.9**(0.4)	15/15
SOO-Derbel	**1.4**(3)	**1.1**(1)	**1.8**(1)	**2.8**(4)	153(161)	15/15

Table 15: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.0**(0.7)	**1.4**(1)	**1.5**(1)	**1.7**(0.5)$^{\star 2}$	∞ *100*	0/15
R-DE-10e2-	**1.3**(0.7)	**2.0**(2)	3.1(2)	4.2(1)	∞ *200*	0/15
R-DE-10e5-	**1.0**(0.7)	**1.1**(1)	**1.9**(0.8)	4.2(0.4)	94(129)	15/15
RL-SHADE-1	**1.2**(0.7)	**2.2**(3)	**2.7**(3)	4.1(2)	∞ *200*	0/15
RL-SHADE-1	**1.3**(0.4)	3.2(1)	4.3(4)	13(6)	**24**(2)	15/15
R-SHADE-10	**1.4**(1)	**2.0**(3)	3.3(3)	5.5(3)	∞ *200*	0/15
R-SHADE-10	**1.3**(0.7)	**2.8**(3)	3.8(4)	4.2(1)	**4.1**(0.4)	15/15
SOO-Derbel	**0.81**(0)	**1.6**(1)	**1.8**(1)	**3.2**(1)	43(19)	15/15

Table 16: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(1)	**1.1**(0.7)	**0.79**(0.8)	**0.85**(0.6)	**1.0**(0.8)	9/15
R-DE-10e2-	**1.7**(0.8)	**1**(0.8)	**1.3**(1)	**1.2**(0.5)	**1.9**(2)	10/15
R-DE-10e5-	**1.8**(1)	**1.4**(0.9)	**1.2**(1)	**0.94**(1)	**2.1**(4)	15/15
RL-SHADE-1	**2.2**(1)	**1.6**(1)	**1.4**(1)	**1.8**(1)	**2.0**(2)	9/15
RL-SHADE-1	**1.8**(1)	**2.0**(4)	**1.4**(2)	**1.5**(2)	**2.7**(2)	15/15
R-SHADE-10	**1.5**(1)	**2.0**(2)	**1.4**(0.8)	**1.7**(1)	4.4(2)	5/15
R-SHADE-10	**1.3**(0.2)	**1.2**(1)	**1.7**(0.8)	**1.0**(0.8)	**1.4**(2)	15/15
SOO-Derbel	**1.1**(0.4)	**0.84**(1)	**0.74**(0.6)	**0.75**(0.4)	**1.1**(0.8)	15/15

Table 17: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(1)	**1.7**(1)	**1.6**(2)	**1.1**(1)	**1.3**(2)	7/15
R-DE-10e2-	**1.4**(0.4)	**2.1**(2)	**1.8**(1)	**1.8**(2)	3.8(8)	5/15
R-DE-10e5-	**1.3**(0.9)	**1.3**(1)	**2.4**(3)	4.3(1)	**2.6**(3)	15/15
RL-SHADE-1	**1.1**(0)	**2.1**(2)	**1.8**(2)	**2.2**(1)	5.0(5)	4/15
RL-SHADE-1	**1.4**(0.9)	**2.4**(2)	3.3(3)	**2.2**(2)	4.5(2)	15/15
R-SHADE-10	**1.4**(0.7)	**2.2**(2)	**2.3**(3)	**1.7**(1)	4.0(5)	5/15
R-SHADE-10	**1.2**(0.4)	**2.3**(2)	**1.8**(2)	**1.9**(3)	**2.2**(1)	15/15
SOO-Derbel	**1.5**(0.4)	**2.3**(2)	**1.9**(1)	**1.0**(0.6)	**0.93**(1)	15/15

Table 18: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.2**(0.6)	**1.5**(1)	**1.2**(0.3)	**0.52**(0.4)	**2.1**(3)	5/15
R-DE-10e2-	**1.7**(1)	3.2(2)	**1.6**(1)	**1.1**(1)	**1.3**(0.9)	14/15
R-DE-10e5-	**1.2**(0.4)	**1.8**(0.8)	**1.5**(2)	**0.85**(0.3)	**1.7**(2)	15/15
RL-SHADE-1	**1.2**(0.4)	**2.3**(3)	**1.8**(1.0)	**1.2**(2)	**1.6**(1)	11/15
RL-SHADE-1	**1.6**(2)	**1.9**(2)	**2.4**(3)	**1.4**(1)	3.4(1)	15/15
R-SHADE-10	**1.5**(1)	**2.8**(1)	**1.9**(2)	**1.0**(1)	4.6(6)	5/15
R-SHADE-10	**1.5**(1)	**2.3**(3)	**1.8**(2)	**1.1**(0.6)	**1.2**(0.5)	15/15
SOO-Derbel	**0.94**(0)	**1.8**(1)	**0.85**(0.5)	**0.48**(0.3)	**0.95**(0.8)	15/15

Table 19: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.1**(0.4)	**0.83**(0.4)	**0.95**(0.4)	**0.86**(0.8)	**1.0**(0.6)	10/15
R-DE-10e2-	**1.3**(0.4)	**1.2**(0.5)	**1.5**(1)	**1.6**(0.7)	**1.0**(0.9)	14/15
R-DE-10e5-	**1.2**(0.2)	**1.2**(0.9)	**1.3**(1)	3.2(0.8)	3.0(7)	15/15
RL-SHADE-1	**1.2**(0.4)	**1.3**(1)	**2.1**(2)	**1.2**(1)	**1.6**(1)	10/15
RL-SHADE-1	**1.4**(0.6)	**1.6**(2)	**1.9**(2)	**2.6**(1)	**2.3**(1)	15/15
R-SHADE-10	**1.2**(0.4)	**0.98**(0.3)	**0.77**(1)	**1.2**(2)	**1.6**(0.8)	12/15
R-SHADE-10	**1.5**(0.6)	**1.8**(0.6)	**1.4**(1)	**1.4**(0.8)	**2.4**(0.4)	15/15
SOO-Derbel	**1.1**(0.4)	**1.6**(1)	**0.94**(0.9)	**0.73**(0.4)	**0.94**(0.5)	15/15

Table 20: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f19	*1.6e-1*:23	*1.0e-1*:26	*6.3e-2*:38	*4.0e-2*:40	*1.0e-2*:216	15/15
MATSUMOTO-	6.9(8)	9.3(4)	11(11)	11(12)	6.6(8)	1/15
R-DE-10e2-	4.3(2)	**4.2**(3)	5.5(2)	6.6(5)	6.7(5)	2/15
R-DE-10e5-	5.2(5)	5.6(6)	6.4(4)	12(14)	11(6)	15/15
RL-SHADE-1	5.4(6)	7.1(7)	6.9(7)	13(10)	**6.5**(7)	2/15
RL-SHADE-1	7.0(11)	7.6(8)	7.0(6)	14(14)	9.1(6)	15/15
R-SHADE-10	9.4(7)	8.4(11)	11(18)	24(33)	∞ *200*	0/15
R-SHADE-10	**2.9**(2)	**3.3**(1)	**3.2**(3)	**4.9**(5)	**5.2**(3)	15/15
SOO-Derbel	**4.1**(2)	5.0(4)	**3.6**(2)	**3.7**(2)	17(15)	15/15

Table 21: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(0.8)	**2.0**(2)	**2.1**(1)	**1.8**(1)	3.4(2)	3/15
R-DE-10e2-	**1**(0.5)	**1.7**(1)	**2.6**(3)	**1.8**(1)	**2.0**(3)	8/15
R-DE-10e5-	**1.2**(1)	**2.2**(2)	**1.9**(1)	**1.8**(1.0)	**2.4**(1)	15/15
RL-SHADE-1	**1.4**(0.9)	**2.7**(2)	3.1(2)	**2.1**(1)	**2.7**(3)	7/15
RL-SHADE-1	**1.7**(1)	3.3(3)	4.7(3)	3.1(3)	4.6(2)	15/15
R-SHADE-10	**2.0**(5)	**2.6**(4)	4.0(2)	**2.4**(1)	6.7(5)	3/15
R-SHADE-10	**1.5**(2)	**2.5**(2)	3.2(2)	**2.1**(1)	**2.3**(1)	15/15
SOO-Derbel	**1.4**(0.3)	5.1(0.2)	**3.0**(0.1)	**1.3**(0.0)	**1.2**(4e-3)	15/15

Table 22: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.1**(1)	**1.8**(2)	**1.7**(2)	**0.92**(1)	**0.64**(1)	12/15
R-DE-10e2-	**1.6**(2)	**2.6**(10)	**1.7**(4)	**2.1**(2)	**2.7**(2)	7/15
R-DE-10e5-	**1.3**(0.6)	**1.3**(1)	**1.5**(0.4)	6.9(2)	5.3(7)	15/15
RL-SHADE-1	**1.1**(0.4)	**1.5**(1)	**1.2**(0.8)	**0.92**(0.2)	**1.1**(0.5)	11/15
RL-SHADE-1	**1.6**(1)	**2.2**(2)	**1.3**(2)	**1.1**(1)	**1.2**(2)	15/15
R-SHADE-10	**1.5**(2)	**1.7**(2)	**1.4**(0.9)	**1.6**(1)	**1.1**(1)	12/15
R-SHADE-10	**1.3**(0.3)	**1.9**(3)	**2.0**(2)	**1.2**(0.7)	3.3(6)	15/15
SOO-Derbel	**0.88**(0.3)	**1.5**(0.7)	**1.3**(0.9)	**0.93**(0.8)	**0.53**(0.6)	15/15

Table 23: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.7**(0.8)	**2.0**(2)	**1.2**(1.0)	**1.1**(0.8)	**0.90**(0.9)	8/15
R-DE-10e2-	**1.9**(2)	**1.7**(2)	**1.5**(0.6)	**1.8**(4)	**0.64**(0.4)	13/15
R-DE-10e5-	**1.6**(1)	**1.5**(1)	**1.5**(2)	11(2)	3.0(6)	15/15
RL-SHADE-1	**1.5**(0.4)	**1.4**(0.9)	**1.4**(1.0)	**1.8**(3)	**0.90**(2)	10/15
RL-SHADE-1	**1.3**(0.4)	**0.58**(0.2)	**0.74**(1)	**2.0**(2)	17(1)	15/15
R-SHADE-10	**1.2**(0.6)	**0.77**(0.7)	**1.3**(2)	**2.5**(3)	**1.5**(2)	8/15
R-SHADE-10	**1.3**(0.8)	**1.2**(0.8)	**0.81**(0.5)	4.2(12)	**1.9**(4)	15/15
SOO-Derbel	**1.3**(0.4)	**1.2**(0.9)	**0.78**(0.8)	**0.89**(1)	**0.71**(0.7)	15/15

Table 24: 02-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.2**(0.7)	**1.4**(1)	**1.2**(2)	**1.4**(2)	**1.9**(2)	6/15
R-DE-10e2-	**1.7**(0.8)	**1.9**(1)	**1.8**(1.0)	**1.3**(1.0)	**1.6**(3)	10/15
R-DE-10e5-	**1.6**(2)	**1.8**(1)	**1.7**(2)	**2.1**(7)	4.5(7)	15/15
RL-SHADE-1	**0.96**(0.7)	**1.5**(2)	**1.9**(2)	**2.0**(4)	3.5(5)	6/15
RL-SHADE-1	**1.2**(1.0)	**1.7**(3)	**2.0**(1)	**2.3**(1)	**2.1**(1)	15/15
R-SHADE-10	**1.3**(0.5)	**1.4**(1)	**2.4**(3)	**2.0**(2)	**3.0**(5)	7/15
R-SHADE-10	**1.3**(2)	**1.4**(0.9)	**2.3**(4)	**2.4**(3)	**2.1**(1)	15/15
SOO-Derbel	**1.4**(2)	**1.3**(0.8)	**2.6**(3)	**1.5**(1)	**1.6**(1)	15/15

Table 25: 02-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(0.9)	**1.7**(2)	**1.1**(0.8)	**1.2**(2)	∞ *100*	0/15
R-DE-10e2-	**1.5**(0.9)	**1.3**(0.9)	**0.77**(0.7)	**0.96**(0.6)	**2.4**(3)	4/15
R-DE-10e5-	**1.5**(0.9)	**1.7**(1)	**1.0**(1)	**1.5**(1)	8.3(7)	15/15
RL-SHADE-1	**1.6**(1)	**1.6**(1)	**1.4**(2)	**1.7**(2)	5.0(6)	2/15
RL-SHADE-1	**1.6**(0.9)	**1.4**(0.9)	**1.4**(2)	**1.3**(1)	**2.9**(2)	15/15
R-SHADE-10	**1.2**(0)	**2.5**(3)	**1.8**(3)	**1.0**(1)	11(13)	1/15
R-SHADE-10	**1.5**(0.9)	**2.0**(2)	**1.2**(1)	**0.75**(0.6)	**1.8**(1)	15/15
SOO-Derbel	**1.1**(0.2)	**1.4**(1)	**1.9**(0.8)	**0.88**(0.7)	**1.6**(3)	15/15

Table 26: 03-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	*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
MATSUMOTO-	**1.6**(2)	**1.9**(1)	∞	∞	∞ *150*	0/15
R-DE-10e2-	**1.6**(0.9)	**2.6**(3)	∞	∞	∞ *300*	0/15
R-DE-10e5-	**2.4**(3)	**2.5**(3)	**72**(33)	**72**(68)	**72**(43)	15/15
RL-SHADE-1	**1.4**(1)	**1.5**(1)	279(150)	279(234)	279(385)	2/15
RL-SHADE-1	**1.4**(0.7)	**2.1**(2)	387(38)	387(29)	387(26)	15/15
R-SHADE-10	**1.8**(2)	**2.9**(3)	∞	∞	∞ *300*	0/15
R-SHADE-10	**2.3**(2)	3.0(2)	**74**(5)	**74**(11)	**74**(8)	15/15
SOO-Derbel	**0.69**(0.5)	**1.1**(1.0)	99(4)	99(7)	99(10)	15/15

Table 27: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(1.0)	**0.89**(0.8)	**1.8**(1)	22(30)	∞ *150*	0/15
R-DE-10e2-	**1.9**(0.8)	**1.8**(2)	**2.1**(3)	**3.3**(1)	∞ *300*	0/15
R-DE-10e5-	**1.8**(6)	**1.6**(3)	**2.1**(2)	**3.6**(3)	**26**(8)	15/15
RL-SHADE-1	**2.1**(2)	**1.7**(0.7)	4.0(2)	3.9(1)	∞ *300*	0/15
RL-SHADE-1	**1.1**(1.0)	**1.3**(0.8)	**2.8**(3)	18(8)	91(6)	15/15
R-SHADE-10	**1.4**(1.0)	**1.5**(0.8)	4.0(3)	4.8(2)	∞ *300*	0/15
R-SHADE-10	**1.3**(0)	**1.2**(1)	**2.7**(1)	4.7(1)	**18**(3)	15/15
SOO-Derbel	**1.5**(1)	**2.5**(2)	3.7(3)	6.0(2)	39(5)	15/15

Table 28: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.9**(1)	**1.1**(1.0)	**1.4**(0.8)	**1.6**(0.8)	**0.95**(1)	6/15
R-DE-10e2-	**2.2**(1)	**2.2**(2)	**2.7**(1.0)	**2.0**(0.8)	**1.1**(1)	10/15
R-DE-10e5-	**2.0**(2)	**1.6**(0.9)	**2.3**(1.0)	**2.9**(0.6)	**1.0**(0.8)	15/15
RL-SHADE-1	**2.1**(2)	**1.8**(2)	**2.5**(1)	**2.2**(1)	**0.44**(0.2)	15/15
RL-SHADE-1	**2.2**(3)	**1.6**(1)	**2.6**(3)	6.5(3)	**2.0**(0.9)	15/15
R-SHADE-10	**2.2**(0.9)	**2.3**(3)	**2.1**(0.8)	**2.3**(0.9)	**0.57**(0.1)	14/15
R-SHADE-10	**2.3**(1)	**1.6**(2)	3.1(1)	**2.1**(2)	**1.7**(2)	15/15
SOO-Derbel	**0.91**(0)	**0.80**(0.5)	**2.2**(1)	**2.2**(1)	**1.4**(1)	15/15

Table 29: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f4	*1.0e+2*:5.4	*6.3e+1*:10	*6.3e+1*:10	*2.5e+1*:36	*4.0e+0*:617	15/15
MATSUMOTO-	**1.6**(2)	**1.8**(1)	**1.8**(1)	**1.4**(1.0)	**1.2**(2)	3/15
R-DE-10e2-	**2.4**(4)	**2.0**(2)	**2.0**(2)	**1.6**(1)	**0.60**(0.4)	10/15
R-DE-10e5-	**1.2**(1)	**2.0**(1)	**2.0**(1)	**1.5**(0.5)	**1.3**(1)	15/15
RL-SHADE-1	**1.0**(1)	**2.2**(1)	**2.2**(2)	**1.5**(1)	**0.58**(0.5)	9/15
RL-SHADE-1	**2.4**(3)	**2.5**(2)	**2.5**(5)	5.4(1)	**1.7**(0.8)	15/15
R-SHADE-10	**1.5**(2)	**1.8**(3)	**1.8**(1)	**1.8**(0.6)	**0.49**(0.4)	11/15
R-SHADE-10	**1.1**(0.9)	**1.2**(1)	**1.2**(1)	**2.0**(3)	**1.3**(1)	15/15
SOO-Derbel	**0.69**(2)	**0.99**(1.0)	**0.99**(1)	**2.1**(1)	**0.94**(0.5)	15/15

Table 30: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(1)	**1.0**(0.8)	**1.9**(0.6)$^{\star 4}$	**1.9**(0.5)$^{\star 4}$	**1.9**(0.5)$^{\star 4}$	15/15
R-DE-10e2-	**2.2**(1)	**3.0**(3)	∞	∞	∞ *300*	0/15
R-DE-10e5-	**2.8**(2)	**3.0**(2)	256(292)	256(170)	256(299)	15/15
RL-SHADE-1	**1.8**(3)	**1.6**(3)	∞	∞	∞ *300*	0/15
RL-SHADE-1	3.8(2)	3.9(4)	431(21)	431(8)	431(15)	15/15
R-SHADE-10	4.7(4)	4.4(3)	∞	∞	∞ *300*	0/15
R-SHADE-10	**1.5**(0.9)	3.6(4)	**150**(17)	**150**(26)	**150**(26)	15/15
SOO-Derbel	**2.0**(0.2)	**1.8**(0.1)	531(0.1)	531(0.1)	531(0.1)	15/15

Table 31: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.2**(2)	**1.1**(1)	3.2(3)	6.3(5)	∞ *150*	0/15
R-DE-10e2-	**2.9**(3)	**2.1**(3)	3.2(1)	**1.9**(3)	∞ *300*	0/15
R-DE-10e5-	**2.8**(6)	3.4(3)	**2.0**(1)	**2.2**(0.5)	26(50)	15/15
RL-SHADE-1	**1.9**(3)	**2.6**(4)	**2.9**(1)	**2.7**(0.4)	28(31)	1/15
RL-SHADE-1	**1.6**(0.6)	4.8(4)	6.3(7)	4.9(5)	**16**(3)	15/15
R-SHADE-10	**2.0**(1)	3.9(8)	**3.0**(2)	**2.6**(1)	∞ *300*	0/15
R-SHADE-10	**1.3**(1)	3.7(8)	3.8(8)	**2.3**(2)	**3.4**(0.3)	15/15
SOO-Derbel	**1.4**(2)	**1.7**(1.0)	**2.0**(2)	**1.9**(2)	1.3e4(8958)	2/15

Table 32: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(0.8)	**1.4**(1)	**2.3**(2)	**1.6**(0.7)	**1.7**(1.0)	7/15
R-DE-10e2-	**1.0**(1.0)	**1.4**(2)	**2.2**(1)	**1.4**(1)	**1.0**(1)	13/15
R-DE-10e5-	**1.2**(1)	**0.94**(1.0)	**1.7**(2)	**1.4**(2)	**2.4**(1)	15/15
RL-SHADE-1	**1.7**(2)	**2.1**(3)	3.5(4)	**2.3**(1)	**1.2**(0.6)	13/15
RL-SHADE-1	**2.0**(0.7)	3.8(2)	6.6(7)	5.0(3)	**2.5**(0.4)	15/15
R-SHADE-10	**1.3**(1.0)	**1.3**(2)	3.4(3)	**2.2**(0.7)	**2.5**(3)	8/15
R-SHADE-10	**2.2**(2)	3.6(2)	3.7(6)	**2.2**(2)	**1.4**(0.7)	15/15
SOO-Derbel	**1.1**(0.7)	**1.1**(2)	**2.3**(0.7)	**1.5**(0.7)	**2.6**(8)	15/15

Table 33: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(2)	**1.1**(0.5)	**1.5**(0.8)	4.1(5)	15(15)	1/15
R-DE-10e2-	**1.7**(1)	3.5(5)	3.2(2)	3.7(1)	5.6(6)	5/15
R-DE-10e5-	**1.3**(1.0)	**1.6**(2)	**2.2**(2)	3.5(1)	14(19)	15/15
RL-SHADE-1	**1.4**(0.8)	3.8(2)	3.2(1)	**3.3**(1)	14(10)	2/15
RL-SHADE-1	**1.1**(0.6)	**2.6**(1)	5.7(3)	11(4)	14(4)	15/15
R-SHADE-10	**1.5**(1)	**2.1**(2)	3.1(2)	5.3(3)	∞ *300*	0/15
R-SHADE-10	**2.0**(2)	**2.0**(2)	**2.2**(2)	3.4(1)	**4.6**(2)	15/15
SOO-Derbel	**1.4**(2)	**1.1**(0.9)	**1.1**(0.8)	**1.8**(0.7)	**4.6**(2)	15/15

Table 34: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f9	*1.0e+1*:21	*6.3e+0*:25	*4.0e+0*:32	*2.5e+0*:48	*6.3e-3*:152	15/15
MATSUMOTO-	**2.5**(2)	**2.5**(1)	**2.7**(3)	**2.4**(3)	∞ *150*	0/15
R-DE-10e2-	4.5(2)	4.4(2)	4.2(2)	4.4(1)	∞ *300*	0/15
R-DE-10e5-	8.1(17)	8.4(24)	15(30)	26(34)	36(23)	15/15
RL-SHADE-1	5.1(4)	5.2(10)	5.0(2)	5.4(5)	∞ *300*	0/15
RL-SHADE-1	14(12)	15(13)	15(8)	13(6)	19(4)	15/15
R-SHADE-10	6.2(3)	8.3(7)	12(12)	9.3(9)	∞ *300*	0/15
R-SHADE-10	3.7(1)	3.6(2)	3.6(2)	3.1(2)	**4.4**(4)	15/15
SOO-Derbel	**2.2**(1)	**2.1**(0.5)	**2.0**(1.0)	**1.7**(0.8)	**12**(5)	15/15

Table 35: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.2**(1)	**2.2**(1)	**1.1**(0.5)	10(18)	∞ *150*	0/15
R-DE-10e2-	**0.88**(0.4)	**2.1**(3)	**1.6**(1)	5.5(5)	29(33)	1/15
R-DE-10e5-	**1.3**(4)	**2.3**(2)	**1.6**(1)	14(7)	52(40)	15/15
RL-SHADE-1	**1.0**(0.8)	3.0(3)	**2.0**(2)	8.9(9)	∞ *300*	0/15
RL-SHADE-1	**1.3**(0.6)	3.7(3)	4.5(4)	11(11)	**14**(4)	15/15
R-SHADE-10	**2.0**(0.9)	**2.8**(2)	**2.3**(1)	20(17)	∞ *300*	0/15
R-SHADE-10	**1.2**(0.6)	**2.6**(3)	**1.8**(2)	**3.0**(1)	**2.2**(0.4)	15/15
SOO-Derbel	**1.0**(1)	**1.8**(2)	**1.2**(0.9)	**3.7**(2)	92(178)	15/15

Table 36: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.6**(2)	**1.3**(1)	**1.4**(0.8)	7.8(12)	∞ *150*	0/15
R-DE-10e2-	**1.2**(1)	**1.3**(1)	**1.4**(1)	13(16)	∞ *300*	0/15
R-DE-10e5-	**2.0**(2)	**1.4**(1)	**1.7**(2)	23(39)	100(176)	15/15
RL-SHADE-1	**2.8**(3)	**2.3**(2)	**2.6**(2)	17(9)	∞ *300*	0/15
RL-SHADE-1	**1.6**(0.8)	**1.3**(1)	**1.3**(0.5)	15(16)	**12**(2)	15/15
R-SHADE-10	**1.2**(1)	**0.99**(0.5)	**1.1**(2)	9.2(13)	∞ *300*	0/15
R-SHADE-10	**1.4**(0.8)	**2.3**(2)	**1.6**(1)	**4.6**(2)	**3.3**(4)	15/15
SOO-Derbel	**1.8**(2)	**1.6**(2)	**1.8**(0.8)	**4.6**(3)	183(927)	14/15

Table 37: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**0.83**(0)	**1.1**(2)	**1.4**(0.5)	**3.4**(1)	14(19)	1/15
R-DE-10e2-	**0.91**(0.7)	**1.6**(1)	**2.4**(2)	6.7(0.9)	∞ *300*	0/15
R-DE-10e5-	**1.1**(1.0)	**2.5**(4)	**2.5**(1)	10(8)	76(126)	15/15
RL-SHADE-1	**1.2**(1.0)	**1.9**(2)	**2.4**(1)	5.4(2)	∞ *300*	0/15
RL-SHADE-1	**1.0**(0.3)	**2.3**(3)	4.2(2)	25(16)	20(7)	15/15
R-SHADE-10	**1.3**(0.7)	**2.0**(2)	**2.3**(2)	11(12)	∞ *300*	0/15
R-SHADE-10	**0.87**(0.3)	**1.3**(1)	**2.7**(2)	5.5(2)	**8.2**(20)	15/15
SOO-Derbel	**0.87**(0)	**0.89**(0.6)	**1.3**(1)	**4.8**(0.8)	**4.3**(0.8)	15/15

Table 38: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(1)	**1.0**(0.9)	**0.86**(0.5)	**1.6**(0.4)	∞ *150*	0/15
R-DE-10e2-	**1.0**(0.9)	**1.2**(1)	**1.1**(0.6)	**2.6**(0.8)	∞ *300*	0/15
R-DE-10e5-	**1.6**(2)	**1.9**(2)	**1.7**(2)	3.9(4)	44(24)	15/15
RL-SHADE-1	**1.2**(1)	**1.6**(2)	**1.7**(2)	3.8(9)	∞ *300*	0/15
RL-SHADE-1	**1.4**(0.3)	**2.4**(2)	3.3(4)	12(3)	21(1)	15/15
R-SHADE-10	**2.0**(1)	**1.5**(2)	**1.8**(2)	3.9(2)	∞ *300*	0/15
R-SHADE-10	**1.3**(0.8)	**2.1**(3)	**2.4**(1)	3.7(1)	**3.8**(1)	15/15
SOO-Derbel	**0.83**(1)	**0.74**(0.7)	**0.87**(0.6)	**2.4**(2)	**20**(17)	15/15

Table 39: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(1)	**1.5**(2)	**1.6**(1)	**2.8**(4)	∞ *150*	0/15
R-DE-10e2-	**1.8**(1)	**1.7**(1)	**2.2**(1)	3.7(1)	∞ *300*	0/15
R-DE-10e5-	**2.8**(7)	**2.0**(2)	**2.2**(2)	4.9(5)	107(113)	15/15
RL-SHADE-1	**1.7**(2)	**2.0**(2)	**2.5**(2)	3.7(2)	∞ *300*	0/15
RL-SHADE-1	**1.3**(1)	**0.90**(0.8)	**2.7**(5)	19(5)	**22**(2)	15/15
R-SHADE-10	3.5(6)	**2.5**(4)	**2.4**(2)	4.9(2)	∞ *300*	0/15
R-SHADE-10	**2.7**(3)	**2.0**(2)	**2.5**(4)	4.5(1)	**3.8**(0.4)	15/15
SOO-Derbel	**1.3**(1)	**0.92**(0.8)	**0.93**(0.5)	**3.4**(1)	225(75)	14/15

Table 40: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(0.9)	**0.96**(0.6)	**1.3**(1)	**0.66**(0.6)	**0.82**(0.6)	9/15
R-DE-10e2-	**2.2**(4)	**1.9**(2)	**2.0**(2)	**0.92**(0.7)	**1.3**(1)	11/15
R-DE-10e5-	**2.1**(4)	**1.9**(2)	**1.5**(1)	**0.97**(0.4)	3.3(3)	15/15
RL-SHADE-1	**1.8**(2)	**2.2**(2)	**2.1**(1.0)	**1.3**(0.4)	**0.91**(0.6)	13/15
RL-SHADE-1	**2.0**(2)	**2.0**(3)	**2.6**(2)	**2.9**(1)	**2.8**(2)	15/15
R-SHADE-10	**0.96**(0.3)	**1.4**(4)	**1.9**(2)	**1.6**(0.9)	**2.1**(2)	8/15
R-SHADE-10	**3.0**(2)	**3.0**(2)	3.1(4)	**1.4**(0.6)	**1.3**(0.6)	15/15
SOO-Derbel	**1.3**(1)	**1.0**(1)	**1**(1)	**1.1**(0.5)	**0.76**(0.3)	15/15

Table 41: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(1)	**1.8**(3)	**2.7**(4)	**1.3**(1.0)	**0.82**(0.7)	10/15
R-DE-10e2-	**2.1**(2)	**1.6**(1)	3.2(4)	**1.8**(0.8)	**1.6**(2)	10/15
R-DE-10e5-	**2.1**(4)	**1.1**(0.5)	**1.3**(0.8)	**1.7**(2)	**2.7**(5)	15/15
RL-SHADE-1	**2.0**(0.8)	**1.9**(2)	3.1(3)	**1.3**(2)	**1.0**(0.8)	12/15
RL-SHADE-1	**1.5**(0.7)	**1.3**(2)	**2.7**(3)	**1.9**(1)	**2.9**(3)	15/15
R-SHADE-10	**1.8**(2)	**1.4**(0.7)	**2.7**(4)	**0.98**(0.7)	**2.3**(2)	7/15
R-SHADE-10	**1.3**(0.7)	**1.3**(2)	**2.2**(3)	**1.0**(0.9)	**1.6**(3)	15/15
SOO-Derbel	**2.0**(2)	**1.2**(1)	**1.6**(0.8)	**0.91**(0.6)	**0.45**(0.3)	15/15

Table 42: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**2.4**(2)	**2.6**(4)	**1.1**(1)	**1.3**(0.4)	**2.8**(4)	4/15
R-DE-10e2-	**2.7**(3)	**2.2**(2)	**1.5**(2)	**1.6**(1)	**1.4**(2)	13/15
R-DE-10e5-	**1.6**(1)	**2.1**(1)	**1.1**(1)	**1.6**(1)	**1.4**(0.7)	15/15
RL-SHADE-1	**2.1**(2)	**2.5**(2)	**1.6**(2)	**1.6**(2)	3.5(5)	6/15
RL-SHADE-1	**2.6**(0.8)	**2.1**(2)	**0.90**(1)	3.6(3)	5.1(2)	15/15
R-SHADE-10	**2.4**(2)	**2.6**(2)	**1.6**(2)	**2.1**(0.8)	4.7(3)	5/15
R-SHADE-10	**1.9**(2)	**1.8**(1)	**0.98**(0.4)	**1.2**(0.8)	**1.1**(0.6)	15/15
SOO-Derbel	**0.67**(0)	**1.2**(1)	**0.72**(0.3)	**0.96**(0.7)	**0.98**(0.1)	15/15

Table 43: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**3.0**(2)	**1.6**(0.9)	**1.2**(1)	**1.1**(0.4)	∞ *150*	0/15
R-DE-10e2-	**1.3**(0.4)	**1.7**(1)	**1.2**(2)	**1.3**(1)	**2.4**(3)	9/15
R-DE-10e5-	**2.1**(3)	**2.1**(0.8)	**1.5**(1)	**1.3**(0.9)	21(0.5)	15/15
RL-SHADE-1	4.0(6)	**2.9**(3)	3.2(3)	**2.2**(1)	24(20)	1/15
RL-SHADE-1	3.1(3)	**2.4**(2)	**2.2**(3)	**2.9**(2)	6.7(0.8)	15/15
R-SHADE-10	3.9(6)	**1.9**(3)	**2.3**(2)	**2.4**(2)	24(12)	1/15
R-SHADE-10	3.1(3)	**1.9**(2)	**1.1**(0.9)	**1.2**(1)	**1.8**(0.4)	15/15
SOO-Derbel	**0.96**(1)	**1.2**(1)	**0.90**(0.5)	**0.97**(0.7)	**1.5**(0.8)	15/15

Table 44: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f19	*1.6e-1*:81	*1.0e-1*:109	*6.3e-2*:109	*4.0e-2*:119	*1.6e-2*:1230	15/15
MATSUMOTO-	∞	∞	∞	∞	∞ *150*	0/15
R-DE-10e2-	25(26)	39(75)	∞	∞	∞ *300*	0/15
R-DE-10e5-	**20**(25)	31(19)	54(67)	103(190)	20(14)	15/15
RL-SHADE-1	27(20)	41(49)	∞	∞	∞ *300*	0/15
RL-SHADE-1	25(11)	25(20)	34(29)	64(61)	11(4)	15/15
R-SHADE-10	52(73)	38(55)	38(64)	**35**(31)	∞ *300*	0/15
R-SHADE-10	24(30)	**24**(41)	**30**(19)	50(60)	**10**(28)	15/15
SOO-Derbel	**2.8**(2)	**2.8**(2)	**3.1**(4)	**4.4**(6)	**0.92**(0.7)	15/15

Table 45: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.1**(0.6)	**0.86**(0.6)	**1.5**(0.8)	17(30)	**2.9**(3)	2/15
R-DE-10e2-	**1.8**(0.7)	**1.6**(2)	**2.9**(2)	3.4(1)	**0.89**(1.0)	10/15
R-DE-10e5-	**1.3**(0.7)	**1.3**(0.8)	18(2)	8.7(40)	**2.7**(3)	15/15
RL-SHADE-1	**1.6**(1)	**1.5**(0.5)	**2.5**(2)	**3.2**(2)	**0.62**(0.5)	12/15
RL-SHADE-1	**1.2**(0.4)	**1.2**(0.8)	4.9(3)	13(7)	**2.7**(1)	15/15
R-SHADE-10	**1.6**(1)	**1.8**(2)	3.3(2)	3.8(1.0)	**1.3**(1)	8/15
R-SHADE-10	**1.6**(2)	**1.6**(2)	**2.8**(2)	6.0(5)	**1.8**(2)	15/15
SOO-Derbel	**0.75**(0.1)	**0.61**(0.1)	3.8(0.0)	**1.8**(6e-3)	**0.19**(1e-3)$^{\star 2}$	15/15

Table 46: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**2.1**(2)	**1.6**(1)	**1.3**(0.8)	**0.70**(0.8)	**0.63**(0.4)	11/15
R-DE-10e2-	**1.4**(2)	**1.3**(0.9)	**1.3**(2)	**1.1**(1)	**1.1**(1)	12/15
R-DE-10e5-	**2.7**(4)	**1.8**(3)	**1.9**(2)	4.2(9)	4.1(8)	15/15
RL-SHADE-1	**2.3**(2)	**1.2**(2)	**1.1**(1)	**0.96**(1)	**1.3**(1)	12/15
RL-SHADE-1	**2.3**(2)	**1.3**(1)	**1.7**(3)	**1.8**(1)	**2.0**(2)	15/15
R-SHADE-10	**2.4**(1)	**2.6**(1)	**2.7**(2)	**2.3**(2)	**2.3**(3)	8/15
R-SHADE-10	**2.5**(3)	**1.4**(1)	**1.5**(1)	**1.5**(2)	**1.4**(3)	15/15
SOO-Derbel	**1.3**(2)	**1.3**(1)	**1.2**(0.4)	**0.90**(0.5)	**0.43**(0.7)	15/15

Table 47: 03-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**2.1**(3)	**2.2**(2)	**1.4**(0.8)	**1.4**(2)	**1.5**(2)	7/15
R-DE-10e2-	**1.8**(0.9)	**1.7**(2)	**1.7**(3)	**1.9**(1)	**1.5**(0.9)	11/15
R-DE-10e5-	**2.1**(6)	**1.9**(1)	8.2(28)	5.6(30)	5.7(7)	15/15
RL-SHADE-1	**1.2**(0.5)	**1.4**(1)	**2.9**(4)	**2.3**(0.7)	**1.5**(1)	11/15
RL-SHADE-1	**1.3**(2)	**0.78**(0.4)	**2.0**(3)	**2.0**(2)	**2.4**(2)	15/15
R-SHADE-10	**1.8**(2)	**2.4**(3)	**2.2**(3)	**2.9**(4)	**1.6**(1)	10/15
R-SHADE-10	**1.3**(0.9)	**1.5**(0.7)	**0.87**(0.5)	**1.2**(1)	**1.9**(3)	15/15
SOO-Derbel	**1.0**(1)	**0.74**(0.6)	**0.71**(0.7)	**0.78**(0.8)	**0.47**(0.2)	15/15

Table 48: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f23	*1.0e+1*:2.6	*6.3e+0*:16	*4.0e+0*:44	*2.5e+0*:79	*1.6e+0*:198	15/15
MATSUMOTO-	4.3(2)	**1.5**(2)	**2.1**(2)	5.0(6)	*∞ 150*	0/15
R-DE-10e2-	**3.1**(2)	**2.1**(3)	**1.8**(2)	3.4(3)	7.0(4)	3/15
R-DE-10e5-	4.6(3)	**2.5**(2)	**1.9**(2)	**2.8**(1)	6.8(6)	15/15
RL-SHADE-1	3.4(4)	**0.90**(0.6)	**1.3**(4)	3.6(3)	7.2(8)	3/15
RL-SHADE-1	3.3(2)	**1.5**(2)	**2.0**(2)	**2.9**(2)	7.9(6)	15/15
R-SHADE-10	4.1(4)	3.6(5)	**2.8**(2)	5.3(6)	6.9(9)	3/15
R-SHADE-10	**2.8**(3)	**1.7**(0.8)	**1.5**(2)	3.3(2)	**3.0**(3)	15/15
SOO-Derbel	4.4(6)	**1.8**(3)	**2.2**(2)	**2.8**(2)	**1.8**(1.0)	15/15

Table 49: 03-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f24	*4.0e+1*:4.6	*2.5e+1*:13	*1.6e+1*:47	*1.6e+1*:47	*6.3e+0*:382	15/15
MATSUMOTO-	**1.6**(1)	**1.9**(2)	**1.7**(1)	**1.7**(4)	5.8(5)	1/15
R-DE-10e2-	**1.9**(0.7)	**2.2**(2)	**1.8**(1)	**1.8**(0.7)	**2.4**(2)	4/15
R-DE-10e5-	**1.1**(1)	**1.6**(0.9)	**2.7**(4)	**2.7**(4)	4.4(7)	15/15
RL-SHADE-1	**1.6**(2)	**2.2**(1)	**1.3**(0.4)	**1.3**(0.5)	**2.6**(3)	4/15
RL-SHADE-1	**1.2**(0.5)	**1.8**(2)	**2.0**(1)	**2.0**(1.0)	3.1(3)	15/15
R-SHADE-10	**1.0**(0.6)	**1.5**(2)	**1.5**(2)	**1.5**(2)	5.4(4)	2/15
R-SHADE-10	**1.7**(1)	**1.9**(2)	**1.5**(0.9)	**1.5**(0.9)	**1.1**(0.9)	15/15
SOO-Derbel	**1.4**(2)	**1.4**(1)	**2.1**(0.7)	**2.1**(3)	**1.4**(1)	15/15

Table 50: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.8**(1)	**1.7**(0.8)	∞	∞	∞ *250*	0/15
R-DE-10e2-	**2.5**(4)	**2.6**(2)	∞	∞	∞ *500*	0/15
R-DE-10e5-	**2.5**(2)	**2.8**(3)	**67**(13)	**67**(8)	**67**(14)	15/15
RL-SHADE-1	**1.7**(1)	**2.7**(1)	614(553)	614(430)	614(932)	1/15
RL-SHADE-1	**2.8**(11)	3.8(5)	632(17)	632(20)	632(36)	15/15
R-SHADE-10	**2.8**(2)	**2.6**(1)	∞	∞	∞ *500*	0/15
R-SHADE-10	**2.2**(2)	3.4(2)	**110**(9)	**110**(9)	**110**(11)	15/15
SOO-Derbel	**0.99**(1)	**1.3**(0.8)	194(12)	194(14)	194(18)	15/15

Table 51: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f2	*1.6e+6*:2.9	*4.0e+5*:11	*4.0e+4*:15	*6.3e+2*:58	*1.0e-8*:95	15/15
MATSUMOTO-	**1.6**(1.0)	**0.68**(0.4)	**4.0**(5)	∞	∞ *250*	0/15
R-DE-10e2-	**2.0**(4)	**1.2**(1)	**4.0**(2)	**2.8**(0.5)	∞ *500*	0/15
R-DE-10e5-	**2.1**(1)	**1.6**(2)	4.4(4)	**2.9**(0.9)	**10**(0.6)	15/15
RL-SHADE-1	**2.5**(3)	**1.4**(0.6)	5.2(3)	**2.9**(0.7)	∞ *500*	0/15
RL-SHADE-1	**1.8**(2)	**0.92**(3)	10(10)	22(5)	113(4)	15/15
R-SHADE-10	**2.9**(8)	**2.2**(3)	4.6(4)	3.7(0.6)	∞ *500*	0/15
R-SHADE-10	**1.7**(0.7)	**0.85**(1)	4.2(6)	5.4(1)	**22**(2)	15/15
SOO-Derbel	6.6(11)	**2.5**(3)	6.7(5)	6.6(2)	872(2643)	13/15

Table 52: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f3	*1.6e+2*:4.1	*1.0e+2*:15	*6.3e+1*:23	*2.5e+1*:73	*1.0e+1*:716	15/15
MATSUMOTO-	**1.8**(2)	**1.2**(1)	**1.8**(1)	**2.7**(2)	**1.1**(1)	4/15
R-DE-10e2-	**2.5**(3)	**1.1**(1)	**1.5**(0.9)	**1.8**(1)	**0.41**(0.1)	15/15
R-DE-10e5-	**1.9**(3)	**1.5**(1)	**1.7**(1)	**1.5**(0.8)	**0.36**(0.1)	15/15
RL-SHADE-1	**2.9**(2)	**2.5**(1)	**2.5**(1)	**1.7**(0.4)	**0.33**(0.4)	14/15
RL-SHADE-1	**1.3**(1.0)	**1.9**(4)	7.8(2)	11(3)	3.1(1)	15/15
R-SHADE-10	4.6(3)	**2.3**(1)	**2.3**(1.0)	**1.6**(0.5)	**0.33**(0.0)	15/15
R-SHADE-10	**1.5**(2)	**1.1**(1)	**2.4**(0.9)	3.0(0.6)	**1.1**(0.5)	15/15
SOO-Derbel	**1.9**(1)	**1.2**(0.5)	**1.9**(1.0)	**2.4**(1)	**1.2**(2)	15/15

Table 53: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f4	*2.5e+2*:2.6	*1.6e+2*:10	*1.0e+2*:19	*4.0e+1*:65	*1.6e+1*:434	15/15
MATSUMOTO-	**2.6**(4)	**1.1**(0.7)	**2.9**(2)	3.0(1)	9.0(12)	1/15
R-DE-10e2-	**2.1**(2)	**1.5**(1)	**1.8**(0.7)	**1.9**(0.5)	**0.92**(0.4)	12/15
R-DE-10e5-	4.3(4)	**2.5**(2)	**2.7**(2)	**2.2**(1)	**1.3**(2)	15/15
RL-SHADE-1	**2.7**(0.8)	**2.0**(2)	**2.5**(2)	**1.9**(0.7)	**0.49**(0.2)	15/15
RL-SHADE-1	3.0(1)	**2.2**(1)	4.2(5)	10(3)	4.1(1)	15/15
R-SHADE-10	3.3(3)	**2.0**(2)	**2.0**(1)	**1.6**(0.6)	**0.54**(0.2)	15/15
R-SHADE-10	**2.4**(2)	**1.5**(0.7)	**2.1**(1)	**2.5**(0.6)	**1.4**(0.5)	15/15
SOO-Derbel	**0.69**(0)	**0.68**(0.7)	**1.3**(1)	**2.3**(1)	3.7(19)	15/15

Table 54: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f5	*6.3e+1*:4.0	*4.0e+1*:10	*1.0e-8*:10	*1.0e-8*:10	*1.0e-8*:10	15/15
MATSUMOTO-	**1.6**(0.8)	**1.2**(0.6)	**1.9**(0.3)$^{\star 4}$	**1.9**(0.3)$^{\star 4}$	**1.9**(0.3)$^{\star 4}$	15/15
R-DE-10e2-	**2.4**(2)	**2.7**(1)	∞	∞	∞ *500*	0/15
R-DE-10e5-	**2.3**(3)	**2.4**(1)	**184**(115)	**184**(44)	**184**(93)	15/15
RL-SHADE-1	**2.0**(2)	3.1(3)	372(325)	372(187)	372(538)	2/15
RL-SHADE-1	3.4(3)	4.4(5)	613(20)	613(21)	613(27)	15/15
R-SHADE-10	3.9(5)	3.9(2)	∞	∞	∞ *500*	0/15
R-SHADE-10	**2.6**(4)	5.6(3)	219(30)	219(20)	219(15)	15/15
SOO-Derbel	3.6(0.1)	**1.4**(0.1)	1054(0.1)	1054(0.1)	1054(0.1)	15/15

Table 55: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(0.5)	**0.90**(0.7)	**1.7**(1)	12(14)	∞ *250*	0/15
R-DE-10e2-	**1.9**(2)	**2.0**(2)	3.5(4)	**2.0**(2)	**6.9**(10)	4/15
R-DE-10e5-	**2.7**(2)	**1.9**(1)	3.5(3)	**2.4**(2)	16(40)	15/15
RL-SHADE-1	3.0(3)	**1.9**(5)	3.8(3)	**2.4**(2)	14(10)	2/15
RL-SHADE-1	**1.9**(3)	**2.6**(3)	11(10)	6.7(5)	13(2)	15/15
R-SHADE-10	**1.9**(2)	**1.6**(1)	3.3(4)	**2.3**(2)	29(35)	1/15
R-SHADE-10	**2.4**(1)	**2.5**(4)	4.2(4)	**2.7**(2)	**2.7**(0.9)	15/15
SOO-Derbel	**1.7**(4)	**1.2**(2)	**1.8**(2)	**2.0**(2)	6105(3636)	4/15

Table 56: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.3**(2)	**1.8**(1)	**1.5**(0.7)	7.6(9)	5.4(4)	2/15
R-DE-10e2-	**2.3**(2)	**2.3**(5)	**2.5**(1)	5.4(5)	5.3(5)	4/15
R-DE-10e5-	**2.0**(0.9)	**2.1**(1)	**2.8**(1)	4.0(3)	**2.7**(1.0)	15/15
RL-SHADE-1	**1.5**(2)	**2.3**(4)	3.5(2)	5.9(5)	**2.8**(1)	7/15
RL-SHADE-1	**2.2**(0.9)	**2.1**(1)	8.2(9)	12(7)	5.5(2)	15/15
R-SHADE-10	**2.9**(3)	3.1(2)	4.0(3)	**3.9**(2)	**1.9**(3)	10/15
R-SHADE-10	**2.3**(2)	**2.6**(5)	**2.4**(2)	**2.6**(2)	**1.3**(2)	15/15
SOO-Derbel	**1.1**(0.9)	**2.2**(1)	**2.1**(2)	4.1(4)	**2.1**(2)	15/15

Table 57: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.7**(1)	**1.9**(1)	**1.4**(0.6)	**2.6**(2)	∞ *250*	0/15
R-DE-10e2-	3.3(4)	**2.9**(2)	**2.9**(2)	**2.4**(1)	**9.0**(19)	3/15
R-DE-10e5-	**2.1**(3)	**2.0**(1)	**2.0**(1)	**2.6**(0.6)	22(57)	15/15
RL-SHADE-1	3.3(2)	3.8(4)	3.9(0.5)	3.6(5)	14(12)	2/15
RL-SHADE-1	**2.5**(2)	3.5(7)	4.9(4)	13(6)	17(3)	15/15
R-SHADE-10	**2.4**(3)	**1.8**(2)	**2.5**(1)	3.4(2)	∞ *500*	0/15
R-SHADE-10	3.6(5)	3.0(3)	3.2(4)	3.9(3)	**4.3**(2)	15/15
SOO-Derbel	**1.4**(0.5)	**1.3**(2)	**1.4**(1.0)	**2.1**(1)	16(42)	15/15

Table 58: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f9	*2.5e+1*:20	*1.6e+1*:26	*1.0e+1*:35	*4.0e+0*:62	*1.6e-2*:256	15/15
MATSUMOTO-	18(20)	34(13)	35(44)	64(65)	∞ *250*	0/15
R-DE-10e2-	**11**(7)	10(5)	**8.1**(3)	11(13)	∞ *500*	0/15
R-DE-10e5-	12(4)	13(12)	11(6)	10(5)	487(633)	15/15
RL-SHADE-1	13(10)	11(3)	10(4)	15(17)	∞ *500*	0/15
RL-SHADE-1	54(20)	51(10)	46(11)	45(4)	**28**(3)	15/15
R-SHADE-10	14(7)	13(8)	10(4)	28(20)	∞ *500*	0/15
R-SHADE-10	11(4)	**10**(1)	8.1(3)	**8.5**(1)	**9.1**(8)	15/15
SOO-Derbel	**6.5**(2)	**6.3**(2)	**5.7**(1)	**5.1**(2)	202(348)	15/15

Table 59: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.5**(2)	**1.9**(1)	**1.4**(0.9)	6.4(9)	∞ *250*	0/15
R-DE-10e2-	**1.4**(2)	**1.2**(0.7)	**1.2**(1)	**2.6**(0.7)	∞ *500*	0/15
R-DE-10e5-	**1.3**(3)	**1.9**(2)	**1.2**(0.4)	7.1(4)	169(117)	15/15
RL-SHADE-1	**1.6**(3)	**1.6**(0.9)	**1.5**(2)	4.5(3)	**12**(15)	2/15
RL-SHADE-1	**1.6**(1)	**2.3**(2)	**1.4**(0.5)	14(6)	15(1)	15/15
R-SHADE-10	**1.7**(1)	**1.7**(3)	**1.2**(2)	6.7(8)	∞ *500*	0/15
R-SHADE-10	**1.8**(3)	**1.2**(1)	**1.3**(2)	**3.5**(1)	**2.2**(0.6)	15/15
SOO-Derbel	**1.4**(0.7)	**0.74**(0.8)	**0.56**(0.6)	4.6(2)	74(254)	15/15

Table 60: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.4**(2)	**2.5**(2)	4.7(1)	8.9(6)	∞ *250*	0/15
R-DE-10e2-	**2.3**(4)	**2.0**(2)	5.1(5)	3.8(3)	∞ *500*	0/15
R-DE-10e5-	**1.9**(2)	3.3(1)	4.6(2)	36(121)	4079(4104)	5/15
RL-SHADE-1	**1.5**(0.8)	**2.2**(5)	**3.7**(2)	**3.1**(3)	∞ *500*	0/15
RL-SHADE-1	**3.0**(6)	3.4(5)	6.6(5)	6.7(3)	**15**(3)	15/15
R-SHADE-10	**2.0**(3)	**1.8**(2)	4.0(5)	4.0(4)	∞ *500*	0/15
R-SHADE-10	**1.2**(1)	4.2(3)	**3.9**(3)	**2.2**(0.9)	**2.7**(1)	15/15
SOO-Derbel	**0.98**(0.3)	**1.6**(3)	5.0(2)	7.5(20)	2683(2985)	6/15

Table 61: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.2**(0.6)	**1.4**(1.0)	**1.7**(0.5)	**3.2**(1)	∞ *250*	0/15
R-DE-10e2-	**1.8**(2)	**2.6**(4)	3.1(1)	**4.5**(2)	∞ *500*	0/15
R-DE-10e5-	**1.6**(2)	**2.3**(2)	**2.4**(2)	5.3(3)	112(66)	15/15
RL-SHADE-1	**1.4**(1)	**2.5**(1.0)	3.3(1)	4.8(0.9)	∞ *500*	0/15
RL-SHADE-1	**2.1**(2)	4.0(1)	6.9(10)	39(11)	29(4)	15/15
R-SHADE-10	**2.1**(0.6)	**2.7**(1)	**2.9**(1)	5.6(2)	∞ *500*	0/15
R-SHADE-10	**0.96**(0.8)	**1.8**(1)	**2.3**(2)	6.0(2)	**7.1**(8)	15/15
SOO-Derbel	**0.50**(0.1)	**0.66**(2)	**1.2**(0.7)	5.4(2)	**6.8**(1)	15/15

Table 62: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**2.0**(1.0)	**1.8**(1)	**1.5**(0.5)	**1.7**(0.3)$^{\star 2}$	∞ *250*	0/15
R-DE-10e2-	**2.5**(2)	**2.7**(0.7)	**2.4**(2)	4.5(1)	∞ *500*	0/15
R-DE-10e5-	**2.0**(2)	**3.0**(2)	3.3(2)	4.6(3)	739(735)	13/15
RL-SHADE-1	**1.7**(0.8)	**2.6**(3)	3.4(2)	3.9(0.9)	∞ *500*	0/15
RL-SHADE-1	3.4(8)	5.0(5)	6.7(5)	23(6)	**28**(2)	15/15
R-SHADE-10	**2.5**(5)	3.4(2)	**2.8**(1)	5.3(5)	∞ *500*	0/15
R-SHADE-10	**1.2**(0.8)	**1.9**(5)	**2.8**(3)	4.2(2)	**4.8**(2)	15/15
SOO-Derbel	**1.0**(2)	**0.96**(0.7)	**1.2**(0.4)	**3.9**(0.3)	29(26)	15/15

Table 63: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f14	*1.6e+1*:3.0	*1.0e+1*:10	*6.3e+0*:15	*2.5e-1*:53	*1.0e-5*:251	15/15
MATSUMOTO-	**2.1**(1)	**1.4**(0.8)	**1.3**(0.5)	**2.3**(1)	∞ *250*	0/15
R-DE-10e2-	3.4(6)	**1.7**(2)	**2.3**(2)	**3.4**(2)	∞ *500*	0/15
R-DE-10e5-	3.5(2)	**2.1**(2)	**1.8**(2)	3.7(2)	579(1050)	14/15
RL-SHADE-1	**2.3**(2)	**2.2**(3)	**2.6**(3)	3.6(1)	∞ *500*	0/15
RL-SHADE-1	5.3(3)	3.1(4)	3.5(3)	20(7)	**29**(2)	15/15
R-SHADE-10	**2.7**(4)	**1.6**(2)	**2.1**(2)	3.9(0.5)	∞ *500*	0/15
R-SHADE-10	3.2(2)	**1.3**(0.7)	**1.4**(1)	4.4(0.6)	**5.3**(0.9)	15/15
SOO-Derbel	**1.0**(1)	**0.59**(0.8)	**0.74**(0.9)	3.6(1)	1342(1364)	12/15

Table 64: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f15	*1.6e+2*:3.0	*1.0e+2*:13	*6.3e+1*:24	*4.0e+1*:55	*1.6e+1*:289	5/5
MATSUMOTO-	**2.4**(2)	**1.3**(0.1)	**1.6**(2)	**1.7**(1.0)	**0.97**(0.8)	10/15
R-DE-10e2-	**2.6**(2)	**1.6**(0.8)	**1.8**(0.8)	**1.9**(2)	**2.1**(0.8)	10/15
R-DE-10e5-	**2.9**(2)	**1.6**(1)	**1.8**(1)	**1.6**(0.4)	3.4(7)	15/15
RL-SHADE-1	**2.5**(2)	**2.8**(2)	3.1(0.8)	**2.0**(0.4)	3.0(3)	7/15
RL-SHADE-1	**1.9**(3)	**2.6**(4)	6.1(5)	7.3(5)	7.6(3)	15/15
R-SHADE-10	4.8(5)	**2.3**(2)	**2.4**(1)	**2.0**(0.4)	**2.8**(3)	8/15
R-SHADE-10	4.2(6)	**2.0**(1)	**2.4**(1)	**2.4**(3)	**2.0**(0.6)	15/15
SOO-Derbel	**1.1**(2)	**0.64**(0.5)	**1.3**(0.7)	**1.8**(2)	**1.6**(0.6)	15/15

Table 65: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f16	*4.0e+1*:4.8	*2.5e+1*:16	*1.6e+1*:46	*1.0e+1*:120	*4.0e+0*:334	15/15
MATSUMOTO-	**1.4**(2)	**0.86**(0.7)	**1.2**(2)	**1.2**(0.9)	3.9(4)	3/15
R-DE-10e2-	**1.4**(0.8)	**1.9**(2)	**1.1**(1)	**0.95**(1)	5.3(4)	4/15
R-DE-10e5-	**1.0**(0.9)	**1.2**(1.0)	**1.7**(2)	**1.7**(2)	4.5(5)	15/15
RL-SHADE-1	**1.8**(0.8)	**1.4**(2)	**1.3**(0.5)	**1.2**(0.9)	**1.7**(1.0)	10/15
RL-SHADE-1	**1.1**(1)	**1.3**(0.6)	**1.6**(3)	**1.6**(2)	7.5(5)	15/15
R-SHADE-10	**1.9**(2)	**1.6**(2)	**1.6**(1)	**1.4**(0.7)	**2.2**(4)	8/15
R-SHADE-10	**1.7**(2)	**0.85**(0.5)	**1.5**(2)	**1.1**(0.8)	**2.7**(1)	15/15
SOO-Derbel	**1.7**(2)	**1.6**(1)	**1.9**(1)	**1.2**(0.7)	**0.85**(0.5)	15/15

Table 66: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f17	*1.0e+1*:5.2	*6.3e+0*:26	*4.0e+0*:57	*2.5e+0*:110	*6.3e-1*:412	15/15
MATSUMOTO-	3.1(2)	**1.2**(1)	**1.1**(0.5)	**1.9**(1)	9.2(8)	1/15
R-DE-10e2-	4.1(4)	**1.8**(1)	**1.5**(0.8)	**1.2**(0.7)	**0.94**(0.2)	12/15
R-DE-10e5-	4.2(2)	**1.8**(2)	**1.7**(1)	5.0(7)	7.4(11)	15/15
RL-SHADE-1	**2.4**(2)	**1.0**(1)	**1.1**(0.5)	**1.3**(0.8)	**1.8**(2)	8/15
RL-SHADE-1	**2.8**(2)	**1.7**(1)	3.1(3)	4.3(2)	4.6(1)	15/15
R-SHADE-10	3.3(4)	**2.1**(2)	**1.9**(1)	**1.8**(2)	**2.7**(2)	6/15
R-SHADE-10	**2.8**(2)	**1.3**(0.4)	**1.2**(0.9)	**1.3**(1)	**1.0**(0.4)	15/15
SOO-Derbel	**1.4**(2)	**0.64**(0.6)	**0.86**(0.7)	**0.93**(0.5)	**0.92**(0.6)	15/15

Table 67: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.9**(2)	**1.4**(2)	**1.1**(0.9)	**0.85**(1)	∞ *250*	0/15
R-DE-10e2-	**2.1**(2)	**2.5**(2)	**1.9**(2)	**1.3**(0.3)	4.2(2)	5/15
R-DE-10e5-	**2.6**(1)	**2.7**(6)	**2.1**(2)	**1.3**(0.5)	**2.9**(3)	15/15
RL-SHADE-1	**1.5**(3)	**1.7**(1)	**2.2**(2)	**1.4**(1)	4.4(3)	5/15
RL-SHADE-1	**1.1**(0.8)	**2.1**(1)	3.1(2)	3.9(3)	8.2(3)	15/15
R-SHADE-10	**1.5**(1)	**2.8**(5)	**1.8**(2)	**1.7**(0.8)	12(13)	2/15
R-SHADE-10	**1.4**(1)	**1.5**(2)	**1.8**(2)	**1.5**(0.9)	**1.5**(1)	15/15
SOO-Derbel	**0.76**(0.8)	**1.0**(1)	**0.76**(0.4)	**0.79**(0.5)	**1.6**(0.7)	15/15

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
MATSUMOTO-	∞	∞	∞	∞	∞ *250*	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ *500*	0/15
R-DE-10e5-	1109(1356)	1951(2356)	3319(4999)	1163(1634)	1470(1238)	1/15
RL-SHADE-1	∞	∞	∞	∞	∞ *500*	0/15
RL-SHADE-1	273(149)	269(150)	134(116)	70(52)	116(136)	9/15
R-SHADE-10	∞	∞	∞	∞	∞ *500*	0/15
R-SHADE-10	**90**(115)	**111**(153)	**85**(75)	**68**(108)	**61**(82)	11/15
SOO-Derbel	**3.7**(1)	**10**(28)	**6.3**(0.3)	**1.4**(0.1)	**0.90**(3)	15/15

Table 69: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.7**(1)	**1.4**(1)	**1.9**(1)	4.1(3)	∞ *250*	0/15
R-DE-10e2-	**2.1**(3)	**1.8**(2)	3.5(3)	**2.0**(0.7)	**1.1**(1)	7/15
R-DE-10e5-	**2.4**(2)	**1.8**(0.8)	**3.3**(2)	**2.2**(0.8)	**0.95**(1)	15/15
RL-SHADE-1	**1.4**(1)	**1.3**(1)	4.3(3)	**2.0**(0.6)	**1.3**(1)	6/15
RL-SHADE-1	**2.0**(2)	**1.8**(2)	12(4)	11(6)	5.6(2)	15/15
R-SHADE-10	**2.5**(3)	**2.2**(3)	4.0(3)	**2.5**(1)	**1.4**(2)	6/15
R-SHADE-10	**2.3**(4)	**1.9**(2)	3.9(2)	3.2(3)	**1.9**(0.8)	15/15
SOO-Derbel	**0.88**(0.1)	**1.5**(0.1)	11(0.0)	3.7(7e-3)	**1.2**(6e-4)	15/15

Table 70: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f21	*4.0e+1*:3.9	*2.5e+1*:11	*1.6e+1*:31	*6.3e+0*:73	*1.6e+0*:347	5/5
MATSUMOTO-	**1.3**(0.6)	**1.3**(2)	**0.73**(0.8)	**1.2**(0.7)	**1.5**(1)	7/15
R-DE-10e2-	**1.8**(1)	**1.2**(0.8)	**0.82**(0.8)	**1.6**(1)	4.7(4)	4/15
R-DE-10e5-	**2.2**(3)	**1.6**(0.7)	**1.1**(0.7)	3.6(6)	33(23)	15/15
RL-SHADE-1	**1**(0.7)	**1.4**(1)	**1.9**(2)	**2.1**(1.0)	**2.3**(3)	7/15
RL-SHADE-1	**1.5**(2)	**1.5**(2)	**1.6**(1)	5.0(3)	4.5(2)	15/15
R-SHADE-10	**1.9**(2)	**1.5**(3)	**1.3**(1)	**2.0**(0.7)	**2.1**(4)	8/15
R-SHADE-10	**2.1**(0.8)	**2.2**(2)	**1.6**(1)	4.5(9)	5.3(10)	15/15
SOO-Derbel	**1.3**(1.0)	**1.1**(2)	**0.88**(0.7)	**0.99**(0.6)	**0.90**(0.8)	15/15

Table 71: 05-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f22	*6.3e+1*:3.6	*4.0e+1*:15	*2.5e+1*:32	*1.0e+1*:71	*1.6e+0*:341	5/5
MATSUMOTO-	**1.6**(1)	**1.1**(0.9)	**1.1**(0.6)	**1.1**(0.7)	**2.3**(3)	5/15
R-DE-10e2-	**2.8**(3)	**1.5**(2)	**1.2**(0.6)	**1.3**(0.9)	**2.3**(3)	7/15
R-DE-10e5-	**1.4**(0.7)	**1.3**(1)	**1.5**(2)	**2.1**(1)	17(29)	15/15
RL-SHADE-1	**1.4**(2)	**1.9**(2)	**1.3**(0.7)	**1.6**(0.7)	**2.9**(4)	6/15
RL-SHADE-1	**2.5**(3)	**2.0**(1.0)	**1.9**(3)	3.7(5)	4.2(2)	15/15
R-SHADE-10	**2.7**(5)	**1.5**(1)	**1.1**(0.4)	**1.6**(2)	6.4(11)	3/15
R-SHADE-10	3.0(2)	**1.8**(1)	**1.2**(1)	**1.5**(0.8)	3.3(7)	15/15
SOO-Derbel	**1.5**(2)	**0.94**(0.6)	**0.77**(0.7)	**1.0**(0.5)	**1.00**(1.0)	15/15

Table 72: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
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
MATSUMOTO-	**1.8**(2)	**1.7**(2)	**2.4**(3)	**3.9**(6)	∞ *250*	0/15
R-DE-10e2-	**1.6**(2)	**1.1**(1)	**2.1**(2)	4.5(3)	14(10)	1/15
R-DE-10e5-	**2.0**(1)	**2.4**(3)	3.8(4)	6.0(7)	34(34)	15/15
RL-SHADE-1	**2.8**(3)	**2.8**(1)	3.1(3)	7.3(9)	∞ *500*	0/15
RL-SHADE-1	**2.2**(2)	**2.3**(0.9)	**2.5**(1)	5.1(6)	18(9)	15/15
R-SHADE-10	3.1(4)	**2.6**(1)	3.5(4)	5.2(6)	∞ *500*	0/15
R-SHADE-10	**2.9**(3)	**2.1**(2)	3.2(3)	6.3(7)	**6.2**(3)	15/15
SOO-Derbel	**1.7**(2)	**2.9**(5)	**2.5**(3)	**3.8**(2)	**1.4**(0.2)	15/15

Table 73: 05-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f24	*6.3e+1*:15	*4.0e+1*:37	*4.0e+1*:37	*2.5e+1*:118	*1.6e+1*:692	15/15
MATSUMOTO-	**1.9**(3)	3.1(2)	3.1(4)	10(8)	5.3(4)	1/15
R-DE-10e2-	**1.7**(0.9)	**2.7**(1)	**2.7**(1)	3.9(4)	11(18)	1/15
R-DE-10e5-	**1.9**(2)	**2.0**(2)	**2.0**(2)	**2.0**(3)	**2.0**(1)	15/15
RL-SHADE-1	**2.4**(3)	3.7(1)	3.7(2)	5.3(4)	11(14)	1/15
RL-SHADE-1	**2.8**(3)	8.0(6)	8.0(7)	10(4)	5.0(3)	15/15
R-SHADE-10	**2.8**(1)	3.3(2)	3.3(2)	5.7(4)	11(16)	1/15
R-SHADE-10	**2.4**(3)	3.1(2)	3.1(2)	**3.0**(2)	**1.8**(1)	15/15
SOO-Derbel	**1.1**(2)	5.1(4)	5.1(3)	5.2(2)	**1.8**(1)	15/15

Table 74: 10-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	*4.0e+1*:8.0	*2.5e+1*:16	*1.0e-8*:23	*1.0e-8*:23	*1.0e-8*:23	15/15
MATSUMOTO-	**2.9**(2)	**2.0**(0.4)	∞	∞	∞ *500*	0/15
R-DE-10e2-	4.7(3)	3.5(1)	∞	∞	∞ *1000*	0/15
R-DE-10e5-	9.0(6)	6.3(3)	**110**(7)	**110**(6)	**110**(7)	15/15
RL-SHADE-1	9.0(5)	7.5(2)	∞	∞	∞ *1000*	0/15
RL-SHADE-1	13(9)	25(16)	884(34)	884(34)	884(20)	15/15
R-SHADE-10	5.5(4)	5.0(3)	∞	∞	∞ *1000*	0/15
R-SHADE-10	7.3(6)	7.2(4)	**191**(10)	**191**(12)	**191**(14)	15/15
SOO-Derbel	**1.5**(1)	**1.8**(2)	457(23)	457(26)	457(17)	15/15

Table 75: 10-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	*2.5e+6*:5.6	*1.0e+6*:17	*1.0e+5*:33	*2.5e+3*:118	*1.0e-8*:196	15/15
MATSUMOTO-	**2.0**(2)	**1.3**(0.6)	4.6(4)	∞	∞ *500*	0/15
R-DE-10e2-	**2.0**(0.9)	**1.7**(1)	**2.9**(2)	**2.1**(0.8)	∞ *1000*	0/15
R-DE-10e5-	**1.7**(2)	**2.2**(3)	5.1(3)	4.1(0.7)	**18**(0.7)	15/15
RL-SHADE-1	**2.8**(2)	3.3(0.7)	5.4(2)	3.1(1)	∞ *1000*	0/15
RL-SHADE-1	**2.0**(2)	**2.6**(2)	26(15)	34(5)	140(3)	15/15
R-SHADE-10	**2.1**(2)	**2.3**(3)	**3.9**(2)	**2.7**(1)	∞ *1000*	0/15
R-SHADE-10	**1.4**(2)	**2.1**(1)	6.5(4)	6.4(2)	**32**(2)	15/15
SOO-Derbel	16(15)	8.7(8)	16(7)	621(5)	6117(2550)	7/15

Table 76: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f3	*4.0e+2*:8.2	*1.6e+2*:37	*1.0e+2*:69	*6.3e+1*:147	*2.5e+1*:1129	15/15
MATSUMOTO-	**1.3**(1.0)	**1.5**(0.3)	**1.8**(0.9)	**2.0**(1.0)	**2.1**(1)	3/15
R-DE-10e2-	**2.1**(2)	**1.6**(0.5)	**1.8**(0.5)	**1.6**(0.5)	**1.5**(1)	8/15
R-DE-10e5-	**1.8**(2)	3.3(1)	3.5(0.8)	**2.8**(0.6)	**0.88**(0.1)	15/15
RL-SHADE-1	**1.9**(2)	3.9(0.5)	**2.9**(0.7)	**1.8**(0.3)	**0.83**(0.7)	10/15
RL-SHADE-1	**2.0**(2)	14(7)	19(3)	18(6)	11(2)	15/15
R-SHADE-10	**2.8**(2)	**2.4**(0.4)	**2.3**(0.6)	**1.7**(0.2)	**0.45**(0.1)	15/15
R-SHADE-10	**1.9**(3)	3.8(2)	5.3(2)	6.2(3)	3.3(0.9)	15/15
SOO-Derbel	**2.3**(4)	**1.7**(0.7)	3.1(1)	3.7(2)	**1.8**(2)	15/15

Table 77: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f4	*2.5e+2*:21	*1.6e+2*:59	*1.6e+2*:59	*6.3e+1*:139	*4.0e+1*:854	15/15
MATSUMOTO-	**2.9**(4)	**2.9**(1)	**2.9**(3)	18(16)	∞ *500*	0/15
R-DE-10e2-	3.0(2)	**2.1**(0.3)	**2.1**(1)	3.1(3)	**0.81**(0.4)	13/15
R-DE-10e5-	4.6(3)	3.4(1)	3.4(0.9)	3.8(1.0)	**0.89**(0.2)	15/15
RL-SHADE-1	6.5(2)	**3.0**(0.7)	**3.0**(0.6)	**2.8**(0.6)	**1.1**(2)	10/15
RL-SHADE-1	13(9)	16(7)	16(6)	29(6)	10(2)	15/15
R-SHADE-10	3.8(3)	**2.3**(0.9)	**2.3**(0.7)	**2.4**(0.7)	**0.56**(0.2)	15/15
R-SHADE-10	4.4(2)	3.9(1)	3.9(2)	10(7)	3.3(0.9)	15/15
SOO-Derbel	**1.4**(2)	**1.5**(0.8)	**1.5**(1.0)	5.9(1)	3.1(0.8)	15/15

Table 78: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f5	*1.0e+2*:16	*6.3e+1*:19	*1.0e-8*:20	*1.0e-8*:20	*1.0e-8*:20	15/15
MATSUMOTO-	**1.2**(0.5)	**1.3**(0.0)$^{\star 2}$	**2.0**(0.2)$^{\star 4}$	**2.0**(0.2)$^{\star 4}$	**2.0**(0.2)$^{\star 4}$	15/15
R-DE-10e2-	**2.1**(0.9)	**4.1**(1)	∞	∞	∞ *1000*	0/15
R-DE-10e5-	3.4(3)	7.9(3)	**203**(9)	**203**(12)	**203**(9)	15/15
RL-SHADE-1	4.2(3)	6.9(3)	∞	∞	∞ *1000*	0/15
RL-SHADE-1	6.9(12)	39(24)	862(10)	862(21)	862(13)	15/15
R-SHADE-10	**2.8**(2)	4.6(3)	367(396)	367(533)	367(644)	2/15
R-SHADE-10	4.8(4)	12(9)	444(43)	444(22)	444(24)	15/15
SOO-Derbel	3.9(0.0)	7.3(0.0)	2290(0.0)	2290(0.0)	2290(0.0)	15/15

Table 79: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f6	*1.6e+5*:7.0	*6.3e+4*:16	*4.0e+2*:36	*1.0e+2*:102	*4.0e+0*:504	15/15
MATSUMOTO-	**1.5**(2)	**1.2**(0.5)	**1.1**(0.3)	5.1(6)	∞ *500*	0/15
R-DE-10e2-	**2.1**(2)	**1.8**(1)	**2.3**(0.8)	**2.1**(1)	4.5(4)	6/15
R-DE-10e5-	**2.2**(3)	3.5(3)	4.5(3)	**3.0**(2)	4.6(0.6)	15/15
RL-SHADE-1	3.1(3)	4.0(4)	3.8(1)	**2.4**(1)	29(70)	1/15
RL-SHADE-1	**2.2**(1)	6.0(5)	18(11)	11(4)	17(3)	15/15
R-SHADE-10	**2.4**(3)	**2.9**(3)	3.3(4)	**1.9**(1)	**4.4**(3)	6/15
R-SHADE-10	3.3(2)	3.9(2)	6.3(5)	3.3(2)	**3.2**(0.6)	15/15
SOO-Derbel	**1.3**(3)	**2.0**(2)	3.4(2)	**2.6**(3)	2.8e4(6e4)	1/15

Table 80: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f7	*2.5e+2*:9.2	*1.6e+2*:18	*1.0e+2*:33	*1.0e+1*:172	*4.0e+0*:678	15/15
MATSUMOTO-	**1.4**(2)	**1.7**(1)	**1.6**(0.6)	45(39)	∞ *500*	0/15
R-DE-10e2-	**2.8**(3)	**2.6**(2)	3.2(1)	4.6(7)	**2.1**(1)	8/15
R-DE-10e5-	3.5(3)	3.5(3)	3.2(3)	7.2(4)	4.7(2)	15/15
RL-SHADE-1	**2.9**(5)	3.5(3)	3.6(3)	**3.7**(5)	**2.5**(1)	7/15
RL-SHADE-1	7.3(6)	7.5(5)	14(8)	19(4)	7.7(1)	15/15
R-SHADE-10	**3.0**(3)	3.1(1)	**2.8**(0.7)	3.9(4)	**2.3**(2)	8/15
R-SHADE-10	5.3(4)	4.1(2)	4.4(2)	**3.9**(1)	**1.5**(0.4)	15/15
SOO-Derbel	**1.8**(2)	**1.8**(2)	**1.7**(2)	6.6(13)	4.7(3)	15/15

Table 81: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f8	*1.6e+4*:15	*1.0e+4*:22	*1.6e+3*:34	*2.5e+2*:103	*4.0e+0*:727	15/15
MATSUMOTO-	**2.6**(2)	**2.1**(0.9)	**2.6**(1)	**2.3**(1)	∞ *500*	0/15
R-DE-10e2-	4.2(2)	3.0(0.9)	**3.5**(2)	**2.1**(0.7)	∞ *1000*	0/15
R-DE-10e5-	5.4(1)	4.2(1)	7.0(2)	4.4(1)	369(391)	13/15
RL-SHADE-1	5.2(4)	5.5(3)	6.2(1)	4.0(2)	∞ *1000*	0/15
RL-SHADE-1	12(12)	15(7)	35(12)	27(5)	21(1)	15/15
R-SHADE-10	4.1(2)	3.1(2)	4.7(1)	**2.9**(0.8)	**4.9**(3)	4/15
R-SHADE-10	3.4(3)	3.7(3)	6.8(1)	4.2(0.8)	**6.9**(6)	15/15
SOO-Derbel	**2.2**(2)	**1.7**(2)	4.2(3)	4.3(1)	944(1254)	11/15

Table 82: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f9	*4.0e+1*:125	*2.5e+1*:148	*1.6e+1*:180	*1.0e+1*:200	*1.6e+0*:563	15/15
MATSUMOTO-	7.3(5)	12(12)	41(38)	∞	∞ *500*	0/15
R-DE-10e2-	7.5(7)	7.4(5)	7.3(11)	**7.3**(4)	∞ *1000*	0/15
R-DE-10e5-	8.4(3)	8.4(4)	8.6(4)	10(7)	2063(1397)	9/15
RL-SHADE-1	17(18)	18(15)	19(22)	74(99)	∞ *1000*	0/15
RL-SHADE-1	42(6)	41(6)	38(4)	44(4)	**57**(180)	15/15
R-SHADE-10	7.3(4)	**6.5**(2)	**6.5**(6)	8.3(5)	∞ *1000*	0/15
R-SHADE-10	**6.5**(1)	**6.2**(0.8)	**5.9**(0.6)	**7.1**(1)	**16**(15)	15/15
SOO-Derbel	**5.9**(2)	7.0(3)	7.4(2)	9.2(4)	1814(2203)	9/15

Table 83: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f10	*2.5e+6*:6.0	*1.0e+6*:21	*4.0e+5*:38	*2.5e+4*:104	*6.3e+2*:512	15/15
MATSUMOTO-	**1.7**(2)	**1.3**(1.0)	**1.5**(0.4)	6.2(1)	∞ *500*	0/15
R-DE-10e2-	**2.3**(3)	**1.3**(1)	**2.0**(2)	7.1(7)	∞ *1000*	0/15
R-DE-10e5-	**2.4**(5)	**1.6**(0.9)	**2.1**(0.9)	12(9)	1606(1431)	10/15
RL-SHADE-1	**1.8**(1)	**1.5**(3)	**2.6**(2)	8.9(13)	∞ *1000*	0/15
RL-SHADE-1	**2.2**(2)	**1.3**(1)	5.4(6)	26(10)	**19**(3)	15/15
R-SHADE-10	**2.5**(2)	**1.8**(2)	**1.9**(2)	5.9(3)	∞ *1000*	0/15
R-SHADE-10	**1.3**(1)	**1.7**(2)	**2.1**(2)	**4.8**(2)	**3.0**(0.8)	15/15
SOO-Derbel	**2.7**(4)	**1.2**(2)	**1.3**(0.8)	**5.3**(5)	360(991)	13/15

Table 84: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f11	*4.0e+4*:6.4	*2.5e+3*:15	*6.3e+1*:217	*4.0e+1*:244	*2.5e+0*:675	15/15
MATSUMOTO-	4.5(4)	3.9(3)	16(22)	**14**(13)	∞ *500*	0/15
R-DE-10e2-	**2.8**(3)	**3.1**(2)	21(25)	∞	∞ *1000*	0/15
R-DE-10e5-	**2.6**(2)	4.5(7)	20(12)	59(55)	∞ *1e6*	0/15
RL-SHADE-1	5.8(5)	5.4(2)	11(13)	58(55)	∞ *1000*	0/15
RL-SHADE-1	**2.2**(2)	3.6(1)	12(12)	22(12)	**19**(2)	15/15
R-SHADE-10	3.5(4)	**3.3**(4)	**7.7**(5)	27(32)	∞ *1000*	0/15
R-SHADE-10	3.9(3)	6.3(3)	**2.0**(1)	**2.7**(2)	**3.7**(3)	15/15
SOO-Derbel	**1.1**(1)	3.6(4)	8.6(3)	334(86)	∞ *1e6*	0/15

Table 85: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f12	*4.0e+7*:15	*2.5e+7*:24	*1.6e+7*:34	*1.0e+6*:103	*1.0e+1*:515	15/15
MATSUMOTO-	**2.0**(0.9)	**1.7**(0.6)	**1.6**(0.5)	**1.5**(0.4)*	∞ *500*	0/15
R-DE-10e2-	3.4(1)	**2.9**(1)	**2.5**(1)	**2.4**(0.9)	**7.0**(4)	4/15
R-DE-10e5-	4.7(3)	4.9(2)	4.9(2)	4.6(1)	25(15)	15/15
RL-SHADE-1	5.5(4)	5.2(0.9)	4.5(0.8)	3.3(2)	∞ *1000*	0/15
RL-SHADE-1	8.9(2)	16(10)	17(9)	32(7)	32(3)	15/15
R-SHADE-10	3.2(2)	3.5(1)	3.4(1)	3.1(0.4)	∞ *1000*	0/15
R-SHADE-10	3.8(2)	3.4(1)	4.2(1)	5.4(1)	**10**(16)	15/15
SOO-Derbel	**1.2**(0.7)	**1.5**(2)	**1.7**(2)	6.3(10)	12(4)	15/15

Table 86: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f13	*1.0e+3*:12	*6.3e+2*:32	*4.0e+2*:40	*6.3e+1*:154	*2.5e+0*:521	15/15
MATSUMOTO-	**2.3**(1)	**1.5**(0.5)	**2.0**(0.4)	**2.3**(3)$^{\star 2}$	**6.9**(7)	2/15
R-DE-10e2-	3.8(2)	**2.8**(1)	4.4(3)	7.1(7)	9.4(15)	3/15
R-DE-10e5-	6.0(6)	5.2(0.9)	7.2(2)	6.6(2)	46(87)	15/15
RL-SHADE-1	7.0(5)	4.6(2)	5.3(1)	7.4(5)	∞ *1000*	0/15
RL-SHADE-1	18(15)	24(10)	36(7)	41(4)	25(1)	15/15
R-SHADE-10	5.1(4)	4.2(1)	5.4(2)	**4.3**(1)	29(22)	1/15
R-SHADE-10	6.3(2)	5.6(0.8)	8.1(4)	6.8(2)	**5.8**(2)	15/15
SOO-Derbel	**1.5**(1)	**2.4**(2)	**4.1**(2)	10(6)	108(62)	15/15

Table 87: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f14	*4.0e+1*:7.7	*1.6e+1*:27	*1.0e+1*:37	*6.3e-1*:107	*1.0e-4*:505	15/15
MATSUMOTO-	**1.1**(0.7)	**1.1**(0.8)	**1.8**(2)	4.6(5)	∞ *500*	0/15
R-DE-10e2-	**1.2**(0.5)	**1.4**(0.7)	**1.8**(1)	**4.1**(3)	∞ *1000*	0/15
R-DE-10e5-	**1.4**(2)	**2.5**(2)	3.2(2)	5.2(0.9)	2.8e4(2e4)	1/15
RL-SHADE-1	**1.1**(0.8)	**2.8**(2)	3.5(2)	4.6(2)	∞ *1000*	0/15
RL-SHADE-1	**1.3**(1)	3.4(3)	10(5)	36(6)	**31**(2)	15/15
R-SHADE-10	**2.1**(2)	**2.1**(1)	**2.8**(2)	**3.8**(0.8)	∞ *1000*	0/15
R-SHADE-10	**1.0**(1)	**2.3**(1.0)	**2.9**(2)	6.3(1)	**7.2**(0.3)	15/15
SOO-Derbel	**0.84**(1)	**0.81**(0.6)	**1.3**(0.7)	6.8(2)	2.9e4(3e4)	1/15

Table 88: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f15	*2.5e+2*:9.0	*1.6e+2*:72	*1.0e+2*:186	*6.3e+1*:450	*4.0e+1*:872	15/15
MATSUMOTO-	**3.5**(1)	**0.98**(0.2)	**0.79**(0.3)	**1.0**(0.7)	**2.7**(2)	3/15
R-DE-10e2-	3.6(2)	**1.1**(0.9)	**1.3**(1)	**2.1**(3)	8.4(14)	2/15
R-DE-10e5-	5.8(5)	**1.8**(0.9)	**1.7**(0.5)	**2.4**(1)	5.2(4)	15/15
RL-SHADE-1	11(5)	**2.2**(0.5)	**1.6**(1.0)	**2.1**(2)	3.2(3)	5/15
RL-SHADE-1	11(10)	8.2(4)	7.2(3)	8.0(2)	14(5)	15/15
R-SHADE-10	4.8(5)	**1.8**(0.7)	**1.5**(1)	**1.7**(1)	5.6(4)	3/15
R-SHADE-10	6.0(6)	**1.8**(1)	**1.6**(0.9)	**2.2**(0.6)	3.2(1.0)	15/15
SOO-Derbel	**1.5**(2)	**0.88**(0.9)	**1.1**(1)	**1.7**(0.9)	**1.5**(0.3)	15/15

Table 89: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f16	*4.0e+1*:12	*2.5e+1*:47	*1.6e+1*:88	*1.0e+1*:425	*4.0e+0*:989	15/15
MATSUMOTO-	**1.9**(4)	**2.8**(1)	8.1(9)	4.3(3)	∞ *500*	0/15
R-DE-10e2-	**1.1**(0.8)	**2.0**(2)	10(8)	11(16)	∞ *1000*	0/15
R-DE-10e5-	**1.3**(2)	**1.6**(2)	8.1(10)	7.9(4)	27(22)	15/15
RL-SHADE-1	**1.1**(1.0)	**2.4**(2)	5.3(4)	**1.9**(2)	15(17)	1/15
RL-SHADE-1	**2.0**(4)	3.6(3)	21(20)	26(12)	34(7)	15/15
R-SHADE-10	**1.9**(1)	**1.9**(1)	**3.3**(4)	**2.5**(0.9)	∞ *1000*	0/15
R-SHADE-10	**1.4**(0.8)	3.2(6)	16(12)	7.5(5)	**8.7**(2)	15/15
SOO-Derbel	**1.4**(2)	**1.6**(2)	**4.0**(2)	**1.1**(0.3)	**1.1**(0.3)	15/15

Table 90: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f17	*1.0e+1*:26	*6.3e+0*:85	*4.0e+0*:155	*2.5e+0*:238	*6.3e-1*:585	15/15
MATSUMOTO-	**1.5**(0.9)	**1.0**(0.4)	**1.9**(2)	6.9(9)	∞ *500*	0/15
R-DE-10e2-	**1.5**(0.6)	**1.7**(0.4)	**2.6**(1)	4.7(4)	26(36)	1/15
R-DE-10e5-	**2.4**(2)	**2.4**(2)	**3.0**(1)	3.6(1)	5.4(3)	15/15
RL-SHADE-1	**2.5**(2)	**1.6**(1)	**1.9**(1.0)	4.3(4)	∞ *1000*	0/15
RL-SHADE-1	3.3(5)	3.9(3)	7.9(3)	11(5)	15(2)	15/15
R-SHADE-10	**2.9**(2)	**2.3**(1.0)	5.3(3)	7.2(6)	26(50)	1/15
R-SHADE-10	**2.0**(1)	**1.5**(0.3)	**1.9**(0.6)	**2.0**(0.4)	**2.2**(0.5)	15/15
SOO-Derbel	**0.69**(1)	**1.2**(1)	**1.4**(0.6)	**2.0**(1)	**4.4**(6)	15/15

Table 91: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f18	*4.0e+1*:11	*2.5e+1*:56	*1.6e+1*:172	*1.6e+1*:172	*2.5e+0*:561	15/15
MATSUMOTO-	**2.2**(3)	**1.3**(2)	**1.7**(2)	**1.7**(2)	∞ *500*	0/15
R-DE-10e2-	3.3(5)	**1.7**(2)	**1.4**(1)	**1.4**(0.5)	13(17)	2/15
R-DE-10e5-	4.6(6)	**2.7**(1)	**2.1**(1.0)	**2.1**(0.6)	10(6)	15/15
RL-SHADE-1	3.9(4)	**2.7**(0.6)	**2.1**(1)	**2.1**(1)	13(15)	2/15
RL-SHADE-1	5.2(4)	8.3(5)	7.0(3)	7.0(2)	15(2)	15/15
R-SHADE-10	11(25)	3.9(2)	**2.3**(0.7)	**2.3**(1)	26(10)	1/15
R-SHADE-10	3.3(3)	**2.1**(0.9)	**1.3**(0.5)	**1.3**(0.4)	**2.1**(0.4)	15/15
SOO-Derbel	**1.9**(3)	**1.3**(1)	**1.1**(0.7)	**1.1**(0.9)	**3.3**(2)	15/15

Table 92: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f19	*1.6e-1*:618	*1.0e-1*:10609	*6.3e-2*:10623	*4.0e-2*:10625	*2.5e-2*:10644	15/15
MATSUMOTO-	∞	∞	∞	∞	∞ *500*	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ *1000*	0/15
R-DE-10e5-	∞	∞	∞	∞	∞ *1e6*	0/15
RL-SHADE-1	∞	∞	∞	∞	∞ *1000*	0/15
RL-SHADE-1	**670**(396)	188(361)	667(1083)	∞	∞ *1e6*	0/15
R-SHADE-10	∞	∞	∞	∞	∞ *1000*	0/15
R-SHADE-10	965(855)	**115**(45)	**309**(378)	**1407**(1200)	∞ *1e6*	0/15
SOO-Derbel	**4.3**(1)	**0.30**(0.1)	**0.38**(0.1)	**7.2**(0.3)	**20**(36)	13/15

Table 93: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f20	*1.0e+4*:17	*6.3e+3*:21	*6.3e+1*:30	*2.5e+0*:122	*1.0e+0*:15426	13/15
MATSUMOTO-	**1.9**(1.0)	**2.0**(0.7)	**3.0**(1)	11(20)	∞ *500*	0/15
R-DE-10e2-	**2.1**(0.5)	**2.0**(0.9)	**4.5**(3)	**2.8**(0.9)	**0.96**(2)	1/15
R-DE-10e5-	3.4(2)	4.3(3)	8.6(1)	5.5(1)	**0.21**(0.1)	15/15
RL-SHADE-1	**2.7**(2)	4.3(3)	6.8(2)	**2.8**(0.6)	**0.22**(0.1)	4/15
RL-SHADE-1	5.8(5)	6.9(4)	40(10)	38(7)	**2.0**(0.4)	15/15
R-SHADE-10	**3.0**(1)	3.0(2)	6.1(2)	**2.7**(0.4)	**0.96**(0.4)	1/15
R-SHADE-10	3.0(1)	3.1(2)	6.8(3)	8.8(4)	**0.71**(0.5)	15/15
SOO-Derbel	**2.3**(0.0)	3.1(0.0)	54(0.0)	163(4e-3)	**2.3**(3e-5)	15/15

Table 94: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f21	*4.0e+1*:30	*2.5e+1*:46	*1.6e+1*:56	*1.0e+1*:130	*6.3e+0*:639	15/15
MATSUMOTO-	**1.9**(2)	**1.8**(0.9)	**2.9**(4)	**2.5**(2)	**0.59**(0.6)	11/15
R-DE-10e2-	4.0(4)	5.4(7)	6.2(2)	3.6(1)	**0.89**(1.0)	13/15
R-DE-10e5-	4.0(3)	5.6(4)	8.3(7)	5.1(3)	3.0(1)	15/15
RL-SHADE-1	4.4(2)	4.1(1)	4.8(1)	**2.8**(2)	**0.93**(0.2)	13/15
RL-SHADE-1	13(8)	19(12)	30(12)	22(12)	6.0(3)	15/15
R-SHADE-10	4.0(3)	4.2(2)	5.4(2)	3.3(2)	**1.4**(0.8)	12/15
R-SHADE-10	4.9(4)	6.3(5)	8.1(5)	4.6(3)	**2.5**(0.7)	15/15
SOO-Derbel	**1.8**(3)	**2.7**(2)	**3.2**(3)	**1.9**(1)	**0.76**(0.4)	15/15

Table 95: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f22	*6.3e+1*:18	*4.0e+1*:30	*4.0e+1*:30	*6.3e+0*:155	*4.0e+0*:631	14/15
MATSUMOTO-	**2.4**(1)	**2.3**(1)	**2.3**(1)	**2.3**(3)	**0.60**(0.7)	11/15
R-DE-10e2-	3.3(1)	4.0(2)	4.0(2)	10(5)	3.7(5)	5/15
R-DE-10e5-	6.4(4)	7.0(5)	7.0(3)	21(38)	22(58)	15/15
RL-SHADE-1	4.7(4)	5.7(4)	5.7(3)	8.1(4)	3.4(2)	6/15
RL-SHADE-1	11(8)	24(19)	24(9)	27(20)	50(4)	15/15
R-SHADE-10	4.0(4)	5.3(5)	5.3(4)	**4.2**(6)	**1.9**(0.9)	9/15
R-SHADE-10	6.7(2)	7.6(4)	7.6(3)	5.3(2)	5.3(9)	15/15
SOO-Derbel	**2.5**(2)	**2.5**(2)	**2.5**(0.9)	5.5(5)	16(2)	15/15

Table 96: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f23	*6.3e+0*:10	*4.0e+0*:62	*2.5e+0*:162	*2.5e+0*:162	*1.0e+0*:915	15/15
MATSUMOTO-	**2.1**(3)	**2.1**(2)	8.6(12)	8.6(12)	∞ *500*	0/15
R-DE-10e2-	**1.6**(2)	**2.6**(2)	8.9(15)	8.9(12)	∞ *1000*	0/15
R-DE-10e5-	**2.0**(3)	**1.9**(3)	8.3(8)	8.3(8)	204(224)	15/15
RL-SHADE-1	**2.0**(1)	**1.2**(1.0)	6.8(6)	6.8(17)	∞ *1000*	0/15
RL-SHADE-1	**2.1**(2)	3.0(4)	**6.8**(3)	**6.8**(7)	68(32)	15/15
R-SHADE-10	**1.6**(2)	**2.3**(2)	8.1(5)	8.1(9)	∞ *1000*	0/15
R-SHADE-10	**1.7**(0.9)	**2.1**(3)	7.8(8)	7.8(11)	**21**(11)	15/15
SOO-Derbel	**1.9**(3)	**2.5**(2)	**5.2**(2)	**5.2**(3)	**2.3**(1)	15/15

Table 97: 10-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f24	*1.0e+2*:66	*6.3e+1*:596	*4.0e+1*:3181	*2.5e+1*:7668	*1.6e+1*:14353	15/15
MATSUMOTO-	7.0(6)	∞	∞	∞	∞ *500*	0/15
R-DE-10e2-	**3.3**(1)	**2.5**(2)	∞	∞	∞ *1000*	0/15
R-DE-10e5-	4.4(2)	3.9(3)	10(7)	71(147)	475(719)	2/15
RL-SHADE-1	5.6(5)	**2.7**(1)	∞	∞	∞ *1000*	0/15
RL-SHADE-1	20(4)	10(6)	10(2)	7.4(1)	9.4(7)	15/15
R-SHADE-10	**4.2**(2)	7.9(5)	∞	∞	∞ *1000*	0/15
R-SHADE-10	4.7(2)	**2.4**(1)	**1.8**(1)	**1.7**(0.9)	**3.6**(2)	15/15
SOO-Derbel	11(8)	3.9(2)	**2.1**(1)	**4.2**(9)	**7.1**(7)	15/15

Table 98: 20-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	*6.3e+1*:24	*4.0e+1*:42	*1.0e-8*:43	*1.0e-8*:43	*1.0e-8*:43	15/15
MATSUMOTO-	**2.5**(0.3)	**2.0**(0.3)$^{\star 2}$	∞	∞	∞ *1000*	0/15
R-DE-10e2-	3.7(2)	**3.2**(1)	∞	∞	∞ *2000*	0/15
R-DE-10e5-	10(3)	10(1)	**230**(6)	**230**(5)	**230**(6)	15/15
RL-SHADE-1	11(0.5)	7.1(0.8)	∞	∞	∞ *2000*	0/15
RL-SHADE-1	53(17)	47(15)	1041(15)	1041(17)	1041(10)	15/15
R-SHADE-10	6.8(3)	5.5(1)	∞	∞	∞ *2000*	0/15
R-SHADE-10	12(4)	12(3)	**294**(18)	**294**(20)	**294**(32)	15/15
SOO-Derbel	**3.6**(2)	4.3(1)	1042(23)	1042(26)	1042(19)	15/15

Table 99: 20-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	*4.0e+6*:29	*2.5e+6*:42	*1.0e+5*:65	*1.0e+4*:207	*1.0e-8*:412	15/15
MATSUMOTO-	**0.79**(1)	**1.3**(1)	9.3(6)	71(145)	∞ *1000*	0/15
R-DE-10e2-	**0.87**(0.8)	**0.95**(0.3)	**3.8**(2)$^{\star 2}$	**2.6**(1)	∞ *2000*	0/15
R-DE-10e5-	**1.4**(2)	**2.0**(2)	12(2)	6.9(0.7)	**32**(0.9)	15/15
RL-SHADE-1	**1.3**(1)	**1.3**(2)	**6.1**(1)	4.2(2)	∞ *2000*	0/15
RL-SHADE-1	**1.5**(1)	3.0(1)	85(10)	61(4)	152(2)	15/15
R-SHADE-10	**1.9**(2)	**2.0**(1)	6.2(1)	**3.6**(0.3)	∞ *2000*	0/15
R-SHADE-10	**1.0**(1)	**2.1**(1)	16(3)	10(2)	**44**(2)	15/15
SOO-Derbel	8.8(3)	12(12)	43(9)	24(7)	3.3e4(2e4)	2/15

Table 100: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f3	*6.3e+2*:33	*4.0e+2*:44	*1.6e+2*:109	*1.0e+2*:255	*2.5e+1*:3277	15/15
MATSUMOTO-	**1.9**(0.6)	**2.5**(0.7)	8.0(5)	10(9)	∞ *1000*	0/15
R-DE-10e2-	**1.4**(0.9)	**2.5**(1)	5.4(9)	**4.3**(1)	9.1(13)	1/15
R-DE-10e5-	4.3(2)	6.8(2)	10(2)	8.7(3)	**2.8**(0.6)	15/15
RL-SHADE-1	5.2(3)	6.4(0.6)	**5.1**(3)	5.9(5)	8.9(10)	1/15
RL-SHADE-1	10(7)	36(8)	75(11)	91(22)	28(2)	15/15
R-SHADE-10	**2.8**(2)	4.0(1)	**5.2**(0.4)	**3.9**(0.7)	**3.0**(5)	3/15
R-SHADE-10	3.8(3)	7.7(2)	20(8)	33(6)	7.7(0.6)	15/15
SOO-Derbel	5.0(3)	5.8(2)	12(5)	12(3)	234(655)	12/15

Table 101: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f4	*6.3e+2*:22	*4.0e+2*:91	*2.5e+2*:250	*1.6e+2*:332	*6.3e+1*:1927	15/15
MATSUMOTO-	7.6(8)	5.1(4)	10(7)	∞	∞ *1000*	0/15
R-DE-10e2-	**4.4**(3)	**2.0**(0.4)	**1.5**(0.4)*	**2.2**(0.7)	∞ *2000*	0/15
R-DE-10e5-	13(4)	6.1(0.8)	4.0(0.4)	5.0(1)	**2.6**(0.3)	15/15
RL-SHADE-1	12(2)	3.6(0.5)	**1.9**(0.4)	4.2(3)	7.4(10)	2/15
RL-SHADE-1	40(19)	30(3)	25(3)	38(6)	29(3)	15/15
R-SHADE-10	7.9(2)	3.2(0.5)	**2.1**(0.3)	**2.4**(0.5)	**0.84**(0.1)	15/15
R-SHADE-10	11(4)	6.6(2)	5.5(0.8)	13(3)	9.0(1)	15/15
SOO-Derbel	**1.1**(2)*	**2.7**(0.8)	3.9(3)	7.0(1.0)	47(1)	15/15

Table 102: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f5	*2.5e+2*:19	*1.6e+2*:34	*1.0e-8*:41	*1.0e-8*:41	*1.0e-8*:41	15/15
MATSUMOTO-	**1.8**(0.6)	**1.3**(0.1)$^{\star 4}$	**2.4**(0.1)$^{\star 4}$	**2.4**(1)$^{\star 4}$	**2.4**(0.1)$^{\star 4}$	15/15
R-DE-10e2-	**2.2**(0.7)	**3.1**(0.8)	∞	∞	∞ *2000*	0/15
R-DE-10e5-	4.3(3)	10(2)	332(7)	332(5)	332(8)	15/15
RL-SHADE-1	6.3(5)	7.9(0.5)	**242**(173)	**242**(333)	**242**(209)	3/15
RL-SHADE-1	8.0(23)	57(13)	1574(45)	1574(34)	1574(44)	15/15
R-SHADE-10	3.5(3)	4.8(0.6)	∞	∞	∞ *2000*	0/15
R-SHADE-10	6.1(5)	19(4)	959(45)	959(33)	959(31)	15/15
SOO-Derbel	10(0.0)	8.7(0.0)	4928(0.0)	4928(0.0)	4928(0.0)	15/15

Table 103: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f6	*2.5e+5*:16	*6.3e+4*:43	*1.6e+4*:62	*1.6e+2*:353	*1.6e+1*:1078	15/15
MATSUMOTO-	**2.4**(0.2)	**1.4**(0.4)	**1.4**(0.7)*	9.5(21)	∞ *1000*	0/15
R-DE-10e2-	**2.9**(2)	**2.2**(1)	**2.4**(1)	**3.4**(2)	27(44)	1/15
R-DE-10e5-	6.8(7)	7.2(5)	8.8(2)	4.9(1.0)	**12**(3)	15/15
RL-SHADE-1	8.0(9)	6.3(1)	5.1(1)	3.9(2)	∞ *2000*	0/15
RL-SHADE-1	16(22)	32(19)	36(23)	19(5)	20(0.7)	15/15
R-SHADE-10	4.0(3)	4.0(2)	4.1(2)	**2.5**(1)	∞ *2000*	0/15
R-SHADE-10	7.7(10)	9.0(7)	11(6)	4.3(1)	**4.3**(0.4)	15/15
SOO-Derbel	**2.9**(3)	4.2(3)	5.3(2)	6.9(3)	1.3e4(2e4)	2/15

Table 104: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f7	*1.0e+3*:11	*4.0e+2*:39	*2.5e+2*:74	*6.3e+1*:319	*1.0e+1*:1351	15/15
MATSUMOTO-	**1.2**(2)	**2.2**(2)	**2.0**(0.5)	5.6(6)	∞ *1000*	0/15
R-DE-10e2-	**1.9**(1)	**2.1**(1)	**2.0**(2)	**3.0**(2)	**11**(6)	2/15
R-DE-10e5-	**1.9**(3)	4.9(3)	5.0(3)	5.0(2)	12(6)	15/15
RL-SHADE-1	**2.0**(4)	5.6(1)	4.1(0.3)	3.5(6)	∞ *2000*	0/15
RL-SHADE-1	**2.5**(2)	22(16)	25(3)	22(6)	11(0.5)	15/15
R-SHADE-10	**2.1**(2)	3.0(1)	3.2(2)	**2.2**(0.7)	11(7)	2/15
R-SHADE-10	3.4(5)	4.5(2)	5.2(1)	3.1(0.4)	**2.0**(0.7)	15/15
SOO-Derbel	**1.3**(0.5)	**2.0**(1)	**1.9**(1)	5.7(5)	59(85)	15/15

Table 105: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f8	*4.0e+4*:19	*2.5e+4*:35	*4.0e+3*:67	*2.5e+2*:231	*1.6e+1*:1470	15/15
MATSUMOTO-	**4.5**(1)	**2.9**(0.8)	**3.3**(1)	**4.0**(7)	∞ *1000*	0/15
R-DE-10e2-	5.1(1)	**3.6**(1)	**4.7**(1)	6.9(2)	∞ *2000*	0/15
R-DE-10e5-	15(5)	11(2)	12(3)	9.1(2)	51(5)	15/15
RL-SHADE-1	14(2)	8.2(0.9)	7.2(2)	22(20)	∞ *2000*	0/15
RL-SHADE-1	67(28)	50(14)	61(11)	48(4)	22(3)	15/15
R-SHADE-10	8.6(2)	5.4(1)	5.8(1)	**3.5**(1)	**20**(25)	1/15
R-SHADE-10	16(5)	9.3(3)	12(3)	8.3(1)	**7.0**(3)	15/15
SOO-Derbel	**3.8**(5)	4.5(5)	8.2(6)	12(4)	2107(3738)	6/15

Table 106: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f9	*1.0e+2*:357	*6.3e+1*:560	*4.0e+1*:684	*2.5e+1*:756	*1.0e+1*:1716	15/15
MATSUMOTO-	7.9(9)	∞	∞	∞	∞ *1000*	0/15
R-DE-10e2-	7.0(3)	12(15)	14(12)	19(19)	∞ *2000*	0/15
R-DE-10e5-	20(7)	18(6)	16(7)	16(3)	3022(3021)	5/15
RL-SHADE-1	20(48)	∞	∞	∞	∞ *2000*	0/15
RL-SHADE-1	37(2)	26(2)	24(1)	25(1)	**33**(2)	15/15
R-SHADE-10	**5.3**(3)	**4.2**(4)	**5.3**(4)	**6.1**(5)	∞ *2000*	0/15
R-SHADE-10	**5.7**(0.5)	**4.3**(0.3)	**4.1**(0.4)	**4.6**(0.3)	**16**(2)	15/15
SOO-Derbel	10(3)	11(3)	15(5)	21(6)	5000(4642)	3/15

Table 107: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f10	*1.6e+6*:15	*1.0e+6*:27	*4.0e+5*:70	*6.3e+4*:231	*4.0e+3*:1015	15/15
MATSUMOTO-	**4.4**(1)	**3.1**(1)	**3.4**(8)	9.1(8)	∞ *1000*	0/15
R-DE-10e2-	4.6(4)	4.2(0.8)	**4.1**(2)	12(13)	∞ *2000*	0/15
R-DE-10e5-	9.2(11)	10(6)	12(9)	70(21)	∞ *2e6*	0/15
RL-SHADE-1	12(4)	8.9(1)	5.9(3)	10(9)	∞ *2000*	0/15
RL-SHADE-1	32(15)	31(15)	38(12)	57(8)	**20**(2)	15/15
R-SHADE-10	6.9(3)	5.8(2)	5.8(3)	**5.6**(2)	29(27)	1/15
R-SHADE-10	10(6)	8.8(4)	8.0(2)	**6.4**(1.0)	**3.8**(2)	15/15
SOO-Derbel	**2.9**(4)	**3.1**(2)	4.8(4)	14(12)	1743(1348)	10/15

Table 108: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f11	*4.0e+4*:11	*2.5e+3*:27	*1.6e+2*:313	*1.0e+2*:481	*1.0e+1*:1002	15/15
MATSUMOTO-	**1.7**(0.7)	**2.5**(2)	21(17)	∞	∞ *1000*	0/15
R-DE-10e2-	**1.7**(1)	**2.0**(1)	13(10)	∞	∞ *2000*	0/15
R-DE-10e5-	**1.7**(1.0)	**2.2**(2)	11(11)	51(83)	∞ *2e6*	0/15
RL-SHADE-1	**2.1**(4)	3.8(2)	11(23)	60(52)	∞ *2000*	0/15
RL-SHADE-1	**1.9**(2)	3.4(2)	14(12)	**22**(17)	**27**(2)	15/15
R-SHADE-10	**1.6**(1)	**1.6**(0.8)	**10**(7)	61(31)	∞ *2000*	0/15
R-SHADE-10	**1.8**(1)	**2.6**(2)	**2.3**(1)	**2.9**(1)	**7.6**(6)	15/15
SOO-Derbel	**1.4**(1.0)	**2.0**(2)	11(0.9)	343(1430)	∞ *2e6*	0/15

Table 109: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f12	*1.0e+8*:23	*6.3e+7*:39	*2.5e+7*:76	*4.0e+6*:209	*1.0e+1*:1042	15/15
MATSUMOTO-	**3.1**(0.7)	**2.9**(0.4)	**3.0**(0.5)	3.6(3)	∞ *1000*	0/15
R-DE-10e2-	**3.1**(2)	**3.1**(2)	**3.3**(1)	**3.0**(2)	∞ *2000*	0/15
R-DE-10e5-	9.4(4)	8.1(2)	9.5(2)	7.1(0.5)	**31**(36)	15/15
RL-SHADE-1	11(1)	7.5(0.8)	5.0(0.4)	6.1(7)	∞ *2000*	0/15
RL-SHADE-1	37(19)	38(16)	44(7)	44(6)	37(9)	15/15
R-SHADE-10	7.8(2)	6.3(2)	5.0(1)	**3.2**(0.5)	∞ *2000*	0/15
R-SHADE-10	11(5)	9.1(3)	10(3)	8.1(1)	**8.5**(0.7)	15/15
SOO-Derbel	3.6(5)	5.0(4)	6.0(3)	11(9)	1110(972)	10/15

Table 110: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f13	*1.6e+3*:28	*1.0e+3*:64	*6.3e+2*:79	*4.0e+1*:211	*2.5e+0*:1724	15/15
MATSUMOTO-	**2.6**(0.6)	**1.8**(0.6)$^{\star 2}$	**2.6**(2)$^{\star 2}$	**8.0**(9)	∞ *1000*	0/15
R-DE-10e2-	**2.9**(1.0)	**3.6**(0.9)	**5.1**(1.0)	∞	∞ *2000*	0/15
R-DE-10e5-	7.1(2)	7.8(2)	12(2)	27(3)	43(46)	15/15
RL-SHADE-1	8.8(0.7)	5.2(0.2)	7.4(0.7)	∞	∞ *2000*	0/15
RL-SHADE-1	30(13)	41(5)	65(5)	93(2)	**18**(0.5)	15/15
R-SHADE-10	5.5(2)	4.8(1)	6.2(0.6)	**11**(3)	∞ *2000*	0/15
R-SHADE-10	8.5(3)	9.3(2)	13(1)	21(4)	**7.4**(5)	15/15
SOO-Derbel	**2.4**(3)	4.9(4)	11(4)	1339(1890)	5103(4830)	3/15

Table 111: 20-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f14	*2.5e+1*:15	*1.6e+1*:42	*1.0e+1*:75	*1.6e+0*:219	*6.3e-4*:1106	15/15
MATSUMOTO-	**5.7**(2)	**3.1**(1)	**2.9**(1)	9.5(11)	∞ *1000*	0/15
R-DE-10e2-	7.3(5)	**4.5**(1)	**4.5**(4)	**6.1**(7)	∞ *2000*	0/15
R-DE-10e5-	13(7)	10(4)	9.4(4)	9.3(2)	∞ *2e6*	0/15
RL-SHADE-1	17(6)	7.7(0.7)	5.8(1)	31(43)	∞ *2000*	0/15
RL-SHADE-1	58(27)	45(10)	40(15)	47(4)	**30**(1)	15/15
R-SHADE-10	12(5)	6.0(2)	4.8(1)	**3.7**(0.6)	∞ *2000*	0/15
R-SHADE-10	13(4)	8.4(1)	8.2(2)	9.0(2)	**10**(2)	15/15
SOO-Derbel	**3.9**(6)	4.7(3)	5.7(3)	34(8)	∞ *2e6*	0/15

Table 112: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f15	*6.3e+2*:15	*4.0e+2*:67	*2.5e+2*:292	*1.6e+2*:846	*1.0e+2*:1671	15/15
MATSUMOTO-	**3.5**(1)	**1.7**(0.4)	**1.1**(1.0)	**2.0**(2)	8.8(12)	1/15
R-DE-10e2-	3.7(2)	**2.3**(2)	**2.2**(3)	11(12)	∞ *2000*	0/15
R-DE-10e5-	7.2(3)	5.0(2)	3.0(0.5)	4.5(3)	32(44)	15/15
RL-SHADE-1	12(4)	4.4(0.4)	**1.9**(0.8)	3.2(1)	18(37)	1/15
RL-SHADE-1	16(25)	24(13)	13(3)	13(3)	23(10)	15/15
R-SHADE-10	6.7(4)	3.5(1)	**1.6**(0.6)	**1.3**(0.4)	**2.8**(6)	6/15
R-SHADE-10	7.1(5)	5.2(1)	**2.6**(0.5)	**2.9**(0.7)	6.3(2)	15/15
SOO-Derbel	**1.4**(3)*	**2.6**(3)	**2.0**(0.8)	3.6(2)	**3.2**(0.8)	15/15

Table 113: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f16	*4.0e+1*:26	*2.5e+1*:127	*1.6e+1*:540	*1.6e+1*:540	*1.0e+1*:1384	15/15
MATSUMOTO-	3.4(3)	4.9(4)	3.3(2)	3.3(2)	11(15)	1/15
R-DE-10e2-	**2.1**(4)	39(29)	∞	∞	∞ *2000*	0/15
R-DE-10e5-	3.4(3)	18(24)	50(30)	50(50)	236(346)	15/15
RL-SHADE-1	3.8(3)	3.8(1)	**2.8**(2)	**2.8**(2)	**6.8**(13)	3/15
RL-SHADE-1	3.8(5)	38(47)	92(45)	92(22)	73(22)	15/15
R-SHADE-10	**1.8**(1)	**3.4**(1)	3.2(1)	3.2(3)	22(20)	1/15
R-SHADE-10	4.1(4)	23(16)	23(7)	23(6)	27(21)	15/15
SOO-Derbel	**2.3**(3)	**3.6**(2)	**2.2**(1.0)	**2.2**(1)	**1.6**(0.4)	15/15

Table 114: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f17	*1.6e+1*:11	*1.0e+1*:63	*6.3e+0*:305	*4.0e+0*:468	*1.0e+0*:1030	15/15
MATSUMOTO-	3.3(2)	**2.2**(0.6)	**1.3**(0.5)	32(46)	∞ *1000*	0/15
R-DE-10e2-	**2.3**(2)	**2.1**(2)	**1.7**(0.8)	4.1(2)	∞ *2000*	0/15
R-DE-10e5-	4.4(4)	4.1(1.0)	**2.7**(1)	4.3(1)	**12**(6)	15/15
RL-SHADE-1	5.5(10)	3.9(2)	**2.2**(2)	11(6)	∞ *2000*	0/15
RL-SHADE-1	4.6(13)	13(8)	10(3)	13(2)	18(2)	15/15
R-SHADE-10	6.0(6)	3.7(1)	**1.7**(0.7)	**3.3**(2)	29(39)	1/15
R-SHADE-10	5.3(13)	3.7(2)	**1.9**(0.9)	**2.3**(0.9)	**3.4**(0.8)	15/15
SOO-Derbel	**0.99**(3)$^{\star 2}$	**1.3**(0.9)	**1.6**(1)	3.5(2)	16(6)	15/15

Table 115: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f18	*4.0e+1*:116	*2.5e+1*:252	*1.6e+1*:430	*1.0e+1*:621	*4.0e+0*:1090	15/15
MATSUMOTO-	**1.0**(0.4)	**2.4**(5)	8.3(12)	∞	∞ *1000*	0/15
R-DE-10e2-	**0.78**(0.6)	**1.2**(0.7)	**2.1**(1)	8.4(13)	∞ *2000*	0/15
R-DE-10e5-	**1.8**(0.7)	**2.4**(0.6)	4.1(2)	6.2(2)	28(44)	15/15
RL-SHADE-1	**2.0**(1)	**2.2**(3)	4.2(0.7)	11(13)	∞ *2000*	0/15
RL-SHADE-1	5.3(2)	10(4)	13(2)	17(2)	18(1)	15/15
R-SHADE-10	**1.4**(0.6)	**2.0**(1)	3.0(1)	5.6(7)	28(44)	1/15
R-SHADE-10	**1.8**(0.6)	**2.0**(0.8)	**2.7**(0.8)	**3.1**(0.7)	**3.4**(0.8)	15/15
SOO-Derbel	**0.54**(0.5)	**1.6**(2)	3.2(2)	**5.0**(2)	**11**(5)	15/15

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
MATSUMOTO-	∞	∞	∞	∞	∞ *1000*	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ *2000*	0/15
R-DE-10e5-	∞	∞	∞	∞	∞ *2e6*	0/15
RL-SHADE-1	∞	∞	∞	∞	∞ *2000*	0/15
RL-SHADE-1	115(99)	∞	∞	∞	∞ *2e6*	0/15
R-SHADE-10	∞	∞	∞	∞	∞ *2000*	0/15
R-SHADE-10	**55**(59)	∞	∞	∞	∞ *2e6*	0/15
SOO-Derbel	**0.55**(2)	**3.2**(3)	**26**(24)	**83**(95)	∞ *2e6*	0/15

Table 117: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f20	*1.6e+4*:38	*1.0e+4*:42	*2.5e+2*:62	*2.5e+0*:250	*1.6e+0*:2536	15/15
MATSUMOTO-	**2.2**(0.5)	**2.4**(0.9)	**4.6**(2)	**4.1**(5)	∞ *1000*	0/15
R-DE-10e2-	**1.9**(0.9)	**2.3**(1)	**4.9**(7)	**2.9**(0.6)	**1.0**(0.6)	9/15
R-DE-10e5-	6.4(4)	8.4(2)	15(2)	13(2)	4.5(2)	15/15
RL-SHADE-1	7.1(0.5)	6.9(1.0)	8.8(2)	7.4(6)	12(16)	1/15
RL-SHADE-1	22(11)	30(11)	68(15)	133(61)	35(5)	15/15
R-SHADE-10	4.4(1)	4.8(2)	7.7(0.5)	4.3(0.4)	**1.8**(2)	6/15
R-SHADE-10	5.0(2)	5.7(2)	10(3)	44(22)	13(3)	15/15
SOO-Derbel	4.8(0.0)	5.6(0.0)	51(8e-3)	13(2e-3)	3.6(2e-4)	15/15

Table 118: 20-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ 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 $\mathrm{ERT}_{\text{best 2009}}$ (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
f21	*6.3e+1*:36	*4.0e+1*:77	*4.0e+1*:77	*1.6e+1*:456	*4.0e+0*:1094	15/15
MATSUMOTO-	**2.7**(0.5)	**1.9**(1)*	**1.9**(0.4)*	**0.56**(0.5)	**0.97**(2)	9/15
R-DE-10e2-	**3.7**(2)	**3.1**(1)	**3.1**(0.9)	**1.2**(0.4)	**2.2**(3)	8/15
R-DE-10e5-	14(7)	12(5)	12(6)	7.3(3)	41(57)	15/15
RL-SHADE-1	9.1(2)	6.6(2)	6.6(3)	4.5(5)	6.1(2)	4/15
RL-SHADE-1	56(16)	52(14)	52(14)	15(3)	169(642)	14/15
R-SHADE-10	8.4(4)	7.0(2)	7.0(2)	3.5(3)	5.0(4)	5/15
R-SHADE-10	13(6)	11(3)	11(3)	**3.0**(0.9)	17(8)	15/15
SOO-Derbel	3.8(1)	4.6(5)	4.6(5)	**2.4**(2)	8.7(40)	15/15

Table 119: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f22	*6.3e+1*:45	*4.0e+1*:68	*4.0e+1*:68	*1.6e+1*:231	*6.3e+0*:1219	15/15
MATSUMOTO-	**2.2**(0.6)	**2.1**(0.6)$^{\star 2}$	**2.1**(2)$^{\star 2}$	**1.5**(2)*	**0.58**(0.7)*	11/15
R-DE-10e2-	**4.4**(2)	**4.7**(2)	**4.7**(2)	6.9(9)	5.2(3)	4/15
R-DE-10e5-	12(8)	13(5)	13(9)	81(66)	25(47)	15/15
RL-SHADE-1	7.7(2)	7.2(4)	7.2(2)	7.2(13)	**2.5**(2)	8/15
RL-SHADE-1	49(17)	59(12)	59(9)	493(1742)	221(2)	14/15
R-SHADE-10	8.1(5)	10(3)	10(3)	**5.1**(4)	**2.1**(2)	9/15
R-SHADE-10	12(7)	12(3)	12(10)	28(161)	14(29)	15/15
SOO-Derbel	4.7(2)	6.8(4)	6.8(5)	180(656)	69(247)	15/15

Table 120: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f23	*6.3e+0*:29	*4.0e+0*:118	*2.5e+0*:306	*2.5e+0*:306	*1.0e+0*:1614	15/15
MATSUMOTO-	**2.0**(2)	10(15)	∞	∞	∞ *1000*	0/15
R-DE-10e2-	**2.1**(2)	6.2(6)	95(126)	95(110)	∞ *2000*	0/15
R-DE-10e5-	**1.1**(2)	6.5(5)	81(17)	81(56)	∞ *2e6*	0/15
RL-SHADE-1	**2.7**(3)	5.4(5)	48(61)	48(39)	∞ *2000*	0/15
RL-SHADE-1	**1.0**(1)	**4.0**(3)	57(11)	57(79)	116(25)	15/15
R-SHADE-10	**2.4**(2)	6.7(6)	**45**(81)	**45**(46)	∞ *2000*	0/15
R-SHADE-10	**1.7**(1)	**4.8**(5)	68(100)	68(88)	**95**(97)	15/15
SOO-Derbel	**1.7**(4)	10(11)	**12**(10)	**12**(3)	**3.9**(2)	15/15

Table 121: 20-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f24	*2.5e+2*:208	*1.6e+2*:918	*1.0e+2*:6628	*6.3e+1*:9885	*4.0e+1*:31629	15/15
MATSUMOTO-	15(10)	∞	∞	∞	∞ *1000*	0/15
R-DE-10e2-	**2.6**(3)	15(13)	∞	∞	∞ *2000*	0/15
R-DE-10e5-	4.8(1)	16(5)	592(369)	∞	∞ *2e6*	0/15
RL-SHADE-1	3.0(1)	**5.7**(6)	∞	∞	∞ *2000*	0/15
RL-SHADE-1	16(5)	16(2)	15(2)	20(3)	**9.4**(1.0)	15/15
R-SHADE-10	3.0(0.7)	32(39)	∞	∞	∞ *2000*	0/15
R-SHADE-10	**2.7**(0.5)	**2.7**(2)	**3.2**(1)	**5.4**(0.6)	**6.1**(3)	15/15
SOO-Derbel	6.9(4)	7.3(5)	**5.6**(3)	**13**(17)	20(9)	15/15

References

[1] 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.

[2] Dimo Brockhoff. Comparison of the matsumoto library for expensive optimization on the noiseless black-box optimization benchmarking testbed. In *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2015, 25-28 May, Sendai, Japan*, 2015.

[3] Bilel Derbel and Philippe Preux. Simultaneous optimistic optimization on the noiseless bbob testbed. In *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2015, 25-28 May, Sendai, Japan*, 2015.

[4] 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.

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

[6] 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.

[7] Ryoji Tanabe and Alex Fukunaga. Parameter tuning for differential evolution for cheap, medium, and expensive computational budgets. In *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2015, 25-28 May, Sendai, Japan*, 2015.