

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 (ERT_{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 name	short	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.6e+1</i> :1.2	<i>4.0e+0</i> :2.6	<i>2.5e-2</i> :6.2	<i>1.0e-8</i> :6.2	<i>1.0e-8</i> :6.2	15/15
MATSUMOTO	1.7 (2)	1.8 (1)	2.8 (0.8) ⁺²	∞	∞ <i>100</i>	0/15
R-DE-10e2-	2.7 (1)	2.3 (3)	10(5)	∞	∞ <i>200</i>	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)	∞	∞ <i>200</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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.

4

#FEs/D	0.5	1.2	3	10	50	#succ
<i>f</i>₂	<i>1.0e+7</i> :1.4	<i>1.6e+6</i> :2.7	<i>1.0e+5</i> :6.1	<i>6.3e-1</i> :20	<i>1.0e-8</i> :30	15/15
MATSUMOTO	1.5 ^(0.7)	1.6 ^(0.5)	1.3 ⁽¹⁾	∞	∞ <i>100</i>	0/15
R-DE-10e2-	1.6 ⁽¹⁾	1.9 ⁽²⁾	2.2 ⁽²⁾	7.3 ⁽³⁾	∞ <i>200</i>	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⁾	∞ <i>200</i>	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⁾	∞ <i>200</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.0e+2</i> :1.4	<i>4.0e+1</i> :4.1	<i>2.5e+1</i> :6.6	<i>6.3e+0</i> :26	<i>2.5e+0</i> :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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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
f_4	<i>6.3e+1:2.4</i>	<i>4.0e+1:5.2</i>	<i>2.5e+1:8.5</i>	<i>1.0e+1:22</i>	<i>2.5e+0:120</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f5	<i>4.0e+1</i> :1.4	<i>1.6e+1</i> :3.5	<i>1.0e-8</i> :4.4	<i>1.0e-8</i> :4.4	<i>1.0e-8</i> :4.4	15/15
MATSUMOTO	1.5 (1)	1.2 (0.9)	2.5 (0.9) ^{*4}	2.5 (1) ^{*4}	2.5 (0.6) ^{*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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+4</i> :1.4	<i>1.0e+2</i> :2.8	<i>1.6e+1</i> :10	<i>1.0e+0</i> :23	<i>2.5e-6</i> :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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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
<i>f7</i>	<i>4.0e+2</i> :1.6	<i>1.0e+1</i> :3.2	<i>2.5e+0</i> :14	<i>1.6e+0</i> :21	<i>1.6e-2</i> :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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>2.5e+3:1.2</i>	<i>1.0e+2:3.2</i>	<i>6.3e+0:7.0</i>	<i>1.6e-1:27</i>	<i>1.6e-6:100</i>	15/15
MATSUMOTO	1.7 ⁽¹⁾	2.1 ⁽¹⁾	2.6 ⁽¹⁾	7.7 ⁽¹¹⁾	∞ <i>100</i>	0/15
R-DE-10e2-	1.5 ^(0.8)	4.2 ⁽⁴⁾	4.1 ⁽⁴⁾	16 ⁽⁵⁾	∞ <i>200</i>	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⁾	∞ <i>200</i>	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⁾	∞ <i>200</i>	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 11: 02-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+6:2.0</i>	<i>4.0e+5:3.2</i>	<i>6.3e+2:8.8</i>	<i>1.0e+1:30</i>	<i>2.5e-8:101</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.0e+7</i> :1.1	<i>1.6e+6</i> :3.2	<i>1.0e+4</i> :6.6	<i>4.0e+1</i> :23	<i>4.0e-8</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	<i>2.5e+8:1.3</i>	<i>6.3e+6:2.7</i>	<i>6.3e+5:6.3</i>	<i>4.0e+1:21</i>	<i>1.6e-3:101</i>	15/15
MATSUMOTO	1.1 ^(0.8)	0.93 ^(0.7)	0.77 ^(0.4)	4.6 ⁽³⁾	∞ <i>100</i>	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⁾	∞ <i>200</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+2</i> :1.6	<i>2.5e+2</i> :3.1	<i>6.3e+1</i> :8.7	<i>1.0e+1</i> :23	<i>4.0e-6</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+1:1.4</i>	<i>2.5e+0:4.2</i>	<i>1.0e+0:7.4</i>	<i>2.5e-2:21</i>	<i>1.0e-8:101</i>	15/15
MATSUMOTO	1.0 (0.7)	1.4 (1)	1.5 (1)	1.7 (0.5) ^{*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 $ERT/ERT_{\text{best } 2009}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	<i>1.6e+2:1.2</i>	<i>4.0e+1:4.7</i>	<i>2.5e+1:10</i>	<i>1.0e+1:37</i>	<i>2.5e+0:118</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+2</i> :1.1	<i>2.5e+1</i> :3.9	<i>1.6e+1</i> :6.5	<i>4.0e+0</i> :31	<i>2.5e-1</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	<i>4.0e+1:1.2</i>	<i>1.0e+1:2.7</i>	<i>4.0e+0:10</i>	<i>2.5e+0:28</i>	<i>1.6e-1:119</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+2:1.2</i>	<i>1.0e+2:3.2</i>	<i>4.0e+1:7.2</i>	<i>6.3e+0:32</i>	<i>1.6e+0:104</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.6e-1:23</i>	<i>1.0e-1:26</i>	<i>6.3e-2:38</i>	<i>4.0e-2:40</i>	<i>1.0e-2:216</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_{20}	<i>4.0e+3</i> :1.9	<i>2.5e+2</i> :2.8	<i>4.0e+0</i> :6.3	<i>2.5e+0</i> :21	<i>6.3e-1</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f21</i>	<i>1.0e+1:1.7</i>	<i>6.3e+0:2.6</i>	<i>2.5e+0:7.9</i>	<i>1.6e+0:30</i>	<i>4.0e-1:105</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	<i>4.0e+1:1.3</i>	<i>1.6e+1:3.2</i>	<i>6.3e+0:9.3</i>	<i>1.6e+0:25</i>	<i>1.0e-1:168</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f23</i>	<i>4.0e+1:1.5</i>	<i>2.5e+1:2.6</i>	<i>1.0e+1:7.8</i>	<i>4.0e+0:55</i>	<i>2.5e+0:103</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f</i>24	<i>4.0e+1</i> :1.1	<i>2.5e+1</i> :2.7	<i>1.6e+1</i> :7.7	<i>6.3e+0</i> :44	<i>2.5e+0</i> :275	5/5
MATSUMOTO	1.6 ^(0.9)	1.7 ⁽²⁾	1.1 ^(0.8)	1.2 ⁽²⁾	∞ <i>100</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+1:3.0</i>	<i>1.0e+1:3.6</i>	<i>1.0e-8:8.0</i>	<i>1.0e-8:8.0</i>	<i>1.0e-8:8.0</i>	15/15
MATSUMOTO	1.6 (2)	1.9 (1)	∞	∞	∞ <i>150</i>	0/15
R-DE-10e2-	1.6 (0.9)	2.6 (3)	∞	∞	∞ <i>300</i>	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)	∞	∞	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_2	<i>6.3e+6</i> :1.5	<i>6.3e+5</i> :4.3	<i>4.0e+4</i> :10	<i>1.0e+2</i> :32	<i>1.0e-8</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+2:2.2</i>	<i>6.3e+1:6.1</i>	<i>4.0e+1:10</i>	<i>1.6e+1:32</i>	<i>4.0e+0:319</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+2:5.4</i>	<i>6.3e+1:10</i>	<i>6.3e+1:10</i>	<i>2.5e+1:36</i>	<i>4.0e+0:617</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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 $\text{ERT}_{\text{best } 2009}$ (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
f5	<i>4.0e+1</i> :2.2	<i>2.5e+1</i> :4.8	<i>1.0e-8</i> :6.6	<i>1.0e-8</i> :6.6	<i>1.0e-8</i> :6.6	15/15
MATSUMOTO	1.5 (1)	1.0 (0.8)	1.9 (0.6) ^{*4}	1.9 (0.5) ^{*4}	1.9 (0.5) ^{*4}	15/15
R-DE-10e2-	2.2 (1)	3.0 (3)	∞	∞	∞ <i>300</i>	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)	∞	∞	∞ <i>300</i>	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)	∞	∞	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>6.3e+4</i> :1.8	<i>6.3e+3</i> :3.7	<i>4.0e+1</i> :13	<i>1.0e+1</i> :34	<i>6.3e-4</i> :159	15/15
MATSUMOTO	1.2 (2)	1.1 (1)	3.2(3)	6.3(5)	∞ <i>150</i>	0/15
R-DE-10e2-	2.9 (3)	2.1 (3)	3.2(1)	1.9 (3)	∞ <i>300</i>	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)	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_7	<i>2.5e+2</i> :1.5	<i>6.3e+1</i> :4.2	<i>1.0e+1</i> :11	<i>2.5e+0</i> :38	<i>4.0e-1</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	<i>1.0e+4</i> :1.8	<i>1.6e+3</i> :4.0	<i>1.0e+2</i> :15	<i>6.3e+0</i> :31	<i>1.0e-1</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*.

Table 35: 03-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>6.3e+6:1.7</i>	<i>1.6e+5:4.4</i>	<i>4.0e+4:12</i>	<i>4.0e+2:37</i>	<i>1.0e+0:152</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>2.5e+6</i> :1.9	<i>4.0e+5</i> :4.5	<i>6.3e+4</i> :9.4	<i>2.5e+1</i> :36	<i>2.5e-1</i> :174	15/15
MATSUMOTO	1.6 (2)	1.3 (1)	1.4 (0.8)	7.8(12)	∞ <i>150</i>	0/15
R-DE-10e2-	1.2 (1)	1.3 (1)	1.4 (1)	13(16)	∞ <i>300</i>	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)	∞ <i>300</i>	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)	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+8</i> :1.5	<i>1.0e+7</i> :3.6	<i>6.3e+5</i> :13	<i>6.3e+2</i> :31	<i>1.0e+0</i> :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)	∞ <i>300</i>	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⁾	∞ <i>300</i>	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⁾	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+3</i> :1.6	<i>4.0e+2</i> :6.8	<i>2.5e+2</i> :11	<i>4.0e+1</i> :30	<i>2.5e-3</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+1:2.2</i>	<i>6.3e+0:4.2</i>	<i>2.5e+0:10</i>	<i>6.3e-2:31</i>	<i>2.5e-6:160</i>	15/15
MATSUMOTO	1.5 (1)	1.5 (2)	1.6 (1)	2.8 (4)	∞ <i>150</i>	0/15
R-DE-10e2-	1.8 (1)	1.7 (1)	2.2 (1)	3.7(1)	∞ <i>300</i>	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)	∞ <i>300</i>	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)	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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.

Table 41: 03-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>6.3e+1:1.5</i>	<i>2.5e+1:8.2</i>	<i>1.6e+1:10</i>	<i>1.0e+1:41</i>	<i>2.5e+0:208</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	<i>1.6e+1:1.8</i>	<i>1.0e+1:3.6</i>	<i>6.3e+0:14</i>	<i>2.5e+0:34</i>	<i>2.5e-1:189</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	<i>6.3e+1:1.8</i>	<i>4.0e+1:4.8</i>	<i>2.5e+1:13</i>	<i>1.0e+1:40</i>	<i>6.3e-1:184</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e-1:81</i>	<i>1.0e-1:109</i>	<i>6.3e-2:109</i>	<i>4.0e-2:119</i>	<i>1.6e-2:1230</i>	15/15
MATSUMOTO	∞	∞	∞	∞	∞ <i>150</i>	0/15
R-DE-10e2-	25(26)	39(75)	∞	∞	∞ <i>300</i>	0/15
R-DE-10e5-	20 (25)	31(19)	54(67)	103(190)	20(14)	15/15
RL-SHADE-1	27(20)	41(49)	∞	∞	∞ <i>300</i>	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)	∞ <i>300</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+3</i> :3.5	<i>2.5e+3</i> :4.3	<i>4.0e+0</i> :13	<i>1.6e+0</i> :41	<i>1.0e+0</i> :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)*2	15/15

Table 46: 03-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f21</i>	<i>1.6e+1:2.5</i>	<i>1.0e+1:5.9</i>	<i>6.3e+0:14</i>	<i>2.5e+0:41</i>	<i>1.6e+0:167</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>4.0e+1:2.9</i>	<i>2.5e+1:5.2</i>	<i>1.0e+1:18</i>	<i>6.3e+0:33</i>	<i>1.0e+0:170</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+1:2.6</i>	<i>6.3e+0:16</i>	<i>4.0e+0:44</i>	<i>2.5e+0:79</i>	<i>1.6e+0:198</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f</i>24	<i>4.0e+1:4.6</i>	<i>2.5e+1:13</i>	<i>1.6e+1:47</i>	<i>1.6e+1:47</i>	<i>6.9e+0:382</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>2.5e+1:4.8</i>	<i>1.6e+1:7.6</i>	<i>1.0e-8:12</i>	<i>1.0e-8:12</i>	<i>1.0e-8:12</i>	15/15
MATSUMOTO	1.8 ⁽¹⁾	1.7 ^(0.8)	∞	∞	∞ <i>250</i>	0/15
R-DE-10e2-	2.5 ⁽⁴⁾	2.6 ⁽²⁾	∞	∞	∞ <i>500</i>	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⁾	∞	∞	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+6:2.9</i>	<i>4.0e+5:11</i>	<i>4.0e+4:15</i>	<i>6.3e+2:58</i>	<i>1.0e-8:95</i>	15/15
MATSUMOTO	1.6 (1.0)	0.68 (0.4)	4.0 (5)	∞	∞ <i>250</i>	0/15
R-DE-10e2-	2.0 (4)	1.2 (1)	4.0 (2)	2.8 (0.5)	∞ <i>500</i>	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)	∞ <i>500</i>	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)	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_3	<i>1.6e+2</i> :4.1	<i>1.0e+2</i> :15	<i>6.3e+1</i> :23	<i>2.5e+1</i> :73	<i>1.0e+1</i> :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/ERT_{best} 2009 on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function

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

#FEs/D	0.5	1.2	3	10	50	#succ
f5	<i>6.3e+1</i> :4.0	<i>4.0e+1</i> :10	<i>1.0e-8</i> :10	<i>1.0e-8</i> :10	<i>1.0e-8</i> :10	15/15
MATSUMOTO	1.6 (0.8)	1.2 (0.6)	1.9 (0.3) ^{*4}	1.9 (0.3) ^{*4}	1.9 (0.3) ^{*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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.0e+5</i> :3.0	<i>2.5e+4</i> :8.4	<i>1.0e+2</i> :16	<i>2.5e+1</i> :54	<i>2.5e-1</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function

Table 57: 05-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+4</i> :4.6	<i>6.3e+3</i> :6.8	<i>1.0e+3</i> :18	<i>6.3e+1</i> :54	<i>1.6e+0</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f</i>₉	<i>2.5e+1:20</i>	<i>1.6e+1:26</i>	<i>1.0e+1:35</i>	<i>4.0e+0:62</i>	<i>1.6e-2:256</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>2.5e+6:2.9</i>	<i>6.3e+5:7.0</i>	<i>2.5e+5:17</i>	<i>6.3e+3:54</i>	<i>2.5e+1:297</i>	15/15
MATSUMOTO	1.5 (2)	1.9 (1)	1.4 (0.9)	6.4(9)	∞ <i>250</i>	0/15
R-DE-10e2-	1.4 (2)	1.2 (0.7)	1.2 (1)	2.6 (0.7)	∞ <i>500</i>	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)	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+6:3.0</i>	<i>6.3e+4:6.2</i>	<i>6.3e+2:16</i>	<i>6.3e+1:74</i>	<i>6.3e-1:298</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+7:3.6</i>	<i>1.6e+7:7.6</i>	<i>4.0e+6:19</i>	<i>1.6e+4:52</i>	<i>1.0e+0:268</i>	15/15
MATSUMOTO	1.2 ^(0.6)	1.4 ^(1.0)	1.7 ^(0.5)	3.2 ⁽¹⁾	∞ <i>250</i>	0/15
R-DE-10e2-	1.8 ⁽²⁾	2.6 ⁽⁴⁾	3.1 ⁽¹⁾	4.5 ⁽²⁾	∞ <i>500</i>	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)	∞ <i>500</i>	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⁾	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+3:2.8</i>	<i>6.3e+2:8.4</i>	<i>4.0e+2:17</i>	<i>6.3e+1:52</i>	<i>6.3e-2:264</i>	15/15
MATSUMOTO	2.0 _(1.0)	1.8 ₍₁₎	1.5 _(0.5)	1.7 _(0.3) ^{*2}	∞ <i>250</i>	0/15
R-DE-10e2-	2.5 ₍₂₎	2.7 _(0.7)	2.4 ₍₂₎	4.5 ₍₁₎	∞ <i>500</i>	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)	∞ <i>500</i>	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₎	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+1:3.0</i>	<i>1.0e+1:10</i>	<i>6.3e+0:15</i>	<i>2.5e-1:53</i>	<i>1.0e-5:251</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+2:3.0</i>	<i>1.0e+2:13</i>	<i>6.3e+1:24</i>	<i>4.0e+1:55</i>	<i>1.6e+1:289</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+1:4.8</i>	<i>2.5e+1:16</i>	<i>1.6e+1:46</i>	<i>1.0e+1:120</i>	<i>4.0e+0:334</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.0e+1:5.2</i>	<i>6.3e+0:26</i>	<i>4.0e+0:57</i>	<i>2.5e+0:110</i>	<i>6.3e-1:412</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>6.3e+1:3.4</i>	<i>4.0e+1:7.2</i>	<i>2.5e+1:20</i>	<i>1.6e+1:58</i>	<i>1.6e+0:318</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e-1:172</i>	<i>1.0e-1:242</i>	<i>6.3e-2:675</i>	<i>4.0e-2:3078</i>	<i>2.5e-2:4946</i>	15/15
MATSUMOTO	∞	∞	∞	∞	∞ <i>250</i>	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ <i>500</i>	0/15
R-DE-10e5-	1109(1356)	1951(2356)	3319(4999)	1163(1634)	1470(1238)	1/15
RL-SHADE-1	∞	∞	∞	∞	∞ <i>500</i>	0/15
RL-SHADE-1	273(149)	269(150)	134(116)	70(52)	116(136)	9/15
R-SHADE-10	∞	∞	∞	∞	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_{20}	<i>6.3e+3:5.1</i>	<i>4.0e+3:8.4</i>	<i>4.0e+1:15</i>	<i>2.5e+0:69</i>	<i>1.0e+0:851</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<i>f</i>21	<i>4.0e+1:3.9</i>	<i>2.5e+1:11</i>	<i>1.6e+1:31</i>	<i>6.3e+0:73</i>	<i>1.6e+0:347</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+1</i> :3.6	<i>4.0e+1</i> :15	<i>2.5e+1</i> :32	<i>1.0e+1</i> :71	<i>1.6e+0</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f23</i>	<i>1.0e+1:3.0</i>	<i>6.3e+0:9.0</i>	<i>4.0e+0:33</i>	<i>2.5e+0:84</i>	<i>1.0e+0:518</i>	15/15
MATSUMOTO	1.8 (2)	1.7 (2)	2.4 (3)	3.9 (6)	∞ <i>250</i>	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)	∞ <i>500</i>	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)	∞ <i>500</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+1:15</i>	<i>4.0e+1:37</i>	<i>4.0e+1:37</i>	<i>2.5e+1:118</i>	<i>1.6e+1:692</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+1:8.0</i>	<i>2.5e+1:16</i>	<i>1.0e-8:23</i>	<i>1.0e-8:23</i>	<i>1.0e-8:23</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f_2	<i>2.5e+6:5.6</i>	<i>1.0e+6:17</i>	<i>1.0e+5:33</i>	<i>2.5e+3:118</i>	<i>1.0e-8:196</i>	15/15
MATSUMOTO-	2.0 (2)	1.3 (0.6)	4.6(4)	∞	∞ <i>500</i>	0/15
R-DE-10e2-	2.0 (0.9)	1.7 (1)	2.9 (2)	2.1 (0.8)	∞ <i>1000</i>	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)	∞ <i>1000</i>	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)	∞ <i>1000</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>4.0e+2:8.2</i>	<i>1.6e+2:37</i>	<i>1.0e+2:69</i>	<i>6.3e+1:147</i>	<i>2.5e+1:1129</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	<i>2.5e+2:21</i>	<i>1.6e+2:59</i>	<i>1.6e+2:59</i>	<i>6.3e+1:139</i>	<i>4.0e+1:854</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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 $\text{ERT}_{\text{best } 2009}$ (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
f5	<i>1.0e+2:16</i>	<i>6.3e+1:19</i>	<i>1.0e-8:20</i>	<i>1.0e-8:20</i>	<i>1.0e-8:20</i>	15/15
MATSUMOTO	1.2 _(0.5)	1.3 _(0.0) ^{*2}	2.0 _(0.2) ^{*4}	2.0 _(0.2) ^{*4}	2.0 _(0.2) ^{*4}	15/15
R-DE-10e2-	2.1 _(0.9)	4.1 ₍₁₎	∞	∞	∞ <i>1000</i>	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₎	∞	∞	∞ <i>1000</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>1.6e+5</i> :7.0	<i>6.3e+4</i> :16	<i>4.0e+2</i> :36	<i>1.0e+2</i> :102	<i>4.0e+0</i> :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 $ERT/ERT_{\text{best } 2009}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f_7	<i>2.5e+2:9.2</i>	<i>1.6e+2:18</i>	<i>1.0e+2:33</i>	<i>1.0e+1:172</i>	<i>4.0e+0:678</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.6e+4:15</i>	<i>1.0e+4:22</i>	<i>1.6e+3:34</i>	<i>2.5e+2:103</i>	<i>4.0e+0:727</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_9	<i>4.0e+1:125</i>	<i>2.5e+1:148</i>	<i>1.6e+1:180</i>	<i>1.0e+1:200</i>	<i>1.6e+0:563</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>2.5e+6:6.0</i>	<i>1.0e+6:21</i>	<i>4.0e+5:38</i>	<i>2.5e+4:104</i>	<i>6.3e+2:512</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+4</i> :6.4	<i>2.5e+3</i> :15	<i>6.3e+1</i> :217	<i>4.0e+1</i> :244	<i>2.5e+0</i> :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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+7:15</i>	<i>2.5e+7:24</i>	<i>1.6e+7:34</i>	<i>1.0e+6:103</i>	<i>1.0e+1:515</i>	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

Table 86: 10-D, running time excess $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.0e+3:12</i>	<i>6.3e+2:32</i>	<i>4.0e+2:40</i>	<i>6.3e+1:154</i>	<i>2.5e+0:521</i>	15/15
MATSUMOTO-	2.3 (1)	1.5 (0.5)	2.0 (0.4)	2.3 (3) ⁺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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+1:7.7</i>	<i>1.6e+1:27</i>	<i>1.0e+1:37</i>	<i>6.3e-1:107</i>	<i>1.0e-4:505</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>2.5e+2:9.0</i>	<i>1.6e+2:72</i>	<i>1.0e+2:186</i>	<i>6.3e+1:450</i>	<i>4.0e+1:872</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>4.0e+1:12</i>	<i>2.5e+1:47</i>	<i>1.6e+1:88</i>	<i>1.0e+1:425</i>	<i>4.0e+0:989</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	<i>1.0e+1:26</i>	<i>6.3e+0:85</i>	<i>4.0e+0:155</i>	<i>2.5e+0:238</i>	<i>6.3e-1:585</i>	15/15
MATSUMOTO	1.5 ^(0.9)	1.0 ^(0.4)	1.9 ⁽²⁾	6.9 ⁽⁹⁾	<i>∞ 500</i>	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⁾	<i>∞ 13533 . -0.24256.7385</i>	<i>27.381 - -0.58124 .0224 334.26 5660.0</i>

10 0 0 10 0 0

Table 91: 10-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+1:11</i>	<i>2.5e+1:56</i>	<i>1.6e+1:172</i>	<i>1.6e+1:172</i>	<i>2.5e+0:561</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.6e-1:618</i>	<i>1.0e-1:10609</i>	<i>6.3e-2:10623</i>	<i>4.0e-2:10625</i>	<i>2.5e-2:10644</i>	15/15
MATSUMOTO	∞	∞	∞	∞	∞ <i>500</i>	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ <i>1000</i>	0/15
R-DE-10e5-	∞	∞	∞	∞	∞ <i>1e6</i>	0/15
RL-SHADE-1	∞	∞	∞	∞	∞ <i>1000</i>	0/15
RL-SHADE-1	670 ₍₃₉₆₎	188 ₍₃₆₁₎	667 ₍₁₀₈₃₎	∞	∞ <i>1e6</i>	0/15
R-SHADE-10	∞	∞	∞	∞	∞ <i>1000</i>	0/15
R-SHADE-10	965 ₍₈₅₅₎	115 ₍₄₅₎	309 ₍₃₇₈₎	1407 ₍₁₂₀₀₎	∞ <i>1e6</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries w

Table 94: 10-D, running time excess $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
f_{21}	<i>4.0e+1:30</i>	<i>2.5e+1:46</i>	<i>1.6e+1:56</i>	<i>1.0e+1:130</i>	<i>6.3e+0:639</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>6.3e+1:18</i>	<i>4.0e+1:30</i>	<i>4.0e+1:30</i>	<i>6.3e+0:155</i>	<i>4.0e+0:631</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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
<i>f23</i>	<i>6.3e+0:10</i>	<i>4.0e+0:62</i>	<i>2.5e+0:162</i>	<i>2.5e+0:162</i>	<i>1.0e+0:915</i>	15/15
MATSUMOTO	2.1 (3)	2.1 (2)	8.6(12)	8.6(12)	∞ <i>500</i>	0/15
R-DE-10e2-	1.6 (2)	2.6 (2)	8.9(15)	8.9(12)	∞ <i>1000</i>	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)	∞ <i>1000</i>	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)	∞ <i>1000</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.0e+2:66</i>	<i>6.3e+1:596</i>	<i>4.0e+1:3181</i>	<i>2.5e+1:7668</i>	<i>1.6e+1:14353</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+1:24</i>	<i>4.0e+1:42</i>	<i>1.0e-8:43</i>	<i>1.0e-8:43</i>	<i>1.0e-8:43</i>	15/15
MATSUMOTO-	2.5 _(0.3)	2.0 _{(0.3)⁺²}	∞	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	3.7 ₍₂₎	3.2 ₍₁₎	∞	∞	∞ <i>2000</i>	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)	∞	∞	∞ <i>2000</i>	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₎	∞	∞	∞ <i>2000</i>	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 $\text{ERT}/\text{ERT}_{\text{best } 2009}$ 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	<i>4.0e+6:29</i>	<i>2.5e+6:42</i>	<i>1.0e+5:65</i>	<i>1.0e+4:207</i>	<i>1.0e-8:412</i>	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₎ *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 $ERT/ERT_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<i>f3</i>	<i>6.3e+2:33</i>	<i>4.0e+2:44</i>	<i>1.6e+2:109</i>	<i>1.0e+2:255</i>	<i>2.5e+1:3277</i>	15/15
MATSUMOTO-	1.9 _(0.6)	2.5 _(0.7)	8.0 ₍₅₎	10 ₍₉₎	∞ <i>1000</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+2:22</i>	<i>4.0e+2:91</i>	<i>2.5e+2:250</i>	<i>1.6e+2:332</i>	<i>6.3e+1:1927</i>	15/15
MATSUMOTO	7.6(8)	5.1(4)	10(∞)			

Table 102: 20-D, running time excess $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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 $\text{ERT}_{\text{best 2009}}$ (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
f5	<i>2.5e+2</i> :19	<i>1.6e+2</i> :34	<i>1.0e-8</i> :41	<i>1.0e-8</i> :41	<i>1.0e-8</i> :41	15/15
MATSUMOTO-	1.8 (0.6)	1.3 (0.1) ^{*4}	2.4 (0.1) ^{*4}	2.4 (1) ^{*4}	2.4 (0.1) ^{*4}	15/15
R-DE-10e2-	2.2 (0.7)	3.1 (0.8)	∞	∞	∞ <i>2000</i>	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)	∞	∞	∞ <i>2000</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>2.5e+5:16</i>	<i>6.3e+4:43</i>	<i>1.6e+4:62</i>	<i>1.6e+2:353</i>	<i>1.6e+1:1078</i>	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 $ERT/ERT_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<i>f7</i>	<i>1.0e+3:11</i>	<i>4.0e+2:39</i>	<i>2.5e+2:74</i>	<i>6.3e+1:319</i>	<i>1.0e+1:1351</i>	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(1.6)	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 $ERT/ERT_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	<i>4.0e+4:19</i>	<i>2.5e+4:35</i>	<i>4.0e+3:67</i>	<i>2.5e+2:231</i>	<i>1.6e+1:1470</i>	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 $ERT/ERT_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	<i>1.0e+2:357</i>	<i>6.3e+1:560</i>	<i>4.0e+1:684</i>	<i>2.5e+1:756</i>	<i>1.0e+1:1716</i>	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 $ERT/ERT_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	<i>1.6e+6:15</i>	<i>1.0e+6:27</i>	<i>4.0e+5:70</i>	<i>6.3e+4:231</i>	<i>4.0e+3:1015</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>4.0e+4</i> :11	<i>2.5e+3</i> :27	<i>1.6e+2</i> :313	<i>1.0e+2</i> :481	<i>1.0e+1</i> :1002	15/15
MATSUMOTO	1.7 ^(0.7)	2.5 ⁽²⁾	21 ⁽¹⁷⁾	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	1.7 ⁽¹⁾	2.0 ⁽¹⁾	13 ⁽¹⁰⁾	∞	∞ <i>2000</i>	0/15
R-DE-10e5-	1.7 ^(1.0)	2.2 ⁽²⁾	11 ⁽¹¹⁾	51 ⁽⁸³⁾	∞ <i>2e6</i>	0/15
RL-SHADE-1	2.1 ⁽⁴⁾	3.8 ⁽²⁾	11 ⁽²³⁾	60 ⁽⁵²⁾	∞ <i>2000</i>	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⁾	∞ <i>2000</i>	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⁾	∞ <i>2e6</i>	0/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f12	<i>1.0e+8:23</i>	<i>6.3e+7:39</i>	<i>2.5e+7:76</i>	<i>4.0e+6:209</i>	<i>1.0e+1:1042</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>1.6e+3:28</i>	<i>1.0e+3:64</i>	<i>6.3e+2:79</i>	<i>4.0e+1:211</i>	<i>2.5e+0:1724</i>	15/15
MATSUMOTO-	2.6 _(0.6)	1.8 _(0.6) ^{*2}	2.6 ₍₂₎ ^{*2}	8.0 ₍₉₎	∞ <i>1000</i>	0/15
R-DE-10e2-	2.9 _(1.0)	3.6 _(0.9)	5.1 _(1.0)	∞	∞ <i>2000</i>	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)	∞	∞ <i>2000</i>	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₎	∞ <i>2000</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>2.5e+1:15</i>	<i>1.6e+1:42</i>	<i>1.0e+1:75</i>	<i>1.6e+0:219</i>	<i>6.3e-4:1106</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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 |

Table 113: 20-D, running time excess $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>4.0e+1:26</i>	<i>2.5e+1:127</i>	<i>1.6e+1:540</i>	<i>1.6e+1:540</i>	<i>1.0e+1:1384</i>	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)	∞	∞	∞ <i>2000</i>	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 $ERT/ERT_{\text{best } 2009}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best } 2009}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<i>f17</i>	<i>1.6e+1:11</i>	<i>1.0e+1:63</i>	<i>6.3e+0:305</i>	<i>4.0e+0:468</i>	<i>1.0e+0:1030</i>	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)*2	1.3 (0.9)	1.6 (1)	3.5(2)	16(6)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f18	<i>4.0e+1:116</i>	<i>2.5e+1:252</i>	<i>1.6e+1:430</i>	<i>1.0e+1:621</i>	<i>4.0e+0:1090</i>	15/15
MATSUMOTO	1.0 ^(0.4)	2.4 ⁽⁵⁾	8.3 ⁽¹²⁾	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	0.78 ^(0.6)	1.2 ^(0.7)	2.1 ⁽¹⁾	8.4 ⁽¹³⁾	∞ <i>2000</i>	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⁾	∞ <i>2000</i>	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 $ERT/ERT_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	<i>1.6e-1:2.5e5</i>	<i>1.0e-1:3.4e5</i>	<i>6.3e-2:3.4e5</i>	<i>4.0e-2:3.4e5</i>	<i>2.5e-2:3.4e5</i>	3/15
MATSUMOTO	∞	∞	∞	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ <i>2000</i>	0/15
R-DE-10e5-	∞	∞	∞	∞	∞ <i>2e6</i>	0/15
RL-SHADE-1	∞	∞	∞	∞	∞ <i>2000</i>	0/15
RL-SHADE-1	115 ⁽⁹⁹⁾	∞	∞	∞	∞ <i>2e6</i>	0/15
R-SHADE-10	∞	∞	∞	∞	∞ <i>2000</i>	0/15
R-SHADE-10	55 ⁽⁵⁹⁾	∞	∞	∞	∞ <i>2e6</i>	0/15
SOO-Derbel	0.55 ⁽²⁾	3.2 ⁽³⁾	26 ⁽²⁴⁾	83 ⁽⁹⁵⁾	∞ <i>2e6</i>	0/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f_{20}	<i>1.6e+4</i> :38	<i>1.0e+4</i> :42	<i>2.5e+2</i> :62	<i>2.5e+0</i> :250	<i>1.6e+0</i> :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 $ERT/ERT_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	<i>6.3e+1:36</i>	<i>4.0e+1:77</i>	<i>4.0e+1:77</i>	<i>1.6e+1:456</i>	<i>4.0e+0:1094</i>	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 $\text{ERT}/\text{ERT}_{\text{best 2009}}$ 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	<i>6.3e+1:45</i>	<i>4.0e+1:68</i>	<i>4.0e+1:68</i>	<i>1.6e+1:231</i>	<i>6.3e+0:1219</i>	15/15
MATSUMOTO-	2.2 _(0.6)	2.1 _(0.6) ^{*2}	2.1 ₍₂₎ ^{*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 $ERT/ERT_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	<i>6.3e+0:29</i>	<i>4.0e+0:118</i>	<i>2.5e+0:306</i>	<i>2.5e+0:306</i>	<i>1.0e+0:1614</i>	15/15
MATSUMOTO	2.0 (2)	10(15)	∞	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	2.1 (2)	6.2(6)	95(126)	95(110)	∞ <i>2000</i>	0/15
R-DE-10e5-	1.1 (2)	6.5(5)	81(17)	81(56)	∞ <i>2e6</i>	0/15
RL-SHADE-1	2.7 (3)	5.4(5)	48(61)	48(39)	∞ <i>2000</i>	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)	∞ <i>2000</i>	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 $ERT/ERT_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $ERT_{\text{best 2009}}$ (preceded by the target Δf -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	<i>2.5e+2:208</i>	<i>1.6e+2:918</i>	<i>1.0e+2:6628</i>	<i>6.3e+1:9885</i>	<i>4.0e+1:31629</i>	15/15
MATSUMOTO	15(10)	∞	∞	∞	∞ <i>1000</i>	0/15
R-DE-10e2-	2.6 (3)	15(13)	∞	∞	∞ <i>2000</i>	0/15
R-DE-10e5-	4.8(1)	16(5)	592(369)	∞	∞ <i>2e6</i>	0/15
RL-SHADE-1	3.0(1)	5.7 (6)	∞	∞	∞ <i>2000</i>	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)	∞	∞	∞ <i>2000</i>	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*.