

# Comparison Tables: BBOB 2013 Testbed in 5-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 ( $\text{ERT}_{\text{best}}$  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: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best 2009}}$  on  $f_1$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best 2009}}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f1</b>	<i>2.5e+1:4.8</i>	<i>1.6e+1:7.6</i>	<i>1.0e-8:12</i>	<i>1.0e-8:12</i>	<i>1.0e-8:12</i>	15/15
BIPOP-aCMA	3.2(3)	3.5(2)	15(0.6)	15(0.6)	15(0.6)	15/15
BIPOP-saAC	4.3(3)	3.7(3)	21(2)	21(2)	21(2)	15/15
CMAES hut	<b>2.4</b> (2)	<b>2.1</b> (2)	$\infty$	$\infty$	$\infty$ 506	0/15
DE pal	<b>1.8</b> (2)	3.9(4)	180(7)	180(7)	180(7)	15/15
HCMA los	<b>1.9</b> (1)	<b>1.5</b> (0)	<b>0.98</b> (0)*4	<b>0.98</b> (0)*4	<b>0.98</b> (0)*4	15/15
HMLSL pal	<b>1.0</b> (0.6)	<b>0.76</b> (0.4)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	15/15
IPOP-10DDr	<b>2.1</b> (2)	<b>2.9</b> (2)	58(5)	58(5)	58(5)	15/15
IPOP-500 l	<b>2.1</b> (2)	<b>2.9</b> (2)	58(5)	58(5)	58(5)	15/15
IPOP-tany	<b>1.4</b> (2)	<b>2.9</b> (3)	62(4)	62(4)	62(4)	15/15
IPOP-texp	<b>2.3</b> (3)	<b>2.4</b> (3)	62(5)	62(5)	62(5)	15/15
IPOP lia	<b>2.1</b> (2)	<b>2.9</b> (2)	58(5)	58(5)	58(5)	15/15
MLSL pal	<b>1.0</b> (0.6)	<b>0.76</b> (0.4)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	15/15
OQNLP pal	<b>2.5</b> (2)	<b>2.0</b> (1)	<b>2.6</b> (0.0)	<b>2.6</b> (0.0)	<b>2.6</b> (0.0)	15/15
P-DCN tra	<b>2.5</b> (4)	5.5(10)	712(331)	712(331)	712(331)	15/15
P-zero tra	8.7(7)	10(13)	2005(871)	2005(871)	2005(871)	15/15
SMAC hut	<b>0.79</b> (0.7)	<b>0.84</b> (0.5)	$\infty$	$\infty$	$\infty$ 500	0/15
U-DCN tra	<b>2.4</b> (2)	<b>2.6</b> (3)	3045(1610)	3045(1610)	3045(1610)	15/15
U-zero tra	<b>2.3</b> (2)	<b>2.3</b> (2)	9.8e4(6e4)	9.8e4(6e4)	9.8e4(6e4)	15/15
fmincon pa	<b>1.0</b> (0.6)	<b>0.76</b> (0.4)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	<b>4.5</b> (0.7)	15/15
fminunc pa	<b>1.1</b> (0.6)	<b>1.2</b> (0.8)	<b>1.1</b> (0)	<b>1.1</b> (0)	<b>1.1</b> (0)	15/15
ga100 hol	<b>2.1</b> (2)	3.2(2)	$\infty$	$\infty$	$\infty$ 2e5	0/15
grid100 ho	3.6(4)	3.9(5)	$\infty$	$\infty$	$\infty$ 2e5	0/15
grid16 hol	6.2(10)	11(13)	$\infty$	$\infty$	$\infty$ 2e5	0/15
hill hol	4.8(5)	4.1(4)	$\infty$	$\infty$	$\infty$ 2e5	0/15
lmmCMA aug	<b>1.2</b> (2)	<b>1.5</b> (1)	9.1(0.6)	9.1(0.6)	9.1(0.6)	15/15
memPSODE v	4.2(5)	5.1(6)	7.6(0.2)	7.6(0.2)	7.6(0.2)	15/15
prcga saw	<b>1.8</b> (2)	<b>2.3</b> (3)	2503(3948)	2503(3948)	2503(3948)	15/15
ring100 ho	<b>2.3</b> (3)	5.7(9)	$\infty$	$\infty$	$\infty$ 2e5	0/15
ring16 hol	<b>2.4</b> (2)	4.4(4)	$\infty$	$\infty$	$\infty$ 2e5	0/15
simplex pa	12(8)	17(17)	96(12)	96(12)	96(12)	15/15

Table 3: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best 2009}}$  on  $f_2$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best 2009}}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f2</b>	<i>1.6e+6:2.9</i>	<i>4.0e+5:11</i>	<i>4.0e+4:15</i>	<i>6.3e+2:58</i>	<i>1.0e-8:95</i>	15/15
BIPOP-aCMA	<b>1.9</b> (2)	<b>1.8</b> (1)	3.8(1)	<b>1.6</b> (0.2)	<b>3.0</b> (0.2)	15/15
BIPOP-saAC	3.6(4)	<b>2.3</b> (2)	6.1(4)	3.4(0.7)	5.6(0.7)	15/15
CMAES hut	3.7(5)	<b>1.9</b> (2)	6.1(3)	8.2(5)	$\infty$ 506	0/15
DE pal	<b>1.8</b> (2)	<b>2.2</b> (2)	7.2(6)	8.0(2)	32(1)	15/15
HCMA los	3.7(2)	<b>1.0</b> (0)	<b>0.97</b> (0.2)	<b>1.2</b> (0.7)	<b>3.8</b> (0.2)	15/15
HMLSL pal	10(12)	3.7(4)	5.2(5)	<b>2.2</b> (2)	10(9)	15/15
IPOP-10DDr	<b>2.6</b> (2)	<b>1.4</b> (1)	5.7(4)	9.3(5)	24(3)	15/15
IPOP-500 l	<b>2.6</b> (2)	<b>1.4</b> (1)	5.7(4)	9.3(5)	24(3)	15/15
IPOP-tany	3.3(4)	<b>1.6</b> (1)	5.3(3)	10(4)	24(1)	15/15
IPOP-texp	4.8(4)	<b>1.6</b> (1)	6.3(7)	9.4(3)	25(2)	15/15
IPOP lia	<b>2.6</b> (2)	<b>1.4</b> (1)	5.7(4)	9.3(5)	24(3)	15/15
MLSL pal	10(12)	3.7(4)	5.2(5)	<b>2.2</b> (2)	7.3(6)	15/15
OQNLP pal	5.9(3)	<b>1.9</b> (1)	<b>3.1</b> (2)	<b>1.3</b> (0.1)	456(460)	3/15
P-DCN tra	<b>1.1</b> (1)	<b>2.8</b> (4)	10(7)	11(8)	881(819)	15/15
P-zero tra	<b>1.1</b> (1)	<b>1.9</b> (2)	7.4(7)	10(8)	8.9e4(9e4)	7/15
SMAC hut	<b>1.0</b> (0.9)	<b>0.74</b> (0.7)	<b>1.7</b> (2)	8.0(7)	$\infty$ 500	0/15
U-DCN tra	<b>1.1</b> (1)	<b>1.0</b> (2)	6.2(6)	19(24)	1.7e4(3e4)	12/15
U-zero tra	<b>1.1</b> (1)	<b>1.0</b> (2)	7.8(8)	23(26)	$\infty$ 5e6	0/15
fmincon pa	10(12)	3.7(4)	4.8(5)	<b>2.0</b> (1)	<b>4.7</b> (2)	15/15
fminunc pa	5.2(4)	<b>2.0</b> (1)	<b>3.1</b> (2)	<b>2.7</b> (1)	7.7(2)	15/15
ga100 hol	<b>1.6</b> (2)	<b>1.4</b> (1)	16(16)	31(17)	$\infty$ 2e5	0/15
grid100 ho	<b>1.7</b> (2)	<b>1.4</b> (2)	24(31)	76(35)	$\infty$ 2e5	0/15
grid16 hol	<b>1.3</b> (0.9)	<b>1.2</b> (1)	11(9)	22(10)	$\infty$ 2e5	0/15
hill hol	8.7(9)	3.4(3)	6.1(6)	13(11)	$\infty$ 2e5	0/15
lmmCMA aug	<b>1.6</b> (2)	<b>0.83</b> (0.9)	<b>2.4</b> (2)	<b>2.5</b> (0.7)	<b>5.5</b> (1)	15/15
memPSODE v	<b>1.7</b> (2)	<b>1.3</b> (1)	7.4(4)	3.5(2)	<b>5.2</b> (3)	15/15
prcga saw	<b>1.5</b> (1)	<b>1.0</b> (1)	5.3(5)	8.3(2)	1122(1443)	15/15
ring100 ho	<b>1.1</b> (1)	<b>1.1</b> (0.9)	25(24)	54(21)	$\infty$ 2e5	0/15
ring16 hol	<b>1.5</b> (1)	<b>1.7</b> (2)	10(7)	17(12)	$\infty$ 2e5	0/15
simplex pa	22(13)	6.6(3)	15(16)	11(9)	28(3)	15/15

Table 4: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best 2009}}$  on  $f_3$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best 2009}}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f3</b>	<i>1.6e+2:4.1</i>	<i>1.0e+2:15</i>	<i>6.3e+1:23</i>	<i>2.5e+1:73</i>	<i>1.0e+1:716</i>	15/15
BIPOP-aCMA	<b>2.9</b> (4)	<b>1.8</b> (1)	<b>1.8</b> (0.5)	<b>1.1</b> (0.6)	<b>0.27</b> (0.1)	15/15
BIPOP-saAC	4.1(2)	<b>2.7</b> (2)	<b>2.6</b> (1)	3.1(1)	<b>1.1</b> (0.9)	15/15
CMAES hut	<b>2.9</b> (3)	<b>1.7</b> (1)	<b>2.3</b> (2)	<b>2.6</b> (1)	<b>1.2</b> (0.8)	8/15
DE pal	<b>1.7</b> (2)	<b>1.7</b> (2)	4.8(3)	4.5(0.9)	<b>1.2</b> (0.4)	15/15
HCMA los	<b>2.6</b> (2)	<b>1.8</b> (2)	<b>1.6</b> (1)	<b>1.1</b> (0.6)	<b>0.29</b> (0.1)	15/15
HMLSL pal	7.1(18)	6.4(7)	11(10)	8.0(5)	<b>1.3</b> (0.2)	15/15
IPOP-10DDr	<b>2.5</b> (3)	<b>2.0</b> (1)	<b>2.5</b> (1)	<b>2.4</b> (2)	<b>0.99</b> (1)	15/15
IPOP-500 l	<b>2.5</b> (3)	<b>2.0</b> (1)	<b>2.5</b> (1)	<b>2.4</b> (2)	<b>0.99</b> (1)	15/15
IPOP-tany	<b>2.0</b> (2)	<b>1.5</b> (1)	<b>2.0</b> (1.0)	<b>2.3</b> (1)	<b>0.94</b> (1)	15/15
IPOP-texp	<b>1.5</b> (1)	<b>1.1</b> (0.9)	<b>1.6</b> (0.9)	<b>2.5</b> (2)	<b>1.1</b> (1)	15/15
IPOP lia	<b>2.5</b> (3)	<b>2.0</b> (1)	<b>2.5</b> (1)	<b>2.4</b> (2)	<b>0.99</b> (1)	15/15
MLSL pal	7.5(20)	7.7(8)	12(9)	18(20)	5.6(5)	15/15
OQNLP pal	13(26)	9.3(12)	12(9)	25(30)	8.6(12)	13/15
P-DCN tra	<b>0.98</b> (0.8)	7.3(2)	7.3(1)	7.5(11)	7.5(11)	15/15
P-zero tra	<b>0.98</b> (0.8)	<b>1.6</b> (2)	3.3(3)	42(40)	9.4(9)	15/15
SMAC hut	<b>0.73</b> (0.6)	<b>0.74</b> (1.0)	<b>2.6</b> (4)	4.4(5)	5.1(5)	2/15
U-DCN tra	<b>0.98</b> (0.8)	<b>1.4</b> (1)	<b>2.6</b> (1)	3.3(1)	<b>0.85</b> (0.5)	15/15
U-zero tra	<b>0.98</b> (0.8)	<b>1.3</b> (1)	<b>1.6</b> (0.7)	<b>1.5</b> (0.5)	<b>0.62</b> (0.5)	15/15
fmincon pa	7.4(19)	7.1(8)	9.1(9)	12(15)	5.2(6)	15/15
fminunc pa	9.1(20)	6.9(7)	9.0(6)	13(9)	5.3(5)	15/15
ga100 hol	<b>1.6</b> (2)	<b>0.99</b> (1)	5.4(6)	11(3)	<b>2.3</b> (0.6)	15/15
grid100 ho	<b>1.1</b> (1)	4.3(4)	17(13)	25(12)	5.3(4)	15/15
grid16 hol	<b>2.2</b> (2)	3.5(3)	5.3(5)	6.1(3)	<b>1.1</b> (0.5)	15/15
hill hol	7.3(5)	<b>2.6</b> (1)	<b>2.0</b> (1)	<b>1.5</b> (0.8)	<b>0.36</b> (0.2)	15/15
lmmCMA aug	<b>1.4</b> (1)	<b>0.92</b> (0.8)	<b>1.4</b> (1)	<b>2.4</b> (1)	<b>0.45</b> (0.1)	15/15
memPSODE v	3.2(3)	8.4(14)	20(18)	15(12)	3.6(3)	15/15
prcga saw	<b>1.6</b> (2)	<b>1.2</b> (1)	<b>2.5</b> (3)	5.0(2)	<b>1.3</b> (0.6)	15/15
ring100 ho	<b>1.6</b> (2)	<b>2.2</b> (2)	12(10)	20(7)	4.4(1)	15/15
ring16 hol	<b>2.0</b> (1)	3.0(3)	4.7(4)	4.6(2)	<b>0.98</b> (0.3)	15/15
simplex pa	91(112)	50(34)	43(13)	25(14)	12(7)	15/15

Table 5: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b><math>f_4</math></b>	<i>2.5e+2:2.6</i>	<i>1.6e+2:10</i>	<i>1.0e+2:19</i>	<i>4.0e+1:65</i>	<i>1.6e+1:434</i>	15/15
BIPOP-aCMA	3.6(4)	<b>2.1</b> (2)	<b>2.6</b> (1)	<b>1.6</b> (0.5)	<b>0.38</b> (0.1)	15/15
BIPOP-saAC	4.3(5)	<b>2.2</b> (2)	<b>2.3</b> (3)	3.3(0.8)	<b>1.9</b> (2)	15/15
CMAES hut	4.1(4)	<b>2.5</b> (4)	<b>2.5</b> (3)	<b>2.5</b> (1)	<b>1.2</b> (0.7)	11/15
DE pal	<b>3.0</b> (3)	<b>2.6</b> (3)	3.6(2)	4.2(2)	<b>1.5</b> (0.6)	15/15
HCMA los	6.3(6)	3.2(2)	3.0(2)	<b>1.8</b> (0.4)	<b>0.45</b> (0.3)	15/15
HMLSL pal	<b>0.69</b> (0)	<b>0.94</b> (1)	5.4(11)	7.3(6)	<b>1.9</b> (1)	15/15
IPOP-10DDr	<b>1.0</b> (0.2)	<b>1.5</b> (2)	<b>2.6</b> (2)	<b>2.4</b> (1.0)	<b>2.9</b> (3)	15/15
IPOP-500 l	<b>1.0</b> (0.2)	<b>1.5</b> (2)	<b>2.6</b> (2)	<b>2.4</b> (1.0)	<b>2.9</b> (3)	15/15
IPOP-tany	<b>0.64</b> (0.2)	<b>1.0</b> (1)	<b>1.7</b> (1)	<b>2.5</b> (1)	<b>1.9</b> (3)	15/15
IPOP-texp	<b>0.64</b> (0.2)	<b>0.86</b> (1.0)	<b>1.3</b> (0.6)	<b>1.3</b> (0.6)	<b>2.5</b> (3)	15/15
IPOP lia	<b>1.0</b> (0.2)	<b>1.5</b> (2)	<b>2.6</b> (2)	<b>2.4</b> (1.0)	<b>2.9</b> (3)	15/15
MLSL pal	<b>0.69</b> (0)	<b>0.94</b> (1)	4.9(9)	18(26)	15(18)	15/15
OQNLP pal	<b>0.72</b> (0)	4.1(12)	12(11)	44(77)	34(47)	9/15
P-DCN tra	4.7(2)	3.9(3)	5.3(5)	22(38)	77(86)	15/15
P-zero tra	<b>1.6</b> (2)	<b>1.5</b> (1)	<b>1.7</b> (1)	11(13)	43(14)	15/15
SMAC hut	<b>0.56</b> (0.4)	<b>0.54</b> (0.6)	<b>1.8</b> (3)	14(13)	$\infty$ 500	0/15
U-DCN tra	<b>1.8</b> (2)	<b>1.7</b> (2)	<b>2.7</b> (1)	<b>2.7</b> (1)	<b>1.2</b> (0.6)	15/15
U-zero tra	<b>1.7</b> (2)	<b>1.3</b> (1)	<b>1.5</b> (0.6)	<b>1.8</b> (0.8)	<b>0.84</b> (0.4)	15/15
fmincon pa	<b>0.69</b> (0)	<b>0.94</b> (1)	4.9(9)	12(20)	13(18)	15/15
fminunc pa	<b>0.69</b> (0)	<b>1.6</b> (2)	4.6(4)	13(14)	19(20)	15/15
ga100 hol	<b>2.5</b> (3)	<b>2.9</b> (2)	6.3(5)	12(3)	3.5(0.9)	15/15
grid100 ho	4.1(5)	<b>2.7</b> (3)	8.6(8)	28(16)	10(4)	15/15
grid16 hol	<b>2.6</b> (2)	<b>2.9</b> (4)	5.6(5)	6.7(4)	<b>2.1</b> (1.0)	15/15
hill hol	7.9(8)	3.9(2)	<b>2.7</b> (2)	<b>1.5</b> (0.7)	<b>0.61</b> (0.3)	15/15
lmmCMA aug	<b>0.51</b> (0.4)	<b>0.79</b> (1)	<b>0.93</b> (1)	<b>2.2</b> (0.9)	<b>1.3</b> (1)	15/15
memPSODE v	<b>2.9</b> (2)	3.2(3)	19(16)	14(10)	4.9(3)	15/15
prcga saw	<b>1.6</b> (2)	<b>0.77</b> (1.0)	<b>1.6</b> (2)	5.2(2)	<b>2.0</b> (1)	15/15
ring100 ho	3.1(3)	<b>1.9</b> (2)	5.7(6)	20(8)	7.0(2)	15/15
ring16 hol	<b>1.7</b> (2)	<b>2.8</b> (4)	5.2(4)	5.2(2)	<b>1.6</b> (0.5)	15/15
simplex pa	<b>1.4</b> (0)	19(52)	41(32)	43(44)	29(26)	15/15

Table 6: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f5</b>	<i>6.3e+1:4.0</i>	<i>4.0e+1:10</i>	<i>1.0e-8:10</i>	<i>1.0e-8:10</i>	<i>1.0e-8:10</i>	15/15
BIPOP-aCMA	<b>1.4</b> (1)	<b>1.0</b> (0.3)	<b>1.2</b> (0)	<b>1.2</b> (0)	<b>1.2</b> (0)	15/15
BIPOP-saAC	3.9(4)	<b>2.4</b> (2)	6.4(2)	6.4(2)	6.4(2)	15/15
CMAES hut	3.2(2)	<b>1.9</b> (2)	5.6(1)	5.6(1)	5.6(1)	15/15
DE pal	<b>1.6</b> (2)	3.3(3)	859(45)	859(45)	859(45)	15/15
HCMA los	<b>2.4</b> (0.9)	<b>1.2</b> (0.1)	<b>1.5</b> (0.3)	<b>1.5</b> (0.3)	<b>1.5</b> (0.3)	15/15
HMLSL pal	<b>1.8</b> (0)	<b>0.71</b> (0)	503(698)	503(698)	503(698)	15/15
IPOP-10DDr	<b>1.7</b> (1)	<b>2.1</b> (1)	46(43)	46(43)	46(43)	15/15
IPOP-500 l	<b>1.7</b> (1)	<b>2.1</b> (1)	46(43)	46(43)	46(43)	15/15
IPOP-tany	<b>2.6</b> (3)	<b>2.3</b> (1)	67(60)	67(60)	67(60)	15/15
IPOP-texp	7.3(2)	4.2(1)	89(71)	89(71)	89(71)	15/15
IPOP lia	<b>1.7</b> (1)	<b>2.1</b> (1)	46(43)	46(43)	46(43)	15/15
MLSL pal	<b>1.8</b> (0)	<b>0.71</b> (0)	527(461)	527(461)	527(461)	14/15
OQNLP pal	5.0(0)	<b>2.1</b> (0)	<b>2.4</b> (0)	<b>2.4</b> (0)	<b>2.4</b> (0)	15/15
P-DCN tra	<b>1.5</b> (2)	3.2(4)	326(66)	326(66)	326(66)	15/15
P-zero tra	<b>1.2</b> (1)	4.0(9)	171(79)	171(79)	171(79)	15/15
SMAC hut	<b>1.3</b> (0.2)	<b>0.63</b> (0.2)	<b>0.95</b> (0.1)* <sup>3</sup>	<b>0.95</b> (0.1)* <sup>3</sup>	<b>0.95</b> (0.1)* <sup>3</sup>	15/15
U-DCN tra	<b>1.4</b> (2)	<b>2.1</b> (2)	3.3e6(4e6)	3.3e6(4e6)	3.3e6(4e6)	2/15
U-zero tra	<b>1.3</b> (1)	<b>2.0</b> (3)	$\infty$	$\infty$	$\infty$ <i>5e6</i>	0/15
fmincon pa	<b>1.8</b> (0)	<b>0.71</b> (0)	1250(2807)	1250(2794)	1250(2794)	13/15
fminunc pa	3.2(0)	<b>1.3</b> (0)	<b>3.1</b> (0)	<b>3.1</b> (0)	<b>3.1</b> (0)	15/15
ga100 hol	<b>2.3</b> (2)	5.0(5)	124(27)	124(27)	124(27)	15/15
grid100 ho	<b>2.2</b> (2)	4.5(4)	264(76)	264(76)	264(76)	15/15
grid16 hol	3.8(7)	10(6)	51(14)	51(14)	51(14)	15/15
hill hol	6.8(6)	4.1(2)	11(4)	11(4)	11(4)	15/15
lmmCMA aug	3.1(2)	<b>1.9</b> (0.8)	5.0(1)	5.0(1)	5.0(1)	15/15
memPSODE v	<b>1.8</b> (2)	<b>2.5</b> (1.0)	14(6)	14(6)	14(6)	15/15
prcga saw	3.5(4)	11(6)	$\infty$	$\infty$	$\infty$ <i>3e5</i>	0/15
ring100 ho	<b>1.8</b> (2)	6.1(8)	201(42)	201(42)	201(42)	15/15
ring16 hol	4.6(10)	6.3(6)	48(12)	48(12)	48(12)	15/15
simplex pa	19(0.4)	12(13)	90(97)	90(97)	90(97)	15/15

Table 7: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best 2009}}$  on  $f_6$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best 2009}}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FES/D	0.5	1.2	3	10	50	#succ
<b>f6</b>	<i>1.0e+5</i> :3.0	<i>2.5e+4</i> :8.4	<i>1.0e+2</i> :16	<i>2.5e+1</i> :54	<i>2.5e-1</i> :254	15/15
BIPOP-aCMA	3.8(6)	<b>2.4</b> (4)	5.1(4)	3.4(3)	<b>2.2</b> (0.5)	15/15
BIPOP-saAC	5.4(7)	<b>2.7</b> (2)	3.5(4)	3.1(2)	<b>2.5</b> (1.0)	15/15
CMAES hut	<b>2.9</b> (4)	<b>1.6</b> (2)	3.2(4)	<b>2.7</b> (2)	<b>2.4</b> (1)	11/15
DE pal	<b>2.3</b> (3)	<b>1.8</b> (2)	5.8(7)	3.2(2)	6.8(2)	15/15
HCMA los	3.2(1)	<b>1.4</b> (0.7)	<b>1.5</b> (1)	<b>2.7</b> (3)	<b>2.4</b> (1.0)	15/15
HMLSL pal	<b>2.8</b> (4)	<b>2.3</b> (2)	3.4(2)	<b>1.7</b> (1)	<b>1.2</b> (0.6)	15/15
IPOP-10DDr	<b>2.4</b> (2)	<b>2.0</b> (2)	3.1(3)	<b>2.0</b> (0.9)	<b>2.0</b> (0.3)	15/15
IPOP-500 l	<b>2.4</b> (2)	<b>2.0</b> (2)	3.1(3)	<b>2.0</b> (0.9)	<b>2.0</b> (0.3)	15/15
IPOP-tany	<b>2.6</b> (3)	<b>2.3</b> (3)	<b>2.8</b> (2)	<b>1.8</b> (0.9)	<b>2.1</b> (0.4)	15/15
IPOP-texp	4.1(6)	3.2(4)	3.6(2)	<b>1.7</b> (1.0)	<b>2.1</b> (0.6)	15/15
IPOP lia	<b>2.4</b> (2)	<b>2.0</b> (2)	3.1(3)	<b>2.0</b> (0.9)	<b>2.0</b> (0.3)	15/15
MLSL pal	<b>2.8</b> (4)	<b>2.3</b> (2)	3.4(2)	<b>1.7</b> (1)	<b>1.2</b> (0.6)	15/15
OQNLP pal	3.6(3)	<b>2.3</b> (2)	<b>2.2</b> (0.8)	<b>1.1</b> (0.6)	<b>0.98</b> (0.4)	15/15
P-DCN tra	<b>2.4</b> (4)	<b>1.9</b> (2)	7.7(7)	8.8(8)	8.0(4)	15/15
P-zero tra	<b>2.2</b> (3)	<b>2.4</b> (3)	10(13)	8.2(8)	7.2(7)	15/15
SMAC hut	<b>1.4</b> (1)	<b>1.1</b> (1)	<b>1.5</b> (2)	<b>1.9</b> (2)	$\infty$ 500	0/15
U-DCN tra	3.2(3)	<b>2.5</b> (4)	7.7(4)	6.9(10)	272(433)	15/15
U-zero tra	<b>2.1</b> (2)	<b>1.3</b> (1)	6.1(4)	3.8(3)	1830(1556)	14/15
fmincon pa	<b>2.8</b> (4)	<b>2.3</b> (2)	3.2(2)	<b>1.6</b> (0.8)	<b>1.2</b> (0.6)	15/15
fminunc pa	<b>2.5</b> (3)	<b>1.8</b> (1)	<b>2.4</b> (1)	<b>1.5</b> (1)	<b>2.5</b> (1)	15/15
ga100 hol	3.6(2)	4.1(5)	13(14)	11(7)	60(34)	15/15
grid100 ho	<b>2.8</b> (3)	<b>2.1</b> (2)	24(33)	22(18)	657(586)	12/15
grid16 hol	4.7(10)	4.4(8)	9.1(6)	8.2(8)	207(165)	14/15
hill hol	4.7(7)	<b>2.5</b> (3)	4.1(5)	3.5(2)	31(29)	15/15
lmmCMA aug	<b>1.6</b> (2)	<b>1.4</b> (1)	4.7(5)	3.8(4)	4.5(3)	15/15
memPSODE v	<b>2.0</b> (1)	<b>1.5</b> (2)	7.8(8)	5.4(4)	<b>2.8</b> (2)	15/15
prcga saw	<b>1.5</b> (2)	3.5(8)	8.3(8)	6.3(5)	218(297)	15/15
ring100 ho	<b>1.9</b> (3)	3.6(7)	10(12)	10(7)	96(50)	15/15
ring16 hol	<b>1.9</b> (2)	<b>2.6</b> (3)	7.0(8)	5.0(4)	23(14)	15/15
simplex pa	21(30)	13(13)	34(55)	22(14)	14(8)	15/15



Table 8: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b><math>f_7</math></b>	<i>1.6e+2:4.2</i>	<i>1.0e+2:6.2</i>	<i>2.5e+1:20</i>	<i>4.0e+0:54</i>	<i>1.0e+0:324</i>	15/15
BIPOP-aCMA	<b>2.0</b> (2)	<b>2.8</b> (3)	4.7(3)	4.0(2)	<b>1.5</b> (1)	15/15
BIPOP-saAC	3.3(3)	3.6(3)	3.0(2)	<b>2.6</b> (0.6)	<b>1.2</b> (1)	15/15
CMAES hut	<b>2.5</b> (3)	<b>2.8</b> (3)	<b>2.7</b> (2)	3.4(2)	<b>1.6</b> (2)	10/15
DE pal	<b>2.0</b> (2)	<b>1.6</b> (2)	3.7(3)	8.3(6)	<b>2.8</b> (1)	15/15
HCMA los	<b>2.8</b> (1)	<b>2.2</b> (0.6)	<b>2.3</b> (2)	3.6(3)	<b>1.0</b> (0.3)	15/15
HMLSL pal	<b>1.9</b> (2)	<b>2.0</b> (2)	5.5(7)	10(4)	3.8(2)	15/15
IPOP-10DDr	<b>1.2</b> (1)	<b>1.2</b> (1)	<b>2.5</b> (2)	<b>3.3</b> (1)	<b>1.1</b> (0.4)	15/15
IPOP-500 l	<b>1.2</b> (1)	<b>1.2</b> (1)	<b>2.5</b> (2)	<b>3.3</b> (1)	<b>1.1</b> (0.4)	15/15
IPOP-tany	<b>1.2</b> (2)	<b>1.5</b> (2)	<b>1.9</b> (2)	<b>2.6</b> (2)	<b>1.7</b> (2)	15/15
IPOP-texp	<b>1.3</b> (1)	<b>1.8</b> (2)	<b>1.7</b> (1)	4.8(10)	<b>2.1</b> (2)	15/15
IPOP lia	<b>1.2</b> (1)	<b>1.2</b> (1)	<b>2.5</b> (2)	<b>3.3</b> (1)	<b>1.1</b> (0.4)	15/15
MLSL pal	<b>1.8</b> (2)	<b>2.1</b> (3)	14(22)	219(203)	1097(1236)	2/15
OQNLP pal	4.0(4)	7.1(12)	11(13)	29(28)	48(57)	2/15
P-DCN tra	6.0(5)	7.6(4)	46(58)	6812(560)	1695(2334)	14/15
P-zero tra	21(11)	74(40)	61(61)	6831(378)	2798(7771)	13/15
SMAC hut	<b>1.3</b> (1)	<b>1.1</b> (0.9)	<b>1.5</b> (1)	<b>1.6</b> (0.8)	<b>0.88</b> (0.9)	13/15
U-DCN tra	3.0(5)	<b>2.8</b> (3)	4.6(4)	31(28)	1132(41)	14/15
U-zero tra	3.2(3)	<b>3.0</b> (2)	3.9(3)	48(37)	1488(2964)	15/15
fmincon pa	<b>2.7</b> (3)	3.4(3)	8.8(8)	164(139)	1086(1244)	2/15
fminunc pa	<b>1.5</b> (1)	<b>2.6</b> (4)	8.1(10)	184(193)	$\infty$ 5e4	0/15
ga100 hol	<b>1.8</b> (2)	3.2(4)	6.9(7)	24(17)	11(10)	15/15
grid100 ho	<b>2.8</b> (4)	<b>2.4</b> (3)	9.3(10)	44(52)	74(83)	15/15
grid16 hol	<b>0.73</b> (0.6)	<b>2.8</b> (6)	7.6(6)	26(32)	97(110)	14/15
hill hol	8.6(6)	7.5(6)	5.4(3)	141(132)	117(200)	15/15
lmmCMA aug	<b>1.2</b> (1)	<b>1.3</b> (1)	<b>1.5</b> (1)	<b>2.3</b> (4)	<b>0.92</b> (1.0)	15/15
memPSODE v	3.7(4)	3.5(3)	10(14)	18(20)	7.5(6)	15/15
prcga saw	<b>1.4</b> (1.0)	<b>1.8</b> (2)	4.1(4)	9.1(5)	17(17)	15/15
ring100 ho	<b>1.6</b> (1)	<b>2.8</b> (2)	13(17)	30(19)	19(11)	15/15
ring16 hol	<b>2.5</b> (3)	3.3(4)	7.3(6)	10(8)	104(276)	15/15
simplex pa	<b>1.6</b> (2)	<b>2.9</b> (3)	11(11)	93(69)	155(137)	10/15

Table 9: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best } 2009}$  on  $f_8$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best } 2009}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f8</b>	<i>1.0e+4</i> :4.6	<i>6.3e+3</i> :6.8	<i>1.0e+3</i> :18	<i>6.3e+1</i> :54	<i>1.6e+0</i> :258	15/15
BIPOP-aCMA	4.8(4)	3.8(3)	4.5(2)	4.3(2)	4.0(4)	15/15
BIPOP-saAC	<b>2.7</b> (4)	<b>2.4</b> (3)	<b>2.6</b> (2)	<b>2.3</b> (0.8)	<b>2.2</b> (2)	15/15
CMAES hut	3.7(3)	<b>2.8</b> (3)	<b>2.9</b> (2)	<b>2.6</b> (1)	3.2(2)	8/15
DE pal	3.6(3)	<b>2.9</b> (2)	4.7(2)	5.4(2)	5.7(2)	15/15
HCMA los	<b>1.8</b> (1)	<b>1.4</b> (0.9)	<b>0.85</b> (0.3)	<b>1.2</b> (2)	<b>2.3</b> (3)	15/15
HMLSL pal	<b>1.2</b> (2)	<b>1.1</b> (1)	<b>1.2</b> (1)	<b>0.92</b> (0.5)	<b>1.3</b> (2)	15/15
IPOP-10DDr	<b>2.9</b> (3)	<b>2.5</b> (2)	<b>2.6</b> (1)	<b>2.6</b> (0.9)	4.0(1)	15/15
IPOP-500 l	<b>2.9</b> (3)	<b>2.5</b> (2)	<b>2.6</b> (1)	<b>2.6</b> (0.9)	4.0(1)	15/15
IPOP-tany	<b>2.6</b> (3)	<b>2.5</b> (2)	<b>2.2</b> (1)	<b>2.1</b> (0.6)	3.3(1)	15/15
IPOP-texp	<b>2.9</b> (3)	<b>2.4</b> (2)	<b>2.1</b> (1)	<b>1.7</b> (0.5)	4.3(3)	15/15
IPOP lia	<b>2.9</b> (3)	<b>2.5</b> (2)	<b>2.6</b> (1)	<b>2.6</b> (0.9)	4.0(1)	15/15
MLSL pal	<b>1.2</b> (2)	<b>1.1</b> (1)	<b>1.2</b> (1)	<b>0.92</b> (0.5)	<b>1.0</b> (1)	15/15
OQNLP pal	<b>1.9</b> (2)	<b>1.6</b> (1)	<b>1.2</b> (0.3)	<b>0.84</b> (0.5)	<b>0.68</b> (0.3)	15/15
P-DCN tra	5.5(4)	5.5(6)	8.2(11)	27(30)	7096(9725)	11/15
P-zero tra	3.0(3)	3.9(7)	6.7(7)	22(24)	3003(9688)	13/15
SMAC hut	<b>0.99</b> (1.0)	<b>0.91</b> (1)	<b>1.2</b> (1)	3.3(2)	$\infty$ 500	0/15
U-DCN tra	3.7(5)	3.3(5)	3.5(3)	10(9)	279(545)	15/15
U-zero tra	<b>2.2</b> (2)	<b>1.9</b> (2)	<b>2.7</b> (2)	7.3(5)	3004(7911)	14/15
fmincon pa	<b>1.2</b> (2)	<b>1.1</b> (1)	<b>1.2</b> (1)	<b>0.92</b> (0.5)	<b>1.0</b> (1)	15/15
fminunc pa	<b>0.83</b> (0.7)	<b>0.79</b> (0.9)	<b>0.84</b> (0.5)	<b>1.1</b> (0.8)	<b>0.88</b> (0.4)	15/15
ga100 hol	3.1(3)	3.3(3)	9.4(9)	16(7)	295(493)	12/15
grid100 ho	<b>1.7</b> (2)	3.1(2)	12(13)	53(36)	426(525)	11/15
grid16 hol	4.0(6)	4.7(6)	11(8)	15(12)	504(969)	10/15
hill hol	4.4(5)	3.4(4)	<b>2.7</b> (2)	5.7(5)	250(442)	14/15
lmmCMA aug	<b>1.0</b> (1)	<b>0.96</b> (1)	<b>1.6</b> (1)	<b>1.5</b> (0.6)	<b>1.7</b> (2)	15/15
memPSODE v	3.8(5)	4.7(4)	4.8(3)	3.5(2)	3.5(5)	15/15
prcga saw	<b>2.1</b> (2)	<b>1.5</b> (2)	4.4(4)	7.5(3)	267(652)	15/15
ring100 ho	<b>2.8</b> (3)	3.2(3)	7.0(6)	28(16)	54(24)	15/15
ring16 hol	3.0(2)	3.8(4)	6.1(5)	9.1(6)	272(484)	13/15
simplex pa	7.9(7)	7.5(10)	7.6(5)	10(7)	5.1(3)	15/15

Table 10: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best 2009}}$  on  $f_9$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best 2009}}$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FES/D	0.5	1.2	3	10	50	#succ
<b>f9</b>	<i>2.5e+1:20</i>	<i>1.6e+1:26</i>	<i>1.0e+1:35</i>	<i>4.0e+0:62</i>	<i>1.6e-2:256</i>	15/15
BIPOP-aCMA	10(3)	8.4(2)	7.0(2)	7.0(2)	5.9(1)	15/15
BIPOP-saAC	7.1(2)	5.8(1)	4.6(0.9)	3.7(2)	<b>2.7</b> (0.6)	15/15
CMAES hut	8.4(3)	7.3(3)	6.1(3)	6.0(4)	$\infty$ <i>506</i>	0/15
DE pal	31(13)	27(9)	24(10)	22(13)	33(8)	15/15
HCMA los	7.0(6)	6.4(6)	5.7(5)	5.6(3)	3.6(3)	15/15
HMLSL pal	<b>0.66</b> (0)	<b>1.1</b> (0.2)	<b>0.94</b> (0)	<b>0.80</b> (0.1)	<b>0.60</b> (0.0)	15/15
IPOP-10DDr	8.8(4)	7.5(3)	6.1(2)	7.3(5)	8.1(5)	15/15
IPOP-500 l	8.8(4)	7.5(3)	6.1(2)	7.3(5)	8.1(5)	15/15
IPOP-tany	6.9(3)	5.9(2)	5.2(2)	7.3(8)	8.2(7)	15/15
IPOP-texp	3.6(2)	3.3(1)	3.0(2)	4.6(4)	7.0(2)	15/15
IPOP lia	8.8(4)	7.5(3)	6.1(2)	7.3(5)	8.1(5)	15/15
MLSL pal	<b>0.66</b> (0)	<b>1.1</b> (0.2)	<b>0.94</b> (0)	<b>0.80</b> (0.1)	<b>0.60</b> (0.0)	15/15
OQNLP pal	<b>0.71</b> (0)	<b>0.62</b> (0)	<b>0.71</b> (0.0)	<b>0.43</b> (8e-3)	<b>0.54</b> (0.0)	15/15
P-DCN tra	776(1712)	597(1314)	446(980)	2.1e4(4e4)	1.7e4(2e4)	9/15
P-zero tra	1194(3320)	918(2547)	685(1900)	482(1068)	1.2e4(7875)	14/15
SMAC hut	14(6)	12(4)	12(8)	120(130)	$\infty$ <i>500</i>	0/15
U-DCN tra	64(43)	64(49)	65(59)	9769(8356)	2.5e4(3e4)	8/15
U-zero tra	648(98)	501(74)	378(55)	2.1e4(4e4)	$\infty$ <i>5e6</i>	0/15
fmincon pa	<b>0.66</b> (0)	<b>1.1</b> (0.2)	<b>0.94</b> (0)	<b>0.80</b> (0.1)	<b>0.60</b> (0.0)	15/15
fminunc pa	<b>0.66</b> (0)	<b>0.50</b> (0)* <sup>4</sup>	<b>0.55</b> (0)* <sup>4</sup>	<b>0.41</b> (0)* <sup>4</sup>	<b>0.51</b> (0.0)* <sup>2</sup>	15/15
ga100 hol	72(30)	68(26)	58(23)	159(57)	$\infty$ <i>2e5</i>	0/15
grid100 ho	268(214)	380(315)	315(218)	1823(2190)	$\infty$ <i>2e5</i>	0/15
grid16 hol	71(36)	62(41)	57(42)	1802(2245)	$\infty$ <i>2e5</i>	0/15
hill hol	14(9)	14(12)	13(11)	1159(2050)	$\infty$ <i>2e5</i>	0/15
lmmCMA aug	3.7(2)	3.3(1)	<b>2.7</b> (0.9)	<b>2.7</b> (2)	<b>2.4</b> (1)	15/15
memPSODE v	10(4)	7.7(3)	5.9(2)	6.2(2)	4.2(6)	15/15
prcga saw	8.8(3)	11(6)	11(5)	11(4)	$\infty$ <i>5e5</i>	0/15
ring100 ho	136(41)	133(38)	138(47)	166(61)	$\infty$ <i>2e5</i>	0/15
ring16 hol	32(15)	30(11)	30(11)	327(16)	$\infty$ <i>2e5</i>	0/15
simplex pa	<b>1.7</b> (0.1)	4.0(3)	5.1(3)	4.1(2)	3.2(0.8)	15/15

Table 11: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{10}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f10</b>	<i>2.5e+6:2.9</i>	<i>6.3e+5:7.0</i>	<i>2.5e+5:17</i>	<i>6.3e+3:54</i>	<i>2.5e+1:297</i>	15/15
BIPOP-aCMA	<b>1.6</b> (2)	<b>1.7</b> (2)	<b>1.5</b> (1)	3.5(2)	<b>2.9</b> (0.8)	15/15
BIPOP-saAC	<b>2.2</b> (2)	<b>2.1</b> (3)	<b>1.4</b> (1)	<b>2.4</b> (1)	<b>0.79</b> (0.1)	15/15
CMAES hut	<b>2.7</b> (3)	<b>2.4</b> (3)	<b>1.5</b> (1)	3.8(2)	$\infty$ 506	0/15
DE pal	<b>1.3</b> (1)	<b>1.3</b> (1)	<b>0.94</b> (0.8)	6.2(4)	20(9)	15/15
HCMA los	<b>1.1</b> (0.7)	<b>1.1</b> (0.9)	<b>1.00</b> (0.6)	<b>2.1</b> (2)	<b>1.2</b> (0.1)	15/15
HMLSL pal	<b>2.9</b> (2)	<b>1.9</b> (3)	<b>1.1</b> (1)	<b>0.88</b> (0.3)	<b>0.31</b> (0.1)	15/15
IPOP-10DDr	<b>1.1</b> (0.9)	<b>1.5</b> (2)	<b>1.2</b> (0.8)	5.3(3)	3.4(0.9)	15/15
IPOP-500 l	<b>1.1</b> (0.9)	<b>1.5</b> (2)	<b>1.2</b> (0.8)	5.3(3)	3.4(0.9)	15/15
IPOP-tany	<b>1.1</b> (0.9)	<b>1.2</b> (1)	<b>0.76</b> (0.7)	3.7(2)	3.4(1)	15/15
IPOP-texp	<b>2.6</b> (4)	<b>2.6</b> (3)	<b>1.9</b> (2)	3.9(3)	3.5(1)	15/15
IPOP lia	<b>1.1</b> (0.9)	<b>1.5</b> (2)	<b>1.2</b> (0.8)	5.3(3)	3.4(0.9)	15/15
MLSL pal	<b>2.9</b> (2)	<b>1.9</b> (3)	<b>1.1</b> (1)	<b>0.88</b> (0.3)	<b>0.31</b> (0.1)	15/15
OQNLP pal	4.3(3)	<b>2.0</b> (1)	<b>1.2</b> (0.9)	<b>0.93</b> (0.5)	<b>0.43</b> (0.3)	15/15
P-DCN tra	<b>1.5</b> (1)	<b>2.4</b> (3)	3.4(6)	23(16)	4834(8682)	12/15
P-zero tra	<b>1.5</b> (1)	4.3(2)	4.7(6)	27(26)	1.1e4(1e4)	11/15
SMAC hut	<b>1.3</b> (0.9)	<b>0.80</b> (0.6)	<b>0.58</b> (0.6)	<b>2.5</b> (2)	$\infty$ 500	0/15
U-DCN tra	<b>1.3</b> (1)	<b>1.2</b> (0.9)	<b>2.2</b> (1)	20(8)	1.2e4(2e4)	11/15
U-zero tra	<b>1.3</b> (1)	<b>1.7</b> (2)	<b>1.2</b> (0.9)	75(69)	7.3e4(8e4)	3/15
fmincon pa	<b>2.9</b> (2)	<b>1.9</b> (3)	<b>1.1</b> (1)	<b>0.88</b> (0.3)	<b>0.30</b> (0.1)	15/15
fminunc pa	<b>2.4</b> (2)	<b>1.6</b> (2)	<b>1.1</b> (0.9)	<b>1.6</b> (1)	<b>0.97</b> (0.5)	15/15
ga100 hol	<b>1.4</b> (1)	<b>1.8</b> (3)	<b>1.2</b> (1)	21(19)	1562(1903)	6/15
grid100 ho	<b>1.7</b> (1)	<b>1.2</b> (0.8)	<b>1.6</b> (2)	49(80)	$\infty$ 2e5	0/15
grid16 hol	<b>1.7</b> (2)	<b>2.9</b> (2)	<b>2.3</b> (2)	32(33)	1.2e4(1e4)	1/15
hill hol	4.7(5)	4.0(5)	<b>2.3</b> (3)	42(65)	1.2e4(1e4)	1/15
lmmCMA aug	<b>2.3</b> (3)	<b>1.5</b> (1)	<b>1.1</b> (0.9)	<b>1.5</b> (0.4)	<b>0.83</b> (0.4)	15/15
memPSODE v	<b>1.8</b> (1)	<b>2.0</b> (2)	<b>1.8</b> (1)	3.2(1)	4.1(8)	15/15
prcga saw	<b>1.1</b> (0.9)	<b>0.86</b> (0.9)	<b>0.92</b> (0.8)	272(136)	3384(4085)	5/15
ring100 ho	1(2)	<b>0.71</b> (0.7)	<b>1.1</b> (2)	27(18)	2047(2172)	5/15
ring16 hol	<b>1.6</b> (1)	<b>1.8</b> (2)	<b>2.4</b> (2)	10(9)	3927(4278)	3/15
simplex pa	15(13)	8.2(5)	5.0(6)	7.0(6)	5.8(2)	15/15

Table 12: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{11}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f11</b>	<i>1.0e+6</i> :3.0	<i>6.3e+4</i> :6.2	<i>6.3e+2</i> :16	<i>6.3e+1</i> :74	<i>6.3e-1</i> :298	15/15
BIPOP-aCMA	<b>2.0</b> (4)	3.0(3)	3.8(4)	7.0(5)	3.4(0.5)	15/15
BIPOP-saAC	3.5(3)	<b>2.8</b> (2)	5.0(7)	3.4(0.6)	<b>1.1</b> (0.1)	15/15
CMAES hut	<b>2.9</b> (2)	<b>2.6</b> (3)	5.5(5)	5.3(5)	$\infty$ 506	0/15
DE pal	<b>1.6</b> (2)	<b>2.9</b> (3)	9.0(8)	5.4(6)	28(8)	15/15
HCMA los	<b>2.5</b> (2)	<b>1.8</b> (1.0)	<b>1.4</b> (0.6)	3.1(2)	<b>1.4</b> (0.2)	15/15
HMLSL pal	3.2(2)	<b>2.5</b> (2)	<b>1.4</b> (0.3)	<b>0.36</b> (0.1)	<b>0.15</b> (0.0)	15/15
IPOP-10DDr	<b>2.0</b> (2)	3.2(4)	6.2(3)	4.8(6)	4.9(0.4)	15/15
IPOP-500 l	<b>2.0</b> (2)	3.2(4)	6.2(3)	4.8(6)	4.9(0.4)	15/15
IPOP-tany	<b>1.9</b> (2)	<b>2.4</b> (2)	5.3(4)	6.0(6)	4.9(0.5)	15/15
IPOP-texp	5.7(6)	4.3(3)	3.7(2)	4.0(6)	4.7(0.8)	15/15
IPOP lia	<b>2.0</b> (2)	3.2(4)	6.2(3)	4.8(6)	4.9(0.4)	15/15
MLSL pal	3.2(2)	<b>2.5</b> (2)	<b>1.4</b> (0.3)	<b>0.36</b> (0.1)	<b>0.15</b> (0.0)	15/15
OQNLP pal	4.4(3)	<b>2.8</b> (2)	<b>2.1</b> (1)	<b>0.56</b> (0.4)	<b>0.65</b> (1.0)	15/15
P-DCN tra	<b>2.3</b> (2)	4.4(4)	4.5(5)	171(238)	646(402)	15/15
P-zero tra	<b>2.1</b> (2)	3.4(6)	3.8(3)	112(216)	967(578)	15/15
SMAC hut	<b>0.73</b> (0.5)	<b>0.94</b> (0.9)	<b>1.9</b> (2)	<b>0.94</b> (0.8)	$\infty$ 500	0/15
U-DCN tra	<b>1.8</b> (2)	<b>2.6</b> (2)	4.6(4)	12(8)	5350(4218)	14/15
U-zero tra	<b>2.5</b> (1)	<b>2.5</b> (3)	5.9(3)	27(34)	4759(4840)	14/15
fmincon pa	3.2(2)	<b>2.5</b> (2)	<b>1.4</b> (0.3)	<b>0.36</b> (0.1)	<b>0.16</b> (0.0)	15/15
fminunc pa	3.5(3)	<b>2.4</b> (1)	<b>1.7</b> (0.8)	<b>0.46</b> (0.1)	<b>1.7</b> (2)	15/15
ga100 hol	<b>1.7</b> (2)	<b>3.0</b> (3)	14(14)	9.5(12)	2161(2359)	5/15
grid100 ho	<b>2.4</b> (3)	3.4(3)	11(9)	210(299)	$\infty$ 2e5	0/15
grid16 hol	<b>2.6</b> (2)	3.5(4)	7.0(5)	148(354)	5991(6296)	2/15
hill hol	3.1(4)	3.2(4)	4.9(5)	45(83)	1433(1428)	7/15
lmmCMA aug	<b>1.4</b> (2)	<b>2.1</b> (2)	<b>2.6</b> (1.0)	<b>1.9</b> (1)	<b>1.3</b> (0.3)	15/15
memPSODE v	<b>1.5</b> (1)	<b>3.0</b> (3)	5.9(2)	<b>2.0</b> (0.7)	5.9(7)	15/15
prcga saw	<b>2.3</b> (1)	<b>2.8</b> (3)	6.7(5)	3.8(3)	$\infty$ 3e5	0/15
ring100 ho	<b>1.6</b> (2)	5.3(4)	8.4(10)	16(14)	2746(2938)	4/15
ring16 hol	<b>1.9</b> (2)	<b>2.4</b> (3)	8.8(8)	12(8)	2695(2938)	4/15
simplex pa	16(12)	10(6)	5.4(2)	<b>1.4</b> (0.6)	8.5(2)	15/15

Table 13: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{12}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f12</b>	<i>4.0e+7:3.6</i>	<i>1.6e+7:7.6</i>	<i>4.0e+6:19</i>	<i>1.6e+4:52</i>	<i>1.0e+0:268</i>	15/15
BIPOP-aCMA	3.3(5)	5.1(4)	3.7(2)	5.4(0.8)	4.7(3)	15/15
BIPOP-saAC	<b>3.0</b> (3)	3.5(3)	<b>2.9</b> (2)	<b>2.5</b> (0.3)	<b>1.9</b> (2)	15/15
CMAES hut	<b>2.0</b> (3)	<b>1.7</b> (2)	<b>2.1</b> (2)	4.2(1)	28(30)	1/15
DE pal	4.0(5)	3.1(3)	3.8(3)	16(4)	50(20)	15/15
HCMA los	<b>2.7</b> (2)	<b>2.0</b> (1)	<b>1.2</b> (0.4)	3.3(2)	4.4(8)	15/15
HMLSL pal	<b>1.4</b> (1.0)	<b>1.6</b> (2)	<b>1.5</b> (0.8)	<b>1.5</b> (0.4)	<b>0.94</b> (0.6)	15/15
IPOP-10DDr	<b>2.3</b> (2)	<b>2.9</b> (2)	<b>2.6</b> (1)	4.1(1)	5.9(3)	15/15
IPOP-500 l	<b>2.3</b> (2)	<b>2.9</b> (2)	<b>2.6</b> (1)	4.1(1)	5.9(3)	15/15
IPOP-tany	<b>1.7</b> (2)	<b>2.4</b> (2)	<b>2.4</b> (1)	4.8(1.0)	7.1(5)	15/15
IPOP-texp	<b>2.1</b> (2)	3.4(2)	<b>2.6</b> (1)	4.2(0.6)	8.7(9)	15/15
IPOP lia	<b>2.3</b> (2)	<b>2.9</b> (2)	<b>2.6</b> (1)	4.1(1)	5.9(3)	15/15
MLSL pal	<b>1.4</b> (1.0)	<b>1.6</b> (2)	<b>1.5</b> (0.8)	<b>1.5</b> (0.4)	<b>0.94</b> (0.6)	15/15
OQNLP pal	3.1(2)	<b>2.3</b> (1)	<b>1.3</b> (0.4)	<b>1.5</b> (0.9)	<b>0.95</b> (0.8)	15/15
P-DCN tra	<b>2.0</b> (2)	<b>2.8</b> (3)	7.2(7)	10(4)	2.8e4(4e4)	6/15
P-zero tra	<b>1.2</b> (1)	5.3(9)	9.4(11)	10(4)	1.2e5(1e5)	2/15
SMAC hut	<b>0.57</b> (0.4)	<b>1.3</b> (2)	3.6(5)	34(34)	$\infty$ 500	0/15
U-DCN tra	<b>1.3</b> (2)	<b>2.8</b> (3)	4.4(4)	33(10)	3.7e4(5e4)	5/15
U-zero tra	<b>1.3</b> (2)	<b>2.3</b> (2)	3.0(1)	21(16)	3.8e4(5e4)	5/15
fmincon pa	<b>1.4</b> (1.0)	<b>1.6</b> (2)	<b>1.5</b> (0.8)	<b>1.5</b> (0.4)	<b>0.95</b> (0.5)	15/15
fminunc pa	<b>2.1</b> (2)	<b>1.9</b> (2)	<b>1.3</b> (0.9)	<b>1.0</b> (0.9)	<b>1.1</b> (1.0)	15/15
ga100 hol	<b>1.6</b> (1)	<b>2.9</b> (3)	7.6(8)	57(27)	2353(2457)	5/15
grid100 ho	<b>1.4</b> (1.0)	<b>2.0</b> (2)	9.3(14)	159(67)	$\infty$ 2e5	0/15
grid16 hol	<b>1.4</b> (2)	6.3(11)	8.1(9)	31(10)	6623(7909)	2/15
hill hol	6.1(8)	4.2(5)	<b>2.7</b> (2)	15(8)	6174(7001)	2/15
ImmCMA aug	<b>0.80</b> (0.6)	<b>1.6</b> (2)	<b>1.6</b> (0.8)	<b>1.7</b> (0.6)	<b>1.4</b> (1)	15/15
memPSODE v	<b>2.1</b> (4)	<b>2.8</b> (3)	6.1(5)	3.9(3)	4.8(5)	15/15
prcga saw	<b>0.98</b> (1)	<b>1.7</b> (2)	3.9(4)	12(4)	296(609)	14/15
ring100 ho	<b>1.6</b> (2)	<b>2.7</b> (3)	8.2(9)	109(40)	1999(1867)	6/15
ring16 hol	<b>1.2</b> (1.0)	4.0(4)	8.1(4)	29(9)	1172(1372)	8/15
simplex pa	16(10)	18(27)	15(14)	12(3)	6.7(3)	15/15

Table 14: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{13}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f13</b>	<i>1.0e+3:2.8</i>	<i>6.3e+2:8.4</i>	<i>4.0e+2:17</i>	<i>6.3e+1:52</i>	<i>6.3e-2:264</i>	15/15
BIPOP-aCMA	<b>1.1</b> (1)	<b>2.4</b> (3)	3.4(2)	4.8(2)	4.1(0.8)	15/15
BIPOP-saAC	<b>2.7</b> (2)	<b>2.0</b> (2)	<b>2.0</b> (2)	<b>2.3</b> (0.5)	<b>1.3</b> (0.3)	15/15
CMAES hut	4.0(4)	<b>2.8</b> (2)	3.0(2)	3.3(1.0)	$\infty$ 506	0/15
DE pal	<b>1.5</b> (1)	<b>1.6</b> (2)	3.2(3)	10(3)	46(20)	15/15
HCMA los	<b>2.5</b> (2)	<b>1.5</b> (1)	<b>1.1</b> (0.3)	<b>1.4</b> (2)	<b>1.6</b> (0.6)	15/15
HMLSL pal	<b>1.7</b> (1)	<b>1.0</b> (0.8)	<b>1.2</b> (0.6)	<b>1.2</b> (0.2)	<b>0.79</b> (0.1)	15/15
IPOP-10DDr	<b>2.5</b> (2)	<b>2.4</b> (2)	<b>2.6</b> (0.7)	3.5(1)	5.7(3)	15/15
IPOP-500 l	<b>2.5</b> (2)	<b>2.4</b> (2)	<b>2.6</b> (0.7)	3.5(1)	5.7(3)	15/15
IPOP-tany	3.0(4)	<b>2.4</b> (3)	<b>2.5</b> (2)	3.5(2)	6.2(2)	15/15
IPOP-texp	4.4(5)	3.0(3)	<b>2.5</b> (2)	3.1(1)	6.4(2)	15/15
IPOP lia	<b>2.5</b> (2)	<b>2.4</b> (2)	<b>2.6</b> (0.7)	3.5(1)	5.7(3)	15/15
MLSL pal	<b>1.7</b> (1)	<b>1.0</b> (0.8)	<b>1.2</b> (0.6)	<b>1.2</b> (0.2)	<b>0.79</b> (0.1)	15/15
OQNLP pal	3.7(3)	<b>2.0</b> (1)	<b>1.4</b> (0.4)	<b>1.5</b> (0.2)	<b>0.98</b> (0.2)	15/15
P-DCN tra	<b>2.5</b> (4)	4.8(5)	10(10)	847(1396)	2.6e5(3e5)	1/15
P-zero tra	<b>2.9</b> (3)	5.2(7)	7.8(10)	1.2e4(4e4)	2.7e5(3e5)	1/15
SMAC hut	1(1)	<b>1.1</b> (1)	<b>0.96</b> (0.5)	<b>1.1</b> (0.5)	$\infty$ 500	0/15
U-DCN tra	<b>1.8</b> (2)	<b>2.6</b> (2)	3.5(3)	20(18)	$\infty$ 5e6	0/15
U-zero tra	<b>2.0</b> (2)	<b>2.5</b> (3)	<b>2.6</b> (2)	6891(21)	2.7e5(3e5)	1/15
fmincon pa	<b>1.7</b> (1)	<b>1.0</b> (0.8)	<b>1.2</b> (0.6)	<b>1.2</b> (0.2)	<b>0.79</b> (0.1)	15/15
fminunc pa	<b>2.4</b> (3)	<b>1.5</b> (1)	<b>1.1</b> (0.4)	<b>1.0</b> (0.2)	<b>0.98</b> (0.1)	15/15
ga100 hol	<b>2.7</b> (4)	<b>2.4</b> (4)	7.1(10)	36(19)	$\infty$ 2e5	0/15
grid100 ho	<b>1.6</b> (2)	<b>2.9</b> (5)	9.2(10)	557(388)	$\infty$ 2e5	0/15
grid16 hol	3.1(2)	5.0(7)	5.9(5)	20(16)	$\infty$ 2e5	0/15
hill hol	7.3(8)	3.5(5)	<b>2.4</b> (3)	1218(2409)	$\infty$ 2e5	0/15
lmmCMA aug	<b>1.4</b> (2)	<b>1.6</b> (2)	<b>1.7</b> (1.0)	<b>1.8</b> (0.5)	<b>1.6</b> (0.4)	15/15
memPSODE v	<b>2.9</b> (2)	3.4(3)	4.7(3)	3.0(3)	<b>2.4</b> (6)	15/15
prcga saw	<b>1.4</b> (1)	4.0(6)	5.0(7)	10(4)	2241(2765)	6/15
ring100 ho	<b>2.2</b> (2)	<b>2.8</b> (3)	12(15)	59(17)	$\infty$ 2e5	0/15
ring16 hol	<b>1.7</b> (2)	<b>2.9</b> (5)	4.9(4)	16(8)	$\infty$ 2e5	0/15
simplex pa	13(13)	14(13)	18(15)	20(12)	16(13)	15/15

Table 15: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{14}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f14</b>	<i>1.6e+1:3.0</i>	<i>1.0e+1:10</i>	<i>6.3e+0:15</i>	<i>2.5e-1:53</i>	<i>1.0e-5:251</i>	15/15
BIPOP-aCMA	<b>1.9</b> (2)	<b>2.9</b> (4)	3.6(4)	4.0(1)	4.3(0.8)	15/15
BIPOP-saAC	3.2(3)	<b>1.5</b> (1)	<b>2.1</b> (1)	<b>2.9</b> (1.0)	<b>1.8</b> (0.3)	15/15
CMAES hut	3.6(6)	<b>1.7</b> (2)	<b>1.7</b> (2)	3.2(1)	$\infty$ 506	0/15
DE pal	<b>2.4</b> (3)	<b>2.2</b> (2)	<b>2.6</b> (3)	9.0(3)	36(12)	15/15
HCMA los	3.7(3)	<b>1.6</b> (0.6)	<b>1.2</b> (0.6)	<b>2.1</b> (2)	<b>2.2</b> (0.3)	15/15
HMLSL pal	<b>1.2</b> (1)	<b>0.68</b> (0.4)	<b>0.59</b> (0.4)	<b>0.60</b> (0.1)	<b>0.72</b> (0.1)	15/15
IPOP-10DDr	<b>2.0</b> (2)	<b>1.2</b> (1)	<b>2.0</b> (2)	3.1(1)	5.7(0.4)	15/15
IPOP-500 l	<b>2.0</b> (2)	<b>1.2</b> (1)	<b>2.0</b> (2)	3.1(1)	5.7(0.4)	15/15
IPOP-tany	<b>1.6</b> (2)	<b>0.94</b> (1)	<b>1.3</b> (0.9)	3.1(1)	5.5(0.7)	15/15
IPOP-texp	<b>1.5</b> (2)	<b>0.93</b> (1)	<b>1.2</b> (0.9)	<b>2.2</b> (0.7)	4.9(0.7)	15/15
IPOP lia	<b>2.0</b> (2)	<b>1.2</b> (1)	<b>2.0</b> (2)	3.1(1)	5.7(0.4)	15/15
MLSL pal	<b>1.2</b> (1)	<b>0.68</b> (0.4)	<b>0.59</b> (0.4)	<b>0.60</b> (0.1)	<b>0.72</b> (0.1)	15/15
OQNLP pal	<b>2.2</b> (3)	<b>1.4</b> (0.9)	<b>1.2</b> (1.0)	<b>0.99</b> (0.3)	7.0(9)	15/15
P-DCN tra	3.2(4)	<b>2.0</b> (2)	3.6(4)	10(3)	5.4e4(5e4)	5/15
P-zero tra	4.0(5)	3.1(4)	5.5(6)	11(6)	$\infty$ 5e6	0/15
SMAC hut	<b>1.2</b> (2)	<b>0.62</b> (0.6)	<b>0.76</b> (1)	4.9(2)	$\infty$ 500	0/15
U-DCN tra	<b>2.7</b> (2)	<b>1.2</b> (0.7)	<b>1.9</b> (2)	14(8)	$\infty$ 5e6	0/15
U-zero tra	3.0(3)	<b>1.5</b> (2)	<b>1.7</b> (2)	6.4(4)	$\infty$ 5e6	0/15
fmincon pa	<b>1.2</b> (1)	<b>0.68</b> (0.4)	<b>0.59</b> (0.4)	<b>0.60</b> (0.1)	<b>0.72</b> (0.1)	15/15
fminunc pa	1(1)	<b>0.71</b> (0.6)	<b>0.71</b> (0.6)	1(0.5)	<b>0.84</b> (0.2)	15/15
ga100 hol	<b>2.4</b> (3)	<b>1.6</b> (1)	<b>1.9</b> (2)	27(8)	$\infty$ 2e5	0/15
grid100 ho	<b>2.5</b> (4)	<b>1.9</b> (2)	3.3(2)	84(42)	$\infty$ 2e5	0/15
grid16 hol	<b>1.2</b> (1)	<b>0.81</b> (0.7)	<b>2.3</b> (2)	18(7)	$\infty$ 2e5	0/15
hill hol	4.1(4)	<b>1.9</b> (2)	<b>2.0</b> (2)	4.0(2)	$\infty$ 2e5	0/15
lmmCMA aug	<b>1.1</b> (1)	<b>0.62</b> (0.6)	<b>0.81</b> (0.7)	<b>1.6</b> (0.5)	<b>1.8</b> (0.2)	15/15
memPSODE v	5.4(5)	<b>2.6</b> (2)	3.3(2)	<b>2.5</b> (0.3)	<b>2.2</b> (0.1)	15/15
prcga saw	<b>2.7</b> (3)	<b>1.1</b> (1)	<b>1.7</b> (2)	10(3)	6773(7089)	3/15
ring100 ho	<b>2.6</b> (2)	<b>1.7</b> (1)	4.3(5)	53(14)	$\infty$ 2e5	0/15
ring16 hol	3.0(4)	<b>1.8</b> (3)	3.8(3)	12(4)	$\infty$ 2e5	0/15
simplex pa	7.4(12)	6.4(4)	7.0(9)	17(7)	5.6(1)	15/15



Table 16: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f15</b>	<i>1.6e+2:3.0</i>	<i>1.0e+2:13</i>	<i>6.3e+1:24</i>	<i>4.0e+1:55</i>	<i>1.6e+1:289</i>	5/5
BIPOP-aCMA	5.8(6)	<b>2.8</b> (2)	3.0(2)	<b>2.5</b> (1)	<b>1.7</b> (1)	15/15
BIPOP-saAC	4.7(4)	<b>1.9</b> (2)	<b>2.1</b> (2)	<b>2.1</b> (1)	<b>1.5</b> (0.9)	15/15
CMAES hut	4.1(4)	<b>1.9</b> (2)	<b>2.0</b> (0.9)	<b>1.7</b> (0.9)	<b>1.7</b> (1)	11/15
DE pal	<b>2.6</b> (2)	<b>1.9</b> (1)	<b>2.9</b> (3)	3.9(2)	3.8(3)	15/15
HCMA los	5.3(2)	<b>2.3</b> (2)	<b>2.1</b> (3)	<b>1.9</b> (2)	<b>2.0</b> (0.8)	15/15
HMLSL pal	16(31)	6.8(7)	10(9)	9.1(4)	3.3(0.6)	15/15
IPOP-10DDr	4.2(4)	<b>1.9</b> (2)	<b>2.0</b> (1)	<b>1.5</b> (0.7)	<b>1.2</b> (0.6)	15/15
IPOP-500 l	4.2(4)	<b>1.9</b> (2)	<b>2.0</b> (1)	<b>1.5</b> (0.7)	<b>1.2</b> (0.6)	15/15
IPOP-tany	3.6(4)	<b>1.6</b> (2)	<b>2.0</b> (1)	<b>1.8</b> (0.7)	<b>2.2</b> (0.8)	15/15
IPOP-texp	6.8(8)	<b>2.8</b> (2)	<b>2.4</b> (1)	<b>2.1</b> (2)	<b>2.1</b> (3)	15/15
IPOP lia	4.2(4)	<b>1.9</b> (2)	<b>2.0</b> (1)	<b>1.5</b> (0.7)	<b>1.2</b> (0.6)	15/15
MLSL pal	15(30)	6.3(9)	10(8)	10(3)	5.5(4)	15/15
OQNLP pal	30(43)	14(17)	20(14)	16(19)	11(11)	15/15
P-DCN tra	3.1(2)	3.4(3)	242(14)	1.3e4(4e4)	7.0e4(8e4)	3/15
P-zero tra	3.2(2)	1275(2705)	3676(5912)	2.8e4(4e4)	4.1e4(4e4)	5/15
SMAC hut	<b>1.1</b> (2)	<b>0.83</b> (0.6)	<b>1.6</b> (1)	<b>2.2</b> (2)	8.1(8)	3/15
U-DCN tra	<b>3.1</b> (3)	<b>2.2</b> (2)	<b>2.9</b> (2)	4.9(4)	10(12)	15/15
U-zero tra	3.8(4)	<b>2.0</b> (1)	<b>2.4</b> (2)	7.0(5)	713(218)	15/15
fmincon pa	16(30)	7.2(10)	10(8)	7.2(3)	5.5(5)	15/15
fminunc pa	15(26)	8.0(9)	11(8)	7.4(3)	10(10)	15/15
ga100 hol	3.3(4)	<b>3.0</b> (2)	7.9(5)	10(5)	7.9(4)	15/15
grid100 ho	4.0(4)	4.9(4)	13(15)	25(26)	71(144)	15/15
grid16 hol	5.7(8)	4.7(6)	8.2(10)	19(10)	49(38)	15/15
hill hol	4.3(6)	<b>1.9</b> (2)	3.7(5)	3.1(2)	65(107)	15/15
lmmCMA aug	<b>1.3</b> (2)	<b>1.3</b> (1)	<b>1.5</b> (0.9)	<b>1.6</b> (1)	<b>1.4</b> (2)	15/15
memPSODE v	15(10)	11(18)	16(10)	12(6)	15(17)	15/15
prcga saw	<b>2.5</b> (2)	<b>1.8</b> (2)	3.5(3)	4.1(3)	<b>3.0</b> (2)	15/15
ring100 ho	4.8(5)	3.3(3)	11(8)	12(10)	18(13)	15/15
ring16 hol	3.7(7)	3.4(3)	6.0(4)	6.2(2)	25(9)	15/15
simplex pa	153(191)	51(40)	41(14)	21(5)	20(20)	15/15

Table 17: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{16}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f16</b>	<i>4.0e+1</i> :4.8	<i>2.5e+1</i> :16	<i>1.6e+1</i> :46	<i>1.0e+1</i> :120	<i>4.0e+0</i> :334	15/15
BIPOP-aCMA	<b>1.8</b> (1)	<b>1.8</b> (3)	3.2(4)	4.6(10)	3.8(5)	15/15
BIPOP-saAC	<b>1.8</b> (2)	<b>2.1</b> (2)	<b>1.9</b> (2)	<b>2.1</b> (2)	<b>2.7</b> (4)	15/15
CMAES hut	<b>1.9</b> (2)	<b>2.3</b> (4)	<b>2.8</b> (2)	<b>2.5</b> (3)	<b>2.6</b> (2)	7/15
DE pal	<b>1.8</b> (2)	<b>1.1</b> (1.0)	<b>2.1</b> (2)	3.3(3)	6.4(5)	15/15
HCMA los	4.7(5)	<b>2.2</b> (2)	<b>1.6</b> (1)	<b>2.9</b> (2)	3.0(4)	15/15
HMLSL pal	14(37)	10(12)	6.5(5)	4.4(2)	7.2(5)	15/15
IPOP-10DDr	<b>1.4</b> (1)	<b>1.3</b> (1)	<b>1.4</b> (1)	<b>2.0</b> (2)	<b>2.3</b> (1)	15/15
IPOP-500 l	<b>1.4</b> (1)	<b>1.3</b> (1)	<b>1.4</b> (1)	<b>2.0</b> (2)	<b>2.3</b> (1)	15/15
IPOP-tany	<b>1.7</b> (2)	<b>1.7</b> (2)	<b>1.9</b> (2)	<b>1.9</b> (1)	<b>1.6</b> (1)	15/15
IPOP-texp	<b>2.1</b> (2)	<b>2.0</b> (2)	<b>1.3</b> (1)	<b>1.1</b> (1)	<b>0.96</b> (1.0)	15/15
IPOP lia	<b>1.4</b> (1)	<b>1.3</b> (1)	<b>1.4</b> (1)	<b>2.0</b> (2)	<b>2.3</b> (1)	15/15
MLSL pal	14(38)	10(11)	5.9(5)	4.7(3)	9.4(8)	15/15
OQNLP pal	48(52)	18(17)	10(4)	12(19)	16(16)	15/15
P-DCN tra	<b>2.2</b> (2)	<b>1.6</b> (1)	<b>1.2</b> (1)	<b>1.1</b> (0.7)	73(9)	15/15
P-zero tra	<b>1.7</b> (1)	<b>1.7</b> (2)	5.1(4)	5.5(16)	5592(7745)	11/15
SMAC hut	<b>1.7</b> (2)	<b>0.77</b> (0.5)	<b>0.53</b> (0.5)	<b>0.42</b> (0.3)	<b>0.45</b> (0.6)	15/15
U-DCN tra	<b>1.8</b> (1)	<b>1.9</b> (2)	<b>2.3</b> (2)	<b>1.7</b> (1.0)	3.6(4)	15/15
U-zero tra	<b>1.9</b> (2)	<b>1.7</b> (2)	<b>1.4</b> (2)	<b>1.3</b> (0.9)	16(21)	15/15
fmincon pa	16(43)	11(14)	6.5(5)	4.1(2)	8.2(11)	15/15
fminunc pa	51(42)	20(15)	10(4)	12(12)	39(36)	15/15
ga100 hol	<b>1.2</b> (1)	<b>1.1</b> (1)	<b>1.5</b> (2)	3.2(3)	6.3(7)	15/15
grid100 ho	<b>1.8</b> (2)	<b>1.1</b> (1)	<b>1.9</b> (3)	<b>2.6</b> (2)	5.6(6)	15/15
grid16 hol	<b>1.5</b> (1)	<b>2.1</b> (3)	<b>1.6</b> (2)	<b>2.1</b> (3)	5.5(8)	15/15
hill hol	3.4(5)	3.1(2)	<b>2.0</b> (2)	<b>2.9</b> (2)	14(30)	15/15
lmmCMA aug	<b>1.7</b> (2)	<b>1.8</b> (2)	<b>2.7</b> (3)	<b>2.0</b> (2)	<b>1.3</b> (0.8)	15/15
memPSODE v	<b>1.4</b> (1)	<b>1.7</b> (1)	3.6(3)	3.0(2)	5.6(12)	15/15
prcga saw	<b>2.2</b> (3)	<b>2.2</b> (3)	<b>2.0</b> (4)	3.0(2)	7.2(3)	15/15
ring100 ho	<b>1.5</b> (2)	<b>1.8</b> (3)	<b>0.94</b> (1)	3.1(4)	8.4(8)	15/15
ring16 hol	<b>1.3</b> (1)	<b>1.5</b> (1)	<b>2.0</b> (2)	<b>1.5</b> (1.0)	<b>2.9</b> (2)	15/15
simplex pa	114(45)	38(12)	14(4)	6.3(1)	3.3(2)	15/15

Table 18: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{17}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f17</b>	<i>1.0e+1:5.2</i>	<i>6.3e+0:26</i>	<i>4.0e+0:57</i>	<i>2.5e+0:110</i>	<i>6.3e-1:412</i>	15/15
BIPOP-aCMA	5.5(6)	<b>2.1</b> (1)	<b>1.6</b> (0.7)	<b>1.2</b> (0.5)	<b>0.63</b> (0.3)	15/15
BIPOP-saAC	4.9(3)	<b>2.8</b> (2)	5.3(3)	3.1(1)	<b>1.4</b> (1)	15/15
CMAES hut	4.1(6)	<b>1.3</b> (1)	<b>1.3</b> (1.0)	<b>0.93</b> (0.4)	<b>0.62</b> (0.3)	14/15
DE pal	5.4(5)	<b>2.3</b> (2)	<b>2.1</b> (1.0)	<b>2.1</b> (0.9)	<b>1.8</b> (0.6)	15/15
HCMA los	<b>2.2</b> (2)	<b>1.6</b> (1)	<b>1.7</b> (1)	<b>1.5</b> (1.0)	<b>0.90</b> (0.2)	15/15
HMLSL pal	26(34)	7.3(7)	20(34)	12(17)	6.3(5)	15/15
IPOP-10DDr	3.2(3)	<b>1.5</b> (1)	<b>1.0</b> (0.5)	<b>0.84</b> (0.3)	<b>0.57</b> (0.2)	15/15
IPOP-500 l	3.2(3)	<b>1.5</b> (1)	<b>1.0</b> (0.5)	<b>0.84</b> (0.3)	<b>0.57</b> (0.2)	15/15
IPOP-tany	3.6(6)	<b>2.1</b> (2)	<b>1.6</b> (1)	<b>0.98</b> (0.5)	<b>0.67</b> (0.3)	15/15
IPOP-texp	47(7)	11(2)	5.9(2)	3.6(3)	<b>1.9</b> (4)	15/15
IPOP lia	3.2(3)	<b>1.5</b> (1)	<b>1.0</b> (0.5)	<b>0.84</b> (0.3)	<b>0.57</b> (0.2)	15/15
MLSL pal	24(35)	7.5(9)	22(34)	16(33)	173(163)	12/15
OQNLP pal	15(19)	15(11)	27(54)	34(51)	108(117)	6/15
P-DCN tra	4.3(4)	3.1(4)	6637(1051)	1.7e4(2e4)	4.9e4(5e4)	3/15
P-zero tra	5.6(8)	5339(1863)	1.6e4(4e4)	2.7e4(5e4)	2.7e4(4e4)	5/15
SMAC hut	<b>2.5</b> (4)	<b>1.6</b> (2)	<b>1.9</b> (2)	<b>2.1</b> (2)	<b>2.5</b> (3)	6/15
U-DCN tra	4.1(4)	<b>1.9</b> (2)	<b>2.0</b> (1)	<b>2.0</b> (1)	32(19)	15/15
U-zero tra	<b>2.6</b> (2)	<b>1.5</b> (1)	<b>2.5</b> (3)	<b>2.5</b> (4)	158(66)	15/15
fmincon pa	20(25)	6.4(7)	9.0(10)	25(66)	161(92)	14/15
fminunc pa	11(16)	6.4(6)	5.4(4)	10(9)	47(67)	14/15
ga100 hol	3.5(6)	<b>2.7</b> (3)	4.7(4)	5.0(3)	5.2(1)	15/15
grid100 ho	3.7(3)	3.3(4)	7.7(9)	15(6)	54(33)	15/15
grid16 hol	4.2(6)	<b>2.7</b> (2)	4.3(5)	18(9)	58(41)	14/15
hill hol	39(19)	11(10)	8.0(12)	8.2(11)	219(312)	12/15
ImmCMA aug	<b>1.7</b> (2)	<b>0.91</b> (0.8)	<b>0.70</b> (0.5)	<b>0.55</b> (0.2)	<b>0.62</b> (0.3)	15/15
memPSODE v	6.9(4)	12(11)	10(12)	12(6)	10(14)	15/15
prcga saw	<b>1.9</b> (2)	<b>1.5</b> (1)	<b>2.0</b> (2)	<b>2.5</b> (2)	6.3(1)	15/15
ring100 ho	3.8(6)	3.9(4)	7.8(5)	10(5)	13(5)	15/15
ring16 hol	5.9(8)	<b>2.9</b> (4)	<b>3.0</b> (2)	<b>2.6</b> (2)	3.6(2)	15/15
simplex pa	63(71)	28(25)	16(10)	20(22)	293(273)	8/15

Table 19: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{18}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f18</b>	<i>6.3e+1:3.4</i>	<i>4.0e+1:7.2</i>	<i>2.5e+1:20</i>	<i>1.6e+1:58</i>	<i>1.6e+0:318</i>	15/15
BIPOP-aCMA	<b>2.4</b> (4)	3.5(2)	<b>2.9</b> (3)	<b>1.8</b> (1)	<b>1.0</b> (0.4)	15/15
BIPOP-saAC	4.4(4)	4.7(4)	3.0(2)	<b>1.5</b> (0.9)	3.5(6)	15/15
CMAES hut	<b>1.6</b> (2)	<b>2.7</b> (3)	<b>1.9</b> (1)	<b>1.2</b> (0.7)	<b>0.92</b> (0.3)	14/15
DE pal	3.5(4)	3.8(5)	<b>2.6</b> (3)	<b>1.6</b> (1)	4.2(1)	15/15
HCMA los	<b>3.0</b> (2)	3.0(2)	<b>1.9</b> (2)	<b>1.5</b> (1)	<b>2.2</b> (0.8)	15/15
HMLSL pal	<b>2.6</b> (3)	17(16)	11(10)	11(13)	10(7)	15/15
IPOP-10DDr	<b>2.2</b> (2)	<b>2.2</b> (2)	<b>1.9</b> (2)	<b>1.1</b> (0.6)	<b>1.9</b> (0.4)	15/15
IPOP-500 l	<b>2.2</b> (2)	<b>2.2</b> (2)	<b>1.9</b> (2)	<b>1.1</b> (0.6)	<b>1.9</b> (0.4)	15/15
IPOP-tany	<b>1.7</b> (2)	<b>1.8</b> (2)	<b>1.4</b> (0.8)	<b>0.94</b> (0.7)	<b>1.0</b> (0.5)	15/15
IPOP-texp	4.2(5)	5.4(5)	<b>3.0</b> (2)	<b>1.4</b> (0.8)	<b>1.8</b> (0.7)	15/15
IPOP lia	<b>2.2</b> (2)	<b>2.2</b> (2)	<b>1.9</b> (2)	<b>1.1</b> (0.6)	<b>1.9</b> (0.4)	15/15
MLSL pal	<b>2.6</b> (3)	17(19)	10(8)	10(12)	781(851)	5/15
OQNLP pal	6.5(2)	13(24)	12(10)	12(1)	136(157)	6/15
P-DCN tra	<b>1.6</b> (1)	<b>1.9</b> (2)	<b>1.8</b> (1)	27(52)	2.6e4(3e4)	6/15
P-zero tra	<b>1.8</b> (2)	<b>2.5</b> (3)	452(1)	533(36)	3.5e4(4e4)	5/15
SMAC hut	<b>1.1</b> (1)	<b>0.85</b> (0.6)	<b>0.97</b> (1.0)	<b>1.1</b> (1)	11(12)	2/15
U-DCN tra	<b>1.2</b> (1)	<b>1.7</b> (2)	3.0(2)	<b>1.7</b> (1)	3684(7109)	14/15
U-zero tra	<b>1.2</b> (1)	<b>1.6</b> (1)	<b>2.8</b> (3)	<b>1.8</b> (2)	4474(8009)	12/15
fmincon pa	<b>2.6</b> (3)	15(15)	11(11)	11(12)	400(342)	9/15
fminunc pa	<b>2.2</b> (3)	12(17)	7.7(7)	5.0(4)	242(322)	8/15
ga100 hol	<b>1.0</b> (0.7)	<b>2.6</b> (3)	3.2(3)	3.3(4)	11(7)	15/15
grid100 ho	<b>1.3</b> (1)	3.0(4)	5.3(7)	7.0(7)	504(786)	10/15
grid16 hol	<b>2.4</b> (3)	3.9(5)	4.8(5)	3.9(3)	251(395)	13/15
hill hol	10(20)	17(23)	12(26)	4.8(9)	269(403)	12/15
lmmCMA aug	<b>1.3</b> (1)	<b>1.6</b> (2)	<b>0.92</b> (0.8)	<b>0.73</b> (0.4)	<b>0.52</b> (0.2)* <sup>2</sup>	15/15
memPSODE v	3.2(3)	5.1(3)	19(17)	13(10)	17(17)	15/15
prcga saw	<b>1.3</b> (2)	<b>1.6</b> (2)	<b>2.8</b> (4)	<b>1.9</b> (3)	<b>2.8</b> (1)	15/15
ring100 ho	<b>1.4</b> (2)	<b>1.7</b> (0.8)	4.7(4)	6.6(5)	27(15)	15/15
ring16 hol	<b>1.8</b> (2)	3.1(3)	4.3(3)	3.8(2)	144(394)	13/15
simplex pa	21(13)	38(45)	28(20)	14(3)	374(357)	8/15

Table 20: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f19</b>	<i>1.6e-1:172</i>	<i>1.0e-1:242</i>	<i>6.3e-2:675</i>	<i>4.0e-2:3078</i>	<i>2.5e-2:4946</i>	15/15
BIPOP-aCMA	159(150)	139(107)	55(38)	18(19)	18(20)	15/15
BIPOP-saAC	51(40)	76(97)	61(78)	15(19)	18(20)	15/15
CMAES hut	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$ <i>506</i>	0/15
DE pal	571(475)	1087(1107)	1053(1111)	238(241)	148(170)	2/15
HCMA los	95(112)	99(104)	41(60)	14(21)	13(12)	15/15
HMLSL pal	<b>0.11</b> (0) $\downarrow_4$	<b>0.17</b> (0.0) $\downarrow_4$	<b>0.07</b> (0.0) $\downarrow$	<b>0.02</b> (3e-3) $\downarrow_4$	<b>0.01</b> (4e-3) $\downarrow_4$	15/15
IPOP-10DDr	614(657)	519(658)	263(299)	72(77)	46(48)	15/15
IPOP-500 l	563(657)	483(576)	251(256)	57(60)	37(37)	15/15
IPOP-tany	493(546)	391(388)	192(136)	57(102)	36(66)	15/15
IPOP-texp	895(958)	746(669)	284(240)	68(62)	51(44)	15/15
IPOP lia	563(657)	483(576)	251(256)	57(60)	37(37)	15/15
MLSL pal	<b>0.11</b> (0) $\downarrow_4$	<b>0.17</b> (0.0) $\downarrow_4$	<b>0.07</b> (0.0) $\downarrow$	<b>0.02</b> (3e-3) $\downarrow_4$	<b>0.01</b> (4e-3) $\downarrow_4$	15/15
OQNLP pal	<b>0.10</b> (3e-3) $\downarrow_4^{\star 4}$	<b>0.09</b> (0) $\downarrow_4^{\star 4}$	<b>0.03</b> (0) $\downarrow^{\star 4}$	<b>7.5e-3</b> (0) $\downarrow_4^{\star 4}$	<b>4.7e-3</b> (0) $\downarrow_4^{\star 4}$	15/15
P-DCN tra	5.2e4(6e4)	6.5e4(7e4)	1.1e5(1e5)	2.3e4(3e4)	1.4e4(2e4)	1/15
P-zero tra	4.3e5(5e5)	3.1e5(3e5)	$\infty$	$\infty$	$\infty$ <i>5e6</i>	0/15
SMAC hut	$\infty$	$\infty$	$\infty$	$\infty$	$\infty$ <i>500</i>	0/15
U-DCN tra	2.6e4(3e4)	4.9e4(5e4)	5.1e4(6e4)	2.4e4(3e4)	1.5e4(2e4)	1/15
U-zero tra	1.7e4(2e4)	3.2e4(4e4)	4.8e4(6e4)	$\infty$	$\infty$ <i>5e6</i>	0/15
fmincon pa	<b>0.11</b> (0) $\downarrow_4$	<b>0.17</b> (0.0) $\downarrow_4$	<b>0.07</b> (0.0) $\downarrow$	<b>0.02</b> (3e-3) $\downarrow_4$	<b>0.01</b> (4e-3) $\downarrow_4$	15/15
fminunc pa	<b>0.16</b> (0.0) $\downarrow_4$	<b>0.13</b> (0.0) $\downarrow_4$	<b>0.05</b> (0.0) $\downarrow$	<b>0.01</b> (3e-3) $\downarrow_4$	<b>7.8e-3</b> (2e-3) $\downarrow_4$	15/15
ga100 hol	807(922)	1121(1120)	724(813)	1156(1381)	$\infty$ <i>2e5</i>	0/15
grid100 ho	1.0e4(1e4)	1.5e4(2e4)	5534(5924)	$\infty$	$\infty$ <i>2e5</i>	0/15
grid16 hol	2686(3001)	4684(5161)	5264(5739)	$\infty$	$\infty$ <i>2e5</i>	0/15
hill hol	3635(3925)	6883(7742)	2472(2777)	1159(1178)	738(809)	1/15
ImmCMA aug	55(57)	56(58)	30(33)	14(14)	$\infty$ <i>2805</i>	0/15
memPSODE v	157(105)	154(80)	96(106)	41(33)	33(23)	15/15
prcga saw	56(7)	55(5)	20(2)	4.8(3)	3.2(2)	15/15
ring100 ho	1435(1554)	3410(4030)	2468(2962)	541(609)	341(404)	2/15
ring16 hol	2015(2001)	3361(3613)	1635(1837)	1149(1340)	715(834)	1/15
simplex pa	<b>0.17</b> (0.1) $\downarrow_4$	<b>0.18</b> (0.0) $\downarrow_4$	<b>0.08</b> (0.0) $\downarrow$	<b>0.02</b> (4e-3) $\downarrow_4$	<b>0.01</b> (3e-3) $\downarrow_4$	15/15

Table 21: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{20}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f20</b>	<i>6.3e+3:5.1</i>	<i>4.0e+3:8.4</i>	<i>4.0e+1:15</i>	<i>2.5e+0:69</i>	<i>1.0e+0:851</i>	15/15
BIPOP-aCMA	<b>2.3</b> (4)	<b>2.2</b> (2)	4.0(2)	<b>2.9</b> (2)	7.6(9)	15/15
BIPOP-saAC	3.5(4)	<b>2.5</b> (2)	3.5(3)	<b>2.5</b> (2)	5.5(7)	15/15
CMAES hut	3.4(3)	<b>2.2</b> (2)	3.7(2)	3.1(1)	$\infty$ 506	0/15
DE pal	<b>1.8</b> (2)	<b>1.6</b> (2)	7.7(4)	5.7(2)	<b>2.3</b> (2)	15/15
HCMA los	<b>1.7</b> (1)	<b>1.3</b> (0.8)	<b>0.90</b> (0.2)	<b>1.2</b> (2)	5.2(7)	15/15
HMLSL pal	<b>1.6</b> (0)	<b>0.95</b> (0)	<b>1.5</b> (0)	<b>1.2</b> (0)	<b>2.3</b> (1)	15/15
IPOP-10DDr	<b>1.9</b> (2)	<b>1.4</b> (2)	3.4(1)	<b>2.9</b> (2)	12(11)	15/15
IPOP-500 l	<b>1.9</b> (2)	<b>1.4</b> (2)	3.4(1)	<b>2.9</b> (2)	12(11)	15/15
IPOP-tany	<b>1.6</b> (2)	<b>1.1</b> (1.0)	<b>2.4</b> (1)	<b>2.9</b> (2)	12(10)	15/15
IPOP-texp	<b>1.1</b> (1)	<b>0.72</b> (0.8)	<b>1.6</b> (0.6)	<b>1.7</b> (1)	17(19)	15/15
IPOP lia	<b>1.9</b> (2)	<b>1.4</b> (2)	3.4(1)	<b>2.9</b> (2)	12(11)	15/15
MLSL pal	<b>1.6</b> (0)	<b>0.95</b> (0)	<b>1.5</b> (0)	<b>1.2</b> (0)	7.2(10)	15/15
OQNLP pal	<b>2.7</b> (0)	<b>2.0</b> (0)	<b>1.2</b> (0)	<b>0.51</b> (0)	5.8(7)	12/15
P-DCN tra	<b>2.4</b> (2)	<b>2.9</b> (1)	8.2(8)	3.6(2)	336(965)	15/15
P-zero tra	7.9(14)	6.4(10)	14(9)	4.9(2)	371(909)	15/15
SMAC hut	<b>0.57</b> (0.2)	<b>0.44</b> (0.2)	<b>0.73</b> (0.3)	4.3(3)	$\infty$ 500	0/15
U-DCN tra	<b>1.6</b> (1)	<b>1.7</b> (2)	3.6(3)	4.6(5)	5.6(8)	15/15
U-zero tra	<b>2.4</b> (2)	<b>2.3</b> (2)	3.7(2)	3.4(2)	17(29)	15/15
fmincon pa	<b>1.6</b> (0)	<b>0.95</b> (0)	<b>1.5</b> (0)	<b>1.2</b> (0)	6.4(9)	15/15
fminunc pa	<b>1.4</b> (0)	<b>0.83</b> (0)	<b>0.46</b> (0)*	<b>1.8</b> (0)	4.6(4)	15/15
ga100 hol	<b>1.9</b> (2)	<b>2.3</b> (2)	17(13)	13(7)	<b>4.1</b> (1)	15/15
grid100 ho	<b>2.4</b> (3)	<b>1.6</b> (2)	28(37)	30(14)	14(7)	15/15
grid16 hol	<b>2.9</b> (2)	<b>2.6</b> (6)	11(9)	7.4(4)	8.1(3)	15/15
hill hol	7.1(8)	4.7(5)	5.9(5)	<b>2.2</b> (1)	4.4(8)	15/15
lmmCMA aug	<b>1.2</b> (0.9)	<b>0.83</b> (0.5)	<b>1.8</b> (1)	4.0(6)	15(15)	3/15
memPSODE v	3.1(3)	3.5(3)	5.3(2)	<b>2.5</b> (0.9)	<b>3.6</b> (5)	15/15
prcga saw	<b>1.2</b> (2)	<b>0.91</b> (1)	4.7(3)	5.5(3)	25(44)	15/15
ring100 ho	<b>2.1</b> (2)	<b>2.3</b> (2)	31(20)	24(7)	7.4(2)	15/15
ring16 hol	<b>1.5</b> (1)	<b>2.3</b> (2)	8.0(6)	5.6(2)	<b>3.3</b> (0.5)	15/15
simplex pa	7.8(0.3)	6.9(0.2)	8.4(4)	10(9)	5.4(3)	15/15

Table 22: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b><i>f21</i></b>	<i>4.0e+1</i> :3.9	<i>2.5e+1</i> :11	<i>1.6e+1</i> :31	<i>6.3e+0</i> :73	<i>1.6e+0</i> :347	5/5
BIPOP-aCMA	<b>1.3</b> (2)	<b>1.5</b> (2)	<b>2.0</b> (2)	<b>2.2</b> (1.0)	10(13)	15/15
BIPOP-saAC	<b>2.4</b> (3)	<b>1.5</b> (2)	<b>1.2</b> (1)	3.2(5)	3.1(5)	15/15
CMAES hut	<b>2.3</b> (3)	<b>1.3</b> (1)	<b>1.3</b> (2)	<b>2.0</b> (1)	3.4(4)	5/15
DE pal	<b>1.7</b> (2)	<b>1.8</b> (2)	<b>2.1</b> (2)	4.7(4)	4.4(4)	15/15
HCMA los	3.3(2)	<b>2.0</b> (0.9)	<b>0.89</b> (0.7)	<b>2.5</b> (6)	4.6(5)	15/15
HMLSL pal	3.6(5)	3.5(5)	<b>1.8</b> (2)	4.5(5)	<b>2.6</b> (2)	15/15
IPOP-10DDr	3.0(3)	<b>2.4</b> (3)	<b>1.2</b> (1)	6.1(11)	15(11)	15/15
IPOP-500 l	3.0(3)	<b>2.4</b> (3)	<b>1.2</b> (1)	6.1(11)	14(11)	15/15
IPOP-tany	<b>2.4</b> (2)	<b>1.7</b> (2)	<b>1.1</b> (1)	3.8(9)	6.6(9)	15/15
IPOP-texp	<b>1.4</b> (1)	<b>1.6</b> (2)	<b>1.0</b> (0.9)	<b>2.0</b> (1)	23(43)	15/15
IPOP lia	3.0(3)	<b>2.4</b> (3)	<b>1.2</b> (1)	6.1(11)	14(11)	15/15
MLSL pal	3.6(5)	3.8(6)	<b>2.0</b> (3)	3.4(3)	<b>2.1</b> (2)	15/15
OQNLP pal	4.9(3)	3.6(5)	<b>1.8</b> (2)	<b>1.9</b> (2)	3.0(6)	13/15
P-DCN tra	<b>2.1</b> (2)	<b>1.8</b> (2)	1.2e4(3)	2.5e4(3e4)	2.2e4(3e4)	6/15
P-zero tra	<b>2.3</b> (2)	<b>2.4</b> (5)	<b>2.9</b> (3)	1.7e4(3e4)	4.0e4(5e4)	4/15
SMAC hut	<b>1.8</b> (2)	<b>1.8</b> (2)	<b>1.0</b> (0.9)	<b>1.0</b> (0.9)	<b>1.0</b> (1)	11/15
U-DCN tra	<b>2.1</b> (3)	<b>1.8</b> (2)	<b>2.2</b> (2)	3.0(2)	358(645)	15/15
U-zero tra	<b>1.8</b> (2)	<b>1.2</b> (0.9)	<b>0.91</b> (0.9)	3.7(7)	697(1323)	15/15
fmincon pa	3.6(5)	3.6(5)	<b>2.2</b> (2)	3.6(3)	<b>1.5</b> (1)	15/15
fminunc pa	4.1(2)	3.6(6)	<b>1.8</b> (3)	<b>2.6</b> (3)	<b>2.1</b> (2)	15/15
ga100 hol	<b>1.7</b> (2)	<b>2.6</b> (3)	<b>2.1</b> (2)	7.0(3)	58(6)	14/15
grid100 ho	<b>1.9</b> (2)	<b>1.9</b> (2)	<b>1.8</b> (2)	12(14)	379(718)	11/15
grid16 hol	<b>1.4</b> (1)	<b>1.2</b> (1)	<b>2.0</b> (2)	6.8(8)	558(729)	9/15
hill hol	<b>2.7</b> (3)	<b>2.9</b> (4)	<b>2.7</b> (4)	13(27)	883(1083)	7/15
lmmCMA aug	<b>1.9</b> (2)	<b>1.4</b> (2)	<b>0.93</b> (0.7)	<b>1.6</b> (0.8)	<b>2.7</b> (4)	13/15
memPSODE v	<b>2.3</b> (2)	<b>2.2</b> (2)	4.3(7)	6.0(4)	12(13)	15/15
prcga saw	<b>1.3</b> (1)	<b>2.2</b> (2)	<b>2.3</b> (2)	4.3(4)	44(68)	15/15
ring100 ho	<b>1.8</b> (2)	<b>1.4</b> (1)	<b>2.3</b> (3)	10(8)	7.5(3)	15/15
ring16 hol	<b>2.1</b> (4)	<b>2.4</b> (3)	<b>2.2</b> (2)	4.0(2)	116(361)	13/15
simplex pa	47(63)	35(57)	18(19)	14(8)	5.1(4)	15/15

Table 23: 05-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  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b><i>f22</i></b>	<i>6.3e+1</i> :3.6	<i>4.0e+1</i> :15	<i>2.5e+1</i> :32	<i>1.0e+1</i> :71	<i>1.6e+0</i> :341	5/5
BIPOP-aCMA	<b>2.6</b> (2)	<b>2.4</b> (2)	<b>1.9</b> (1)	8.1(11)	22(40)	15/15
BIPOP-saAC	3.8(5)	4.4(3)	3.8(8)	9.1(13)	<b>7.2</b> (10)	15/15
CMAES hut	<b>2.5</b> (3)	<b>1.8</b> (2)	<b>1.3</b> (2)	<b>1.9</b> (1)	4.6(5)	4/15
DE pal	<b>2.0</b> (2)	<b>2.2</b> (2)	<b>1.8</b> (2)	3.5(2)	77(149)	12/15
HCMA los	<b>2.9</b> (2)	<b>2.4</b> (2)	<b>1.5</b> (1)	3.3(6)	6.1(6)	15/15
HMLSL pal	<b>2.4</b> (2)	3.2(5)	<b>1.9</b> (3)	3.6(3)	<b>2.3</b> (2)	15/15
IPOP-10DDr	<b>2.7</b> (2)	<b>2.2</b> (2)	<b>1.8</b> (1)	7.4(11)	47(60)	15/15
IPOP-500 l	<b>2.7</b> (2)	<b>2.2</b> (2)	<b>1.8</b> (1)	7.4(11)	104(31)	15/15
IPOP-tany	<b>1.9</b> (2)	<b>1.6</b> (1)	<b>1.4</b> (1)	12(9)	32(57)	15/15
IPOP-texp	<b>2.1</b> (2)	<b>1.5</b> (1)	<b>1.0</b> (1)	22(9)	14(27)	15/15
IPOP lia	<b>2.7</b> (2)	<b>2.2</b> (2)	<b>1.8</b> (1)	7.4(11)	37(31)	15/15
MLSL pal	<b>2.4</b> (2)	3.3(6)	<b>2.0</b> (4)	<b>2.7</b> (3)	<b>2.0</b> (3)	15/15
OQNLP pal	4.5(3)	3.0(3)	<b>2.7</b> (3)	<b>2.8</b> (2)	<b>2.8</b> (4)	14/15
P-DCN tra	4.5(6)	4.0(4)	2.4e4(8e4)	2.6e4(4e4)	2.2e4(3e4)	6/15
P-zero tra	3.3(4)	4.0(5)	2.4e4(8e4)	2.6e4(4e4)	4.0e4(5e4)	4/15
SMAC hut	<b>1.8</b> (3)	<b>1.5</b> (1)	<b>1.0</b> (0.8)	<b>0.90</b> (0.8)	<b>1.0</b> (1)	11/15
U-DCN tra	<b>2.9</b> (2)	<b>2.3</b> (2)	<b>1.6</b> (0.8)	<b>1.9</b> (1)	368(186)	15/15
U-zero tra	<b>2.8</b> (2)	<b>1.7</b> (1)	<b>2.5</b> (2)	4.8(6)	1248(519)	14/15
fmincon pa	<b>2.4</b> (2)	3.4(6)	<b>1.9</b> (3)	<b>2.2</b> (3)	<b>1.4</b> (1)	15/15
fminunc pa	<b>2.8</b> (2)	<b>2.5</b> (4)	<b>2.4</b> (3)	3.3(3)	<b>1.4</b> (1)	15/15
ga100 hol	3.2(5)	<b>2.2</b> (3)	<b>2.1</b> (2)	5.9(5)	272(369)	11/15
grid100 ho	3.0(2)	<b>2.0</b> (1)	<b>2.5</b> (3)	5.5(6)	139(368)	13/15
grid16 hol	<b>1.9</b> (2)	<b>1.3</b> (2)	<b>1.00</b> (1.0)	5.1(7)	771(1103)	8/15
hill hol	4.7(5)	<b>2.0</b> (2)	7.5(11)	6.7(10)	516(735)	9/15
lmmCMA aug	<b>1.6</b> (2)	<b>1.4</b> (1)	<b>1.3</b> (1)	<b>2.5</b> (3)	4.2(5)	12/15
memPSODE v	4.0(3)	4.4(3)	4.3(6)	4.6(3)	8.7(12)	15/15
prcga saw	<b>1.3</b> (2)	<b>2.4</b> (4)	<b>2.1</b> (3)	3.3(3)	123(275)	15/15
ring100 ho	<b>2.1</b> (2)	<b>1.4</b> (2)	3.4(4)	5.5(3)	8.7(5)	15/15
ring16 hol	<b>2.0</b> (2)	<b>1.3</b> (1)	<b>2.0</b> (2)	<b>2.5</b> (2)	104(347)	14/15
simplex pa	22(25)	25(37)	16(21)	11(10)	5.7(3)	15/15



Table 24: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{23}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b><i>f23</i></b>	<i>1.0e+1:3.0</i>	<i>6.3e+0:9.0</i>	<i>4.0e+0:33</i>	<i>2.5e+0:84</i>	<i>1.0e+0:518</i>	15/15
BIPOP-aCMA	<b>1.8</b> (2)	<b>2.0</b> (3)	3.6(7)	6.8(7)	8.6(6)	15/15
BIPOP-saAC	<b>2.0</b> (2)	<b>2.3</b> (3)	4.2(4)	7.3(12)	12(14)	15/15
CMAES hut	<b>1.6</b> (1)	<b>2.9</b> (3)	5.4(5)	6.7(7)	$\infty$ 506	0/15
DE pal	<b>1.9</b> (2)	<b>2.5</b> (4)	3.2(2)	7.0(8)	31(27)	15/15
HCMA los	8.1(8)	4.6(5)	3.0(2)	6.8(7)	13(10)	15/15
HMLSL pal	9.2(11)	5.1(5)	<b>2.9</b> (2)	<b>2.4</b> (1)	<b>2.4</b> (2)	15/15
IPOP-10DDr	<b>2.1</b> (2)	<b>1.9</b> (2)	<b>2.3</b> (2)	3.4(3)	7.2(5)	15/15
IPOP-500 l	<b>2.1</b> (2)	<b>1.9</b> (2)	<b>2.3</b> (2)	3.4(3)	7.2(5)	15/15
IPOP-tany	<b>1.7</b> (1)	<b>2.3</b> (2)	3.3(4)	6.7(6)	7.5(6)	15/15
IPOP-texp	<b>3.0</b> (3)	<b>2.7</b> (4)	3.1(3)	3.7(4)	7.0(7)	15/15
IPOP lia	<b>2.1</b> (2)	<b>1.9</b> (2)	<b>2.3</b> (2)	3.4(3)	7.2(5)	15/15
MLSL pal	9.2(11)	5.1(5)	<b>2.9</b> (2)	<b>2.2</b> (1)	<b>1.9</b> (2)	15/15
OQNLP pal	29(47)	14(18)	6.5(7)	5.8(4)	5.6(4)	15/15
P-DCN tra	3.6(3)	<b>2.9</b> (3)	<b>2.1</b> (2)	<b>2.6</b> (3)	<b>3.2</b> (4)	15/15
P-zero tra	3.6(3)	<b>2.1</b> (2)	<b>2.7</b> (3)	<b>2.6</b> (3)	29(62)	15/15
SMAC hut	<b>1.6</b> (2)	<b>2.9</b> (3)	<b>2.6</b> (3)	3.3(3)	$\infty$ 500	0/15
U-DCN tra	3.4(2)	<b>2.9</b> (3)	<b>2.2</b> (2)	3.2(4)	10(17)	15/15
U-zero tra	3.2(2)	<b>2.2</b> (3)	<b>2.0</b> (1)	6.4(8)	43(50)	15/15
fmincon pa	7.2(6)	5.6(4)	3.4(6)	<b>2.4</b> (3)	<b>1.6</b> (2)	15/15
fminunc pa	9.1(16)	10(12)	6.4(7)	5.5(5)	10(10)	15/15
ga100 hol	<b>2.3</b> (2)	<b>2.7</b> (4)	<b>2.3</b> (3)	5.2(5)	45(46)	15/15
grid100 ho	<b>3.0</b> (2)	<b>2.1</b> (2)	4.7(3)	8.9(13)	60(74)	15/15
grid16 hol	<b>2.2</b> (3)	<b>1.8</b> (2)	<b>3.0</b> (4)	6.5(5)	40(30)	15/15
hill hol	<b>2.5</b> (3)	<b>1.7</b> (2)	4.1(4)	11(9)	49(45)	15/15
lmmCMA aug	<b>1.9</b> (2)	<b>2.6</b> (2)	<b>2.6</b> (3)	7.1(6)	10(11)	6/15
memPSODE v	3.4(3)	4.1(2)	5.4(5)	4.0(3)	3.4(3)	15/15
prcga saw	<b>1.9</b> (2)	3.0(2)	3.0(4)	6.2(7)	39(50)	15/15
ring100 ho	<b>1.8</b> (2)	<b>2.3</b> (2)	<b>2.8</b> (2)	4.6(4)	23(27)	15/15
ring16 hol	<b>3.0</b> (2)	<b>2.5</b> (3)	4.1(4)	7.3(10)	26(29)	15/15
simplex pa	63(104)	47(51)	19(16)	11(3)	<b>3.0</b> (1)	15/15

Table 25: 05-D, running time excess  $\text{ERT}/\text{ERT}_{\text{best}} 2009$  on  $f_{24}$  for given run-length based budgets (0.5D, 1.2D, 3D, 10D, and 50D function evaluations). The ERT and in braces, as dispersion measure, the half difference between 90 and 10%-tile of bootstrapped run lengths appear for each algorithm and run-length based target, the corresponding  $\text{ERT}_{\text{best}} 2009$  (preceded by the target  $\Delta f$ -value in *italics*) in the first row. #succ is the number of trials that reached the target value of the last column. The median number of conducted function evaluations is additionally given in *italics*, if the target in the last column was never reached. Entries with succeeding star are statistically significantly better (according to the rank-sum test) compared to all other algorithms in the table, with  $p = 0.05$  or  $p = 10^{-k}$  when the number  $k$  following the star is larger than 1, with Bonferroni correction by the number of instances.

#FEs/D	0.5	1.2	3	10	50	#succ
<b>f24</b>	<i>6.3e+1</i> :15	<i>4.0e+1</i> :37	<i>4.0e+1</i> :37	<i>2.5e+1</i> :118	<i>1.6e+1</i> :692	15/15
BIPOP-aCMA	<b>2.6</b> (1)	3.3(2)	3.3(2)	<b>2.2</b> (1)	<b>1.1</b> (0.7)	15/15
BIPOP-saAC	<b>1.9</b> (2)	<b>2.6</b> (1)	<b>2.6</b> (1)	<b>2.6</b> (2)	<b>0.98</b> (1.0)	15/15
CMAES hut	<b>1.6</b> (2)	<b>1.9</b> (2)	<b>1.9</b> (2)	<b>2.2</b> (2)	<b>1.4</b> (2)	6/15
DE pal	3.6(3)	6.2(3)	6.2(3)	5.3(3)	4.4(3)	15/15
HCMA los	<b>1.0</b> (0.3)	<b>1.4</b> (1)	<b>1.4</b> (1)	<b>2.0</b> (2)	<b>1.3</b> (1)	15/15
HMLSL pal	<b>2.1</b> (0.9)	<b>2.9</b> (4)	<b>2.9</b> (4)	6.0(4)	<b>1.9</b> (0.9)	15/15
IPOP-10DDr	<b>2.1</b> (2)	<b>2.1</b> (0.7)	<b>2.1</b> (0.7)	<b>1.6</b> (1)	<b>0.99</b> (0.6)	15/15
IPOP-500 l	<b>2.1</b> (2)	<b>2.1</b> (0.7)	<b>2.1</b> (0.7)	<b>1.6</b> (1)	<b>0.99</b> (0.6)	15/15
IPOP-tany	<b>1.3</b> (1)	<b>2.1</b> (1)	<b>2.1</b> (1)	<b>1.4</b> (0.6)	<b>1.2</b> (0.9)	15/15
IPOP-texp	<b>0.49</b> (0.4)	<b>1.3</b> (1)	<b>1.3</b> (1)	<b>1.2</b> (1.0)	<b>1.2</b> (1)	15/15
IPOP lia	<b>2.1</b> (2)	<b>2.1</b> (0.7)	<b>2.1</b> (0.7)	<b>1.6</b> (1)	<b>0.99</b> (0.6)	15/15
MLSL pal	<b>1.9</b> (0.9)	<b>2.7</b> (5)	<b>2.7</b> (5)	3.5(3)	<b>2.0</b> (2)	15/15
OQNLP pal	<b>1.1</b> (0.7)	<b>1.6</b> (0.6)	<b>1.6</b> (0.6)	4.4(1)	<b>1.2</b> (0.8)	15/15
P-DCN tra	<b>1.9</b> (1)	4.7(4)	4.7(4)	422(1096)	534(553)	15/15
P-zero tra	<b>2.1</b> (2)	1691(1763)	1691(1763)	6787(2e4)	1.2e4(1e4)	7/15
SMAC hut	<b>0.51</b> (0.5)	<b>2.6</b> (3)	<b>2.6</b> (3)	7.4(7)	$\infty$ 500	0/15
U-DCN tra	<b>2.4</b> (3)	5.1(5)	5.1(5)	5.8(3)	7.2(8)	15/15
U-zero tra	<b>1.5</b> (0.8)	3.3(5)	3.3(5)	7.5(9)	10(9)	15/15
fmincon pa	<b>1.7</b> (0.9)	<b>2.3</b> (2)	<b>2.3</b> (2)	3.4(2)	<b>2.2</b> (2)	15/15
fminunc pa	<b>1.9</b> (1)	<b>1.9</b> (0.7)	<b>1.9</b> (0.7)	6.0(1)	<b>2.5</b> (2)	15/15
ga100 hol	<b>2.1</b> (3)	10(6)	10(6)	10(6)	6.2(6)	15/15
grid100 ho	4.2(4)	18(18)	18(18)	23(16)	38(55)	15/15
grid16 hol	3.3(4)	6.4(4)	6.4(4)	6.0(4)	6.1(7)	15/15
hill hol	5.2(6)	6.6(7)	6.6(7)	17(35)	12(8)	15/15
ImmCMA aug	<b>0.77</b> (1.0)	<b>2.0</b> (2)	<b>2.0</b> (2)	<b>1.7</b> (1)	<b>0.91</b> (0.6)	15/15
memPSODE v13	13(16)	18(7)	18(7)	11(9)	8.9(8)	15/15
prcga saw	<b>1.2</b> (1)	<b>2.5</b> (2)	<b>2.5</b> (2)	<b>2.8</b> (2)	<b>2.0</b> (1)	15/15
ring100 ho	4.6(5)	14(11)	14(11)	20(13)	8.7(7)	15/15
ring16 hol	<b>2.6</b> (2)	5.1(6)	5.1(6)	7.7(6)	3.7(2)	15/15
simplex pa	<b>2.9</b> (4)	6.9(4)	6.9(4)	20(19)	6.2(6)	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.