Comparison Tables: CEC BBOB 2015 Testbed in 10-D (Expensive Setting)

The BBOBies

May 27, 2015

Abstract

This document provides tabular results of the special session on Black-Box Optimization Benchmarking at CEC 2015 with a focus on benchmarking black-box algorithms for small function evaluation budgets ("expensive setting"), see `http://coco.gforge.inria.fr/doku.php?id=cec-bbob-2015`. Overall, eight algorithms have been tested on 24 benchmark functions in dimensions between 2 and 20. A description of the used objective functions can be found in [6, 4]. The experimental set-up is described in [5].

The performance measure provided in the following tables is the expected number of objective function evaluations to reach a given target function value (ERT, expected running time), divided by the respective value for the best algorithm in BBOB-2009 (see [1]) if an algorithm from BBOB-2009 reached the given target function value. The ERT value is given otherwise ($\mathrm{ERT}_{\mathrm{best}}$ is noted as infinite). See [5] for details on how ERT is obtained. Bold entries in the table correspond to values below 3 or the top-three best values. Table 1 gives an overview on all algorithms submitted to the noise-free testbed at CEC 2015.

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

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

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

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*4.0e+1*:8.0	*2.5e+1*:16	*1.0e-8*:23	*1.0e-8*:23	*1.0e-8*:23	15/15
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 3: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_2 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*2.5e+6*:5.6	*1.0e+6*:17	*1.0e+5*:33	*2.5e+3*:118	*1.0e-8*:196	15/15
MATSUMOTO-	**2.0**(2)	**1.3**(0.6)	4.6(4)	∞	∞ *500*	0/15
R-DE-10e2-	**2.0**(0.9)	**1.7**(1)	**2.9**(2)	**2.1**(0.8)	∞ *1000*	0/15
R-DE-10e5-	**1.7**(2)	**2.2**(3)	5.1(3)	4.1(0.7)	**18**(0.7)	15/15
RL-SHADE-1	**2.8**(2)	3.3(0.7)	5.4(2)	3.1(1)	∞ *1000*	0/15
RL-SHADE-1	**2.0**(2)	**2.6**(2)	26(15)	34(5)	140(3)	15/15
R-SHADE-10	**2.1**(2)	**2.3**(3)	**3.9**(2)	**2.7**(1)	∞ *1000*	0/15
R-SHADE-10	**1.4**(2)	**2.1**(1)	6.5(4)	6.4(2)	**32**(2)	15/15
SOO-Derbel	16(15)	8.7(8)	16(7)	621(5)	6117(2550)	7/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*4.0e+2*:8.2	*1.6e+2*:37	*1.0e+2*:69	*6.3e+1*:147	*2.5e+1*:1129	15/15
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 5: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*2.5e+2*:21	*1.6e+2*:59	*1.6e+2*:59	*6.3e+1*:139	*4.0e+1*:854	15/15
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 6: 10-D, running time excess 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	*1.0e+2*:16	*6.3e+1*:19	*1.0e-8*:20	*1.0e-8*:20	*1.0e-8*:20	15/15
MATSUMOTO-	**1.2**(0.5)	**1.3**(0.0)$^{\star 2}$	**2.0**(0.2)$^{\star 4}$	**2.0**(0.2)$^{\star 4}$	**2.0**(0.2)$^{\star 4}$	15/15
R-DE-10e2-	**2.1**(0.9)	**4.1**(1)	∞	∞	∞ *1000*	0/15
R-DE-10e5-	3.4(3)	7.9(3)	**203**(9)	**203**(12)	**203**(9)	15/15
RL-SHADE-1	4.2(3)	6.9(3)	∞	∞	∞ *1000*	0/15
RL-SHADE-1	6.9(12)	39(24)	862(10)	862(21)	862(13)	15/15
R-SHADE-10	**2.8**(2)	4.6(3)	367(396)	367(533)	367(644)	2/15
R-SHADE-10	4.8(4)	12(9)	444(43)	444(22)	444(24)	15/15
SOO-Derbel	3.9(0.0)	7.3(0.0)	2290(0.0)	2290(0.0)	2290(0.0)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f6	*1.6e+5*:7.0	*6.3e+4*:16	*4.0e+2*:36	*1.0e+2*:102	*4.0e+0*:504	15/15
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 8: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*2.5e+2*:9.2	*1.6e+2*:18	*1.0e+2*:33	*1.0e+1*:172	*4.0e+0*:678	15/15
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 9: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.6e+4*:15	*1.0e+4*:22	*1.6e+3*:34	*2.5e+2*:103	*4.0e+0*:727	15/15
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 10: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_9 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*4.0e+1*:125	*2.5e+1*:148	*1.6e+1*:180	*1.0e+1*:200	*1.6e+0*:563	15/15
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 11: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ on f_{10} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\text{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*2.5e+6*:6.0	*1.0e+6*:21	*4.0e+5*:38	*2.5e+4*:104	*6.3e+2*:512	15/15
MATSUMOTO-	**1.7**(2)	**1.3**(1.0)	**1.5**(0.4)	6.2(1)	∞ *500*	0/15
R-DE-10e2-	**2.3**(3)	**1.3**(1)	**2.0**(2)	7.1(7)	∞ *1000*	0/15
R-DE-10e5-	**2.4**(5)	**1.6**(0.9)	**2.1**(0.9)	12(9)	1606(1431)	10/15
RL-SHADE-1	**1.8**(1)	**1.5**(3)	**2.6**(2)	8.9(13)	∞ *1000*	0/15
RL-SHADE-1	**2.2**(2)	**1.3**(1)	5.4(6)	26(10)	**19**(3)	15/15
R-SHADE-10	**2.5**(2)	**1.8**(2)	**1.9**(2)	5.9(3)	∞ *1000*	0/15
R-SHADE-10	**1.3**(1)	**1.7**(2)	**2.1**(2)	**4.8**(2)	**3.0**(0.8)	15/15
SOO-Derbel	**2.7**(4)	**1.2**(2)	**1.3**(0.8)	**5.3**(5)	360(991)	13/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*4.0e+4*:6.4	*2.5e+3*:15	*6.3e+1*:217	*4.0e+1*:244	*2.5e+0*:675	15/15
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 13: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*4.0e+7*:15	*2.5e+7*:24	*1.6e+7*:34	*1.0e+6*:103	*1.0e+1*:515	15/15
MATSUMOTO-	**2.0**(0.9)	**1.7**(0.6)	**1.6**(0.5)	**1.5**(0.4)*	∞ *500*	0/15
R-DE-10e2-	3.4(1)	**2.9**(1)	**2.5**(1)	**2.4**(0.9)	**7.0**(4)	4/15
R-DE-10e5-	4.7(3)	4.9(2)	4.9(2)	4.6(1)	25(15)	15/15
RL-SHADE-1	5.5(4)	5.2(0.9)	4.5(0.8)	3.3(2)	∞ *1000*	0/15
RL-SHADE-1	8.9(2)	16(10)	17(9)	32(7)	32(3)	15/15
R-SHADE-10	3.2(2)	3.5(1)	3.4(1)	3.1(0.4)	∞ *1000*	0/15
R-SHADE-10	3.8(2)	3.4(1)	4.2(1)	5.4(1)	**10**(16)	15/15
SOO-Derbel	**1.2**(0.7)	**1.5**(2)	**1.7**(2)	6.3(10)	12(4)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*1.0e+3*:12	*6.3e+2*:32	*4.0e+2*:40	*6.3e+1*:154	*2.5e+0*:521	15/15
MATSUMOTO-	**2.3**(1)	**1.5**(0.5)	**2.0**(0.4)	**2.3**(3)$^{\star 2}$	**6.9**(7)	2/15
R-DE-10e2-	3.8(2)	**2.8**(1)	4.4(3)	7.1(7)	9.4(15)	3/15
R-DE-10e5-	6.0(6)	5.2(0.9)	7.2(2)	6.6(2)	46(87)	15/15
RL-SHADE-1	7.0(5)	4.6(2)	5.3(1)	7.4(5)	∞ *1000*	0/15
RL-SHADE-1	18(15)	24(10)	36(7)	41(4)	25(1)	15/15
R-SHADE-10	5.1(4)	4.2(1)	5.4(2)	**4.3**(1)	29(22)	1/15
R-SHADE-10	6.3(2)	5.6(0.8)	8.1(4)	6.8(2)	**5.8**(2)	15/15
SOO-Derbel	**1.5**(1)	**2.4**(2)	**4.1**(2)	10(6)	108(62)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*4.0e+1*:7.7	*1.6e+1*:27	*1.0e+1*:37	*6.3e-1*:107	*1.0e-4*:505	15/15
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 16: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*2.5e+2*:9.0	*1.6e+2*:72	*1.0e+2*:186	*6.3e+1*:450	*4.0e+1*:872	15/15
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 17: 10-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:12	*2.5e+1*:47	*1.6e+1*:88	*1.0e+1*:425	*4.0e+0*:989	15/15
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 18: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\text{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.0e+1*:26	*6.3e+0*:85	*4.0e+0*:155	*2.5e+0*:238	*6.3e-1*:585	15/15
MATSUMOTO-	**1.5**(0.9)	**1.0**(0.4)	**1.9**(2)	6.9(9)	∞ *500*	0/15
R-DE-10e2-	**1.5**(0.6)	**1.7**(0.4)	**2.6**(1)	4.7(4)	26(36)	1/15
R-DE-10e5-	**2.4**(2)	**2.4**(2)	**3.0**(1)	3.6(1)	5.4(3)	15/15
RL-SHADE-1	**2.5**(2)	**1.6**(1)	**1.9**(1.0)	4.3(4)	∞ *1000*	0/15
RL-SHADE-1	3.3(5)	3.9(3)	7.9(3)	11(5)	15(2)	15/15
R-SHADE-10	**2.9**(2)	**2.3**(1.0)	5.3(3)	7.2(6)	26(50)	1/15
R-SHADE-10	**2.0**(1)	**1.5**(0.3)	**1.9**(0.6)	**2.0**(0.4)	**2.2**(0.5)	15/15
SOO-Derbel	**0.69**(1)	**1.2**(1)	**1.4**(0.6)	**2.0**(1)	**4.4**(6)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*4.0e+1*:11	*2.5e+1*:56	*1.6e+1*:172	*1.6e+1*:172	*2.5e+0*:561	15/15
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 20: 10-D, running time excess ERT/$\text{ERT}_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\text{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:618	*1.0e-1*:10609	*6.3e-2*:10623	*4.0e-2*:10625	*2.5e-2*:10644	15/15
MATSUMOTO-	∞	∞	∞	∞	∞ *500*	0/15
R-DE-10e2-	∞	∞	∞	∞	∞ *1000*	0/15
R-DE-10e5-	∞	∞	∞	∞	∞ *1e6*	0/15
RL-SHADE-1	∞	∞	∞	∞	∞ *1000*	0/15
RL-SHADE-1	**670**(396)	188(361)	667(1083)	∞	∞ *1e6*	0/15
R-SHADE-10	∞	∞	∞	∞	∞ *1000*	0/15
R-SHADE-10	965(855)	**115**(45)	**309**(378)	**1407**(1200)	∞ *1e6*	0/15
SOO-Derbel	**4.3**(1)	**0.30**(0.1)	**0.38**(0.1)	**7.2**(0.3)	**20**(36)	13/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*1.0e+4*:17	*6.3e+3*:21	*6.3e+1*:30	*2.5e+0*:122	*1.0e+0*:15426	13/15
MATSUMOTO-	**1.9**(1.0)	**2.0**(0.7)	**3.0**(1)	11(20)	∞ *500*	0/15
R-DE-10e2-	**2.1**(0.5)	**2.0**(0.9)	**4.5**(3)	**2.8**(0.9)	**0.96**(2)	1/15
R-DE-10e5-	3.4(2)	4.3(3)	8.6(1)	5.5(1)	**0.21**(0.1)	15/15
RL-SHADE-1	**2.7**(2)	4.3(3)	6.8(2)	**2.8**(0.6)	**0.22**(0.1)	4/15
RL-SHADE-1	5.8(5)	6.9(4)	40(10)	38(7)	**2.0**(0.4)	15/15
R-SHADE-10	**3.0**(1)	3.0(2)	6.1(2)	**2.7**(0.4)	**0.96**(0.4)	1/15
R-SHADE-10	3.0(1)	3.1(2)	6.8(3)	8.8(4)	**0.71**(0.5)	15/15
SOO-Derbel	**2.3**(0.0)	3.1(0.0)	54(0.0)	163(4e-3)	**2.3**(3e-5)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*4.0e+1*:30	*2.5e+1*:46	*1.6e+1*:56	*1.0e+1*:130	*6.3e+0*:639	15/15
MATSUMOTO-	**1.9**(2)	**1.8**(0.9)	**2.9**(4)	**2.5**(2)	**0.59**(0.6)	11/15
R-DE-10e2-	4.0(4)	5.4(7)	6.2(2)	3.6(1)	**0.89**(1.0)	13/15
R-DE-10e5-	4.0(3)	5.6(4)	8.3(7)	5.1(3)	3.0(1)	15/15
RL-SHADE-1	4.4(2)	4.1(1)	4.8(1)	**2.8**(2)	**0.93**(0.2)	13/15
RL-SHADE-1	13(8)	19(12)	30(12)	22(12)	6.0(3)	15/15
R-SHADE-10	4.0(3)	4.2(2)	5.4(2)	3.3(2)	**1.4**(0.8)	12/15
R-SHADE-10	4.9(4)	6.3(5)	8.1(5)	4.6(3)	**2.5**(0.7)	15/15
SOO-Derbel	**1.8**(3)	**2.7**(2)	**3.2**(3)	**1.9**(1)	**0.76**(0.4)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:18	*4.0e+1*:30	*4.0e+1*:30	*6.3e+0*:155	*4.0e+0*:631	14/15
MATSUMOTO-	**2.4**(1)	**2.3**(1)	**2.3**(1)	**2.3**(3)	**0.60**(0.7)	11/15
R-DE-10e2-	3.3(1)	4.0(2)	4.0(2)	10(5)	3.7(5)	5/15
R-DE-10e5-	6.4(4)	7.0(5)	7.0(3)	21(38)	22(58)	15/15
RL-SHADE-1	4.7(4)	5.7(4)	5.7(3)	8.1(4)	3.4(2)	6/15
RL-SHADE-1	11(8)	24(19)	24(9)	27(20)	50(4)	15/15
R-SHADE-10	4.0(4)	5.3(5)	5.3(4)	**4.2**(6)	**1.9**(0.9)	9/15
R-SHADE-10	6.7(2)	7.6(4)	7.6(3)	5.3(2)	5.3(9)	15/15
SOO-Derbel	**2.5**(2)	**2.5**(2)	**2.5**(0.9)	5.5(5)	16(2)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*6.3e+0*:10	*4.0e+0*:62	*2.5e+0*:162	*2.5e+0*:162	*1.0e+0*:915	15/15
MATSUMOTO-	**2.1**(3)	**2.1**(2)	8.6(12)	8.6(12)	∞ *500*	0/15
R-DE-10e2-	**1.6**(2)	**2.6**(2)	8.9(15)	8.9(12)	∞ *1000*	0/15
R-DE-10e5-	**2.0**(3)	**1.9**(3)	8.3(8)	8.3(8)	204(224)	15/15
RL-SHADE-1	**2.0**(1)	**1.2**(1.0)	6.8(6)	6.8(17)	∞ *1000*	0/15
RL-SHADE-1	**2.1**(2)	3.0(4)	**6.8**(3)	**6.8**(7)	68(32)	15/15
R-SHADE-10	**1.6**(2)	**2.3**(2)	8.1(5)	8.1(9)	∞ *1000*	0/15
R-SHADE-10	**1.7**(0.9)	**2.1**(3)	7.8(8)	7.8(11)	**21**(11)	15/15
SOO-Derbel	**1.9**(3)	**2.5**(2)	**5.2**(2)	**5.2**(3)	**2.3**(1)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*1.0e+2*:66	*6.3e+1*:596	*4.0e+1*:3181	*2.5e+1*:7668	*1.6e+1*:14353	15/15
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

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.