Comparison Tables: BBOB 2013 Testbed in 40-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 ($\mathrm{ERT}_{\mathrm{best}}$ 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 in 2013.

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

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

#FEs/D	0.5	1.2	3	10	50	#succ
f1	*2.5e+2*:48	*1.6e+2*:82	*1.0e-8*:83	*1.0e-8*:83	*1.0e-8*:83	15/15
BIPOP-aCMA	3.4(0.9)	**2.6**(0)	**19**(0)	**19**(0)	**19**(0)	15/15
BIPOP-saAC	**1.7**(1.0)	**1.7**(0.6)	**14**(0.8)	**14**(0.8)	**14**(0.8)	15/15
CMAES hut	**1.7**(1)	**1.9**(0.5)	∞	∞	∞ *4007*	0/15
IPOP-10DDr	**0.44**(1)	**2.4**(0.3)	63(2)	63(2)	63(2)	15/15
IPOP-500 l	**0.44**(1)	**2.4**(0.3)	63(2)	63(2)	63(2)	15/15
IPOP-tany	**0.39**(1)	**1.8**(0.5)	67(2)	67(2)	67(2)	15/15
IPOP-texp	**0.11**(0.2)	**0.69**(0.4)$^{\star 2}$	60(2)	60(2)	60(2)	15/15
IPOP lia	**0.44**(1)	**2.4**(0.3)	63(2)	63(2)	63(2)	15/15
ga100 hol	7.9(5)	12(4)	∞	∞	∞ *2e6*	0/15
grid100 ho	43(19)	82(24)	∞	∞	∞ *2e6*	0/15
grid16 hol	10(6)	14(5)	∞	∞	∞ *2e6*	0/15
hill hol	3.7(2)	3.3(1)	∞	∞	∞ *2e6*	0/15
memPSODE v	**2.9**(0)	**1.8**(0)	**1.8**(0)$^{\star 4}$	**1.8**(0)$^{\star 4}$	**1.8**(0)$^{\star 4}$	15/15
prcga saw	**1.8**(3)	5.6(5)	6936(4284)	6936(4284)	6936(4284)	15/15
ring100 ho	14(11)	28(7)	∞	∞	∞ *2e6*	0/15
ring16 hol	5.0(2)	6.2(2)	∞	∞	∞ *2e6*	0/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f2	*1.0e+7*:39	*6.3e+6*:71	*4.0e+5*:121	*2.5e+4*:499	*1.0e-8*:1188	15/15
BIPOP-aCMA	**1.9**(2)	**2.6**(0.6)	**3.5**(0.4)$^{\star 4}$	**1.1**(0.1)$^{\star 4}$	**41**(0.8)	15/15
BIPOP-saAC	**1.2**(0.9)	**1.3**(0.9)	11(5)	**4.3**(0.8)	**7.5**(0.9)	15/15
CMAES hut	**1.2**(0.8)	**1.0**(0.7)	15(3)	∞	∞ *4007*	0/15
IPOP-10DDr	**1.8**(2)	**2.1**(2)	11(3)	11(2)	68(1)	15/15
IPOP-500 l	**1.8**(2)	**2.1**(2)	11(3)	11(2)	68(1)	15/15
IPOP-tany	**0.90**(0.7)	**1.1**(0.9)	11(2)	12(1.0)	66(1)	15/15
IPOP-texp	**0.45**(0.5)	**0.49**(0.5)	**7.2**(2)	10(2)	62(1)	15/15
IPOP lia	**1.8**(2)	**2.1**(2)	11(3)	11(2)	68(1)	15/15
ga100 hol	3.9(4)	6.9(3)	33(6)	21(3)	∞ *2e6*	0/15
grid100 ho	9.0(7)	24(20)	142(19)	91(19)	∞ *2e6*	0/15
grid16 hol	6.1(6)	7.8(4)	24(5)	16(4)	∞ *2e6*	0/15
hill hol	7.6(4)	5.0(2)	**5.9**(2)	**4.9**(3)	∞ *2e6*	0/15
memPSODE v	3.0(2)	**2.2**(1)	8.9(4)	6.5(0.5)	**16**(0.5)	15/15
prcga saw	**1.2**(1)	**2.0**(1)	13(2)	8.4(2)	1.1e4(1e4)	4/15
ring100 ho	4.1(7)	11(10)	66(13)	45(6)	∞ *2e6*	0/15
ring16 hol	4.4(4)	4.5(2)	14(3)	10(1)	∞ *2e6*	0/15

Table 4: 40-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\text{best 2009}}$ on f_3 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\mathrm{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f3	*1.6e+3*:68	*1.0e+3*:222	*6.3e+2*:471	*4.0e+2*:662	*6.3e+1*:6332	15/15
BIPOP-aCMA	**2.5**(0.7)	**0.94**(0)	**0.60**(0)	**0.55**(0)$^{\star 4}$	**0.25**(0.0)$^{\star 4}$	15/15
BIPOP-saAC	**1.5**(0.8)	**0.93**(0.2)	**0.68**(0.1)	**1.5**(0.7)	**1.5**(0.8)	15/15
CMAES hut	**1.2**(0.5)	**1.0**(0.3)	**1.1**(0.3)	**1.7**(0.4)	∞ *4007*	0/15
IPOP-10DDr	**1.2**(1)	**1.1**(0.2)	**0.99**(0.2)	**1.5**(0.3)	3.8(5)	15/15
IPOP-500 l	**1.2**(1)	**1.1**(0.2)	**0.99**(0.2)	**1.5**(0.3)	3.8(5)	15/15
IPOP-tany	**0.65**(0.8)	**0.97**(0.2)	**0.84**(0.2)	**1.5**(0.3)	**2.0**(2)	15/15
IPOP-texp	**0.13**(0.1)	**0.32**(0.2)$^{\star 4}$	**0.50**(0.1)	**1.4**(0.5)	3.3(2)	15/15
IPOP lia	**1.2**(1)	**1.1**(0.2)	**0.99**(0.2)	**1.5**(0.3)	3.8(5)	15/15
ga100 hol	6.7(5)	5.7(1)	5.2(1)	6.0(0.6)	**2.5**(0.3)	15/15
grid100 ho	23(31)	29(17)	25(7)	29(8)	13(3)	15/15
grid16 hol	7.7(6)	6.0(2)	4.7(0.7)	5.1(0.8)	**2.3**(0.3)	15/15
hill hol	**2.9**(2)	**1.3**(0.5)	**0.83**(0.2)	**0.84**(0.2)	**0.44**(0.1)	15/15
memPSODE v	34(2)	11(0.5)	6.7(0.8)	5.6(0.9)	3.8(2)	15/15
prcga saw	**1.9**(1)	**2.6**(1)	3.3(0.9)	4.9(0.9)	**2.4**(0.6)	15/15
ring100 ho	12(7)	13(2)	13(1)	14(1)	5.7(0.4)	15/15
ring16 hol	4.0(2)	**2.9**(0.6)	**2.5**(0.3)	**2.8**(0.3)	**1.1**(0.1)	15/15

Table 5: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_4 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f4	*1.0e+3*:439	*6.3e+2*:670	*4.0e+2*:707	*2.5e+2*:735	*1.0e+2*:5369	15/15
BIPOP-aCMA	**0.56**(0.1)$_{\downarrow 4}$	**0.71**(0.1)	**0.89**(0.1)	**1.2**(0.2)	**0.27**(0.0)$^{\star 4}$	15/15
BIPOP-saAC	**1.0**(0.2)	**1.2**(0.2)	**2.8**(0.6)	**3.6**(0.7)	**2.9**(0.5)	15/15
CMAES hut	**1.1**(0.3)	**1.4**(0.3)	3.0(0.8)	4.3(0.7)	5.5(6)	2/15
IPOP-10DDr	**1.1**(0.2)	**1.4**(0.2)	**2.9**(0.8)	4.5(0.5)	3.3(2)	15/15
IPOP-500 l	**1.1**(0.2)	**1.4**(0.2)	**2.9**(0.8)	4.5(0.5)	3.3(2)	15/15
IPOP-tany	**0.94**(0.2)	**1.4**(0.2)	**3.0**(0.6)	4.7(0.6)	3.4(2)	15/15
IPOP-texp	**0.27**(0.1)$^{\star 2}_{\downarrow 4}$	**0.65**(0.2)	**2.4**(0.9)	3.8(0.6)	3.4(2)	15/15
IPOP lia	**1.1**(0.2)	**1.4**(0.2)	**2.9**(0.8)	4.5(0.5)	3.3(2)	15/15
ga100 hol	4.7(1)	5.3(0.7)	7.6(0.6)	10(1)	**2.7**(0.3)	15/15
grid100 ho	24(6)	26(7)	36(8)	49(7)	14(1)	15/15
grid16 hol	4.3(1.0)	4.5(0.9)	6.3(1.0)	8.8(2)	**2.6**(0.3)	15/15
hill hol	**0.79**(0.1)$_{\downarrow 2}$	**0.69**(0.2)	**0.91**(0.2)	**1.3**(0.3)	**0.43**(0.0)	15/15
memPSODE v	19(10)	13(7)	13(6)	13(6)	3.6(1)	15/15
prcga saw	**1.8**(1)	3.7(0.6)	7.0(1)	11(3)	3.1(0.9)	15/15
ring100 ho	11(2)	13(2)	18(2)	25(2)	6.2(0.3)	15/15
ring16 hol	**2.4**(0.4)	**2.4**(0.3)	3.4(0.4)	4.6(0.5)	**1.2**(0.2)	15/15

Table 6: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_5 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f5	*4.0e+2*:51	*2.5e+2*:81	*1.0e-1*:120	*1.0e-8*:121	*1.0e-8*:121	15/15
BIPOP-aCMA	**1.6**(0)	**1.0**(0)$^{\star 4}$	**0.68**(0)$^{\star 4}$	**0.68**(0)$^{\star 4}$	**0.68**(0)$^{\star 4}$	15/15
BIPOP-saAC	**1.8**(0.3)	**1.8**(0.3)	**2.9**(0.2)	**2.9**(0.2)	**2.9**(0.2)	15/15
CMAES hut	**2.3**(0.7)	**2.2**(0.4)	**4.7**(0.6)	**4.7**(0.6)	**4.7**(0.6)	15/15
IPOP-10DDr	**1.9**(0.9)	3.9(1)	298(415)	657(535)	657(535)	15/15
IPOP-500 l	**1.9**(0.9)	3.9(1)	298(415)	657(535)	657(535)	15/15
IPOP-tany	**1.9**(0.4)	3.3(1)	261(270)	435(305)	435(305)	15/15
IPOP-texp	3.3(0.5)	3.6(1.0)	1.5e4(5e4)	3.3e4(1e5)	3.3e4(1e5)	15/15
IPOP lia	**1.9**(0.9)	3.9(1)	298(415)	657(535)	657(535)	15/15
ga100 hol	19(4)	25(5)	88(8)	91(8)	91(8)	15/15
grid100 ho	83(30)	128(26)	371(62)	376(66)	376(66)	15/15
grid16 hol	23(6)	26(4)	67(8)	67(8)	67(8)	15/15
hill hol	4.2(1)	4.3(0.9)	12(2)	11(2)	11(2)	15/15
memPSODE v	3.8(0.4)	3.3(0.5)	6.0(0.9)	6.1(0.8)	6.1(0.8)	15/15
prcga saw	15(6)	24(14)	∞	∞	∞ *2e6*	0/15
ring100 ho	45(10)	61(11)	190(13)	190(13)	190(13)	15/15
ring16 hol	9.4(2)	11(2)	35(3)	35(5)	35(5)	15/15

Table 7: 40-D, running time excess $\mathrm{ERT}/\mathrm{ERT}_{\mathrm{best\ 2009}}$ on f_6 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\mathrm{ERT}_{\mathrm{best\ 2009}}$ (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
f6	*6.3e+5*:50	*4.0e+5*:82	*4.0e+4*:127	*4.0e+2*:734	*6.3e+1*:2121	15/15
BIPOP-aCMA	**2.9**(1)	**2.7**(1)	3.6(1)	**3.0**(0.5)	**1.8**(0.3)	15/15
BIPOP-saAC	**1.3**(0.9)	**1.2**(0.5)	**1.7**(0.4)	**1.9**(0.5)	**1.4**(0.2)	15/15
CMAES hut	**1.4**(0.8)	**1.2**(0.6)	**2.1**(0.6)	**2.8**(0.5)	**2.4**(1)	11/15
IPOP-10DDr	**1.2**(1)	**1.8**(0.9)	**2.9**(0.9)	**1.3**(0.3)	**1.2**(0.2)	15/15
IPOP-500 l	**1.2**(1)	**1.8**(0.9)	**2.9**(0.9)	**1.3**(0.3)	**1.2**(0.2)	15/15
IPOP-tany	**0.89**(0.8)	**1.1**(0.6)	**2.4**(0.7)	**1.2**(0.3)	**1.3**(0.1)	15/15
IPOP-texp	**0.53**(0.5)	**0.78**(0.5)	**2.0**(0.9)	**1.1**(0.6)	**1.2**(0.2)	15/15
IPOP lia	**1.2**(1)	**1.8**(0.9)	**2.9**(0.9)	**1.3**(0.3)	**1.2**(0.2)	15/15
ga100 hol	5.9(5)	7.7(4)	16(6)	9.4(6)	19(9)	15/15
grid100 ho	16(20)	31(30)	88(41)	105(60)	769(561)	11/15
grid16 hol	4.5(4)	7.1(5)	17(6)	27(18)	193(181)	14/15
hill hol	**2.4**(2)	**2.1**(1)	3.0(1)	6.4(8)	269(480)	12/15
memPSODE v	**2.6**(1)	**2.1**(0.7)	**1.9**(1)	4.1(7)	8.1(4)	15/15
prcga saw	**2.8**(3)	5.7(4)	15(6)	5.4(2)	190(228)	14/15
ring100 ho	12(12)	17(11)	39(16)	27(14)	57(10)	15/15
ring16 hol	4.4(4)	5.3(4)	10(5)	8.5(5)	17(17)	15/15

Table 8: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_7 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f7	*1.6e+3*:35	*1.0e+3*:106	*6.3e+2*:165	*2.5e+2*:489	*2.5e+1*:2987	15/15
BIPOP-aCMA	4.6(1)	3.1(0.7)	3.5(0.7)	**2.1**(0.3)	**1.1**(0.3)	15/15
BIPOP-saAC	**2.6**(0.8)	**1.5**(0.8)	**1.9**(0.4)	**1.1**(0.2)	**2.2**(2)	15/15
CMAES hut	**1.8**(1)	**1.4**(0.7)	**2.1**(0.7)	**1.6**(0.4)	**0.75**(0.2)	15/15
IPOP-10DDr	3.1(3)	**2.0**(0.6)	**2.1**(0.4)	**1.4**(0.2)	**2.8**(5)	15/15
IPOP-500 l	3.1(3)	**2.0**(0.6)	**2.1**(0.4)	**1.4**(0.2)	**2.8**(5)	15/15
IPOP-tany	**1.5**(1)	**1.5**(0.8)	**1.7**(0.6)	**1.2**(0.2)	**0.72**(0.1)	15/15
IPOP-texp	**0.63**(0.6)	**0.57**(0.4)$^{\star 2}$	**0.68**(0.4)$^{\star 3}$	**0.76**(0.2)$^{\star 3}$	3.4(6)	15/15
IPOP lia	3.1(3)	**2.0**(0.6)	**2.1**(0.4)	**1.4**(0.2)	**2.8**(5)	15/15
ga100 hol	7.1(7)	7.0(2)	9.0(1.0)	8.2(2)	∞ *2e6*	0/15
grid100 ho	19(27)	37(24)	57(27)	90(58)	∞ *2e6*	0/15
grid16 hol	7.0(5)	6.8(4)	10(4)	16(20)	∞ *2e6*	0/15
hill hol	4.5(2)	**2.4**(1)	**2.4**(1)	3.6(2)	∞ *2e6*	0/15
memPSODE v	24(14)	11(12)	18(16)	16(3)	41(25)	15/15
prcga saw	**2.0**(2)	3.0(2)	4.1(2)	3.9(1)	3423(4069)	3/15
ring100 ho	12(10)	15(5)	20(4)	23(5)	1968(2343)	4/15
ring16 hol	5.6(4)	4.6(2)	5.3(1)	5.8(2)	9455(1e4)	1/15

Table 9: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_8 for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f8	*1.0e+5*:85	*6.3e+4*:111	*4.0e+4*:125	*2.5e+3*:430	*6.3e+1*:2106	15/15
BIPOP-aCMA	3.1(0.9)	3.2(1)	4.2(2)	**2.5**(0.2)	**1.4**(0.8)	15/15
BIPOP-saAC	**2.2**(0.4)	**2.0**(0.2)	**1.9**(0.2)	**0.86**(0.2)	**0.65**(0.2)	15/15
CMAES hut	**2.5**(0.5)	**2.5**(0.5)	**2.9**(0.6)	**1.8**(0.2)	**2.2**(2)	9/15
IPOP-10DDr	**2.8**(0.5)	**2.7**(0.4)	**2.8**(0.4)	**1.7**(0.1)	**2.1**(0.2)	15/15
IPOP-500 l	**2.8**(0.5)	**2.7**(0.4)	**2.8**(0.4)	**1.7**(0.1)	**2.1**(0.2)	15/15
IPOP-tany	**2.4**(0.5)	**2.4**(0.5)	**2.7**(0.4)	**1.7**(0.2)	3.5(9)	15/15
IPOP-texp	**0.51**(0.4)$^{\star 2}$	**0.76**(0.4)$^{\star 3}$	**1.0**(0.3)	**1.3**(0.2)	**2.4**(1)	15/15
IPOP lia	**2.8**(0.5)	**2.7**(0.4)	**2.8**(0.4)	**1.7**(0.1)	**2.1**(0.2)	15/15
ga100 hol	16(4)	17(2)	19(4)	16(2)	1632(2001)	6/15
grid100 ho	97(40)	90(36)	102(35)	75(10)	6722(7457)	2/15
grid16 hol	15(4)	15(5)	17(5)	12(2)	1.3e4(2e4)	1/15
hill hol	4.2(2)	3.6(1)	3.6(0.6)	**2.3**(0.5)	578(736)	10/15
memPSODE v	**1.4**(0.2)	**1.3**(0.2)	**1.2**(0.2)	**0.58**(0.1)$^{\star 3}$	**1.5**(2)	15/15
prcga saw	3.7(2)	6.6(6)	9.4(5)	10(2)	1458(2009)	9/15
ring100 ho	39(6)	40(5)	46(7)	37(3)	36(3)	15/15
ring16 hol	8.8(2)	8.6(2)	9.3(1)	7.0(0.9)	1734(2158)	6/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f9	*2.5e+2*:676	*1.6e+2*:865	*1.0e+2*:1397	*6.3e+1*:1896	*4.0e+1*:2180	15/15
BIPOP-aCMA	**2.3**(0.3)	**2.0**(0.2)	**1.4**(0.3)	**1.2**(0.2)	**1.3**(0.2)	15/15
BIPOP-saAC	**0.90**(0.3)	**0.79**(0.3)	**0.55**(0.2)	**0.61**(0.3)	**0.60**(0.3)	15/15
CMAES hut	**1.9**(0.2)	**1.7**(0.2)	**1.2**(0.1)	**0.98**(0.1)	**1.2**(0.1)	15/15
IPOP-10DDr	**2.4**(0.3)	**2.3**(0.7)	**1.8**(1)	**2.7**(0.9)	**2.6**(0.7)	15/15
IPOP-500 l	**2.4**(0.3)	**2.3**(0.7)	**1.8**(1)	**2.7**(0.9)	**2.6**(0.7)	15/15
IPOP-tany	**2.1**(0.3)	**1.9**(0.4)	**1.7**(1)	3.8(10)	3.6(9)	15/15
IPOP-texp	**1.1**(0.2)	**1.0**(0.1)	**0.76**(0.1)	**0.69**(0.1)	**0.87**(0.1)	15/15
IPOP lia	**2.4**(0.3)	**2.3**(0.7)	**1.8**(1)	**2.7**(0.9)	**2.6**(0.7)	15/15
ga100 hol	37(26)	77(35)	104(197)	192(315)	193(284)	14/15
grid100 ho	192(98)	423(425)	480(717)	986(1165)	1052(993)	9/15
grid16 hol	139(58)	221(165)	274(671)	426(544)	392(469)	12/15
hill hol	11(12)	14(11)	169(393)	146(329)	136(290)	14/15
memPSODE v	**0.79**(0.3)	**0.83**(0.6)	**1.2**(1)	**1.2**(1.0)	**1.1**(0.9)	15/15
prcga saw	**2.6**(0.7)	29(10)	49(70)	1187(1795)	1049(1091)	11/15
ring100 ho	70(8)	72(10)	61(10)	59(13)	91(25)	15/15
ring16 hol	26(8)	56(39)	262(717)	302(530)	271(463)	12/15

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

#FEs/D	0.5	1.2	3	10	50	#succ
f10	*1.0e+7*:44	*6.3e+6*:80	*2.5e+6*:126	*2.5e+5*:408	*6.3e+3*:2376	15/15
BIPOP-aCMA	**1.9**(2)	**2.1**(2)	4.3(2)	6.1(2)	**3.8**(0.7)	15/15
BIPOP-saAC	**1.2**(0.6)	**1.1**(0.8)	3.2(3)	4.4(1)	**1.3**(0.2)$^{\star 4}$	15/15
CMAES hut	**0.92**(0.5)	**1.0**(0.6)	**2.8**(1)	5.0(1)	∞ *4007*	0/15
IPOP-10DDr	**1.4**(2)	**1.7**(1.0)	3.2(1)	4.4(0.7)	3.9(0.9)	15/15
IPOP-500 l	**1.4**(2)	**1.7**(1.0)	3.2(1)	4.4(0.7)	3.9(0.9)	15/15
IPOP-tany	**0.73**(0.9)	**1.2**(0.9)	**2.5**(0.9)	**4.3**(0.7)	4.0(0.8)	15/15
IPOP-texp	**0.42**(0.6)	**0.50**(0.4)	**0.85**(0.5)$^{\star 2}$	**3.4**(1.0)	**3.3**(0.5)	15/15
IPOP lia	**1.4**(2)	**1.7**(1.0)	3.2(1)	4.4(0.7)	3.9(0.9)	15/15
ga100 hol	**2.2**(2)	4.6(2)	10(3)	25(11)	∞ *2e6*	0/15
grid100 ho	**2.0**(2)	7.7(6)	38(27)	307(267)	∞ *2e6*	0/15
grid16 hol	**1.9**(3)	3.3(3)	8.1(4)	94(83)	∞ *2e6*	0/15
hill hol	**3.0**(2)	**2.1**(1)	**2.4**(1.0)	16(11)	6137(6734)	2/15
memPSODE v	3.9(4)	**2.8**(3)	3.3(2)	**3.4**(0.9)	31(14)	15/15
prcga saw	**0.96**(1)	**2.5**(2)	5.9(3)	14(5)	625(307)	15/15
ring100 ho	**2.5**(3)	7.0(6)	19(7)	92(44)	∞ *2e6*	0/15
ring16 hol	**1.8**(2)	**2.3**(2)	4.8(1)	27(19)	6191(7103)	2/15

Table 12: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{11} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f11	*1.0e+4*:22	*2.5e+3*:52	*2.5e+2*:432	*1.6e+2*:887	*1.6e+1*:2204	15/15
BIPOP-aCMA	**2.8**(3)	**1.9**(2)	23(2)	**12**(1)	**5.3**(0.3)	15/15
BIPOP-saAC	**1.5**(1)	**1.0**(0.7)	**15**(8)	**7.3**(4)	**3.0**(2)	15/15
CMAES hut	**1.5**(1)	**1.4**(0.8)	∞	∞	∞ *4007*	0/15
IPOP-10DDr	**2.8**(4)	**1.6**(1)	39(5)	22(2)	13(0.6)	15/15
IPOP-500 l	**2.8**(4)	**1.6**(1)	39(5)	22(2)	13(0.6)	15/15
IPOP-tany	**2.4**(2)	**1.2**(1)	40(5)	22(2)	12(0.5)	15/15
IPOP-texp	**1.3**(0.9)	**0.92**(0.7)	33(5)	20(2)	11(0.4)	15/15
IPOP lia	**2.8**(4)	**1.6**(1)	39(5)	22(2)	13(0.6)	15/15
ga100 hol	**2.1**(3)	**1.5**(2)	43(44)	52(33)	376(222)	15/15
grid100 ho	3.6(4)	3.2(3)	1693(892)	3337(3384)	∞ *2e6*	0/15
grid16 hol	**2.0**(2)	**1.7**(1)	523(322)	472(228)	∞ *2e6*	0/15
hill hol	**2.7**(3)	**1.3**(1)	156(94)	157(87)	592(182)	14/15
memPSODE v	3.6(6)	**2.8**(3)	**1.9**(0.6)	**0.97**(0.3)$^{\star 3}$	**0.75**(0.1)$^{\star 4}$	15/15
prcga saw	**2.6**(3)	**2.0**(2)	**6.5**(8)	76(85)	1601(1144)	11/15
ring100 ho	**2.5**(3)	**2.0**(3)	174(142)	232(91)	804(469)	13/15
ring16 hol	**1.9**(2)	**1.4**(1)	96(66)	90(45)	432(219)	15/15

Table 13: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{12} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f12	*2.5e+8*:54	*1.6e+8*:218	*1.0e+8*:284	*1.0e+7*:424	*4.0e+1*:2479	15/15
BIPOP-aCMA	7.6(2)	**2.5**(0.5)	**2.3**(0.4)	**2.7**(0.2)	**1.9**(0.1)	15/15
BIPOP-saAC	3.7(1)	**1.3**(0.4)	**1.2**(0.4)	**1.2**(0.2)	**0.43**(0.0)	15/15
CMAES hut	4.2(1)	**1.5**(0.3)	**1.5**(0.4)	**2.1**(0.3)	**1.6**(0.1)	14/15
IPOP-10DDr	4.0(1)	**1.4**(0.2)	**1.4**(0.2)	**2.0**(0.2)	**1.5**(0.1)	15/15
IPOP-500 l	4.0(1)	**1.4**(0.2)	**1.4**(0.2)	**2.0**(0.2)	**1.5**(0.1)	15/15
IPOP-tany	**3.1**(1)	**1.3**(0.2)	**1.3**(0.2)	**2.1**(0.2)	**1.6**(0.1)	15/15
IPOP-texp	**0.85**(0.7)$^{\star 2}$	**0.51**(0.2)$^{\star 2}$	**0.58**(0.2)	**1.6**(0.2)	**1.4**(0.1)	15/15
IPOP lia	4.0(1)	**1.4**(0.2)	**1.4**(0.2)	**2.0**(0.2)	**1.5**(0.1)	15/15
ga100 hol	17(6)	7.5(2)	8.9(1)	17(3)	405(409)	13/15
grid100 ho	108(53)	46(17)	52(15)	86(11)	∞ *2e6*	0/15
grid16 hol	20(16)	7.4(4)	8.1(3)	13(3)	1083(844)	9/15
hill hol	4.8(3)	**1.6**(0.5)	**1.7**(0.4)	**2.5**(0.5)	595(808)	10/15
memPSODE v	**3.6**(2)	**0.97**(0.5)	**0.86**(0.4)	**0.77**(0.3)$^{\star 2}$	**0.24**(0.1)$^{\star 4}$	15/15
prcga saw	6.4(5)	3.7(2)	4.9(1)	10(2)	40(30)	15/15
ring100 ho	42(10)	20(4)	23(2)	44(3)	345(75)	15/15
ring16 hol	10(4)	4.4(1)	4.5(0.8)	7.9(0.7)	212(87)	15/15

Table 14: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{13} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f13	*2.5e+3*:85	*1.6e+3*:121	*1.6e+3*:121	*6.3e+1*:429	*1.0e+1*:2029	15/15
BIPOP-aCMA	3.3(2)	5.1(0.9)	5.1(0.9)	5.4(0.3)	**1.6**(0.1)	15/15
BIPOP-saAC	**1.8**(0.5)	**1.9**(0.2)	**1.9**(0.2)	**1.6**(0.1)	**0.76**(0.8)	15/15
CMAES hut	**1.5**(0.8)	3.2(0.9)	3.2(0.9)	5.1(0.4)	**2.1**(1)	12/15
IPOP-10DDr	**2.4**(0.3)	3.3(0.4)	3.3(0.4)	5.0(0.6)	3.1(4)	15/15
IPOP-500 l	**2.4**(0.3)	3.3(0.4)	3.3(0.4)	5.0(0.6)	3.1(4)	15/15
IPOP-tany	**1.5**(0.5)	**2.8**(0.5)	**2.8**(0.5)	5.1(0.5)	**2.6**(4)	15/15
IPOP-texp	**0.53**(0.5)$^{\star 2}$	**1.5**(0.5)	**1.5**(0.5)	**4.4**(0.4)	**1.5**(0.1)	15/15
IPOP lia	**2.4**(0.3)	3.3(0.4)	3.3(0.4)	5.0(0.6)	3.1(4)	15/15
ga100 hol	10(4)	19(3)	19(3)	141(32)	505(553)	11/15
grid100 ho	46(36)	105(26)	105(26)	916(561)	1.5e4(1e4)	1/15
grid16 hol	10(4)	18(5)	18(5)	135(24)	463(546)	12/15
hill hol	**2.6**(1)	3.9(1)	3.9(1)	76(56)	585(985)	10/15
memPSODE v	**1.5**(0.5)	**1.8**(0.7)	**1.8**(0.7)	**1.8**(0.2)	**1.4**(2)	15/15
prcga saw	4.4(3)	13(4)	13(4)	70(14)	769(1001)	11/15
ring100 ho	24(7)	52(4)	52(4)	293(47)	212(76)	15/15
ring16 hol	5.4(2)	10(1)	10(1)	66(12)	1195(1485)	7/15

Table 15: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{14} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f14	*6.3e+1*:34	*4.0e+1*:137	*2.5e+1*:176	*4.0e+0*:438	*1.0e-3*:2207	15/15
BIPOP-aCMA	8.1(6)	3.8(1)	3.9(1.0)	3.1(0.5)	**3.6**(0.3)	15/15
BIPOP-saAC	4.0(2)	**1.9**(0.7)	**2.4**(0.7)	**1.8**(0.4)	**1.4**(0.1)$^{\star 4}$	15/15
CMAES hut	**3.2**(1)	**1.5**(0.8)	**2.4**(0.5)	**2.5**(0.4)	∞ *4007*	0/15
IPOP-10DDr	4.6(3)	**2.0**(0.3)	**2.2**(0.3)	**2.1**(0.3)	4.5(0.3)	15/15
IPOP-500 l	4.6(3)	**2.0**(0.3)	**2.2**(0.3)	**2.1**(0.3)	4.5(0.3)	15/15
IPOP-tany	**2.5**(2)	**1.5**(0.5)	**1.8**(0.5)	**2.2**(0.3)	4.7(0.2)	15/15
IPOP-texp	**0.59**(0.6)*	**0.42**(0.2)$^{\star 3}$	**0.74**(0.3)$^{\star 3}$	**1.8**(0.3)	**4.2**(0.4)	15/15
IPOP lia	4.6(3)	**2.0**(0.3)	**2.2**(0.3)	**2.1**(0.3)	4.5(0.3)	15/15
ga100 hol	13(4)	7.1(3)	12(4)	20(4)	∞ *2e6*	0/15
grid100 ho	43(34)	43(27)	64(20)	96(19)	∞ *2e6*	0/15
grid16 hol	15(16)	9.1(5)	12(5)	17(3)	∞ *2e6*	0/15
hill hol	5.9(2)	**2.2**(0.5)	**2.5**(0.6)	**3.0**(0.7)	∞ *2e6*	0/15
memPSODE v	6.6(2)	**1.9**(0.7)	**1.6**(0.6)	**1.2**(0.3)$^{\star 3}$	8.2(4)	15/15
prcga saw	4.6(4)	3.4(2)	5.3(2)	10(2)	2279(2073)	9/15
ring100 ho	24(24)	19(7)	31(10)	52(6)	∞ *2e6*	0/15
ring16 hol	8.2(5)	4.8(1)	6.9(2)	10(2)	∞ *2e6*	0/15

Table 16: 40-D, running time excess ERT/$\mathrm{ERT}_{\text{best 2009}}$ on f_{15} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding $\mathrm{ERT}_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f15	*1.0e+3*:192	*6.3e+2*:458	*4.0e+2*:1170	*2.5e+2*:3875	*2.5e+2*:3875	15/15
BIPOP-aCMA	**2.2**(0.9)	**1.6**(0.3)	**1.3**(0.3)	**1.1**(0.3)	**1.1**(0.3)	15/15
BIPOP-saAC	**1.1**(0.4)	**0.68**(0.1)	**1.00**(0.3)	**0.91**(0.3)	**0.91**(0.3)	15/15
CMAES hut	**0.99**(0.3)	**1.0**(0.3)	**1.1**(0.3)	**1.8**(1)	**1.8**(1)	8/15
IPOP-10DDr	**1.3**(0.4)	**0.93**(0.1)	**0.83**(0.2)	**1.1**(0.3)	**1.1**(0.3)	15/15
IPOP-500 l	**1.3**(0.4)	**0.93**(0.1)	**0.83**(0.2)	**1.1**(0.3)	**1.1**(0.3)	15/15
IPOP-tany	**1.0**(0.2)	**0.85**(0.1)	**0.81**(0.2)	**1.1**(0.3)	**1.1**(0.3)	15/15
IPOP-texp	**0.56**(0.3)$^{\star}_{\downarrow 2}$	**0.61**(0.2)	**0.77**(0.2)	**0.90**(0.3)	**0.90**(0.3)	15/15
IPOP lia	**1.3**(0.4)	**0.93**(0.1)	**0.83**(0.2)	**1.1**(0.3)	**1.1**(0.3)	15/15
ga100 hol	6.2(2)	5.6(0.9)	5.6(1)	6.5(2)	6.5(2)	15/15
grid100 ho	62(25)	1069(1180)	7226(8506)	∞	∞ *2e6*	0/15
grid16 hol	13(5)	975(2187)	4926(5982)	∞	∞ *2e6*	0/15
hill hol	4.7(4)	5030(6576)	∞	∞	∞ *2e6*	0/15
memPSODE v	55(30)	55(12)	37(19)	24(8)	24(8)	15/15
prcga saw	**2.9**(2)	3.4(0.5)	3.7(0.6)	3.2(1)	3.2(1)	15/15
ring100 ho	16(4)	19(3)	34(19)	116(224)	116(224)	14/15
ring16 hol	4.0(1)	6.6(6)	438(856)	605(780)	605(777)	7/15

Table 17: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{16} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f16	*4.0e+1*:117	*2.5e+1*:297	*1.6e+1*:4010	*1.6e+1*:4010	*1.0e+1*:5244	15/15
BIPOP-aCMA	21(17)	14(3)	**1.1**(0.3)	**1.1**(0.3)	**1.0**(0.3)	15/15
BIPOP-saAC	22(10)	12(5)	**1.0**(0.4)	**1.0**(0.4)	**0.94**(0.4)	15/15
CMAES hut	23(17)	14(9)	**1.7**(1)	**1.7**(2)	3.7(4)	3/15
IPOP-10DDr	8.9(6)	**7.0**(1)	**0.61**(0.1)	**0.61**(0.1)	**0.55**(0.1)	15/15
IPOP-500 l	8.9(6)	**7.0**(1)	**0.61**(0.1)	**0.61**(0.1)	**0.55**(0.1)	15/15
IPOP-tany	9.3(6)	7.6(2)	**0.67**(0.1)	**0.67**(0.1)	**0.58**(0.1)	15/15
IPOP-texp	5.8(5)	**5.3**(2)	**0.49**(0.2)	**0.49**(0.2)	**0.45**(0.1)*	15/15
IPOP lia	8.9(6)	**7.0**(1)	**0.61**(0.1)	**0.61**(0.1)	**0.55**(0.1)	15/15
ga100 hol	13(15)	24(11)	42(13)	42(13)	339(480)	8/15
grid100 ho	9.0(7)	96(67)	515(608)	515(748)	∞ *2e6*	0/15
grid16 hol	**3.1**(2)	416(300)	622(770)	622(756)	5566(6102)	1/15
hill hol	**1.3**(1)	19(6)	1077(1496)	1077(1291)	5611(6102)	1/15
memPSODE v	22(4)	26(17)	3.8(3)	3.8(3)	7.5(6)	15/15
prcga saw	11(14)	22(14)	**2.3**(0.7)	**2.3**(0.7)	120(382)	13/15
ring100 ho	5.8(6)	27(12)	8.3(2)	8.3(2)	37(30)	15/15
ring16 hol	**2.9**(2)	8.0(4)	253(499)	253(499)	1058(1158)	4/15

Table 18: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{17} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f17	*1.6e+1*:54	*1.0e+1*:399	*6.3e+0*:688	*4.0e+0*:1115	*1.0e+0*:4220	15/15
BIPOP-aCMA	6.7(4)	**1.7**(0.6)	**1.4**(0.3)	**1.1**(0.2)	**0.54**(0.1)	15/15
BIPOP-saAC	3.8(2)	**0.99**(0.4)	**1.00**(0.2)	**0.88**(0.1)	**2.4**(0.4)	15/15
CMAES hut	**2.8**(3)	**1.0**(0.4)	**1.1**(0.3)	**1.00**(0.3)	**0.59**(0.2)	15/15
IPOP-10DDr	**2.2**(2)	**0.81**(0.2)	**0.84**(0.2)	**0.77**(0.2)	**0.49**(0.1)	15/15
IPOP-500 l	**2.2**(2)	**0.81**(0.2)	**0.84**(0.2)	**0.77**(0.2)	**0.49**(0.1)	15/15
IPOP-tany	**1.4**(1)	**0.59**(0.2)	**0.71**(0.2)	**0.72**(0.1)	**0.43**(0.1)	15/15
IPOP-texp	**0.37**(0.3)	**0.31**(0.1)$^{\star 2}$	**0.53**(0.2)	**0.61**(0.2)	**0.46**(0.1)	15/15
IPOP lia	**2.2**(2)	**0.81**(0.2)	**0.84**(0.2)	**0.77**(0.2)	**0.49**(0.1)	15/15
ga100 hol	6.3(3)	3.3(0.8)	4.9(0.9)	5.9(2)	717(948)	6/15
grid100 ho	38(13)	1549(2583)	∞	∞	∞ *2e6*	0/15
grid16 hol	48(27)	975(2505)	4.3e4(5e4)	∞	∞ *2e6*	0/15
hill hol	1.5e4(2e4)	∞	∞	∞	∞ *2e6*	0/15
memPSODE v	67(102)	27(18)	56(57)	188(209)	279(156)	15/15
prcga saw	**2.1**(3)	**1.5**(0.9)	**3.0**(1)	3.3(0.8)	1400(1895)	6/15
ring100 ho	10(8)	11(6)	37(23)	1340(1814)	∞ *2e6*	0/15
ring16 hol	5.1(4)	3.8(3)	1162(1525)	2.5e4(3e4)	∞ *2e6*	0/15

Table 19: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{18} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f18	*6.3e+1*:55	*4.0e+1*:329	*4.0e+1*:329	*2.5e+1*:579	*6.3e+0*:2006	15/15
BIPOP-aCMA	5.1(2)	**1.6**(0.5)	**1.6**(0.5)	**1.4**(0.4)	**0.97**(0.2)	15/15
BIPOP-saAC	**2.5**(1)	**1.1**(0.5)	**1.1**(0.5)	**1.1**(0.3)	**0.82**(0.3)	15/15
CMAES hut	**2.5**(1)	**0.97**(0.4)	**0.97**(0.4)	**1.1**(0.3)	**0.96**(0.3)	15/15
IPOP-10DDr	**1.9**(2)	**0.83**(0.2)	**0.83**(0.2)	**0.84**(0.2)	**0.80**(0.2)	15/15
IPOP-500 l	**1.9**(2)	**0.83**(0.2)	**0.83**(0.2)	**0.84**(0.2)	**0.80**(0.2)	15/15
IPOP-tany	**1.0**(0.9)	**0.67**(0.2)	**0.67**(0.2)	**0.79**(0.2)	**0.77**(0.1)	15/15
IPOP-texp	**0.45**(0.5)	**0.38**(0.2)$^{\star 3}$	**0.38**(0.2)$^{\star 3}$	**0.59**(0.1)	**0.73**(0.1)	15/15
IPOP lia	**1.9**(2)	**0.83**(0.2)	**0.83**(0.2)	**0.84**(0.2)	**0.80**(0.2)	15/15
ga100 hol	6.2(4)	3.9(1)	3.9(1)	4.8(1)	1153(1498)	7/15
grid100 ho	12(10)	218(191)	218(191)	1.5e4(2e4)	∞ *2e6*	0/15
grid16 hol	5.0(5)	638(839)	638(839)	1.5e4(2e4)	∞ *2e6*	0/15
hill hol	6026(2e4)	1.3e4(2e4)	1.3e4(2e4)	∞	∞ *2e6*	0/15
memPSODE v	78(94)	34(13)	34(13)	41(20)	283(222)	15/15
prcga saw	**2.0**(3)	**1.7**(0.7)	**1.7**(0.7)	**3.0**(0.8)	337(997)	13/15
ring100 ho	10(9)	11(5)	11(5)	31(18)	∞ *2e6*	0/15
ring16 hol	3.5(3)	4.3(1)	4.3(1)	541(1169)	∞ *2e6*	0/15

Table 20: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{19} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f19	*1.6e-1*:8.6e5	*1.0e-1*:1.4e6	*6.3e-2*:3.1e6	*4.0e-2*:5.2e6	*2.5e-2*:8.7e6	15/15
BIPOP-aCMA	**1.1**(0.9)	**1.0**(1)	**0.73**(0.6)	**0.77**(0.4)	**0.73**(0.3)	15/15
BIPOP-saAC	**0.75**(0.8)	**0.75**(0.9)	**0.75**(0.7)	**0.89**(0.6)	**1.0**(0.4)	15/15
CMAES hut	∞	∞	∞	∞	∞ *4007*	0/15
IPOP-10DDr	4.4(5)	**2.8**(3)	**1.5**(1)	**1.1**(1)	**0.94**(0.7)	15/15
IPOP-500 l	4.4(5)	**2.8**(3)	**1.5**(1)	**1.2**(1.0)	**1.4**(1)	15/15
IPOP-tany	5.2(8)	3.4(5)	**1.7**(2)	**1.3**(1)	**1.0**(0.9)	15/15
IPOP-texp	**2.4**(4)	**1.6**(3)	**0.94**(1)	**0.97**(0.9)	**0.80**(0.5)	15/15
IPOP lia	4.4(5)	**2.8**(3)	**1.5**(1)	**1.1**(1)	**0.94**(0.7)	15/15
ga100 hol	∞	∞	∞	∞	∞ *2e6*	0/15
grid100 ho	∞	∞	∞	∞	∞ *2e6*	0/15
grid16 hol	∞	∞	∞	∞	∞ *2e6*	0/15
hill hol	∞	∞	∞	∞	∞ *2e6*	0/15
memPSODE v	∞	∞	∞	∞	∞ *2e7*	0/15
prcga saw	**0.43**(0.2)	**0.77**(0.7)	**0.43**(0.4)	**0.31**(0.2)	**0.23**(0.1)	14/15
ring100 ho	∞	∞	∞	∞	∞ *2e6*	0/15
ring16 hol	∞	∞	∞	∞	∞ *2e6*	0/15

Table 21: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{20} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f20	*2.5e+4*:83	*1.6e+4*:86	*1.0e+3*:125	*2.5e+0*:515	*1.6e+0*:5582	15/15
BIPOP-aCMA	5.8(2)	6.4(2)	7.1(0.8)	5.6(0.6)	71(58)	15/15
BIPOP-saAC	**2.9**(0.3)	**3.0**(0.5)	**2.9**(0.5)	**3.3**(1)	52(50)	15/15
CMAES hut	3.4(0.8)	4.0(0.8)	5.0(0.9)	5.1(1.0)	11(11)	1/15
IPOP-10DDr	4.4(0.4)	4.9(0.7)	5.8(0.5)	5.6(1.0)	141(64)	15/15
IPOP-500 l	4.4(0.4)	4.9(0.7)	5.8(0.5)	5.6(1.0)	132(56)	15/15
IPOP-tany	3.2(0.8)	3.8(0.7)	5.1(0.5)	5.8(0.9)	145(112)	15/15
IPOP-texp	**0.44**(0.2)$^{\star 4}_{\downarrow 6}$	**0.67**(0.4)$^{\star 4}$	**2.1**(1)	4.7(0.8)	309(136)	15/15
IPOP lia	4.4(0.4)	4.9(0.7)	5.8(0.5)	5.6(1.0)	141(64)	15/15
ga100 hol	20(4)	24(4)	33(8)	22(3)	3.3(0.4)	15/15
grid100 ho	100(37)	121(39)	174(30)	92(17)	16(5)	15/15
grid16 hol	17(6)	20(6)	27(5)	15(1)	**2.7**(0.6)	15/15
hill hol	4.4(2)	4.7(2)	6.0(2)	**2.9**(0.7)	**0.60**(0.1)*	15/15
memPSODE v	**2.7**(0.5)	**2.6**(0.5)	**2.3**(0.7)	**4.0**(2)	**1.6**(1)	15/15
prcga saw	**2.9**(0.5)	4.6(2)	11(2)	13(3)	4.4(2)	15/15
ring100 ho	51(7)	59(9)	83(8)	48(4)	7.2(0.5)	15/15
ring16 hol	10(2)	12(3)	16(2)	9.0(1)	**1.4**(0.1)	15/15

Table 22: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{21} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f21	*6.3e+1*:160	*4.0e+1*:305	*2.5e+1*:380	*1.6e+1*:784	*6.3e+0*:2510	30/30
BIPOP-aCMA	3.8(0.8)	**2.6**(0.5)	4.0(0.4)	**2.2**(0.4)	3.3(5)	15/15
BIPOP-saAC	**2.4**(0.9)	3.5(5)	3.5(4)	**1.7**(2)	**0.98**(0.7)	15/15
CMAES hut	**2.5**(1)	**2.7**(0.8)	**2.6**(1)	**2.2**(3)	**1.8**(2)	8/15
IPOP-10DDr	**2.1**(0.5)	**1.5**(0.3)	**1.4**(0.4)	**2.7**(0.2)	189(5)	15/15
IPOP-500 l	**2.1**(0.5)	**1.5**(0.3)	**1.4**(0.4)	**2.7**(0.2)	15(5)	15/15
IPOP-tany	**1.8**(0.4)	**1.3**(0.1)	**1.3**(0.3)	**0.79**(0.1)	**2.5**(8)	15/15
IPOP-texp	**1.0**(0.3)$^{\star 3}$	**0.96**(0.3)*	**1.1**(0.2)	**0.69**(0.2)	6.2(5)	15/15
IPOP lia	**2.1**(0.5)	**1.5**(0.3)	**1.4**(0.4)	**2.7**(0.2)	299(5)	15/15
ga100 hol	14(5)	11(2)	12(5)	189(5)	293(400)	11/15
grid100 ho	63(15)	48(13)	51(16)	217(17)	928(1202)	7/15
grid16 hol	11(4)	8.0(2)	9.2(6)	642(1276)	202(399)	12/15
hill hol	**2.6**(1)	470(1)	1917(2635)	1277(1276)	698(798)	8/15
memPSODE v	**2.7**(1)	**2.1**(0.9)	**2.5**(3)	**2.8**(3)	**1.1**(1.0)	15/15
prcga saw	10(3)	7.3(2)	7.2(2)	4.7(1)	411(797)	11/15
ring100 ho	31(10)	25(8)	26(6)	16(4)	8.6(3)	15/15
ring16 hol	7.5(1)	5.9(1)	6.2(1)	3.7(0.8)	125(398)	13/15

Table 23: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{22} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f22	*6.3e+1*:160	*4.0e+1*:231	*2.5e+1*:687	*1.6e+1*:1392	*1.0e+1*:3090	15/15
BIPOP-aCMA	4.0(1)	**3.7**(1)	6.4(7)	10(16)	7.8(16)	15/15
BIPOP-saAC	**2.5**(0.7)	4.8(7)	**2.6**(3)	**2.9**(5)	7.2(4)	15/15
CMAES hut	**2.4**(0.6)	**2.2**(0.7)	**1.4**(0.4)	**1.0**(1)	**0.62**(0.7)	12/15
IPOP-10DDr	**2.3**(0.8)	37(91)	13(31)	9.0(15)	388(848)	15/15
IPOP-500 l	**2.3**(0.8)	37(91)	13(31)	9.0(15)	3262(6474)	12/15
IPOP-tany	**2.0**(0.5)	4.2(0.7)	3.6(6)	135(13)	87(186)	15/15
IPOP-texp	**1.3**(0.9)*	21(15)	17(5)	99(62)	158(294)	15/15
IPOP lia	**2.3**(0.8)	37(91)	13(31)	9.0(15)	1110(924)	14/15
ga100 hol	14(5)	631(5)	214(2)	107(2)	49(2)	14/15
grid100 ho	69(21)	80(47)	755(1458)	543(749)	574(649)	8/15
grid16 hol	908(5)	632(11)	1064(1458)	1261(2155)	742(971)	7/15
hill hol	**2.7**(1.0)	620(1)	729(1455)	719(1437)	740(971)	7/15
memPSODE v	**2.9**(2)	**2.9**(3)	**1.6**(1)	**0.89**(0.7)	**0.77**(0.9)	15/15
prcga saw	11(3)	11(3)	292(2)	340(994)	274(647)	12/15
ring100 ho	37(6)	36(8)	15(3)	10(2)	**5.4**(2)	15/15
ring16 hol	7.8(3)	8.3(4)	451(1455)	361(719)	163(324)	12/15

Table 24: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{23} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f23	*6.3e+0*:68	*4.0e+0*:292	*2.5e+0*:603	*2.5e+0*:603	*1.6e+0*:2487	15/15
BIPOP-aCMA	5.1(7)	110(112)	185(143)	185(143)	**45**(35)	15/15
BIPOP-saAC	13(15)	86(63)	175(183)	175(183)	**43**(44)	15/15
CMAES hut	15(12)	96(108)	∞	∞	∞ *4007*	0/15
IPOP-10DDr	3.5(3)	49(60)	396(378)	396(378)	96(92)	15/15
IPOP-500 l	3.5(3)	49(60)	396(378)	396(378)	96(92)	15/15
IPOP-tany	**2.4**(3)	65(106)	220(185)	220(185)	53(45)	15/15
IPOP-texp	4.6(5)	54(53)	389(379)	389(379)	102(150)	15/15
IPOP lia	3.5(3)	49(60)	396(378)	396(378)	96(92)	15/15
ga100 hol	4.4(4)	56(44)	1216(1129)	1216(1129)	1531(1656)	6/15
grid100 ho	**2.9**(3)	54(41)	447(177)	447(177)	1094(1050)	8/15
grid16 hol	**2.8**(4)	**35**(33)	572(606)	572(606)	510(651)	11/15
hill hol	**2.4**(2)	**10**(7)	300(110)	300(110)	576(811)	9/15
memPSODE v	89(63)	42(10)	**51**(36)	**51**(36)	**29**(16)	15/15
prcga saw	4.6(3)	82(93)	1563(1523)	1563(1464)	1.4e4(2e4)	1/15
ring100 ho	3.6(5)	40(33)	**118**(41)	**118**(41)	98(46)	15/15
ring16 hol	**1.8**(2)	**20**(15)	**95**(76)	**95**(76)	839(1091)	8/15

Table 25: 40-D, running time excess ERT/ERT$_{\text{best 2009}}$ on f_{24} for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding ERT$_{\text{best 2009}}$ (preceded by the target Δf-value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with $p = 0.05$ or $p = 10^{-k}$ when the number k following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
f24	*4.0e+2*:1404	*2.5e+2*:17825	*1.6e+2*:18980	*1.0e+2*:38677	*6.3e+1*:1.6e5	15/15
BIPOP-aCMA	**1.2**(0.8)	7.4(7)	24(32)	13(16)	3.7(4)	15/15
BIPOP-saAC	**0.95**(0.3)	4.9(5)	11(16)	7.9(12)	**2.0**(3)	15/15
CMAES hut	**1.1**(0.7)	∞	∞	∞	∞ *4007*	0/15
IPOP-10DDr	**1.2**(0.5)	**1.4**(0.8)	**1.4**(0.9)	**1.8**(1.0)	**1.3**(0.4)	15/15
IPOP-500 l	**1.2**(0.5)	**1.4**(0.8)	**1.4**(0.9)	**1.8**(1.0)	**1.3**(0.4)	15/15
IPOP-tany	**0.88**(0.4)	**1.5**(1.0)	**1.5**(0.9)	**1.2**(1.0)	**0.84**(0.6)	15/15
IPOP-texp	**0.61**(0.4)	**1.5**(1)	**1.5**(1)	**1.3**(1)	**1.1**(1)	15/15
IPOP lia	**1.2**(0.5)	**1.4**(0.8)	**1.4**(0.9)	**1.8**(1.0)	**1.3**(0.4)	15/15
ga100 hol	5.4(1.0)	5.6(5)	13(5)	28(28)	∞ *2e6*	0/15
grid100 ho	6228(6890)	∞	∞	∞	∞ *2e6*	0/15
grid16 hol	4376(4987)	∞	∞	∞	∞ *2e6*	0/15
hill hol	9458(9974)	∞	∞	∞	∞ *2e6*	0/15
memPSODE v	19(13)	4.0(2)	9.4(5)	19(17)	71(64)	13/15
prcga saw	19(14)	5.7(6)	14(10)	7.7(4)	11(13)	11/15
ring100 ho	45(14)	23(9)	753(876)	∞	∞ *2e6*	0/15
ring16 hol	21(23)	162(181)	∞	∞	∞ *2e6*	0/15

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 agorithms 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.