

Comparison Tables: BBOB 2013 Testbed in 2-D (Expensive Setting)

The BBOBies

May 15, 2014

Abstract

This document provides tabular results of the workshop for Black-Box Optimization Benchmarking at GECCO 2013 with a focus on benchmarking black-box algorithms for small function evaluation budgets (“expensive setting”), see <http://coco.gforge.inria.fr/doku.php?id=bbob-2013>. About 30 algorithms have been tested on 24 benchmark functions in dimensions between 2 and 40. 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 [2]) 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 in 2013.

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

algorithm name	short	paper	reference
BIPOP-aCMA-STEP los		BI-Population CMA-ES Algorithms with Surrogate Models and Line Searches (Page 1177)	[11]
BIPOP-saACM-k los		BI-Population CMA-ES Algorithms with Surrogate Models and Line Searches (Page 1177)	[11]
CMAES hut		An Evaluation of Sequential Model-Based Optimization for Expensive Blackbox Functions (Page 1209)	[8]
DE pal		Benchmarking a Hybrid Multi Level Single Linkage Algorithm on the BBOB Noiseless Testbed	[12]
HCMA los		BI-Population CMA-ES Algorithms with Surrogate Models and Line Searches (Page 1177)	[11]
HMSL pal		Benchmarking a Hybrid Multi Level Single Linkage Algorithm on the BBOB Noiseless Testbed	[12]
IPOP-10DDr lia		Bounding the Population Size of IPOP-CMA-ES on the Noiseless BBOB Testbed (Page 1161)	[9]
IPOP-500 lia		Bounding the Population Size of IPOP-CMA-ES on the Noiseless BBOB Testbed (Page 1161)	[9]
IPOP-tany lia		Testing the Impact of Parameter Tuning on a Variant of IPOP-CMA-ES with a Bounded Maximum Population Size on the Noiseless BBOB Testbed (Page 1169)	[10]
IPOP-texp lia		Testing the Impact of Parameter Tuning on a Variant of IPOP-CMA-ES with a Bounded Maximum Population Size on the Noiseless BBOB Testbed (Page 1169)	[10]
IPOP lia		Bounding the Population Size of IPOP-CMA-ES on the Noiseless BBOB Testbed (Page 1161)	[9]
		Testing the Impact of Parameter Tuning on a Variant of IPOP-CMA-ES with a Bounded Maximum Population Size on the Noiseless BBOB Testbed (Page 1169)	[10]
MLSL pal		Benchmarking a Hybrid Multi Level Single Linkage Algorithm on the BBOB Noiseless Testbed	[12]
OQNLP pal		Comparison of Multistart Global Optimization Algorithms on the BBOB Noiseless Testbed (Page 1153)	[13]
P-DCN tra		Multiobjectivization with NSGA-II on the Noiseless BBOB Testbed (Page 1217)	[15]
P-zero tra		Multiobjectivization with NSGA-II on the Noiseless BBOB Testbed (Page 1217)	[15]
SMAC hut		An Evaluation of Sequential Model-Based Optimization for Expensive Blackbox Functions (Page 1209)	[8]
U-DCN tra		Multiobjectivization with NSGA-II on the Noiseless BBOB Testbed (Page 1217)	[15]
U-zero tra		Multiobjectivization with NSGA-II on the Noiseless BBOB Testbed (Page 1217)	[15]
fmincon pal		Comparison of Multistart Global Optimization Algorithms on the BBOB Noiseless Testbed (Page 1153)	[13]
fminunc pal		Comparison of Multistart Global Optimization Algorithms on the BBOB Noiseless Testbed (Page 1153)	[13]
ga100 hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
grid100 hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
grid16 hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
hill hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
ImmCMA aug		Benchmarking the Local Metamodel CMA-ES on the Noiseless BBOB'2013 Test Bed (Page 1225)	[1]
memPSODE vog		Adapt-MEMPSODE: A Memetic Algorithm with Adaptive Selection of Local Searches (Page 1137)	[16]
prcga saw		Benchmarking Projection-Based Real Coded Genetic Algorithm on BBOB-2013 Noiseless Function Testbed (Page 1193)	[14]
ring100 hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
ring16 hol		Benchmarking Cellular Genetic Algorithms on the BBOB Noiseless Testbed (Page 1201)	[7]
simplex pal		Comparison of Multistart Global Optimization Algorithms on the BBOB Noiseless Testbed (Page 1153)	[13]

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
BIPOP-aCMA	2.4 (1)	4.1(3)	5.2(1.0)	11(0.5)	11(0.5)	15/15
BIPOP-saAC	5.8(5)	6.5(6)	10(4)	21(2)	21(2)	15/15
CMAES hut	4.4(9)	4.1(5)	11(4)	166(150)	166(166)	3/15
DE pal	2.4 (2)	2.1 (2)	14(11)	76(9)	76(9)	15/15
HCMA los	3.3(2)	2.2 (0)	0.97 (0) ^{*4}	0.97 (0) ^{*4}	0.97 (0) ^{*4}	15/15
HMLSL pal	1.5 (1)	1.5 (1)	1.8 (0.5)	3.1 (0.5)	3.1 (0.5)	15/15
IPOP-10DDr	1.7 (0.8)	2.9 (2)	9.3(3)	36(3)	36(3)	15/15
IPOP-500 l	1.7 (0.8)	2.9 (2)	9.3(3)	36(3)	36(3)	15/15
IPOP-tany	1.7 (0.8)	2.7 (2)	10(2)	33(3)	33(3)	15/15
IPOP-texp	4.3(7)	4.8(5)	7.4(4)	35(5)	35(5)	15/15
IPOP lia	1.7 (0.8)	2.9 (2)	9.3(3)	36(3)	36(3)	15/15
MLSL pal	1.5 (1)	1.5 (1)	1.8 (0.5)	3.1 (0.5)	3.1 (0.5)	15/15
OQNLP pal	3.3(5)	3.8(2)	2.8 (0.6)	3.2(0.7)	3.2(0.7)	15/15
P-DCN tra	1.3 (0.8)	3.5(3)	22(14)	281(150)	281(150)	15/15
P-zero tra	1.3 (0.8)	3.3(4)	16(12)	1404(1414)	1404(1414)	15/15
SMAC hut	1.6 (1)	1.9 (1)	2.4 (0.6)	∞	∞ 200	0/15
U-DCN tra	1.3 (0.8)	3.3(5)	29(26)	3261(4538)	3261(4538)	15/15
U-zero tra	1.3 (0.8)	2.2 (2)	32(31)	5.5e4(4e4)	5.5e4(4e4)	15/15
fmincon pa	1.5 (1)	1.5 (1)	1.8 (0.5)	3.1 (0.5)	3.1 (0.5)	15/15
fminunc pa	1.5 (1)	2.0 (1)	1.1 (0)	1.1 (0)	1.1 (0)	15/15
ga100 hol	2.6 (3)	2.8 (3)	65(48)	2.3e5(2e5)	2.3e5(3e5)	1/15
grid100 ho	2.6 (2)	2.5 (2)	103(80)	2.4e5(3e5)	2.4e5(2e5)	1/15
grid16 hol	2.3 (2)	3.7(2)	44(32)	7.3e4(8e4)	7.3e4(8e4)	3/15
hill hol	6.5(8)	4.6(4)	18(12)	2.9e4(3e4)	2.9e4(3e4)	7/15
lmmCMA aug	1.8 (2)	2.9 (3)	3.6(0.9)	8.8(0.8)	8.8(0.8)	15/15
memPSODE v	3.2(3)	3.3(2)	18(8)	78(27)	78(27)	15/15
prcga saw	1.9 (1)	2.3 (2)	31(35)	247(165)	247(165)	15/15
ring100 ho	2.1 (2)	2.7 (3)	77(70)	7.5e4(8e4)	7.5e4(9e4)	3/15
ring16 hol	1.9 (2)	2.4 (3)	45(24)	5.5e4(6e4)	5.5e4(6e4)	4/15
simplex pa	8.0(13)	13(14)	12(4)	20(3)	20(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.

#FEs/D	0.5	1.2	3	10	50	#succ
f2	<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
BIPOP-aCMA	1.7 (1)	3.4(3)	2.7 (1)	2.8 (0.4)	3.7 (0.7)	15/15
BIPOP-saAC	1.7 (1)	2.5 (2)	2.5 (3)	6.9(2)	7.5(2)	15/15
CMAES hut	1.6 (1)	2.1 (3)	3.1(3)	36(36)	∞ 206	0/15
DE pal	1.2 (0.4)	1.4 (1)	1.4 (1)	12(4)	23(4)	15/15
HCMA los	2.4 (2)	1.8 (0.9)	0.98 (0.6)	4.2(4)	6.3(3)	15/15
HMLSL pal	1.4 (1)	3.1(3)	2.0 (1)	1.4 (0.5)	2.8 (0.6)	15/15
IPOP-10DDr	1.1 (1)	2.2 (1)	2.8 (2)	27(2)	28(1)	15/15
IPOP-500 l	1.1 (1)	2.2 (1)	2.8 (2)	27(2)	28(1)	15/15
IPOP-tany	1.2 (1)	2.3 (1)	1.9 (1)	22(10)	27(1)	15/15
IPOP-texp	1.4 (2)	2.8 (2)	2.0 (0.9)	20(11)	26(3)	15/15
IPOP lia	1.1 (1)	2.2 (1)	2.8 (2)	27(2)	28(1)	15/15
MLSL pal	1.4 (1)	3.1(3)	2.0 (1)	1.4 (0.5)	2.8 (0.6)	15/15
OQNLP pal	2.9 (4)	4.0(2)	2.6 (2)	3.5(1)	103(118)	6/15
P-DCN tra	0.95 (0.4)	1.3 (0.9)	3.0 (5)	17(15)	1575(357)	15/15
P-zero tra	0.95 (0.4)	1.6 (0.9)	3.2(4)	27(20)	1.0e5(1e5)	7/15
SMAC hut	1.4 (1)	1.4 (1)	1.1 (0.7)	32(33)	∞ 200	0/15
U-DCN tra	0.95 (0.4)	1.2 (0.9)	1.8 (2)	177(238)	2.8e4(4e4)	11/15
U-zero tra	0.95 (0.4)	1.2 (0.9)	1.9 (2)	1214(824)	∞ 2e6	0/15
fmincon pa	1.4 (1)	3.1(3)	2.0 (1)	1.4 (0.5)	2.8 (0.6)	15/15
fminunc pa	1.4 (1)	2.4 (2)	1.5 (0.7)	2.9 (2)	4.0 (0.8)	15/15
ga100 hol	1.6 (1)	1.5 (1)	2.9 (3)	269(350)	∞ 1e5	0/15
grid100 ho	1.4 (2)	2.0 (2)	3.6(4)	291(145)	∞ 1e5	0/15
grid16 hol	1.3 (0.7)	1.6 (1)	3.9(5)	463(338)	∞ 1e5	0/15
hill hol	5.0(4)	4.5(4)	3.8(3)	247(370)	∞ 1e5	0/15
lmmCMA aug	1.4 (1)	2.3 (3)	1.8 (1)	4.8(1)	5.3(0.6)	15/15
memPSODE v	1.5 (1)	2.6 (5)	2.6 (2)	63(51)	58(21)	15/15
prcga saw	1.8 (2)	1.9 (2)	7.4(13)	18(7)	111(130)	15/15
ring100 ho	1.3 (1)	1.5 (1)	4.9(5)	657(1063)	∞ 1e5	0/15
ring16 hol	1.0 (0.7)	1.3 (1)	6.8(10)	355(490)	∞ 1e5	0/15
simplex pa	6.8(11)	10(7)	5.5(1.0)	4.8(2)	5.7(0.6)	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
BIPOP-aCMA	2.1 (2)	2.5 (2)	2.3 (1)	1.9 (1)	1.7 (3)	15/15
BIPOP-saAC	3.0(4)	2.7 (3)	10(26)	5.6(7)	2.8 (2)	15/15
CMAES hut	3.6(4)	3.4(5)	4.3(4)	3.4(3)	3.6(4)	6/15
DE pal	2.3 (1)	1.7 (2)	2.3 (2)	3.0 (2)	1.3 (0.9)	15/15
HCMA los	2.0 (2)	2.3 (1)	2.1 (2)	1.7 (1)	0.82 (0.5)	15/15
HMLSL pal	8.1(18)	6.0(7)	5.1(5)	4.5(3)	2.0 (1)	15/15
IPOP-10DDr	2.2 (2)	1.7 (2)	1.8 (1)	6.3(14)	4.7(4)	15/15
IPOP-500 l	2.2 (2)	1.7 (2)	1.8 (1)	6.3(14)	4.7(4)	15/15
IPOP-tany	2.2 (2)	1.5 (2)	2.4 (3)	2.2 (2)	3.3(4)	15/15
IPOP-texp	2.4 (2)	1.8 (2)	2.0 (2)	7.9(13)	5.2(3)	15/15
IPOP lia	2.2 (2)	1.7 (2)	1.8 (1)	6.3(14)	4.7(4)	15/15
MLSL pal	8.2(17)	5.9(8)	5.3(6)	5.8(5)	2.7 (2)	15/15
OQNLP pal	28(49)	19(18)	13(11)	10(3)	3.6(3)	15/15
P-DCN tra	1.6 (1)	1.5 (2)	2.3 (2)	543(62)	297(477)	15/15
P-zero tra	1.6 (1)	1.4 (2)	2.6 (2)	524(1632)	434(709)	15/15
SMAC hut	1.5 (2)	1.2 (1)	1.7 (1)	2.4 (2)	1.7 (1)	11/15
U-DCN tra	1.6 (1)	1.4 (1)	2.6 (2)	3.9(4)	2.5 (2)	15/15
U-zero tra	1.6 (1)	1.4 (2)	3.4(3)	3.3(2)	1.8 (2)	15/15
fmincon pa	8.2(17)	6.0(7)	5.2(4)	4.6(4)	2.1 (2)	15/15
fminunc pa	6.7(12)	4.3(6)	6.3(8)	4.4(4)	2.9 (3)	15/15
ga100 hol	1.6 (1)	2.0 (3)	2.9 (2)	7.9(8)	3.5(3)	15/15
grid100 ho	1.7 (1)	2.5 (3)	4.4(6)	14(15)	10(5)	15/15
grid16 hol	1.7 (2)	1.5 (2)	3.6(4)	4.6(4)	2.1 (2)	15/15
hill hol	4.1(6)	3.1(2)	2.5 (2)	2.2 (2)	1.4 (0.9)	15/15
ImmCMA aug	1.2 (1)	1.8 (1)	1.8 (0.9)	2.8 (4)	2.3 (2)	15/15
memPSODE v	2.2 (1)	1.7 (1)	2.2 (3)	8.8(15)	3.4(5)	15/15
prcga saw	1.4 (1)	2.2 (2)	2.9 (5)	2.9 (3)	1.9 (1)	15/15
ring100 ho	1.4 (1)	1.4 (0.7)	2.7 (3)	15(16)	6.1(5)	15/15
ring16 hol	1.2 (1)	0.90 (0.9)	1.3 (1.0)	6.2(6)	2.5 (2)	15/15
simplex pa	45(54)	24(20)	18(11)	11(3)	4.4(4)	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
BIPOP-aCMA	3.0(3)	2.4 (2)	2.6 (2)	1.8 (1)	1.8 (0.6)	15/15
BIPOP-saAC	2.8 (2)	1.7 (2)	2.3 (2)	2.7 (2)	4.5(4)	15/15
CMAES hut	3.6(5)	2.8 (3)	2.7 (3)	2.2 (2)	4.4(5)	5/15
DE pal	1(0.6)	1.3 (2)	2.2 (2)	3.2(3)	1.5 (0.8)	15/15
HCMA los	2.4 (1)	2.1 (0.8)	2.0 (2)	2.0 (1)	1.2 (0.6)	15/15
HMLSL pal	5.6(11)	4.4(5)	3.4(3)	4.7(4)	2.2 (2)	15/15
IPOP-10DDr	1.6 (2)	1.8 (2)	3.1(3)	10(16)	8.8(8)	15/15
IPOP-500 l	1.6 (2)	1.8 (2)	3.1(3)	10(16)	8.9(8)	15/15
IPOP-tany	1.4 (1)	1.1 (0.9)	1.4 (1)	4.3(3)	13(12)	15/15
IPOP-texp	2.3 (2)	1.5 (2)	1.3 (1)	1.0 (1)	11(16)	15/15
IPOP lia	1.6 (2)	1.8 (2)	3.1(3)	10(16)	8.9(8)	15/15
MLSL pal	5.6(10)	4.6(6)	3.5(3)	4.9(5)	2.9 (3)	15/15
OQNLP pal	20(29)	15(15)	10(10)	8.7(6)	11(14)	12/15
P-DCN tra	1.5 (1)	1.3 (1)	1.9 (2)	5.4(4)	596(865)	15/15
P-zero tra	1.5 (1)	1.2 (1)	1.7 (1)	2.5 (1)	400(1111)	15/15
SMAC hut	0.97 (1)	0.95 (0.8)	1.1 (0.9)	2.5 (3)	2.0 (2)	9/15
U-DCN tra	1.5 (1)	1.0 (0.7)	2.1 (2)	3.7(2)	1.8 (1)	15/15
U-zero tra	1.5 (1)	1.1 (0.8)	2.5 (2)	3.3(2)	2.3 (1)	15/15
fmincon pa	5.5(11)	4.5(6)	3.4(4)	4.6(4)	4.1(4)	15/15
fminunc pa	7.1(11)	3.8(5)	3.4(3)	3.6(3)	4.9(3)	15/15
ga100 hol	1.8 (2)	1.7 (2)	4.5(4)	6.9(6)	6.4(2)	15/15
grid100 ho	1.4 (0.6)	1.8 (2)	2.3 (2)	14(12)	13(10)	15/15
grid16 hol	1.6 (1)	2.1 (3)	4.7(6)	6.0(5)	4.3(4)	15/15
hill hol	2.9 (4)	2.2 (2)	2.2 (2)	1.7 (2)	2.0 (1)	15/15
lmmCMA aug	1.6 (2)	1.2 (0.8)	1.4 (1)	2.2 (2)	4.1(7)	13/15
memPSODE v	2.5 (2)	3.6(3)	4.7(4)	3.6(2)	5.2(9)	15/15
prcga saw	1.4 (1)	1.1 (0.9)	2.6 (3)	2.9 (2)	6.1(10)	15/15
ring100 ho	1.7 (1)	2.2 (3)	2.7 (3)	12(10)	10(5)	15/15
ring16 hol	2.7 (1)	3.1(5)	4.3(4)	5.3(5)	2.8 (2)	15/15
simplex pa	33(34)	29(18)	19(11)	11(5)	7.4(6)	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
BIPOP-aCMA	2.0 (1)	1.1 (0.4)	1.4 (0)	1.4 (0)	1.4 (0)	15/15
BIPOP-saAC	3.3(5)	3.1(2)	5.5(2)	5.5(2)	5.5(2)	15/15
CMAES hut	5.0(4)	3.5(3)	5.5(3)	5.5(3)	5.5(3)	15/15
DE pal	1.9 (1)	2.7 (3)	357(34)	357(34)	357(34)	15/15
HCMA los	2.3 (1)	1.2 (0.3)	1.5 (0.2)	1.5 (0.2)	1.5 (0.2)	15/15
HMLSL pal	2.9 (0)	1.2 (0)	46(49)	46(49)	46(49)	15/15
IPOP-10DDr	3.0(1)	1.9 (0.1)	3.8 (0.5)	3.8 (0.5)	3.8 (0.5)	15/15
IPOP-500 l	3.0(1)	1.9 (0.1)	3.8 (0.5)	3.8 (0.5)	3.8 (0.5)	15/15
IPOP-tany	3.0(1)	2.2 (0.3)	8.6(6)	8.6(6)	8.6(6)	15/15
IPOP-texp	6.7(2)	5.0(0.1)	43(27)	43(27)	43(27)	15/15
IPOP lia	3.0(1)	1.9 (0.1)	3.8 (0.5)	3.8 (0.5)	3.8 (0.5)	15/15
MLSL pal	2.9 (0)	1.2 (0)	26(34)	26(34)	26(34)	15/15
OQNLP pal	9.3(0)	4.3(0)	4.1(0)	4.1(0)	4.1(0)	15/15
P-DCN tra	1.1 (0.4)	1.3 (0.7)	226(38)	226(38)	226(38)	15/15
P-zero tra	1.1 (0.4)	1.3 (0.7)	115(43)	115(43)	115(43)	15/15
SMAC hut	2.1 (0.7)	1 (0.1)	1.1 (0.2)	1.1 (0.2)	1.1 (0.2)	15/15
U-DCN tra	1.1 (0.4)	1.3 (0.7)	6.7e6(7e6)	6.7e6(7e6)	6.7e6(7e6)	1/15
U-zero tra	1.1 (0.4)	1.3 (0.7)	∞	∞	∞ 2e6	0/15
fmincon pa	2.9 (0)	1.2 (0)	20(20)	20(20)	20(20)	15/15
fminunc pa	5.0(0)	2.0 (0)	3.0 (0)	3.0 (0)	3.0 (0)	15/15
ga100 hol	2.3 (3)	2.2 (2)	59(30)	59(30)	59(30)	15/15
grid100 ho	1.7 (1)	2.4 (2)	85(76)	85(76)	85(76)	15/15
grid16 hol	2.1 (2)	2.7 (2)	35(23)	35(23)	35(23)	15/15
hill hol	5.7(10)	3.5(4)	7.2(5)	7.2(5)	7.2(5)	15/15
lmmCMA aug	5.5(4)	3.8(2)	5.1(2)	5.1(2)	5.1(2)	15/15
memPSODE v	2.9 (2)	2.6 (1)	8.4(6)	8.4(6)	8.4(6)	15/15
prcga saw	2.0 (2)	11(17)	3.6e5(4e5)	3.6e5(4e5)	3.6e5(4e5)	1/15
ring100 ho	1.9 (1)	2.0 (2)	78(40)	78(40)	78(40)	15/15
ring16 hol	1.8 (2)	3.2(4)	33(20)	33(20)	33(20)	15/15
simplex pa	25(0)	11(0.1)	37(41)	37(41)	37(41)	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
BIPOP-aCMA	1 (0.7)	2.6 (2)	4.3(5)	4.7(3)	4.4(0.8)	15/15
BIPOP-saAC	3.5(4)	3.4(6)	2.3 (4)	4.0(4)	4.3(3)	15/15
CMAES hut	3.1(6)	3.5(4)	2.7 (2)	3.5(2)	∞ 206	0/15
DE pal	1.3 (0.7)	2.9 (4)	1.6 (1)	3.5(2)	10(3)	15/15
HCMA los	1.4 (1)	2.4 (2)	3.0 (3)	5.9(3)	5.0(2)	15/15
HMLSL pal	1.3 (1)	3.3(4)	1.7 (1)	1.6 (0.6)	1.1 (0.4)	15/15
IPOP-10DDr	1.7 (2)	1.6 (1)	1.2 (1)	3.8(3)	4.1(0.9)	15/15
IPOP-500 l	1.7 (2)	1.6 (1)	1.2 (1)	3.8(3)	4.1(0.9)	15/15
IPOP-tany	1.8 (2)	1.9 (1)	1.2 (2)	3.0 (2)	4.2(0.7)	15/15
IPOP-texp	2.8 (4)	3.5(3)	2.1 (2)	3.9(3)	4.5(2)	15/15
IPOP lia	1.7 (2)	1.6 (1)	1.2 (1)	3.8(3)	4.1(0.9)	15/15
MLSL pal	1.3 (1)	3.3(4)	1.7 (1)	1.6 (0.6)	1.1 (0.4)	15/15
OQNLP pal	2.9 (5)	3.9(4)	2.0 (1)	2.5 (2)	2.5 (2)	15/15
P-DCN tra	1.0 (0.4)	4.0(2)	3.7(6)	6.7(6)	43(31)	15/15
P-zero tra	1.0 (0.4)	6.0(3)	4.0(10)	8.2(7)	188(157)	15/15
SMAC hut	0.90 (0.4)	1.3 (1)	0.98 (0.9)	9.2(9)	∞ 200	0/15
U-DCN tra	1.0 (0.4)	1.3 (2)	0.95 (0.8)	6.2(5)	6123(9749)	13/15
U-zero tra	1.0 (0.4)	1.2 (1)	0.93 (1)	12(19)	2.0e4(3e4)	9/15
fmincon pa	1.3 (1)	3.3(4)	1.7 (1)	1.6 (0.6)	1.1 (0.4)	15/15
fminunc pa	1.9 (2)	2.9 (3)	1.6 (1)	2.6 (2)	1.9 (1)	15/15
ga100 hol	1 (0.4)	1.5 (2)	1.8 (2)	16(15)	∞ 1e5	0/15
grid100 ho	1.4 (1)	1.9 (2)	2.1 (3)	22(23)	∞ 1e5	0/15
grid16 hol	1.1 (0.7)	2.9 (3)	2.5 (3)	17(18)	7066(7757)	2/15
hill hol	2.4 (3)	2.1 (2)	1.7 (2)	5.6(4)	7195(7757)	2/15
lmmCMA aug	1 (1)	1.3 (0.9)	0.89 (0.8)	2.6 (2)	2.8 (0.6)	15/15
memPSODE v	1.7 (1)	2.0 (2)	2.8 (2)	4.1(2)	7.3(3)	15/15
prcga saw	1.5 (1)	2.0 (2)	2.0 (2)	26(43)	174(207)	15/15
ring100 ho	1.2 (0.4)	1.5 (2)	2.0 (1)	13(15)	∞ 1e5	0/15
ring16 hol	1.3 (1)	2.4 (2)	2.5 (3)	13(9)	7039(7272)	2/15
simplex pa	6.7(12)	11(15)	5.1(3)	4.4(2)	1.8 (0.5)	15/15

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
f7	<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
BIPOP-aCMA	1.2 (0.6)	4.3(4)	2.8 (4)	3.4(3)	1.2 (0.7)	15/15
BIPOP-saAC	2.5 (4)	5.4(3)	2.8 (3)	2.9 (4)	1.1 (0.9)	15/15
CMAES hut	2.5 (2)	5.9(6)	3.0(3)	2.8 (4)	1.3 (1)	9/15
DE pal	0.87 (0.3)	4.6(6)	3.1(3)	2.5 (2)	1.5 (0.5)	15/15
HCMA los	1.5 (1)	2.8 (1)	2.5 (3)	3.2(3)	1.6 (0.8)	15/15
HMLSL pal	1.6 (1)	3.9(5)	2.3 (2)	2.8 (3)	2.2 (2)	15/15
IPOP-10DDr	0.96 (0.6)	4.1(4)	1.7 (2)	3.9(7)	1.7 (2)	15/15
IPOP-500 l	0.96 (0.6)	4.1(4)	1.7 (2)	3.9(7)	1.7 (2)	15/15
IPOP-tany	1.2 (0.6)	3.6(3)	2.8 (2)	3.8(6)	2.0 (2)	15/15
IPOP-texp	2.0 (3)	3.9(4)	3.5(2)	3.2(5)	1.6 (1.0)	15/15
IPOP lia	0.96 (0.6)	4.1(4)	1.7 (2)	3.9(7)	1.7 (2)	15/15
MLSL pal	2.5 (3)	4.7(5)	3.1(4)	5.0(7)	45(57)	13/15
OQNLP pal	4.0(4)	6.4(5)	2.7 (2)	2.5 (3)	1.5 (0.5)	15/15
P-DCN tra	1.9 (0.6)	22(30)	6610(97)	1.5e4(5e4)	1643(5312)	13/15
P-zero tra	1.3 (0.6)	10(17)	1.0e4(313)	1.5e4(5e4)	1686(5316)	13/15
SMAC hut	1.6 (1)	2.0 (2)	0.86 (0.8)	0.89 (0.5)	2.0 (2)	6/15
U-DCN tra	1.5 (0.6)	5.6(7)	3.2(3)	3.6(3)	72(167)	15/15
U-zero tra	1.3 (0.6)	5.0(5)	3.8(6)	3.0 (4)	43(49)	15/15
fmincon pa	1.8 (2)	3.9(5)	3.4(4)	4.3(4)	33(50)	14/15
fminunc pa	1.9 (2)	4.9(5)	4.0(6)	3.7(4)	38(49)	14/15
ga100 hol	1.1 (0.9)	2.4 (2)	2.2 (2)	2.2 (2)	10(10)	15/15
grid100 ho	1.2 (2)	6.4(7)	3.8(2)	4.4(6)	41(32)	15/15
grid16 hol	1.0 (0.6)	4.2(4)	1.9 (3)	2.4 (3)	7.3(9)	15/15
hill hol	3.5(8)	6.2(10)	4.2(8)	4.6(6)	30(46)	15/15
lmmCMA aug	1.9 (2)	3.5 (3)	2.5 (4)	2.1 (3)	0.56 (0.7)	15/15
memPSODE v	1.5 (0.9)	5.5(8)	13(27)	16(23)	10(8)	15/15
prcga saw	1.5 (2)	3.5 (2)	1.4 (1)	2.2 (2)	3.8(5)	15/15
ring100 ho	1.3 (0.6)	4.1(5)	3.0 (4)	3.5(4)	15(12)	15/15
ring16 hol	1.5 (2)	7.0(9)	4.5(4)	3.8(4)	10(10)	15/15
simplex pa	1.9 (2)	5.5(6)	3.4(3)	3.6(2)	5.7(8)	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</i> :1.2	<i>1.0e+2</i> :3.2	<i>6.3e+0</i> :7.0	<i>1.6e-1</i> :27	<i>1.6e-6</i> :100	15/15
BIPOP-aCMA	2.3 (2)	3.3(4)	7.7(5)	11(9)	5.5(3)	15/15
BIPOP-saAC	1.7 (1)	3.1(3)	3.9(3)	6.5(6)	3.3(2)	15/15
CMAES hut	2.8 (4)	3.4(3)	10(15)	13(12)	∞ 206	0/15
DE pal	1.9 (2)	2.9 (3)	6.5(5)	12(8)	11(2)	15/15
HCMA los	1.8 (0.8)	2.0 (2)	1.8 (2)	9.0(6)	4.2(2)	15/15
HMLSL pal	1.8 (1)	1.6 (2)	2.8 (3)	2.5 (2)	1.0 (0.5)	15/15
IPOP-10DDr	2.4 (0.8)	3.8(3)	2.6 (2)	11(7)	6.2(1)	15/15
IPOP-500 l	2.4 (0.8)	3.8(3)	2.6 (2)	11(7)	6.2(1)	15/15
IPOP-tany	3.7(4)	3.4(2)	4.4(3)	10(3)	6.0(0.9)	15/15
IPOP-texp	7.8(10)	5.2(4)	4.6(3)	7.7(6)	5.9(0.9)	15/15
IPOP lia	2.4 (0.8)	3.8(3)	2.6 (2)	11(7)	6.2(1)	15/15
MLSL pal	1.8 (1)	1.6 (2)	2.8 (3)	2.5 (2)	1.0 (0.5)	15/15
OQNLP pal	3.3(5)	2.9 (3)	2.2 (2)	2.2 (2)	1.1 (0.6)	15/15
P-DCN tra	1.3 (0.8)	8.8(14)	11(8)	51(61)	383(370)	15/15
P-zero tra	1.3 (0.8)	4.3(7)	79(222)	74(123)	927(1121)	15/15
SMAC hut	1.7 (2)	1.6 (1)	3.3(3)	10(11)	∞ 200	0/15
U-DCN tra	1.3 (0.8)	2.6 (2)	8.4(5)	35(55)	∞ 2e6	0/15
U-zero tra	1.3 (0.8)	2.7 (2)	6.1(4)	82(133)	∞ 2e6	0/15
fmincon pa	1.8 (1)	1.6 (2)	2.8 (3)	2.5 (2)	1.0 (0.5)	15/15
fminunc pa	1.7 (1)	1.9 (2)	1.8 (2)	1.9 (2)	0.91 (0.4)	15/15
ga100 hol	2.3 (2)	3.1(3)	7.9(7)	32(32)	∞ 1e5	0/15
grid100 ho	2.4 (2)	3.9(4)	5.6(4)	68(28)	∞ 1e5	0/15
grid16 hol	2.3 (2)	5.9(8)	13(18)	69(109)	∞ 1e5	0/15
hill hol	4.0(6)	4.3(5)	14(22)	49(69)	∞ 1e5	0/15
lmmCMA aug	2.1 (3)	1.7 (2)	2.4 (1)	3.3(2)	1.6 (0.5)	15/15
memPSODE v	3.7(3)	6.9(8)	14(25)	29(34)	18(6)	15/15
prcga saw	1.6 (0.8)	2.0 (3)	4.0(4)	55(146)	170(151)	15/15
ring100 ho	1.7 (2)	2.7 (2)	8.7(10)	54(41)	∞ 1e5	0/15
ring16 hol	2.2 (2)	4.7(6)	10(9)	47(24)	1.4e4(2e4)	1/15
simplex pa	7.5(13)	6.8(5)	6.0(6)	3.7(3)	1.6 (0.7)	15/15

Table 10: 02-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
f9	<i>6.3e+0:13</i>	<i>4.0e+0:15</i>	<i>2.5e+0:15</i>	<i>2.5e-1:21</i>	<i>1.0e-8:94</i>	15/15
BIPOP-aCMA	2.9 (3)	3.5(3)	4.9(7)	12(10)	6.6(3)	15/15
BIPOP-saAC	2.4 (3)	2.5 (3)	4.3(4)	8.8(6)	4.0(2)	15/15
CMAES hut	2.9 (3)	4.1(5)	5.8(8)	13(15)	∞ 206	0/15
DE pal	4.3(4)	4.1(4)	4.1(4)	9.0(8)	15(3)	15/15
HCMA los	2.6 (3)	2.7 (3)	3.1(4)	7.3(8)	3.6(1)	15/15
HMLSL pal	0.87 (0)	0.74 (0)	0.74 (0)	0.53 (0)	0.62 (0.0)	15/15
IPOP-10DDr	2.7 (1)	3.9(2)	4.9(3)	11(11)	7.4(3)	15/15
IPOP-500 l	2.7 (1)	3.9(2)	4.9(3)	11(11)	7.4(3)	15/15
IPOP-tany	2.2 (2)	2.0 (1)	3.4(3)	7.3(8)	7.1(2)	15/15
IPOP-texp	1.2 (0.4)	1.1 (0.4)	1.6 (1)	6.8(6)	6.6(2)	15/15
IPOP lia	2.7 (1)	3.9(2)	4.9(3)	11(11)	7.4(3)	15/15
MLSL pal	0.87 (0)	0.74 (0)	0.74 (0)	0.53 (0)	0.62 (0.0)	15/15
OQNLP pal	0.63 (0)	0.54 (0)	0.61 (0)	0.56 (0.0)	1.4 (2)	15/15
P-DCN tra	16(5)	20(6)	25(21)	89(116)	1.2e4(2e4)	10/15
P-zero tra	27(75)	48(121)	57(128)	96(96)	5.2e4(6e4)	5/15
SMAC hut	2.2 (2)	2.6 (2)	3.6(2)	33(32)	∞ 200	0/15
U-DCN tra	1.6 (2)	1.9 (2)	2.0 (2)	46(76)	3.0e5(3e5)	1/15
U-zero tra	2.1 (2)	2.9 (2)	3.4(2)	95(213)	∞ 2e6	0/15
fmincon pa	0.87 (0)	0.74 (0)	0.74 (0)	0.53 (0)	0.62 (0.0)	15/15
fminunc pa	0.56 (0)	0.47 (0)*	0.47 (0)*	0.34 (0)* ³	0.68 (0.0)	15/15
ga100 hol	4.1(4)	5.6(8)	9.1(9)	20(17)	∞ 1e5	0/15
grid100 ho	7.5(7)	11(14)	13(15)	40(27)	∞ 1e5	0/15
grid16 hol	4.4(4)	4.6(4)	10(6)	33(26)	∞ 1e5	0/15
hill hol	4.6(6)	4.6(5)	5.4(6)	36(43)	∞ 1e5	0/15
lmmCMA aug	1.5 (1)	1.7 (0.6)	1.9 (0.9)	3.1(2)	1.8 (0.5)	15/15
memPSODE v	10(13)	10(11)	12(11)	24(37)	19(7)	15/15
prcga saw	2.9 (2)	2.8 (2)	3.1(2)	47(59)	193(124)	15/15
ring100 ho	3.5(3)	6.0(6)	9.0(10)	37(27)	∞ 1e5	0/15
ring16 hol	4.2(4)	4.6(4)	5.2(5)	25(16)	∞ 1e5	0/15
simplex pa	0.88 (0.1)	1.4 (0.1)	1.6 (0.1)	1.3 (0.2)	1.5 (0.1)	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</i> :2.0	<i>4.0e+5</i> :3.2	<i>6.3e+2</i> :8.8	<i>1.0e+1</i> :30	<i>2.5e-8</i> :101	15/15
BIPOP-aCMA	2.4 (4)	1.9 (2)	6.6(5)	8.0(5)	5.8 (0.7)	15/15
BIPOP-saAC	2.8 (3)	2.2 (4)	5.8(4)	3.8(2)	2.2 (0.3)	15/15
CMAES hut	4.8(5)	3.8(5)	6.5(6)	10(8)	∞ 206	0/15
DE pal	2.4 (3)	2.4 (3)	12(7)	11(4)	15(2)	15/15
HCMA los	2.4 (2)	2.0 (1)	1.4 (0.5)	3.6(2)	2.6 (0.5)	15/15
HMLSL pal	2.8 (3)	2.2 (2)	1.8 (0.6)	0.71 (0.2)	38(38)	15/15
IPOP-10DDr	1.4 (0.8)	2.1 (3)	6.1(4)	12(8)	8.0(0.7)	15/15
IPOP-500 l	1.4 (0.8)	2.1 (3)	6.1(4)	12(8)	8.0(0.7)	15/15
IPOP-tany	1.8 (2)	2.0 (1)	5.4(4)	8.9(8)	7.9(0.6)	15/15
IPOP-texp	1.9 (2)	2.1 (2)	5.1(3)	9.4(6)	7.5(0.5)	15/15
IPOP lia	1.4 (0.8)	2.1 (3)	6.1(4)	12(8)	8.0(0.7)	15/15
MLSL pal	2.8 (3)	2.2 (2)	1.8 (0.6)	0.71 (0.2)	181(314)	10/15
OQNLP pal	4.3(3)	3.1(2)	2.6 (2)	1.2 (0.7)	61(64)	7/15
P-DCN tra	2.3 (4)	4.3(7)	7.3(7)	392(597)	2.8e5(3e5)	1/15
P-zero tra	2.5 (2)	3.6(6)	5.7(3)	385(880)	3.0e5(3e5)	1/15
SMAC hut	1.5 (1)	1.1 (0.8)	1.8 (0.7)	1.3 (1)	∞ 200	0/15
U-DCN tra	2.1 (2)	2.8 (4)	13(9)	164(243)	∞ 2e6	0/15
U-zero tra	1.9 (2)	2.2 (2)	12(12)	269(456)	∞ 2e6	0/15
fmincon pa	2.8 (3)	2.2 (2)	1.8 (0.6)	0.71 (0.2)	141(190)	11/15
fminunc pa	2.4 (2)	1.8 (1)	1.4 (0.7)	1.1 (1.0)	303(387)	8/15
ga100 hol	2.3 (2)	2.0 (2)	29(23)	386(725)	∞ 1e5	0/15
grid100 ho	1.8 (2)	2.7 (3)	25(35)	311(511)	∞ 1e5	0/15
grid16 hol	1.6 (2)	2.5 (2)	19(16)	45(38)	∞ 1e5	0/15
hill hol	2.1 (2)	2.2 (2)	16(22)	265(433)	∞ 1e5	0/15
lmmCMA aug	3.3(3)	2.9 (2)	3.2(1)	2.4 (1)	1.6 (0.2)	15/15
memPSODE v	3.2(4)	4.7(5)	10(3)	22(35)	28(14)	15/15
prcga saw	2.9 (4)	3.3(3)	9.4(6)	88(128)	8766(9765)	2/15
ring100 ho	2.2 (2)	2.4 (3)	18(29)	58(71)	∞ 1e5	0/15
ring16 hol	2.2 (2)	2.4 (2)	13(9)	63(61)	∞ 1e5	0/15
simplex pa	11(9)	8.0(6)	4.6(2)	2.4 (2)	1.7 (0.2)	15/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f11	<i>1.0e+7:1.1</i>	<i>1.6e+6:3.2</i>	<i>1.0e+4:6.6</i>	<i>4.0e+1:23</i>	<i>4.0e-8:100</i>	15/15
BIPOP-aCMA	1.9 (1)	2.2 (2)	5.4(4)	6.2(6)	5.7 (0.6)	15/15
BIPOP-saAC	3.8(5)	2.8 (3)	4.1(3)	3.5(1)	2.2 (0.4)	15/15
CMAES hut	1.9 (2)	2.1 (3)	4.2(4)	5.6(5)	∞ 206	0/15
DE pal	1.8 (0.9)	1.2 (0.8)	6.2(5)	8.7(4)	14(1)	15/15
HCMA los	2.0 (1)	2.0 (1)	1.5 (0.5)	2.2 (3)	2.6 (0.5)	15/15
HMLSL pal	1.7 (1)	2.3 (2)	2.0 (1)	0.84 (0.4)	19(27)	15/15
IPOP-10DDr	1.6 (0.9)	1.6 (1)	4.5(2)	6.5(10)	8.2(0.5)	15/15
IPOP-500 l	1.6 (0.9)	1.6 (1)	4.5(2)	6.5(10)	8.2(0.5)	15/15
IPOP-tany	1.6 (0.9)	2.3 (1)	4.6(3)	5.3(2)	8.1(0.6)	15/15
IPOP-texp	3.2(5)	3.7(3)	4.0(3)	4.5(2)	8.1(1.0)	15/15
IPOP lia	1.6 (0.9)	1.6 (1)	4.5(2)	6.5(10)	8.2(0.5)	15/15
MLSL pal	1.7 (1)	2.3 (2)	2.0 (1)	0.84 (0.4)	97(158)	12/15
OQNLP pal	2.9 (3)	3.1(2)	3.5(2)	1.6 (1)	72(83)	7/15
P-DCN tra	1.4 (0.9)	3.7(8)	10(12)	25(72)	8.6e4(1e5)	3/15
P-zero tra	1.4 (0.9)	2.5 (2)	6.8(7)	6.9(3)	1.4e5(1e5)	2/15
SMAC hut	1.6 (1)	1.2 (0.8)	1.6 (1)	1.4 (0.7)	∞ 200	0/15
U-DCN tra	1.4 (0.9)	1.8 (2)	5.8(7)	13(12)	∞ 2e6	0/15
U-zero tra	1.4 (0.9)	1.8 (2)	6.3(6)	28(27)	∞ 2e6	0/15
fmincon pa	1.7 (1)	2.3 (2)	2.0 (1)	0.84 (0.3)	115(172)	13/15
fminunc pa	1.7 (1)	1.9 (1)	1.8 (0.9)	0.84 (0.3)	66(124)	13/15
ga100 hol	1.6 (0.9)	1.4 (1)	13(23)	17(17)	∞ 1e5	0/15
grid100 ho	1.6 (0.9)	1.4 (1)	11(8)	29(26)	∞ 1e5	0/15
grid16 hol	1.3 (0.5)	1.3 (1)	10(9)	30(39)	∞ 1e5	0/15
hill hol	4.7(6)	3.4(3)	4.2(3)	23(37)	∞ 1e5	0/15
lmmCMA aug	2.5 (2)	2.0 (2)	2.6 (1)	1.6 (0.7)	1.6 (0.2)	15/15
memPSODE v	2.2 (1)	1.7 (2)	7.4(5)	11(5)	34(25)	15/15
prcga saw	2.1 (1)	1.8 (1)	6.1(6)	5.5(2)	5861(6109)	3/15
ring100 ho	1.7 (1)	1.4 (1)	9.0(8)	33(24)	∞ 1e5	0/15
ring16 hol	1.4 (0.9)	1.2 (0.9)	8.1(7)	20(16)	∞ 1e5	0/15
simplex pa	5.9(9)	8.0(6)	5.7(1)	2.2 (0.6)	1.7 (0.1)	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</i> :1.3	<i>6.3e+6</i> :2.7	<i>6.3e+5</i> :6.3	<i>4.0e+1</i> :21	<i>1.6e-3</i> :101	15/15
BIPOP-aCMA	0.95 (0)	1.2 (1)	1.3 (1)	4.7(3)	5.7(1)	15/15
BIPOP-saAC	1.5 (2)	1.9 (2)	1.6 (2)	5.3(2)	4.5(5)	15/15
CMAES hut	0.75 (0)	1.7 (3)	1.7 (2)	9.4(10)	∞ 200	0/15
DE pal	1.3 (0.4)	1(1)	1.7 (2)	7.0(5)	16(10)	15/15
HCMA los	1.1 (0.8)	1.5 (1)	1.2 (0.8)	4.0(4)	5.3(4)	15/15
HMLSL pal	1.1 (1)	1.4 (2)	1.2 (1.0)	1.0 (0.6)	0.99 (2)	15/15
IPOP-10DDr	1.1 (1)	1.5 (0.9)	1.5 (1)	7.3(7)	14(12)	15/15
IPOP-500 l	1.1 (1)	1.5 (0.9)	1.5 (1)	7.3(7)	14(12)	15/15
IPOP-tany	1.2 (1)	1.6 (1)	1.2 (1)	6.2(7)	8.3(4)	15/15
IPOP-texp	1.2 (1)	2.6 (2)	1.9 (1)	4.0(2)	10(10)	15/15
IPOP lia	1.1 (1)	1.5 (0.9)	1.5 (1)	7.3(7)	14(12)	15/15
MLSL pal	1.1 (1)	1.4 (2)	1.2 (1.0)	1.0 (0.6)	0.99 (2)	15/15
OQNLP pal	1.6 (3)	3.2(2)	1.9 (1)	1.9 (1)	4.5(3)	15/15
P-DCN tra	0.95 (0.4)	0.98 (0.5)	0.78 (1)	4.7(4)	446(428)	15/15
P-zero tra	0.95 (0.4)	0.98 (0.5)	0.78 (0.6)	3.8(1)	645(376)	15/15
SMAC hut	0.85 (0.4)	0.73 (0.5)	0.44 (0.3)	3.0(2)	∞ 200	0/15
U-DCN tra	0.95 (0.4)	1.1 (0.5)	0.85 (1)	12(8)	1.9e4(3e4)	8/15
U-zero tra	0.95 (0.4)	0.98 (0.5)	0.83 (1)	29(26)	2.4e4(3e4)	7/15
fmincon pa	1.1 (1)	1.4 (2)	1.2 (1.0)	1.0 (0.6)	0.99 (2)	15/15
fminunc pa	1.1 (1)	2.2 (2)	1.6 (2)	1.0 (0.7)	1.7 (2)	15/15
ga100 hol	0.90 (0.4)	1.2 (0.9)	1.6 (2)	23(25)	2233(2601)	5/15
grid100 ho	1(0.4)	0.88 (0.7)	1.4 (1)	57(52)	3375(3479)	4/15
grid16 hol	0.95 (0.4)	1(0.7)	1.5 (2)	31(30)	1.4e4(2e4)	1/15
hill hol	1.6 (2)	2.6 (3)	1.8 (2)	49(57)	7062(7455)	2/15
lmmCMA aug	0.95 (0.8)	1.4 (2)	1.1 (1)	1.8 (0.5)	3.5(3)	15/15
memPSODE v	1.3 (1)	1.5 (1)	1.8 (2)	24(50)	17(8)	15/15
prcga saw	0.90 (0.4)	1.1 (1)	4.1(2)	17(9)	712(823)	9/15
ring100 ho	0.90 (0.4)	1.0 (0.7)	1.6 (2)	32(29)	3187(3479)	4/15
ring16 hol	0.75 (0)	0.90 (0.5)	1.3 (1)	16(11)	1.4e4(2e4)	1/15
simplex pa	2.7 (5)	8.2(7)	5.6(4)	3.0 (1)	2.1 (2)	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
BIPOP-aCMA	3.2(3)	1.9 (2)	3.2(3)	3.9(2)	6.0 (0.8)	15/15
BIPOP-saAC	2.4 (2)	1.9 (2)	3.2(5)	3.3(2)	2.7 (0.6)	15/15
CMAES hut	2.6 (3)	1.9 (2)	2.4 (2)	3.2(2)	∞ 206	0/15
DE pal	2.3 (2)	2.0 (2)	3.0 (2)	4.9(2)	14(2)	15/15
HCMA los	2.7 (2)	1.7 (1)	2.9 (5)	2.9 (3)	3.2 (0.8)	15/15
HMLSL pal	1.5 (1)	1.2 (0.8)	1.2 (0.6)	0.77 (0.3)	31(24)	15/15
IPOP-10DDr	2.4 (3)	2.0 (2)	3.3(2)	4.3(7)	8.7(2)	15/15
IPOP-500 l	2.4 (3)	2.0 (2)	3.3(2)	4.3(7)	8.7(2)	15/15
IPOP-tany	2.7 (3)	1.9 (2)	3.0 (2)	5.9(8)	8.4(0.7)	15/15
IPOP-texp	3.5(5)	2.4 (3)	2.6 (2)	7.4(10)	8.1(0.8)	15/15
IPOP lia	2.4 (3)	2.0 (2)	3.3(2)	4.3(7)	8.7(2)	15/15
MLSL pal	1.5 (1)	1.2 (0.8)	1.2 (0.6)	0.77 (0.3)	360(456)	6/15
OQNLP pal	3.5(4)	2.7 (2)	1.7 (0.5)	1.2 (0.6)	19(17)	10/15
P-DCN tra	1.9 (2)	2.1 (3)	4.2(4)	21(30)	∞ 2e6	0/15
P-zero tra	1.9 (2)	1.9 (2)	4.0(3)	37(102)	∞ 2e6	0/15
SMAC hut	1.5 (2)	1.0 (1)	1.3 (1.0)	1.6 (1)	∞ 200	0/15
U-DCN tra	1.9 (2)	1.8 (2)	3.2(4)	10(10)	∞ 2e6	0/15
U-zero tra	1.9 (2)	1.6 (2)	2.0 (2)	11(11)	∞ 2e6	0/15
fmincon pa	1.5 (1)	1.2 (0.8)	1.2 (0.6)	0.77 (0.3)	172(237)	9/15
fminunc pa	1.4 (0.9)	1.7 (1)	1.4 (0.9)	5.2(5)	∞ 2e4	0/15
ga100 hol	2.5 (4)	2.1 (2)	3.5(3)	11(12)	∞ 1e5	0/15
grid100 ho	2.0 (2)	1.4 (0.8)	2.3 (2)	24(21)	∞ 1e5	0/15
grid16 hol	1.7 (0.9)	1.9 (2)	4.0(5)	15(25)	∞ 1e5	0/15
hill hol	4.6(5)	2.8 (3)	10(9)	13(18)	∞ 1e5	0/15
lmmCMA aug	1.5 (1)	1.1 (1)	1.4 (0.9)	2.0 (2)	3.0 (0.7)	15/15
memPSODE v	4.1(2)	3.5(2)	3.2(3)	12(19)	15(7)	15/15
prcga saw	1.9 (1)	1.4 (0.8)	2.0 (2)	33(63)	∞ 1e5	0/15
ring100 ho	1.9 (2)	2.0 (2)	3.5(4)	15(17)	∞ 1e5	0/15
ring16 hol	2.5 (2)	2.3 (2)	4.0(4)	10(8)	∞ 1e5	0/15
simplex pa	8.3(10)	7.0(5)	5.5(4)	4.1(2)	2.2 (0.5)	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
BIPOP-aCMA	2.1 (2)	4.7(7)	4.8(6)	5.4(2)	6.5 (1)	15/15
BIPOP-saAC	1.9 (2)	3.0 (5)	3.4(4)	3.5(2)	2.9 (0.3)	15/15
CMAES hut	1.2 (0.7)	2.0 (2)	2.5 (3)	4.2(2)	∞ 206	0/15
DE pal	1.5 (1)	1.8 (2)	3.7(3)	7.7(3)	14(2)	15/15
HCMA los	1.4 (1)	3.6(3)	3.0(5)	4.9(4)	3.2 (0.5)	15/15
HMLSL pal	1.8 (2)	1.3 (1)	1.1 (0.8)	1.0 (0.4)	39(14)	15/15
IPOP-10DDr	1(0.7)	1.4 (2)	1.9 (2)	3.6(1)	8.2(0.7)	15/15
IPOP-500 l	1(0.7)	1.4 (2)	1.9 (2)	3.6(1)	8.2(0.7)	15/15
IPOP-tany	1(0.7)	1.7 (2)	2.0 (2)	3.7(1)	8.2(1.0)	15/15
IPOP-texp	1.9 (2)	2.3 (3)	2.5 (2)	3.4(1)	7.9(0.9)	15/15
IPOP lia	1(0.7)	1.4 (2)	1.9 (2)	3.6(1)	8.2(0.7)	15/15
MLSL pal	1.8 (2)	1.3 (1)	1.1 (0.8)	1.0 (0.4)	629(757)	4/15
OQNLP pal	3.2(4)	2.6 (3)	2.1 (2)	1.5 (0.7)	100(98)	5/15
P-DCN tra	1.1 (0.7)	4.2(3)	6.2(7)	8.0(4)	2.7e4(3e4)	7/15
P-zero tra	1.1 (0.7)	2.4 (2)	5.8(9)	6.2(4)	2.9e5(3e5)	1/15
SMAC hut	0.76 (0)	1.1 (1.0)	1.8 (1)	3.9(1)	∞ 200	0/15
U-DCN tra	1.1 (0.7)	1.8 (3)	3.0(3)	23(14)	∞ 2e6	0/15
U-zero tra	1.1 (0.7)	1.3 (2)	2.4 (2)	31(42)	∞ 2e6	0/15
fmincon pa	1.8 (2)	1.3 (1)	1.1 (0.8)	1.0 (0.4)	2964(3028)	1/15
fminunc pa	1.4 (1)	1.6 (2)	1.7 (2)	1.5 (1)	526(595)	5/15
ga100 hol	1.9 (1)	2.1 (2)	2.3 (2)	41(13)	∞ 1e5	0/15
grid100 ho	1(0.4)	1.4 (2)	3.2(2)	50(44)	∞ 1e5	0/15
grid16 hol	1.2 (0.4)	2.0 (3)	4.5(8)	37(23)	∞ 1e5	0/15
hill hol	3.6(4)	3.5(4)	2.5 (2)	21(38)	∞ 1e5	0/15
lmmCMA aug	0.95 (0)	1.4 (1)	1.8 (2)	1.9 (0.6)	2.4 (0.4)	15/15
memPSODE v	1.7 (1)	3.0(2)	3.5(3)	5.9(3)	34(32)	15/15
prcga saw	1.4 (1)	1.8 (2)	2.3 (2)	7.4(3)	1692(1930)	7/15
ring100 ho	1.3 (0.7)	2.5 (2)	4.2(4)	55(30)	∞ 1e5	0/15
ring16 hol	2.0 (1)	3.0 (3)	3.3(4)	21(16)	∞ 1e5	0/15
simplex pa	7.6(11)	6.9(9)	7.9(6)	4.4(1)	1.7 (0.2)* ³	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</i> :1.2	<i>4.0e+1</i> :4.7	<i>2.5e+1</i> :10	<i>1.0e+1</i> :37	<i>2.5e+0</i> :118	5/5
BIPOP-aCMA	1.2 (0.4)	1.2 (1)	1.8 (2)	1.4 (2)	2.0 (3)	15/15
BIPOP-saAC	1.7 (2)	1.5 (2)	1.3 (1)	1.1 (0.7)	1.4 (2)	15/15
CMAES hut	2.6 (2)	1.7 (2)	1.2 (1)	0.98 (0.8)	1.9 (2)	9/15
DE pal	1.7 (1)	0.77 (0.5)	0.78 (0.9)	1.6 (1)	1.9 (1)	15/15
HCMA los	2.3 (2)	1.9 (1)	1.5 (1)	1.1 (0.8)	1.7 (2)	15/15
HMLSL pal	4.7(1)	5.3(6)	4.8(4)	3.0(2)	2.0 (1)	15/15
IPOP-10DDr	2.4 (2)	2.0 (2)	1.6 (2)	1.1 (0.7)	2.2 (4)	15/15
IPOP-500 l	2.4 (2)	2.0 (2)	1.6 (2)	1.1 (0.7)	2.2 (4)	15/15
IPOP-tany	2.4 (2)	1.7 (1.0)	1.3 (0.8)	0.91 (0.6)	2.6 (3)	15/15
IPOP-texp	3.9(4)	3.2(2)	2.0 (2)	1.1 (0.9)	3.7(6)	15/15
IPOP lia	2.4 (2)	2.0 (2)	1.6 (2)	1.1 (0.7)	2.2 (4)	15/15
MLSL pal	4.7(1)	6.0(7)	4.9(4)	2.6 (2)	2.7 (3)	15/15
OQNLP pal	25(58)	22(17)	11(8)	4.6(4)	2.6 (1)	15/15
P-DCN tra	1.2 (0.4)	1.1 (1.0)	1.1 (0.9)	5.7(13)	3109(8446)	13/15
P-zero tra	1.2 (0.4)	1.1 (1.0)	1.3 (2)	4.9(6)	1847(5461)	15/15
SMAC hut	1.7 (2)	1.2 (1)	0.93 (0.9)	1.4 (1)	2.4 (2)	9/15
U-DCN tra	1.2 (0.4)	1 (1)	0.92 (0.7)	2.0 (3)	5.5(7)	15/15
U-zero tra	1.2 (0.4)	1.0 (1)	1.3 (1)	1.3 (1)	8.9(13)	15/15
fmincon pa	4.3(1)	5.2(6)	4.9(4)	3.3(2)	2.6 (1)	15/15
fminunc pa	3.5(1)	7.6(9)	5.2(5)	3.0 (2)	2.2 (2)	15/15
ga100 hol	1.8 (2)	2.4 (2)	1.8 (2)	2.7 (3)	5.6(5)	15/15
grid100 ho	1.7 (0.8)	1.5 (2)	0.78 (1)	2.7 (4)	12(9)	15/15
grid16 hol	1.3 (0.8)	1.5 (1.0)	1.5 (2)	4.7(5)	5.2(6)	15/15
hill hol	3.6(5)	3.4(5)	2.7 (3)	2.6 (2)	4.5(5)	15/15
lmmCMA aug	1.6 (2)	1.5 (1)	1.1 (0.8)	0.74 (0.5)	1.0 (1)	15/15
memPSODE v	2.3 (2)	1.4 (1)	1.3 (2)	1.6 (1)	4.5(7)	15/15
prcga saw	1.9 (2)	1.2 (1)	1.4 (1)	1.3 (1)	1.5 (0.8)	15/15
ring100 ho	1.3 (0.8)	1.5 (1)	1.8 (2)	3.1(5)	5.8(5)	15/15
ring16 hol	1.7 (0.8)	2.0 (2)	1.9 (3)	3.1(2)	3.7(3)	15/15
simplex pa	32(61)	25(17)	13(8)	4.8(3)	5.0(4)	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+2:1.1</i>	<i>2.5e+1:3.9</i>	<i>1.6e+1:6.5</i>	<i>4.0e+0:31</i>	<i>2.5e-1:127</i>	5/5
BIPOP-aCMA	1.5 (0.9)	2.2 (3)	2.3 (2)	3.5(2)	3.3(4)	15/15
BIPOP-saAC	1.4 (0.4)	2.5 (4)	2.2 (2)	1.5 (2)	1.6 (0.9)	15/15
CMAES hut	1.4 (0.4)	1.7 (1)	2.8 (3)	3.5(3)	2.5 (2)	8/15
DE pal	1.6 (0.9)	2.3 (2)	3.4(5)	2.5 (3)	5.3(4)	15/15
HCMA los	5.1(9)	3.9(2)	2.7 (2)	2.7 (2)	5.7(5)	15/15
HMLSL pal	2.0 (3)	6.0(8)	8.0(7)	3.1(3)	5.4(8)	15/15
IPOP-10DDr	1.5 (0.9)	2.0 (3)	1.8 (2)	1.6 (2)	2.3 (4)	15/15
IPOP-500 l	1.5 (0.9)	2.0 (3)	1.8 (2)	1.6 (2)	2.3 (4)	15/15
IPOP-tany	1.8 (0.9)	2.1 (2)	1.6 (2)	1.7 (1)	1.5 (0.8)	15/15
IPOP-texp	1.5 (0.9)	1.8 (2)	2.1 (2)	1.7 (1)	3.2(4)	15/15
IPOP lia	1.5 (0.9)	2.0 (3)	1.8 (2)	1.6 (2)	2.3 (4)	15/15
MLSL pal	2.0 (3)	6.4(9)	8.5(7)	4.6(4)	7.1(8)	15/15
OQNLP pal	2.6 (3)	20(22)	21(19)	6.4(4)	8.5(6)	14/15
P-DCN tra	1.5 (0.9)	1.3 (1)	1.6 (1)	4583(20)	1126(5)	14/15
P-zero tra	1.5 (0.9)	1.3 (1)	1.5 (2)	7370(2e4)	1836(4792)	14/15
SMAC hut	1.2 (0.4)	1.8 (2)	1.6 (1)	0.88 (0.7)	1.8 (2)	9/15
U-DCN tra	1.5 (0.9)	1.4 (0.9)	1.6 (2)	1.9 (2)	6.3(13)	15/15
U-zero tra	1.5 (0.9)	1.3 (1)	1.3 (1)	1.8 (2)	11(13)	15/15
fmincon pa	2.0 (3)	6.2(8)	8.7(9)	4.6(6)	7.6(9)	15/15
fminunc pa	1.6 (1)	13(17)	14(12)	4.1(3)	17(14)	15/15
ga100 hol	1.3 (0.9)	2.2 (2)	1.8 (1)	3.3(5)	10(7)	15/15
grid100 ho	1.3 (0.9)	2.3 (2)	1.9 (2)	2.4 (2)	11(12)	15/15
grid16 hol	1.5 (0.4)	1.5 (2)	2.0 (2)	2.8 (5)	24(32)	15/15
hill hol	1.5 (0.9)	3.2(7)	2.8 (4)	2.1 (2)	18(45)	15/15
lmmCMA aug	1.6 (0.9)	2.1 (2)	2.1 (3)	1.4 (1)	2.2 (2)	14/15
memPSODE v	1.5 (0.9)	2.1 (2)	2.5 (2)	4.3(5)	5.9(7)	15/15
prcga saw	1.5 (0.9)	1.5 (1)	1.2 (1)	2.3 (2)	4.3(2)	15/15
ring100 ho	1.4 (0.9)	2.6 (3)	2.0 (2)	3.3(6)	13(17)	15/15
ring16 hol	1.2 (0.4)	0.97 (1)	2.0 (3)	2.3 (3)	8.5(7)	15/15
simplex pa	23(56)	33(8)	21(6)	5.5(2)	2.7 (2)	15/15

Table 18: 02-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>4.0e+1</i> :1.2	<i>1.0e+1</i> :2.7	<i>4.0e+0</i> :10	<i>2.5e+0</i> :28	<i>1.6e-1</i> :119	5/5
BIPOP-aCMA	1.7 (0.4)	3.8(2)	6.1(3)	2.4 (1)	1.6 (0.9)	15/15
BIPOP-saAC	1.8 (2)	10(8)	4.6(3)	1.8 (1)	1.5 (0.9)	15/15
CMAES hut	2.0 (3)	3.2(5)	2.2 (3)	0.98 (1)	1.0 (0.6)	14/15
DE pal	1.8 (1)	2.9 (3)	2.1 (2)	1.2 (0.9)	2.1 (0.5)	15/15
HCMA los	1.7 (2)	2.4 (2)	4.9(2)	2.5 (2)	2.4 (2)	15/15
HMLSL pal	2.3 (2)	12(17)	10(15)	4.6(6)	17(18)	15/15
IPOP-10DDr	1.5 (0.8)	2.4 (2)	1.6 (1)	2.4 (0.6)	3.5(4)	15/15
IPOP-500 l	1.5 (0.8)	2.4 (2)	1.6 (1)	2.4 (0.6)	3.5(4)	15/15
IPOP-tany	1.9 (0.8)	3.0(4)	1.6 (0.9)	0.88 (0.7)	2.6 (3)	15/15
IPOP-texp	4.6(10)	29(70)	16(25)	7.2(10)	6.1(6)	15/15
IPOP lia	1.5 (0.8)	2.4 (2)	1.6 (1)	2.4 (0.6)	3.5(4)	15/15
MLSL pal	2.3 (2)	12(16)	10(17)	4.7(6)	59(49)	15/15
OQNLP pal	11(4)	25(27)	13(2)	5.5(2)	4.8(3)	15/15
P-DCN tra	1.3 (0.8)	7.0(6)	18(33)	9.2(15)	6.0(7)	15/15
P-zero tra	1.3 (0.8)	2.6 (2)	5.8(7)	3.2(6)	16(32)	15/15
SMAC hut	1.1 (0.4)	1.5 (0.9)	1.2 (2)	0.74 (0.5)	24(28)	1/15
U-DCN tra	1.3 (0.8)	2.9 (3)	2.6 (2)	1.2 (0.9)	6.3(7)	15/15
U-zero tra	1.3 (0.8)	2.6 (3)	1.9 (1)	1.0 (0.8)	9.1(11)	15/15
fmincon pa	2.3 (2)	12(17)	7.9(11)	3.6(4)	59(49)	15/15
fminunc pa	2.3 (2)	5.3(5)	4.8(4)	2.4 (2)	23(30)	15/15
ga100 hol	1.4 (0.8)	1.7 (2)	1.8 (2)	1.4 (1)	10(5)	15/15
grid100 ho	1.2 (0.8)	1.9 (2)	2.2 (2)	2.2 (2)	21(17)	15/15
grid16 hol	1.8 (0.8)	1.9 (2)	3.4(4)	2.2 (2)	10(11)	15/15
hill hol	1.7 (0.8)	4.1(8)	2.2 (3)	0.97 (1.0)	11(16)	15/15
lmmCMA aug	1.3 (0.8)	2.6 (2)	1.4 (2)	0.80 (0.6)	1.2 (1)	15/15
memPSODE v	2.3 (1)	2.4 (2)	2.5 (2)	1.8 (2)	6.2(14)	15/15
prcga saw	1.1 (0.4)	1.6 (2)	1.4 (1)	0.84 (0.9)	3.6(6)	15/15
ring100 ho	1.2 (0.4)	3.0 (3)	2.3 (2)	2.6 (4)	20(14)	15/15
ring16 hol	1.1 (0.4)	2.5 (3)	3.5(3)	2.2 (1)	9.0(5)	15/15
simplex pa	10(10)	18(18)	11(2)	4.6(1)	13(11)	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</i> :1.2	<i>1.0e+2</i> :3.2	<i>4.0e+1</i> :7.2	<i>6.3e+0</i> :32	<i>1.6e+0</i> :104	5/5
BIPOP-aCMA	1.3 (0.8)	1.6 (2)	1.7 (1)	1.6 (1)	2.5 (4)	15/15
BIPOP-saAC	3.2(3)	4.1(5)	2.4 (3)	3.9(7)	4.6(3)	15/15
CMAES hut	1.4 (0.4)	1.7 (2)	1.1 (1)	1.2 (2)	1.4 (2)	11/15
DE pal	1.4 (0.8)	1.2 (1)	1.5 (2)	1.5 (1)	1.7 (0.7)	15/15
HCMA los	1.6 (2)	3.7(4)	2.2 (2)	3.0 (2)	3.4(3)	15/15
HMLSL pal	9.4(1)	12(23)	11(14)	11(12)	17(22)	15/15
IPOP-10DDr	1.5 (0.8)	1.9 (1)	1.5 (2)	3.4(3)	6.7(11)	15/15
IPOP-500 l	1.5 (0.8)	1.9 (1)	1.5 (2)	3.4(3)	6.7(11)	15/15
IPOP-tany	1.8 (1)	2.2 (2)	1.6 (2)	3.0 (2)	6.2(8)	15/15
IPOP-texp	9.4(2)	18(14)	18(46)	7.1(16)	6.5(8)	15/15
IPOP lia	1.5 (0.8)	1.9 (1)	1.5 (2)	3.4(3)	6.7(11)	15/15
MLSL pal	9.4(1)	12(23)	11(13)	25(29)	30(34)	15/15
OQNLP pal	9.2(4)	19(21)	15(10)	5.7(2)	2.9 (1)	15/15
P-DCN tra	1.2 (0.4)	1.8 (2)	8.8(11)	4.4(5)	276(148)	15/15
P-zero tra	1.2 (0.4)	1.7 (1)	3.7(7)	3.2(4)	2654(5128)	14/15
SMAC hut	1.3 (1)	1.7 (2)	1.6 (2)	1.2 (1)	3.1(3)	7/15
U-DCN tra	1.2 (0.4)	2.0 (3)	2.2 (2)	2.3 (2)	8.5(16)	15/15
U-zero tra	1.2 (0.4)	1.4 (1)	1.2 (0.8)	2.4 (2)	4.2(5)	15/15
fmincon pa	8.7(1)	13(21)	10(13)	14(5)	17(21)	15/15
fminunc pa	3.7(2)	4.8(6)	4.0(4)	3.3(2)	4.8(4)	15/15
ga100 hol	1.6 (0.8)	1.8 (0.9)	1.5 (1)	2.5 (4)	3.6(3)	15/15
grid100 ho	1.1 (0.4)	1.0 (0.6)	1.5 (2)	5.0(6)	11(11)	15/15
grid16 hol	1.3 (0.8)	1.2 (0.9)	0.87 (0.6)	2.5 (2)	5.2(6)	15/15
hill hol	1.1 (0.8)	2.6 (3)	2.0 (2)	3.9(4)	6.4(7)	15/15
lmmCMA aug	1.4 (1)	6.9(2)	3.6(1)	1.3 (0.5)	1.5 (2)	15/15
memPSODE v	2.0 (2)	1.6 (1)	1.3 (1)	4.0(2)	9.3(9)	15/15
prcga saw	1.2 (0.8)	1.5 (2)	1.6 (1)	1.4 (1)	2.4 (1)	15/15
ring100 ho	1.4 (0.8)	1.6 (2)	1.8 (2)	3.3(4)	7.4(7)	15/15
ring16 hol	1.3 (0.4)	1.7 (1)	1.8 (2)	3.0 (3)	3.9(4)	15/15
simplex pa	8.7(10)	16(16)	12(6)	5.6(2)	3.9(2)	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
BIPOP-aCMA	8.2(13)	8.9(11)	13(11)	21(32)	18(27)	15/15
BIPOP-saAC	4.0(3)	4.6(8)	5.1(6)	6.3(9)	13(18)	15/15
CMAES hut	8.3(10)	11(14)	8.6(9)	13(15)	14(15)	1/15
DE pal	8.0(6)	8.4(7)	7.3(5)	10(8)	71(95)	11/15
HCMA los	9.4(16)	10(13)	22(18)	28(30)	17(24)	15/15
HMLS pal	0.47 (0)	0.64 (0)	0.61 (0)	0.58 (0)	0.15 (0)	15/15
IPOP-10DDr	9.3(3)	13(31)	14(23)	16(24)	15(25)	15/15
IPOP-500 l	9.3(3)	13(31)	14(23)	16(24)	16(21)	15/15
IPOP-tany	23(19)	28(26)	25(21)	31(28)	26(35)	15/15
IPOP-texp	13(19)	21(25)	16(18)	19(25)	12(12)	15/15
IPOP lia	9.3(3)	13(31)	14(23)	16(24)	16(21)	15/15
MLSL pal	0.47 (0)	0.64 (0)	0.61 (0)	0.58 (0)	0.15 (0)	15/15
OQNLP pal	0.47 (0)	0.45 (0)	0.32 (0)	0.30 (0)	8.7(8)	13/15
P-DCN tra	11(12)	11(10)	8.8(7)	11(11)	6.0e4(7e4)	2/15
P-zero tra	8.5(15)	10(17)	8.9(12)	13(11)	1.3e5(1e5)	1/15
SMAC hut	5.3(5)	6.5(6)	5.9(6)	8.7(9)	6.5 (7)	2/15
U-DCN tra	10(11)	12(14)	18(28)	19(26)	53(69)	15/15
U-zero tra	11(13)	11(12)	19(28)	28(27)	117(134)	15/15
fmincon pa	0.47 (0)	0.64 (0)	0.61 (0)	0.58 (0)	0.15 (0)	15/15
fminunc pa	0.57 (0.1)	0.55 (0.1)	0.38 (0.1)	0.37 (0.1)	8.9(10)	15/15
ga100 hol	13(12)	15(13)	16(14)	27(28)	20(15)	15/15
grid100 ho	17(18)	22(21)	25(18)	33(33)	22(17)	15/15
grid16 hol	11(10)	10(12)	11(8)	13(10)	24(53)	15/15
hill hol	7.0(8)	10(8)	11(11)	19(21)	32(40)	15/15
lmmCMA aug	5.2(11)	5.1(10)	6.1(8)	5.8(8)	6.1 (7)	11/15
memPSODE v	14(11)	14(11)	17(40)	17(38)	47(106)	15/15
prcga saw	2.4 (2)	2.9 (3)	2.3 (2)	4.5(4)	11(14)	15/15
ring100 ho	13(13)	18(21)	25(24)	35(25)	19(18)	15/15
ring16 hol	10(10)	12(11)	12(9)	24(33)	20(24)	15/15
simplex pa	0.30 (0.0) \downarrow_4^*	0.33 (0.0) \downarrow_4^*	0.26 (0.0) \downarrow_4^{*2}	0.26 (0.0) \downarrow_4^{*2}	9.2(9)	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
f20	<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
BIPOP-aCMA	2.2 (4)	4.5(6)	4.0(3)	2.3 (2)	6.7(12)	15/15
BIPOP-saAC	1.8 (2)	2.3 (3)	2.6 (2)	1.6 (2)	3.2 (2)	15/15
CMAES hut	2.4 (3)	2.3 (2)	2.0 (3)	1.6 (2)	2.5 (3)	7/15
DE pal	1.8 (2)	2.9 (3)	3.3(4)	3.2(2)	23(2)	14/15
HCMA los	2.5 (2)	2.2 (1)	1.6 (0.6)	1.1 (1)	3.4(3)	15/15
HMLSL pal	2.7 (0)	2.9 (0)	2.2 (0)	1.4 (0)	2.5 (1)	15/15
IPOP-10DDr	2.2 (2)	3.9(4)	5.5(3)	1.9 (0.7)	10(20)	15/15
IPOP-500 l	2.2 (2)	3.9(4)	5.5(3)	1.9 (0.7)	7.8(13)	15/15
IPOP-tany	1.8 (1)	1.6 (2)	1.2 (1)	2.0 (1)	10(12)	15/15
IPOP-texp	2.5 (2)	3.1(3)	1.8 (0.9)	1.3 (0.8)	9.3(13)	15/15
IPOP lia	2.2 (2)	3.9(4)	5.5(3)	1.9 (0.7)	7.8(13)	15/15
MLSL pal	2.7 (0)	2.9 (0)	2.2 (0)	1.4 (0)	2.9 (1)	15/15
OQNLP pal	5.4(0)	4.6(0)	5.9(0)	9.2(4)	5.2(4)	15/15
P-DCN tra	1.1 (0.8)	1.4 (0.9)	6.1(6)	5.6(5)	8854(1e4)	12/15
P-zero tra	1.1 (0.8)	1.4 (0.9)	4.1(3)	1.9 (1)	2.2e4(3e4)	6/15
SMAC hut	1.5 (0.5)	1.3 (0.7)	1.9 (3)	1.8 (2)	6.4(7)	3/15
U-DCN tra	1.1 (0.8)	1.4 (0.9)	4.9(4)	3.7(3)	6.5(6)	15/15
U-zero tra	1.1 (0.8)	1.4 (0.9)	3.1(2)	2.4 (3)	6.4(9)	15/15
fmincon pa	2.7 (0)	2.9 (0)	2.2 (0)	1.4 (0)	2.1 (1)	15/15
fminunc pa	2.1 (0)	1.4 (0)	1.1 (0)	6.7(3)	3.5(3)	15/15
ga100 hol	1.2 (0.8)	2.1 (2)	2.3 (2)	3.5(3)	6.0(4)	15/15
grid100 ho	1.8 (1)	3.2(4)	4.5(5)	7.0(8)	13(11)	15/15
grid16 hol	1.3 (1)	2.4 (2)	4.0(6)	2.3 (2)	6.7(9)	15/15
hill hol	2.4 (3)	5.2(6)	4.0(4)	2.3 (2)	4.3(6)	15/15
lmmCMA aug	2.5 (0.8)	2.1 (1)	1.8 (1)	1.8 (0.9)	8.6(10)	11/15
memPSODE v	2.5 (1)	3.6(1)	3.5(3)	3.1(2)	3.8(6)	15/15
prcga saw	1.1 (1)	1.2 (1)	2.2 (3)	1.8 (2)	8.6(7)	15/15
ring100 ho	1.1 (0.5)	2.0 (3)	3.5(5)	5.2(5)	10(8)	15/15
ring16 hol	0.96 (0.5)	1.9 (2)	4.1(4)	2.9 (3)	4.7(4)	15/15
simplex pa	14(0.3)	12(0.2)	5.6(0.5)	3.4(4)	6.3(5)	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
f21	<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
BIPOP-aCMA	1.5 ^(0.9)	1.6 ⁽¹⁾	1.7 ⁽²⁾	6.2 ⁽⁹⁾	10 ⁽¹⁸⁾	15/15
BIPOP-saAC	1.5 ⁽²⁾	2.3 ⁽³⁾	1.8 ⁽¹⁾	5.3 ⁽⁷⁾	6.8 ⁽¹⁰⁾	15/15
CMAES hut	1.2 ⁽¹⁾	2.2 ⁽³⁾	1.7 ⁽¹⁾	2.7 ⁽⁴⁾	1.6 ⁽²⁾	9/15
DE pal	2.0 ⁽²⁾	2.3 ⁽²⁾	1.1 ⁽¹⁾	1.7 ⁽²⁾	96 ⁽¹⁹¹⁾	12/15
HCMA los	2.1 ⁽²⁾	2.6 ⁽²⁾	1.3 ⁽¹⁾	3.0 ⁽⁶⁾	7.8 ⁽⁹⁾	15/15
HMLSL pal	2.3 ⁽³⁾	2.8 ⁽²⁾	1.6 ⁽¹⁾	2.1 ⁽³⁾	2.1 ⁽²⁾	15/15
IPOP-10DDr	2.6 ⁽²⁾	2.6 ⁽³⁾	2.3 ⁽²⁾	4.1 ⁽¹⁰⁾	12 ⁽¹⁴⁾	15/15
IPOP-500 l	2.6 ⁽²⁾	2.6 ⁽³⁾	2.3 ⁽²⁾	4.1 ⁽¹⁰⁾	11 ⁽¹⁴⁾	15/15
IPOP-tany	2.2 ⁽²⁾	2.8 ⁽³⁾	1.9 ⁽¹⁾	6.5 ⁽¹²⁾	12 ⁽¹³⁾	15/15
IPOP-texp	2.7 ⁽²⁾	7.0 ⁽³⁾	7.2 ⁽²⁾	12 ⁽¹²⁾	30 ⁽⁴¹⁾	15/15
IPOP lia	2.6 ⁽²⁾	2.6 ⁽³⁾	2.3 ⁽²⁾	4.1 ⁽¹⁰⁾	11 ⁽¹⁴⁾	15/15
MLSL pal	2.3 ⁽³⁾	2.8 ⁽²⁾	1.6 ⁽¹⁾	1.6 ⁽²⁾	1.7 ⁽²⁾	15/15
OQNLP pal	4.4 ⁽⁶⁾	7.1 ⁽⁴⁾	4.5 ⁽⁷⁾	2.5 ⁽²⁾	1.4 ⁽¹⁾	15/15
P-DCN tra	1.4 ⁽¹⁾	3.2 ⁽⁴⁾	1135 ⁽³⁾	3.3e4 ^(7e4)	2.2e4 ^(3e4)	7/15
P-zero tra	1.4 ⁽¹⁾	2.1 ⁽³⁾	1.8e4 ⁽⁶⁾	4.4e4 ^(7e4)	2.2e4 ^(3e4)	7/15
SMAC hut	0.92 ^(0.9)	1.7 ⁽²⁾	1.6 ⁽²⁾	1.4 ⁽¹⁾	0.74 ^(0.5)	15/15
U-DCN tra	1.4 ⁽¹⁾	1.8 ⁽²⁾	1.4 ^(0.9)	0.85 ^(0.8)	0.67 ⁽¹⁾	15/15
U-zero tra	1.4 ⁽¹⁾	2.0 ⁽²⁾	1.6 ⁽²⁾	1.1 ⁽²⁾	0.82 ⁽¹⁾	15/15
fmincon pa	2.3 ⁽³⁾	2.8 ⁽²⁾	1.6 ⁽¹⁾	1.2 ⁽¹⁾	1.3 ⁽¹⁾	15/15
fminunc pa	3.0 ⁽³⁾	4.2 ⁽⁵⁾	2.8 ⁽³⁾	1.7 ⁽¹⁾	0.88 ^(0.4)	15/15
ga100 hol	1 ^(0.6)	2.3 ⁽³⁾	1.8 ⁽²⁾	1.8 ⁽²⁾	1.4 ⁽²⁾	15/15
grid100 ho	1.1 ^(0.9)	1.9 ⁽²⁾	1.2 ^(0.9)	1.1 ⁽¹⁾	2.3 ⁽⁴⁾	15/15
grid16 hol	1.1 ^(0.6)	1.4 ⁽¹⁾	1.8 ⁽²⁾	1.9 ⁽²⁾	21 ⁽²²⁾	15/15
hill hol	1.7 ⁽²⁾	2.9 ⁽³⁾	2.0 ⁽²⁾	5.2 ⁽⁷⁾	161 ⁽⁴⁷⁶⁾	13/15
ImmCMA aug	1 ^(0.9)	0.92 ^(0.6)	0.88 ^(0.9)	2.4 ⁽⁴⁾	4.1 ⁽⁸⁾	13/15
memPSODE v	1.9 ⁽²⁾	2.8 ⁽²⁾	2.1 ⁽²⁾	1.8 ⁽²⁾	2.8 ⁽⁴⁾	15/15
prcga saw	1.6 ⁽¹⁾	1.6 ⁽¹⁾	1.5 ⁽²⁾	2.4 ⁽⁴⁾	14 ⁽²⁰⁾	15/15
ring100 ho	1.8 ^(0.9)	1.8 ⁽¹⁾	1.3 ⁽¹⁾	1.3 ⁽²⁾	1.7 ⁽²⁾	15/15
ring16 hol	1.1 ^(0.9)	1.6 ⁽¹⁾	0.97 ^(0.7)	3.3 ⁽⁴⁾	2.6 ⁽⁴⁾	15/15
simplex pa	15 ⁽²²⁾	17 ⁽²²⁾	10 ⁽⁹⁾	4.8 ⁽⁴⁾	1.9 ^(0.9)	15/15

Table 23: 02-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
<i>f22</i>	<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
BIPOP-aCMA	1.6 (2)	2.5 (3)	2.1 (2)	5.9(11)	4.3(4)	15/15
BIPOP-saAC	1.5 (1)	1.5 (2)	1.5 (2)	10(24)	5.9(9)	15/15
CMAES hut	1.1 (0.4)	1.4 (2)	1.7 (2)	8.3(12)	3.6(4)	4/15
DE pal	1.5 (1)	1.6 (2)	1.1 (1)	2.7 (3)	18(2)	14/15
HCMA los	2.2 (2)	2.4 (2)	1.8 (2)	4.5(7)	3.5(5)	15/15
HMLSL pal	3.0 (4)	2.5 (2)	2.2 (3)	3.5(3)	1.2 (0.8)	15/15
IPOP-10DDr	1.8 (2)	1.7 (1)	1.1 (0.8)	11(14)	10(12)	15/15
IPOP-500 l	1.8 (2)	1.7 (1)	1.1 (0.8)	11(14)	8.5(10)	15/15
IPOP-tany	1.8 (2)	1.3 (2)	1.1 (1)	2.2 (1)	5.9(7)	15/15
IPOP-texp	1.8 (1)	2.3 (2)	4.7(2)	21(35)	11(16)	15/15
IPOP lia	1.8 (2)	1.7 (1)	1.1 (0.8)	11(14)	8.5(10)	15/15
MLSL pal	3.0 (4)	2.5 (2)	2.2 (3)	2.8 (3)	0.96 (0.8)	15/15
OQNLP pal	5.6(5)	4.5(4)	3.4(3)	4.3(5)	1.4 (0.8)	15/15
P-DCN tra	1.2 (0.8)	1.7 (1)	2.4 (5)	1.3e4(4e4)	1.1e4(1e4)	8/15
P-zero tra	1.2 (0.8)	1.3 (1)	1.5e4(5)	1.5e4(4e4)	1.1e4(2e4)	8/15
SMAC hut	1.3 (0.8)	1.3 (1)	1.3 (1)	0.74 (0.9)	0.91 (0.9)	11/15
U-DCN tra	1.2 (0.8)	1.1 (1)	1.3 (1)	2.2 (3)	1.3 (2)	15/15
U-zero tra	1.2 (0.8)	1.1 (1)	0.88 (0.9)	2.0 (2)	2.4 (2)	15/15
fmincon pa	3.0 (4)	2.5 (2)	2.5 (4)	3.9(3)	0.87 (0.7)	15/15
fminunc pa	2.8 (2)	2.2 (1)	3.0 (4)	3.0 (2)	0.88 (0.7)	15/15
ga100 hol	1.5 (0.8)	1.0 (0.5)	1.5 (1)	1.5 (1)	2.3 (3)	15/15
grid100 ho	1.7 (0.8)	1.3 (0.9)	0.59 (0.4)	1.5 (2)	2.8 (3)	15/15
grid16 hol	1.9 (2)	1.5 (1)	1.6 (2)	3.0(3)	2.0 (2)	15/15
hill hol	3.4(2)	2.0 (2)	1.8 (2)	287(8)	45(3)	14/15
lmmCMA aug	1.8 (1)	1.9 (2)	2.0 (0.9)	3.0 (5)	2.5 (4)	14/15
memPSODE v	2.4 (2)	1.9 (2)	2.6 (3)	2.9 (2)	2.3 (3)	15/15
prcga saw	1.2 (0.8)	0.98 (1)	1.2 (1)	1.8 (2)	4.3(8)	15/15
ring100 ho	0.95 (0.4)	1.1 (0.8)	1.5 (1)	2.8 (3)	2.4 (3)	15/15
ring16 hol	1.3 (1)	1.7 (1)	1.8 (2)	5.5(10)	3.3(3)	15/15
simplex pa	15(13)	16(25)	10(8)	7.2(4)	1.5 (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</i> :1.5	<i>2.5e+1</i> :2.6	<i>1.0e+1</i> :7.8	<i>4.0e+0</i> :55	<i>2.5e+0</i> :103	5/5
BIPOP-aCMA	1.3 (1)	1.2 (1)	1.8 (2)	4.0(6)	4.4(7)	15/15
BIPOP-saAC	1.1 (0.7)	1.1 (1.0)	2.0 (2)	1.7 (2)	1.8 (2)	15/15
CMAES hut	1.3 (1.0)	1.5 (2)	2.0 (2)	2.5 (3)	8.4(10)	3/15
DE pal	1.6 (1.0)	1.6 (1)	2.3 (2)	1.8 (2)	2.1 (2)	15/15
HCMA los	1.9 (2)	2.4 (2)	2.4 (2)	2.0 (2)	3.3(5)	15/15
HMLSL pal	2.7 (2)	11(24)	4.6(8)	2.8 (3)	2.1 (2)	15/15
IPOP-10DDr	1.3 (1.0)	1.1 (0.8)	1.8 (2)	1.3 (2)	3.1(4)	15/15
IPOP-500 l	1.3 (1.0)	1.1 (0.8)	1.8 (2)	1.3 (2)	3.1(4)	15/15
IPOP-tany	1.5 (1.0)	1.4 (1)	2.4 (2)	1.9 (2)	1.7 (1)	15/15
IPOP-texp	1.7 (2)	1.3 (1.0)	2.1 (1)	2.6 (3)	3.4(4)	15/15
IPOP lia	1.3 (1.0)	1.1 (0.8)	1.8 (2)	1.3 (2)	3.1(4)	15/15
MLSL pal	2.7 (2)	11(24)	4.5(8)	2.1 (2)	1.6 (1)	15/15
OQNLP pal	3.3(3)	19(34)	9.1(11)	4.1(3)	2.9 (2)	15/15
P-DCN tra	1.3 (1.0)	0.97 (0.8)	1.7 (2)	1.7 (2)	1.4 (1)	15/15
P-zero tra	1.3 (1.0)	0.97 (0.8)	1.1 (0.8)	2.1 (3)	1.5 (2)	15/15
SMAC hut	1.3 (1.0)	1.6 (2)	2.0 (2)	1.8 (2)	2.5 (3)	8/15
U-DCN tra	1.3 (1.0)	0.97 (0.8)	1.7 (2)	1.5 (2)	2.2 (2)	15/15
U-zero tra	1.3 (1.0)	0.97 (0.8)	1.7 (2)	1.1 (1)	3.0 (3)	15/15
fmincon pa	2.7 (2)	8.3(22)	4.5(9)	2.2 (2)	2.4 (2)	15/15
fminunc pa	1.8 (2)	12(18)	5.2(8)	2.0 (2)	2.2 (3)	15/15
ga100 hol	2.1 (2)	1.9 (2)	1.7 (2)	1.8 (2)	2.9 (4)	15/15
grid100 ho	1.6 (1)	2.8 (3)	2.5 (3)	1.8 (2)	1.9 (2)	15/15
grid16 hol	1.3 (0.7)	1.4 (1)	1.7 (1)	3.4(7)	3.7(5)	15/15
hill hol	1.3 (1.0)	1.6 (1)	3.4(3)	3.5(5)	3.3(4)	15/15
lmmCMA aug	1.7 (1)	1.4 (1)	2.1 (2)	2.8 (2)	4.3(4)	15/15
memPSODE v	1.4 (1.0)	1.2 (1.0)	1.6 (2)	3.5(6)	3.0 (3)	15/15
prcga saw	1.6 (1.0)	1.4 (1.0)	2.5 (3)	2.2 (2)	2.3 (2)	15/15
ring100 ho	1.5 (1.0)	1.8 (2)	2.3 (2)	1.9 (2)	2.9 (2)	15/15
ring16 hol	1.2 (1)	1.2 (1.0)	2.0 (2)	2.5 (4)	3.3(4)	15/15
simplex pa	64(84)	54(50)	23(19)	4.8(3)	2.9 (2)	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>f24</i>	<i>4.0e+1:1.1</i>	<i>2.5e+1:2.7</i>	<i>1.6e+1:7.7</i>	<i>6.3e+0:44</i>	<i>2.5e+0:275</i>	5/5
BIPOP-aCMA	1.4 (0.9)	2.7 (3)	1.9 (2)	1.6 (1)	3.0(4)	15/15
BIPOP-saAC	1.5 (0.9)	2.8 (4)	1.7 (2)	1.4 (1)	3.6(2)	15/15
CMAES hut	2.2 (2)	2.6 (2)	3.0(3)	1.9 (3)	1.2 (1)	7/15
DE pal	1.5 (0.9)	1.4 (1)	1.3 (1)	1.8 (2)	1.8 (2)	15/15
HCMA los	2.7 (3)	2.1 (2)	1.3 (1)	1.6 (2)	2.5 (2)	15/15
HMLSL pal	2.1 (2)	4.3(4)	3.8(5)	2.2 (2)	1.7 (2)	15/15
IPOP-10DDr	1.2 (0.9)	0.95 (0.8)	1.6 (2)	1.4 (1)	6.7(7)	15/15
IPOP-500 l	1.2 (0.9)	0.95 (0.8)	1.6 (2)	1.4 (1)	6.6(7)	15/15
IPOP-tany	1.2 (0.9)	0.90 (0.9)	1.1 (0.6)	1.8 (0.8)	7.1(11)	15/15
IPOP-texp	1.2 (0.9)	1.0 (0.9)	1.9 (1.0)	1.4 (1)	5.2(7)	15/15
IPOP lia	1.2 (0.9)	0.95 (0.8)	1.6 (2)	1.4 (1)	6.6(7)	15/15
MLSL pal	2.1 (2)	4.2(4)	3.4(4)	2.6 (2)	1.6 (2)	15/15
OQNLP pal	2.8 (3)	4.0(3)	2.3 (2)	5.1(2)	2.6 (2)	15/15
P-DCN tra	1.6 (0.9)	2.3 (2)	1.9 (2)	2.4 (2)	26(29)	15/15
P-zero tra	1.6 (0.9)	2.2 (2)	2.3 (3)	9.2(1)	42(54)	15/15
SMAC hut	1.1 (0.5)	1.5 (1)	1.6 (1)	2.1 (2)	3.6(4)	3/15
U-DCN tra	1.6 (0.9)	2.3 (2)	1.9 (2)	1.9 (3)	3.2(3)	15/15
U-zero tra	1.6 (0.9)	2.3 (2)	1.5 (1.0)	1.6 (1)	8.5(6)	15/15
fmincon pa	2.1 (2)	4.2(4)	3.3(4)	2.1 (2)	1.9 (3)	15/15
fminunc pa	2.6 (4)	3.3(3)	2.6 (2)	2.3 (2)	1.9 (2)	15/15
ga100 hol	1.6 (0.9)	1.2 (1)	1.1 (1)	2.5 (3)	8.9(11)	15/15
grid100 ho	1.5 (0.5)	1.4 (1)	1.6 (1)	2.0 (1)	6.4(8)	15/15
grid16 hol	1.8 (0.9)	1.8 (1)	1.5 (2)	2.1 (2)	8.8(13)	15/15
hill hol	1.2 (0.9)	2.4 (3)	2.0 (2)	1.7 (1)	5.5(7)	15/15
lmmCMA aug	1.2 (0.9)	1.5 (1)	1.2 (0.8)	1.6 (2)	1.2 (0.8)	15/15
memPSODE v	3.7(4)	2.6 (2)	3.1(4)	1.9 (1)	5.6(10)	15/15
prcga saw	1.6 (1)	1.7 (2)	1.0 (0.9)	1.8 (2)	7.8(14)	15/15
ring100 ho	2.1 (2)	1.6 (1)	1.4 (1)	2.1 (2)	5.9(6)	15/15
ring16 hol	1.4 (0.5)	1.1 (1)	1.5 (2)	2.4 (2)	3.8(6)	15/15
simplex pa	4.2(7)	9.1(10)	5.2(3)	5.0(1)	3.3(3)	15/15

References

- [1] Anne Auger, Dimo Brockhoff, and Nikolaus Hansen. Benchmarking the local metamodel CMA-ES on the noiseless BBOB'2013 test bed. In Blum and Alba [3], pages 1225–1232.
- [2] Anne Auger, Steffen Finck, Nikolaus Hansen, and Raymond Ros. BBOB 2009: Comparison tables of all algorithms on all noiseless functions. Technical Report RT-0383, INRIA, April 2010.
- [3] Christian Blum and Enrique Alba, editors. *Genetic and Evolutionary Computation Conference, GECCO '13, Amsterdam, The Netherlands, July 6-10, 2013, Companion Material Proceedings*. ACM, 2013.
- [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] Neal J. Holtschulte and Melanie Moses. Benchmarking cellular genetic algorithms on the BBOB noiseless testbed. In Blum and Alba [3], pages 1201–1208.
- [8] Frank Hutter, Holger Hoos, and Kevin Leyton-Brown. An evaluation of sequential model-based optimization for expensive blackbox functions. In Blum and Alba [3], pages 1209–1216.
- [9] Tianjun Liao and Thomas Stützle. Bounding the population size of IPOP-CMA-ES on the noiseless BBOB testbed. In Blum and Alba [3], pages 1161–1168.
- [10] Tianjun Liao and Thomas Stützle. Testing the impact of parameter tuning on a variant of IPOP-CMA-ES with a bounded maximum population size on the noiseless BBOB testbed. In Blum and Alba [3], pages 1169–1176.
- [11] Ilya Loshchilov, Marc Schoenauer, and Michèle Sebag. Bi-population CMA-ES algorithms with surrogate models and line searches. In Blum and Alba [3], pages 1177–1184.
- [12] László Pál. Benchmarking a hybrid multi level single linkage algorithm on the BBOB noiseless testbed. In Blum and Alba [3], pages 1145–1152.
- [13] László Pál. Comparison of multistart global optimization algorithms on the BBOB noiseless testbed. In Blum and Alba [3], pages 1153–1160.

- [14] Babatunde A. Sawyerr, Aderemi Oluyinka Adewumi, and Montaz M. Ali. Benchmarking projection-based real coded genetic algorithm on BBOB-2013 noiseless function testbed. In Blum and Alba [3], pages 1193–1200.
- [15] Thanh-Do Tran, Dimo Brockhoff, and Bilel Derbel. Multiobjectivization with NSGA-II on the noiseless BBOB testbed. In Blum and Alba [3], pages 1217–1224.
- [16] Costas Voglis. Adapt-MEMPSODE: a memetic algorithm with adaptive selection of local searches. In Blum and Alba [3], pages 1137–1144.