Probability Matching-based Adaptive Strategy Selection Compared with Uniform Strategy Selection within Differential Evolution on the Noiseless Testbed

Draft version *

Álvaro Fialho
Microsoft Research – INRIA Joint Centre
Parc Orsay Université
91893 Orsay, France
alvaro.fialho@inria.fr

Wenyin Gong
School of Computer Science
China University of Geosciences
Wuhan, 430074 P.R. China
cug11100304@yahoo.com.cn

Zhihua Cai
School of Computer Science
China University of Geosciences
Wuhan, 430074 P.R. China
zhcai@cug.edu.cn

ABSTRACT

The decision of which of the several existent strategies should be applied for the offspring generation is critical for the performance of the Differential Evolution algorithm, besides being problem-dependent. In this paper, we use the BBOB noiseless benchmarking suite to better empirically validate the Probability Matching-based Adaptive Strategy Selection, a technique used to automatically select between the available strategies while solving the problem, recently proposed in [2], referred to as *PM-AdapSS-DE*. It is compared with what would be a possible choice for a naïve user, the uniform strategy selection within the same sub-set of strategies.

Categories and Subject Descriptors

G.1.6 [**Numerical Analysis**]: Optimization—*global optimization, unconstrained optimization*; F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems

General Terms

Algorithms

Keywords

Benchmarking, Black-box optimization

1. INTRODUCTION

The decision of which of the several available strategies should be applied for the offspring generation in Differential Evolution is critical for its performance, besides being problem-dependent. Although the best strategy for a given problem might be found by means of statistics over an expensive set of experiments, a subsequent use of different strategies during the optimization process should achieve better performance, following its intuitive migration from a global (early) exploration of the landscape to a more focused, exploitation-like behavior.

*Submission deadline: March 25th.

2. ALGORITHM PRESENTATION

3. RESULTS

Results from experiments according to [3] on the benchmark functions given in [1, 4] are presented in Figures 1, 2 and 3 and in Table 1. The **expected running time (ERT)**, used in the figures and table, depends on a given target function value, $f_t = f_{opt} + \Delta f$, and is computed over all relevant trials as the number of function evaluations executed during each trial while the best function value did not reach f_t, summed over all trials and divided by the number of trials that actually reached f_t [3, 5]. **Statistical significance** is tested with the rank-sum test for a given target Δf_t (10^{-8} in Figure 1) using, for each trial, either the number of needed function evaluations to reach Δf_t (inverted and multiplied by -1), or, if the target was not reached, the best Δf-value achieved, measured only up to the smallest number of overall function evaluations for any unsuccessful trial under consideration.

4. REFERENCES

[1] 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.

[2] W. Gong, A. Fialho, and Z. Cai. Adaptive strategy selection in differential evolution. In J. B. et al., editor, *GECCO'10: Proceedings of the 12th Annual Conference on Genetic and Evolutionary Computation*. ACM Press, July 2010. to appear.

[3] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2010: Experimental setup. Technical Report RR-7215, INRIA, 2010.

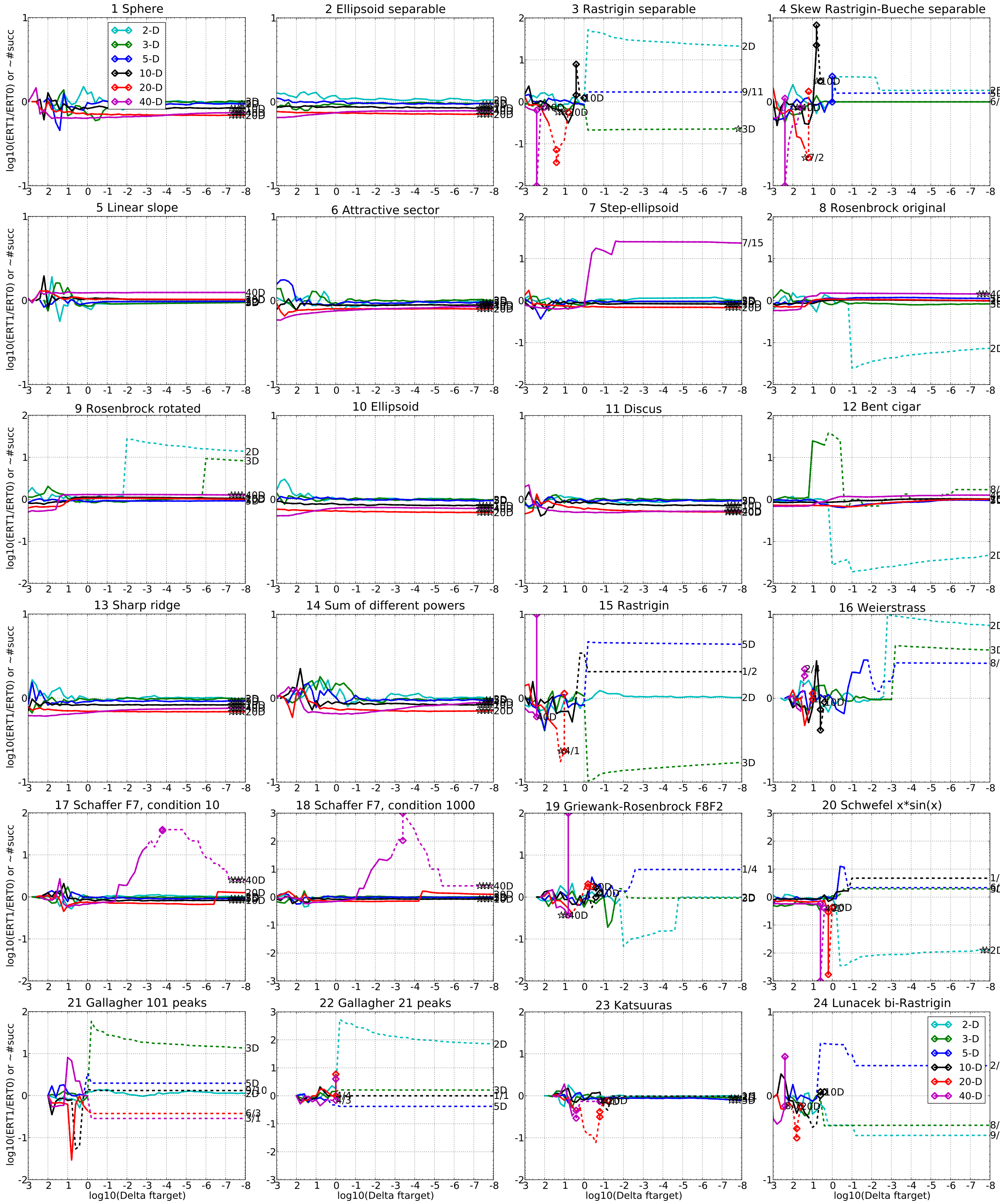

Figure 1: ERT ratio of ALG1-acronym divided by ALG0-acronym versus $\log_{10}(\Delta f)$ for f_1–f_{24} in 2, 3, 5, 10, 20, 40-D. Ratios $< 10^0$ indicate an advantage of ALG1-acronym, smaller values are always better. The line gets dashed when for any algorithm the ERT exceeds thrice the median of the trial-wise overall number of f-evaluations for the same algorithm on this function. Symbols indicate the best achieved Δf-value of one algorithm (ERT gets undefined to the right). The dashed line continues as the fraction of successful trials of the other algorithm, where 0 means 0% and the y-axis limits mean 100%, values below zero for ALG1-acronym. The line ends when no algorithm reaches Δf anymore. The number of successful trials is given, only if it was in $\{1 \dots 9\}$ for ALG1-acronym (1st number) and non-zero for ALG0-acronym (2nd number). Results are significant with $p = 0.05$ for one star and $p = 10^{-\#\star}$ otherwise, with Bonferroni correction within each figure.

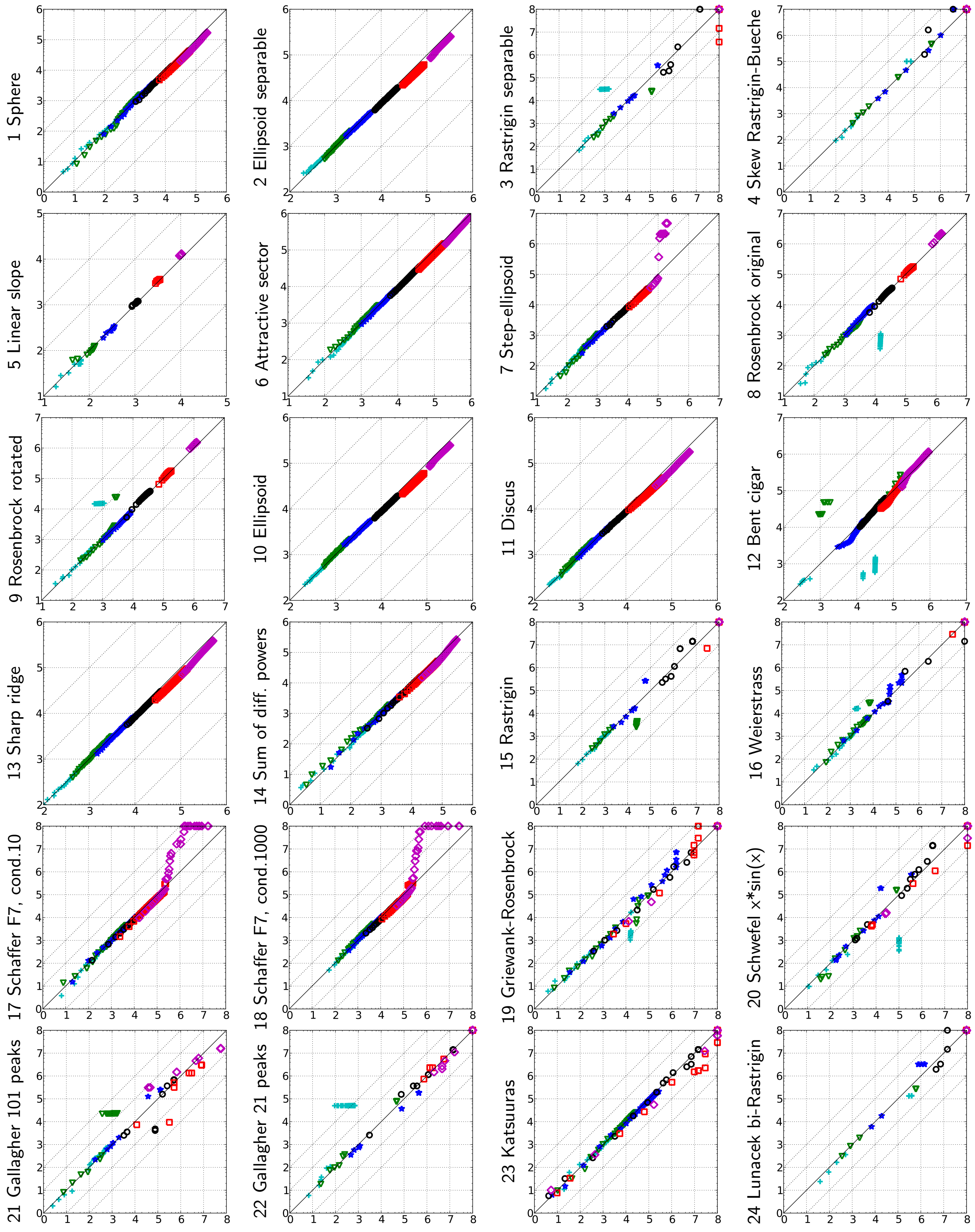

Figure 2: Expected running time (ERT in log10 of number of function evaluations) of ALG1-acronym versus ALG0-acronym for 46 target values $\Delta f \in [10^{-8}, 10]$ in each dimension for functions f_1–f_{24}. Markers on the upper or right egde indicate that the target value was never reached by ALG1-acronym or ALG0-acronym respectively. Markers represent dimension: 2:+, 3:▽, 5:⋆, 10:○, 20:□, 40:◇.

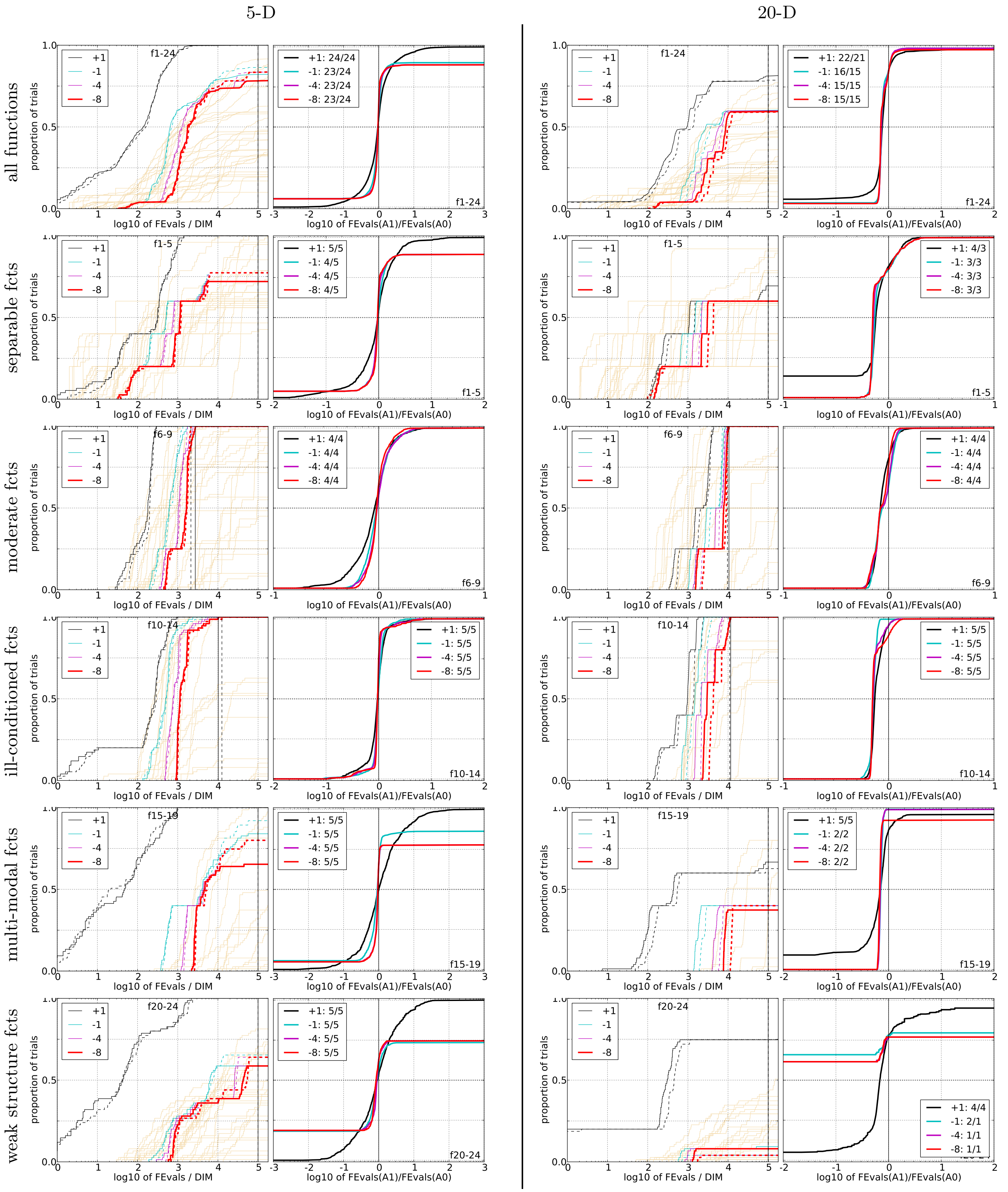

Figure 3: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/D) to reach a target value $f_{\mathrm{opt}} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ is given by the first value in the legend, for ALG1-acronym (solid) and ALG0-acronym (dashed). Light beige lines show the ECDF of FEvals for target value $\Delta f = 10^{-8}$ of algorithms benchmarked during BBOB-2009. Right sub-columns: ECDF of FEval ratios of ALG1-acronym divided by ALG0-acronym, all trial pairs for each function. Pairs where both trials failed are disregarded, pairs where one trial failed are visible in the limits being > 0 or < 1. The legends indicate the number of functions that were solved in at least one trial (ALG1-acronym first).

5-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	11	12	12	12	12	12	15/15
0: stG	8.6	42	80	160	240	310	15/15
1: wen	7.4	41	80	150	220	300	15/15
$\mathbf{f_2}$	83	87	88	90	92	94	15/15
0: stG	20	24	30	40	50	58	15/15
1: wen	20	25	29	38	47	**55***	15/15
$\mathbf{f_3}$	720	1600	1600	1600	1700	1700	15/15
0: stG	3.5	12	130	130	130	130	11/15
1: wen	3.9	11	220	210	210	210	9/15
$\mathbf{f_4}$	810	1600	1700	1800	1900	1900	15/15
0: stG	5.1	630	1.9e3	1.8e3	1.7e3	1.7e3	2/15
1: wen	4.8	620	∞	∞	∞	∞ *5.0e5*	0/15
$\mathbf{f_5}$	10	10	10	10	10	10	15/15
0: stG	20	34	35	35	35	35	15/15
1: wen	19	32	34	34	34	34	15/15
$\mathbf{f_6}$	110	210	280	580	1000	1300	15/15
0: stG	9	9.2	10	7.8	6.1	6.1	15/15
1: wen	8.2	8.2	9.4	7.3	5.8	5.8	15/15
$\mathbf{f_7}$	24	320	1200	1600	1600	1600	15/15
0: stG	14	2.7	1.2	1.4	1.4	1.6	15/15
1: wen	11	2.5	1.1	1.4	1.4	1.5	15/15
$\mathbf{f_8}$	73	270	340	390	410	420	15/15
0: stG	15	10	14	17	18	20	15/15
1: wen	14	12	16	20	21	22	15/15
$\mathbf{f_9}$	35	130	210	300	340	370	15/15
0: stG	30	23	21	21	22	22	15/15
1: wen	28	20	20	20	20	20	15/15
$\mathbf{f_{10}}$	350	500	570	630	830	880	15/15
0: stG	4.5	4.2	4.6	5.5	5.4	6.2	15/15
1: wen	4.8	4.3	4.5	5.6	5.2	6	15/15
$\mathbf{f_{11}}$	140	200	760	1200	1500	1700	15/15
0: stG	6.2	6.6	2.4	2.4	2.6	2.8	15/15
1: wen	6	6.4	2.4	2.3	2.5	2.7	15/15
$\mathbf{f_{12}}$	110	270	370	460	1300	1500	15/15
0: stG	28	21	21	22	10	11	15/15
1: wen	27	14	14	18	9.3	10	15/15
$\mathbf{f_{13}}$	130	190	250	1300	1800	2300	15/15
0: stG	11	12	13	3.7	3.8	3.7	15/15
1: wen	**10**$^{\star 2}$	**11***	**12**$^{\star 2}$	**3.4**$^{\star 3}$	**3.5**$^{\star 3}$	**3.4**$^{\star 3}$	15/15
$\mathbf{f_{14}}$	9.8	41	58	140	250	480	15/15
0: stG	2.3	10	16	15	13	9.3	15/15
1: wen	1.8	10	18	15	**12**$^{\star 2}$	9	15/15
$\mathbf{f_{15}}$	510	9300	1.9e4	2.0e4	2.1e4	2.1e4	14/15
0: stG	4.8	2.2	3	3	2.9	2.9	14/15
1: wen	5.3	1.9	14	13	13	13	10/15
$\mathbf{f_{16}}$	120	610	2700	1.0e4	1.2e4	1.2e4	15/15
0: stG	4.3	44	20	17	16	15	12/15
1: wen	5.6	43	41	27	41	39	8/15
$\mathbf{f_{17}}$	5.2	210	900	3700	6400	7900	15/15
0: stG	3.8	4.3	2.7	1.6	1.5	1.6	15/15
1: wen	3	4.3	2.5	1.5	**1.4**$^{\star 2}$	**1.5**$^{\star 2}$	15/15
$\mathbf{f_{18}}$	100	380	4000	9300	1.1e4	1.2e4	15/15
0: stG	4	4.3	0.81	0.73	0.95	1.1	15/15
1: wen	3.5	4.2	0.78	0.72	0.92	1	15/15
$\mathbf{f_{19}}$	1	1	240	1.2e5	1.2e5	1.2e5	15/15
0: stG	35	3.4e3	1.6e3	13	13	13	4/15
1: wen	41	3.2e3	1.6e3	60	60	59	1/15
$\mathbf{f_{20}}$	16	850	3.8e4	5.4e4	5.5e4	5.5e4	14/15
0: stG	11	10	9.2	6.4	6.4	6.4	9/15
1: wen	8.5	8.9	20	14	14	14	6/15
$\mathbf{f_{21}}$	41	1200	1700	1700	1700	1800	14/15
0: stG	4.5	33	76	75	74	74	12/15
1: wen	5.4	110	150	150	150	140	10/15
$\mathbf{f_{22}}$	71	390	940	1000	1000	1100	14/15
0: stG	6.6	200	470	440	420	410	8/15
1: wen	5	96	200	180	180	170	11/15
$\mathbf{f_{23}}$	3	520	1.4e4	3.2e4	3.3e4	3.4e4	15/15
0: stG	2	11	2.5	3.6	5.5	7.2	15/15
1: wen	2.1	9.7	2.1	**3.1**$^{\star 2}$	**4.8**$^{\star 3}$	**5.9**$^{\star 3}$	15/15
$\mathbf{f_{24}}$	1600	2.2e5	6.4e6	9.6e6	1.3e7	1.3e7	3/15
0: stG	4.3	3.7	0.17	0.15	0.11	0.11	4/15
1: wen	3.7	15	0.52	0.34	0.26	0.26	2/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	43	43	43	43	43	43	15/15
0: stG	150	300	440	730	1.0e3	1.3e3	15/15
1: wen	**110**$^{\star 3}$	**210**$^{\star 3}$	**310**$^{\star 3}$	**510**$^{\star 3}$	**710**$^{\star 3}$	**910**$^{\star 3}$	15/15
$\mathbf{f_2}$	380	390	390	390	390	390	15/15
0: stG	77	92	110	140	170	200	15/15
1: wen	**57**$^{\star 3}$	**68**$^{\star 3}$	**80**$^{\star 3}$	**100**$^{\star 3}$	**120**$^{\star 3}$	**140**$^{\star 3}$	15/15
$\mathbf{f_3}$	5100	7600	7600	7600	7600	7700	15/15
0: stG	∞	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	**730**$^{\star 2}$	∞	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_4}$	4700	7600	7700	7700	7800	1.4e5	9/15
0: stG	∞	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	∞	∞	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_5}$	41	41	41	41	41	41	15/15
0: stG	69	83	85	86	86	86	15/15
1: wen	72	87	88	88	88	88	15/15
$\mathbf{f_6}$	1300	2300	3400	5200	6700	8400	15/15
0: stG	29	23	20	20	20	20	15/15
1: wen	**23**$^{\star 3}$	**18**$^{\star 3}$	**16**$^{\star 3}$	**16**$^{\star 3}$	**16**$^{\star 3}$	**16**$^{\star 3}$	15/15
$\mathbf{f_7}$	1400	4300	9500	1.7e4	1.7e4	1.7e4	15/15
0: stG	8.4	5.2	3.3	2.7	2.7	2.7	15/15
1: wen	**6.3**$^{\star 3}$	**3.7**$^{\star 3}$	**2.3**$^{\star 3}$	**1.8**$^{\star 3}$	**1.8**$^{\star 3}$	**1.9**$^{\star 3}$	15/15
$\mathbf{f_8}$	2000	3900	4000	4200	4400	4500	15/15
0: stG	34	31	33	35	37	39	15/15
1: wen	35	33	35	37	38	39	15/15
$\mathbf{f_9}$	1700	3100	3300	3500	3600	3700	15/15
0: stG	41	39	41	43	45	46	15/15
1: wen	38	40	43	45	46	47	15/15
$\mathbf{f_{10}}$	7400	8700	1.1e4	1.5e4	1.7e4	1.7e4	15/15
0: stG	3.9	4.1	3.9	3.6	3.9	4.6	15/15
1: wen	**2.9**$^{\star 3}$	**3**$^{\star 3}$	**2.8**$^{\star 3}$	**2.6**$^{\star 3}$	**2.8**$^{\star 3}$	**3.2**$^{\star 3}$	15/15
$\mathbf{f_{11}}$	1000	2200	6300	9800	1.2e4	1.5e4	15/15
0: stG	11	8.1	3.8	3.8	4	4.2	15/15
1: wen	**9.6**$^{\star 3}$	**6.2**$^{\star 3}$	**2.9**$^{\star 3}$	**2.8**$^{\star 3}$	**2.9**$^{\star 3}$	**3**$^{\star 3}$	15/15
$\mathbf{f_{12}}$	1000	1900	2700	4100	1.2e4	1.4e4	15/15
0: stG	44	27	25	26	12	13	15/15
1: wen	**31**$^{\star 3}$	**19**$^{\star 3}$	**17**$^{\star 3}$	**22***	12	13	15/15
$\mathbf{f_{13}}$	650	2000	2800	1.9e4	2.4e4	3.0e4	15/15
0: stG	43	20	19	4.1	4.2	4.2	15/15
1: wen	**30**$^{\star 3}$	**14**$^{\star 3}$	**13**$^{\star 3}$	**2.8**$^{\star 3}$	**2.9**$^{\star 3}$	**2.9**$^{\star 3}$	15/15
$\mathbf{f_{14}}$	75	240	300	930	1600	1.6e4	15/15
0: stG	53	50	64	38	31	4.2	15/15
1: wen	46	**37**$^{\star 3}$	**47**$^{\star 3}$	**27**$^{\star 3}$	**22**$^{\star 3}$	**3**$^{\star 3}$	15/15
$\mathbf{f_{15}}$	3.0e4	1.5e5	3.1e5	3.2e5	4.5e5	4.6e5	15/15
0: stG	980	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	230	∞	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{16}}$	1400	2.7e4	7.7e4	1.9e5	2.0e5	2.2e5	15/15
0: stG	2.1e4	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	2.1e4	∞	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{17}}$	63	1000	4000	3.1e4	5.6e4	8.0e4	15/15
0: stG	38	19	10	3	3.2	2.8	15/15
1: wen	23	**14**$^{\star 3}$	**7.5**$^{\star 3}$	**2.1**$^{\star 3}$	**2.1**$^{\star 3}$	3.6	14/15
$\mathbf{f_{18}}$	620	4000	2.0e4	6.8e4	1.3e5	1.5e5	15/15
0: stG	18	7.8	2.8	1.7	1.5	1.6	15/15
1: wen	**14**$^{\star 3}$	**5.5**$^{\star 3}$	**2**$^{\star 3}$	**1.2**$^{\star 3}$	2.1	2.1	14/15
$\mathbf{f_{19}}$	1	1	3.4e5	6.2e6	6.7e6	6.7e6	15/15
0: stG	2.8e3	9.5e6	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	**1.9e3***	1.4e7	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{20}}$	82	4.6e4	3.1e6	5.5e6	5.6e6	5.6e6	14/15
0: stG	76	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	**51**$^{\star 3}$	300	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{21}}$	560	6500	1.4e4	1.5e4	1.6e4	1.8e4	15/15
0: stG	21	460	570	550	520	460	3/15
1: wen	**13**$^{\star 2}$	210	210	210	190	170	6/15
$\mathbf{f_{22}}$	470	5600	2.3e4	2.5e4	2.7e4	1.3e5	12/15
0: stG	1.6e3	990	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	1.6e3	990	∞	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{23}}$	3.2	1600	6.7e4	4.9e5	8.1e5	8.4e5	15/15
0: stG	3	6.0e3	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	2.4	**940***	**430***	∞	∞	∞ *2.0e6*	0/15
$\mathbf{f_{24}}$	1.3e6	7.5e6	5.2e7	5.2e7	5.2e7	5.2e7	3/15
0: stG	∞	∞	∞	∞	∞	∞ *2.0e6*	0/15
1: wen	∞	∞	∞	∞	∞	∞ *2.0e6*	0/15

Table 1: Expected running time (ERT in number of function evaluations) divided by the best ERT measured during BBOB-2009 (given in the respective first row) for different Δf values for functions f_1–f_{24}. The median number of conducted function evaluations is additionally given in *italics*, if $\mathrm{ERT}(10^{-7}) = \infty$. #succ is the number of trials that reached the final target $f_{\mathrm{opt}} + 10^{-8}$. 0: stG is ALG0-acronym and 1: wen is ALG1-acronym. Bold entries are statistically significantly better compared to the other algorithm, with $p = 0.05$ or $p = 10^{-k}$ where $k > 1$ is the number following the $\star$ symbol, with Bonferroni correction of 48.

Algorithm 1 Probability matching-based DE with adaptive strategy selection: PM-AdapSS-DE

```
Set CR = 0.9, F = 0.5 and NP = 10 × D
Generate the initial population
Evaluate the fitness for each individual
Set the generation counter t = 1
Set K = 4, p_min = 0.05, and α = 0.3                                  ⇐
For each strategy a, set q_a(t) = 0 and p_a(t) = 1/K                  ⇐
while The halting criterion is not satisfied do
    for i = 1 to NP do
        Select the strategy SI_i based on its probability             ⇐
        Select uniform randomly r_1 ≠ r_2 ≠ r_3 ≠ r_4 ≠ r_5 ≠ i
        j_rand = rndint(1, D)
        for j = 1 to D do
            if rndreal_j[0, 1) < CR or j == j_rand then
                if SI_i == 1 then
                    u_{i,j} is generated by "DE/rand/1" strategy
                else if SI_i == 2 then
                    u_{i,j} is generated by "DE/rand/2" strategy
                else if SI_i == 3 then
                    u_{i,j} is generated by "DE/rand-to-best/2" strategy
                else if SI_i == 4 then
                    u_{i,j} is generated by "DE/current-to-rand/1"
                end if
            else
                u_{i,j} = x_{i,j}
            end if
        end for
    end for
    for i = 1 to NP do
        Evaluate the offspring u_i
        if f(u_i) is better than or equal to f(x_i) then
            Calculate η_i using Eqn. (??)                             ⇐
            Replace x_i with u_i
        else
            Set η_i = 0                                               ⇐
        end if
        S_{SI_i} ← η_i                                                ⇐
    end for
    Calculate the reward r_a(t) for each strategy                     ⇐
    Update the quality q_a(t) for each strategy                       ⇐
    Update the probability p_a(t) for each strategy                   ⇐
    t = t + 1
end while
```


[4] 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.

[5] K. Price. Differential evolution vs. the functions of the second ICEO. In *Proceedings of the IEEE International Congress on Evolutionary Computation*, pages 153–157, 1997.