Example paper: Black-Box Optimization Benchmarking Template for Noisy Function Testbed

Draft version *

BBOBies

ABSTRACT

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. RESULTS

Results of NEWUOA from experiments according to [**?**] on the benchmark functions given in [**?**, **?**] are presented in Figures 1, 2 and 3 and in Tables 1, and 2.

Table 2: ERT loss ratio compared to the respective best result from BBOB-2009 for budgets given in the first column (see also Figure 3). The last row $\mathrm{RL_{US}}$/D gives the number of function evaluations in unsuccessful runs divided by dimension. Shown are the smallest, 10%-ile, 25%-ile, 50%-ile, 75%-ile and 90%-ile value (smaller values are better). The ERT Loss ratio equals to one for the respective best algorithm from BBOB-2009. Typical median values are between ten and hundred.

	f_{101}–f_{130} **in 5-D**, maxFE/D=6802					
#FEs/D	best	10%	25%	**med**	75%	90%
2	2.3	3.9	5.4	10	10	10
10	0.66	1.2	4.0	18	42	50
100	1.7	2.9	4.9	18	1.3e2	5.0e2
1e3	3.1	5.5	18	47	1.1e2	2.7e3
1e4	3.1	15	35	65	1.5e2	2.5e4
$\mathrm{RL_{US}}$/D	4e3	4e3	5e3	5e3	5e3	5e3
	f_{101}–f_{130} **in 20-D**, maxFE/D=7998					
#FEs/D	best	10%	25%	**med**	75%	90%
2	1.0	2.4	40	40	40	40
10	0.38	0.49	1.0	1.3e2	2.0e2	2.0e2
100	0.75	1.3	3.2	1.6e2	2.0e3	2.0e3
1e3	1.7	8.2	30	1.1e2	1.4e4	2.0e4
1e4	1.7	45	1.8e2	2.8e2	5.7e2	2.0e5
1e5	0.85	65	3.3e2	1.7e3	2.8e3	1.0e6
$\mathrm{RL_{US}}$/D	4e3	4e3	4e3	4e3	4e3	5e3

*Submission deadline: March 28th.

GECCO'12, July 7–11, 2012, Philadelphia, USA.

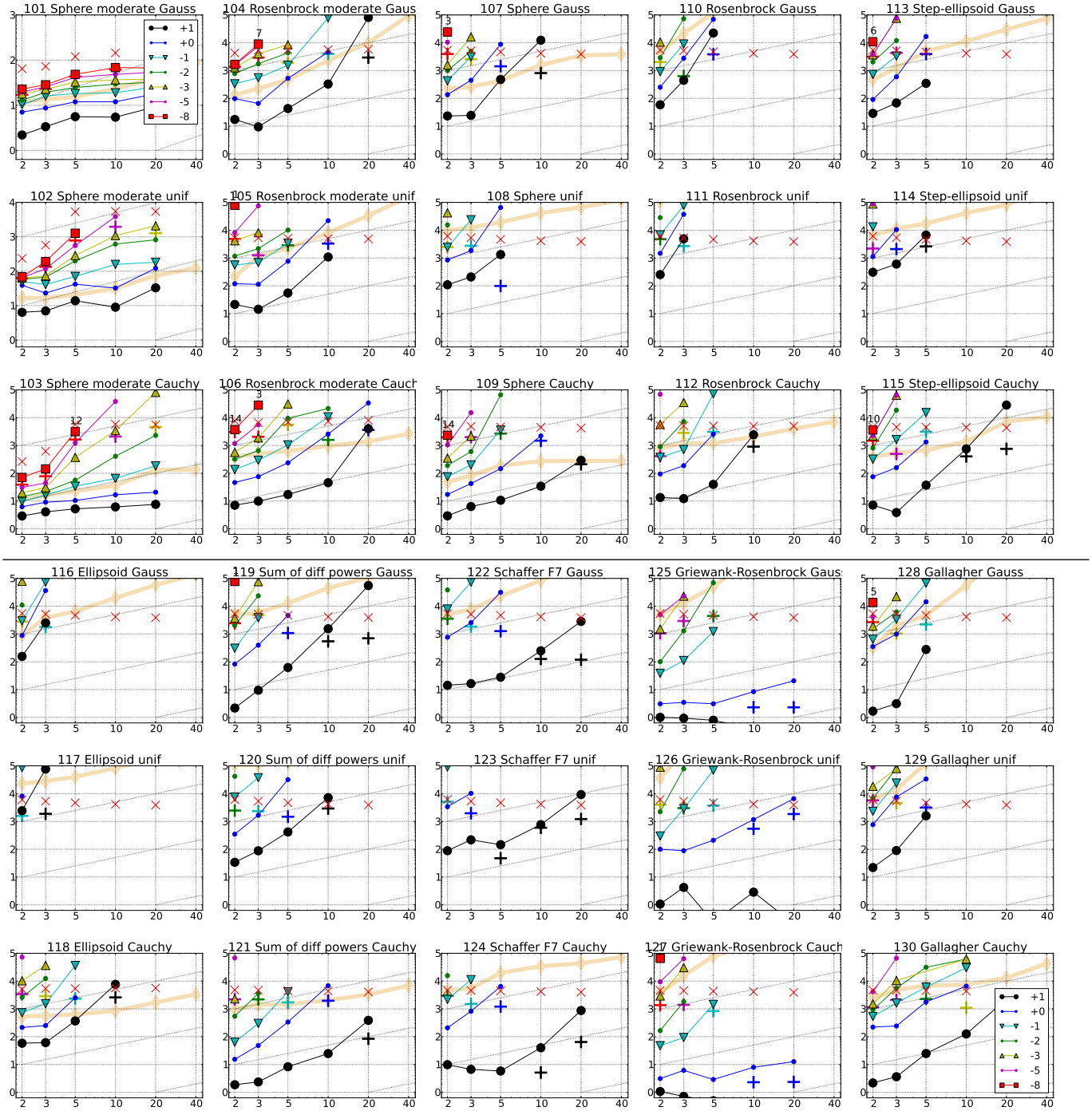

Figure 1: Expected number of f-evaluations (ERT, with lines, see legend) **to reach $f_{\mathrm{opt}} + \Delta f$, median number of f-evaluations to reach the most difficult target that was reached at least once (+) and maximum number of f-evaluations in any trial (×), all divided by dimension and plotted as $\log_{10}$ values versus dimension. Shown are $\Delta f = 10^{\{1,0,-1,-2,-3,-5,-8\}}$. Numbers above ERT-symbols indicate the number of successful trials. The light thick line with diamonds indicates the respective best result from BBOB-2009 for $\Delta f = 10^{-8}$. Horizontal lines mean linear scaling, slanted grid lines depict quadratic scaling.**

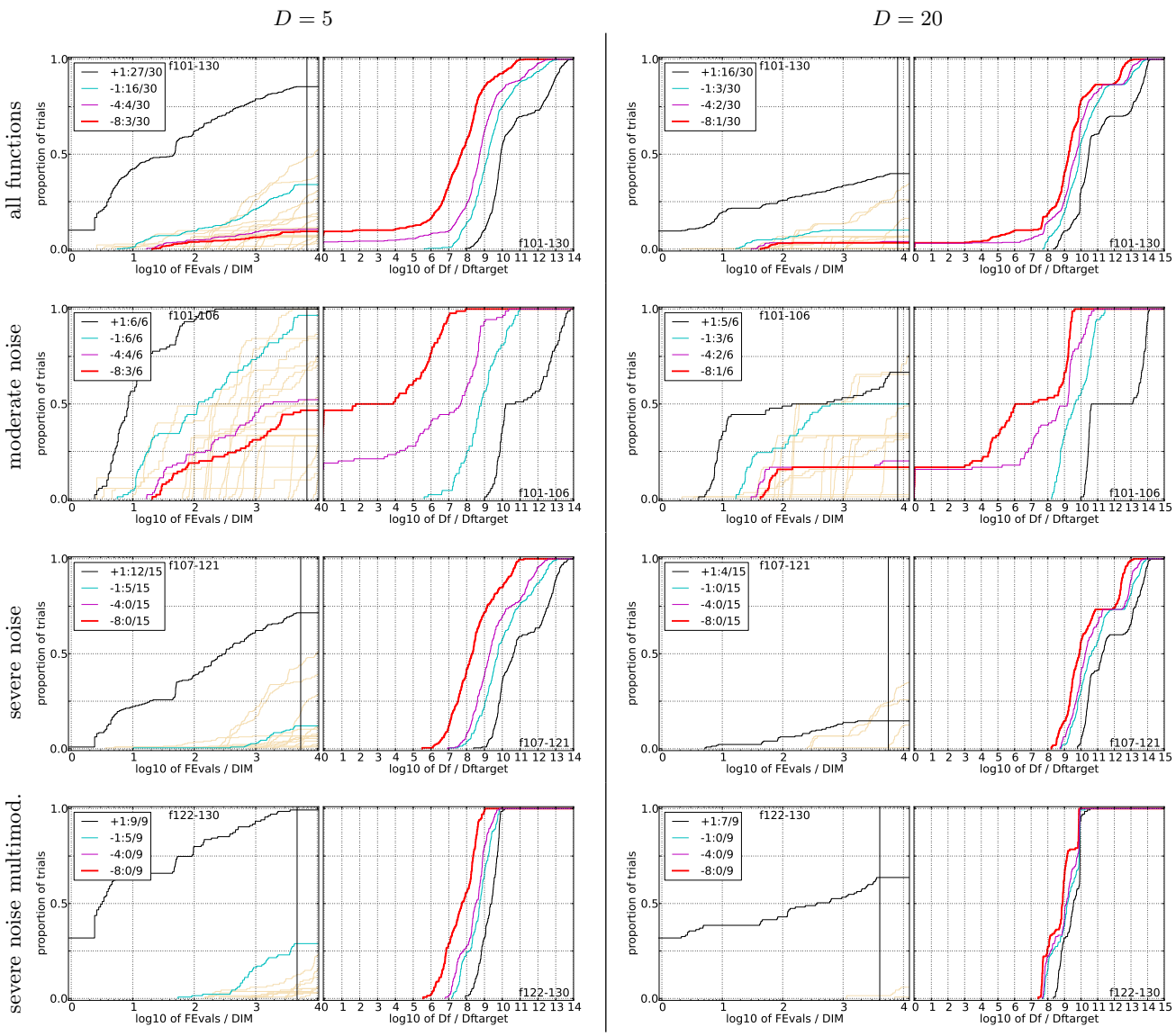

Figure 2: Empirical cumulative distribution functions (ECDFs), plotting the fraction of trials versus running time (left subplots) or versus Δf (right subplots). The thick red line represents the best achieved results. Left subplots: ECDF of the running time (number of function evaluations), divided by search space dimension D, to fall below $f_{\mathrm{opt}} + \Delta f$ with $\Delta f = 10^k$, where k is the first value in the legend. Right subplots: ECDF of the best achieved Δf divided by 10^k (upper left lines in continuation of the left subplot), and best achieved Δf divided by 10^{-8} for running times of $D, 10\,D, 100\,D \ldots$ function evaluations (from right to left cycling black-cyan-magenta). The legends indicate the number of functions that were solved in at least one trial. FEvals denotes number of function evaluations, D and DIM denote search space dimension, and Δf and Df denote the difference to the optimal function value. Light brown lines in the background show ECDFs for target value 10^{-8} of all algorithms benchmarked during BBOB-2009.

5-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f_{101}	11	37	44	62	69	75	15/15
	2.5(1)	1.6(0.7)	2.1(0.9)	2.6(2)	3.0(2)	3.1(2)	15/15
f_{102}	11	35	50	72	86	99	15/15
	6.3(11)	6.0(7)	7.0(9)	20(19)	33(32)	41(57)	15/15
f_{103}	11	28	30	31	35	115	15/15
	2.4(1.0)	1.9(0.7)	5.7(9)	60(66)	178(172)	136(174)	12/15
f_{104}	173	773	1287	1768	2040	2284	15/15
	1.2(2)	3.4(4)	6.0(8)	24(26)	∞	∞*2.7e4*	0/15
f_{105}	167	1436	5174	10388	10824	11202	15/15
	1.7(2)	2.7(3)	3.3(4)	∞	∞	∞*2.5e4*	0/15
f_{106}	92	529	1050	2666	2887	3087	15/15
	0.93(0.7)	2.2(3)	5.0(6)	59(60)	∞	∞*3.3e4*	0/15
f_{107}	40	228	453	940	1376	1850	15/15
	60(50)	194(213)	∞	∞	∞	∞*2.3e4*	0/15
f_{108}	87	5144	14469	30935	58628	80667	15/15
	77(89)	64(73)	∞	∞	∞	∞*2.3e4*	0/15
f_{109}	11	57	216	572	873	946	15/15
	4.8(10)	13(10)	83(101)	∞	∞	∞*2.3e4*	0/15
f_{110}	949	33625	1.2e5	5.9e5	6.0e5	6.1e5	15/15
	118(124)	10(11)	∞	∞	∞	∞*2.3e4*	0/15
f_{111}	6856	6.1e5	8.8e6	2.3e7	3.1e7	3.1e7	3/15
	∞	∞	∞	∞	∞	∞*2.3e4*	0/15
f_{112}	107	1684	3421	4502	5132	5596	15/15
	1.9(3)	7.7(9)	105(115)	∞	∞	∞*2.5e4*	0/15
f_{113}	133	1883	8081	24128	24128	24402	15/15
	13(16)	44(49)	∞	∞	∞	∞*2.3e4*	0/15
f_{114}	767	14720	56311	83272	83272	84949	15/15
	43(41)	∞	∞	∞	∞	∞*2.3e4*	0/15
f_{115}	64	485	1829	2550	2550	2970	15/15
	2.9(3)	14(20)	42(45)	∞	∞	∞*2.2e4*	0/15
f_{116}	5730	14472	22311	26868	30329	31661	15/15
	∞	∞	∞	∞	∞	∞*2.3e4*	0/15
f_{117}	26686	76052	1.1e5	1.4e5	1.7e5	1.9e5	15/15
	∞	∞	∞	∞	∞	∞*2.3e4*	0/15
f_{118}	429	1217	1555	1998	2430	2913	15/15
	4.3(6)	10(12)	116(129)	∞	∞	∞*2.6e4*	0/15
f_{119}	12	657	1136	10372	35296	49747	15/15
	26(40)	35(41)	∞	∞	∞	∞*2.3e4*	0/15
f_{120}	16	2900	18698	72438	3.3e5	5.5e5	15/15
	130(196)	55(65)	∞	∞	∞	∞*2.3e4*	0/15
f_{121}	8.6	111	273	1583	3870	6195	15/15
	4.8(13)	15(24)	76(82)	∞	∞	∞*2.2e4*	0/15
f_{122}	10	1727	9190	30087	53743	1.1e5	15/15
	14(23)	91(104)	∞	∞	∞	∞*2.3e4*	0/15
f_{123}	11	16066	81505	3.4e5	6.7e5	2.2e6	15/15
	65(85)	∞	∞	∞	∞	∞*2.3e4*	0/15
f_{124}	10	202	1040	20478	45337	95200	15/15
	3.0(1)	158(177)	∞	∞	∞	∞*2.2e4*	0/15
f_{125}	1	1	1	2.4e5	2.4e5	2.5e5	15/15
	3.9(6)	15(6)	6088(8822)	∞	∞	∞*2.3e4*	0/15
f_{126}	1	1	1	∞	∞	∞	0
	1.2	1053(1172)	3.5e5 (4e5)	∞	∞	∞	0/15
f_{127}	1	1	1	3.4e5	3.9e5	4.0e5	15/15
	2.5(6)	14(4)	7248(10060)	∞	∞	∞*2.2e4*	0/15
f_{128}	111	4248	7808	12447	17217	21162	15/15
	12(22)	17(20)	43(46)	∞	∞	∞*2.3e4*	0/15
f_{129}	64	10710	59443	2.8e5	5.1e5	5.8e5	15/15
	124(124)	16(16)	∞	∞	∞	∞*2.3e4*	0/15
f_{130}	55	812	3034	32823	33889	34528	10/15
	2.3(4)	11(10)	10(11)	∞	∞	∞*2.2e4*	0/15

20-D

Δf	1e+1	1e+0	1e-1	1e-3	1e-5	1e-7	#succ
f_{101}	59	425	571	700	739	783	15/15
	3.1(0.8)	0.85(0.2)	0.90(0.2)	1.1(0.2)	1.5(0.3)	1.6(0.4)	15/15
f_{102}	231	399	579	921	1157	1407	15/15
	2.9(4)	6.1(9)	6.3(7)	45(49)	∞	∞*1.0e5*	0/15
f_{103}	65	417	629	1313	1893	2464	14/15
	2.3(0.9)	1.00(0.2)	5.9(6)	1231(1406)	∞	∞*1.1e5*	0/15
f_{104}	23690	85656	1.7e5	1.8e5	1.9e5	2.0e5	15/15
	68(72)	∞	∞	∞	∞	∞*1.1e5*	0/15
f_{105}	1.9e5	6.1e5	6.3e5	6.5e5	6.6e5	6.7e5	15/15
	∞	∞	∞	∞	∞	∞*9.4e4*	0/15
f_{106}	11480	21668	23746	25470	26492	27360	15/15
	7.0(5)	31(32)	∞	∞	∞	∞*1.5e5*	0/15
f_{107}	8571	13582	16226	27357	52486	65052	15/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{108}	58063	97228	2.0e5	4.5e5	6.3e5	9.0e5	15/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{109}	333	632	1138	2287	3583	4952	15/15
	17(23)	∞	∞	∞	∞	∞*8.4e4*	0/15
f_{110}	∞	∞	∞	∞	∞	∞	0
	∞	∞	∞	∞	∞	∞	0/15
f_{111}	∞	∞	∞	∞	∞	∞	0
	∞	∞	∞	∞	∞	∞	0/15
f_{112}	25552	64124	69621	73557	76137	78238	15/15
	∞	∞	∞	∞	∞	∞*9.8e4*	0/15
f_{113}	50123	3.6e5	5.6e5	5.9e5	5.9e5	5.9e5	15/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{114}	2.1e5	1.1e6	1.4e6	1.6e6	1.6e6	1.6e6	15/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{115}	2405	30268	91749	1.3e5	1.3e5	1.3e5	15/15
	236(247)	∞	∞	∞	∞	∞*8.5e4*	0/15
f_{116}	5.0e5	6.9e5	8.9e5	1.0e6	1.1e6	1.1e6	15/15
	∞	∞	∞	∞	∞	∞*7.7e4*	0/15
f_{117}	1.8e6	2.5e6	2.6e6	2.9e6	3.2e6	3.6e6	15/15
	∞	∞	∞	∞	∞	∞*7.7e4*	0/15
f_{118}	6908	11786	17514	26342	30062	32659	15/15
	∞	∞	∞	∞	∞	∞*1.1e5*	0/15
f_{119}	2771	29365	35930	4.1e5	1.4e6	1.9e6	15/15
	398(449)	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{120}	36040	1.8e5	2.8e5	1.6e6	6.7e6	1.4e7	13/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{121}	249	769	1426	9304	34434	57404	15/15
	31(63)	∞	∞	∞	∞	∞*8.0e4*	0/15
f_{122}	692	52008	1.4e5	7.9e5	2.0e6	5.8e6	15/15
	82(114)	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{123}	1063	5.3e5	1.5e6	5.3e6	2.7e7	1.6e8	0
	174(189)	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{124}	192	1959	40840	1.3e5	3.9e5	8.0e5	15/15
	91(187)	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{125}	1	1	1	2.5e7	8.0e7	8.1e7	4/15
	1	414(426)	∞	∞	∞	∞*7.8e4*	0/15
f_{126}	1	1	1	∞	∞	∞	0
	4.2	1.3e5(1e5)	∞	∞	∞	∞	0/15
f_{127}	1	1	1	4.4e6	7.3e6	7.4e6	15/15
	3.7	253(389)	∞	∞	∞	∞*7.8e4*	0/15
f_{128}	1.4e5	1.3e7	1.7e7	1.7e7	1.7e7	1.7e7	9/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{129}	7.8e6	4.1e7	4.2e7	4.2e7	4.2e7	4.2e7	5/15
	∞	∞	∞	∞	∞	∞*7.8e4*	0/15
f_{130}	4904	93149	2.5e5	2.5e5	2.6e5	2.6e5	7/15
	9.1(9)	∞	∞	∞	∞	∞*8.0e4*	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 $f_{\mathrm{opt}} + 10^{-8}$.

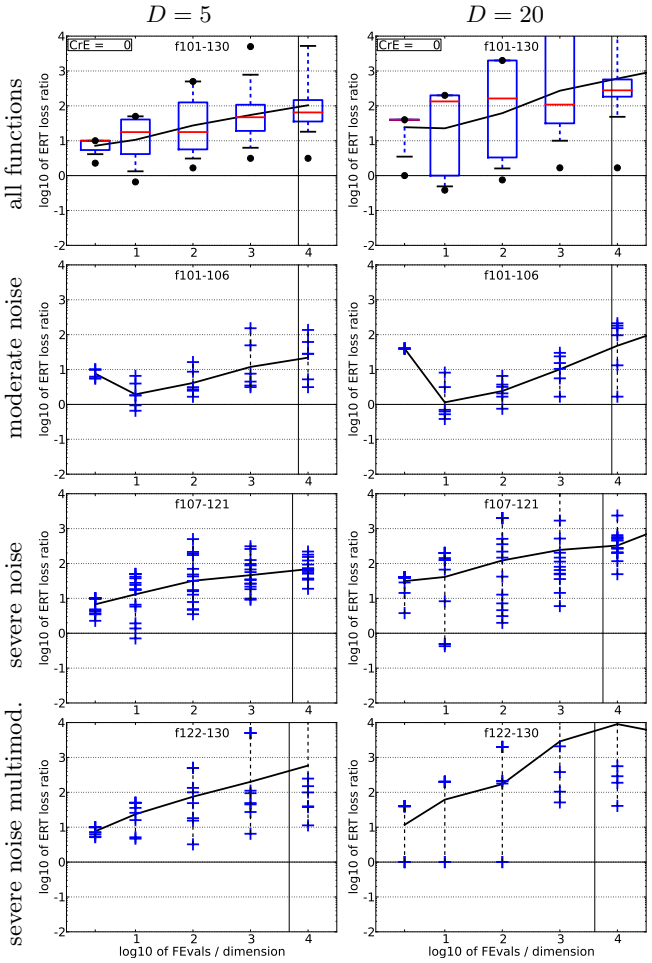

Figure 3: ERT **loss ratio vs. a given budget** FEvals**. Each cross (+) represents a single function. The target value f_t used for a given** FEvals **is the smallest (best) recorded function value such that** $\mathrm{ERT}(f_t) \leq$ FEvals **for the presented algorithm. Shown is** FEvals **divided by the respective best** $\mathrm{ERT}(f_t)$ **from BBOB-2009 for functions f_{101}–f_{130} in 5-D and 20-D. Line: geometric mean. Box-Whisker error bar: 25-75%-ile with median (box), 10-90%-ile (caps), and minimum and maximum** ERT **loss ratio (points). The vertical line gives the maximal number of function evaluations in a single trial in this function subset.**