

BBOB-Benchmarking Two Variants of the Line-Search Algorithm

Petr Pošík

Czech Technical University in Prague, Faculty of Electrical Engineering, Dept. of Cybernetics
Technická 2, 166 27 Prague 6
posik@labe.felk.cvut.cz

ABSTRACT

The restarted line search, or coordinate-wise search, algorithm is tested on the BBOB 2009 testbed. Two different univariate search algorithms (`fminbnd` from MATLAB and STEP) were tried and compared. The results are as expected: line search method can optimize only separable functions, for other functions it fails. The STEP method is slightly slower, however, is more robust in the multimodal case. The line search algorithms also identified 2 functions of the test suite (in addition to separable problems) which might be easy for algorithms exploiting separability.

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, Experimentation, Performance, Reliability

Keywords

Benchmarking, Black-box optimization, Evolutionary computation, Brent's method, `fminbnd`, STEP

1. INTRODUCTION

The line search algorithm is one of the basic and simplest optimization algorithms. In any comparison, it should serve as a baseline algorithm. It is pretty efficient for separable functions. The results of the line search algorithm should thus discover test functions which are separable or functions which are easy for algorithms exploiting separability.

It is not expected that the line search algorithm will be effective on non-separable functions. This should demonstrate the need to construct algorithms which can solve non-separable functions.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'09, July 8–12, 2009, Montréal Québec, Canada.
Copyright 2009 ACM 978-1-60558-505-5/09/07 ...\$5.00.

2. ALGORITHM DESCRIPTION

The line search heuristic is pretty simple. It starts from a randomly selected point. Then it iterates through individual directions and optimizes the function with respect to the chosen direction, keeping the other solution components fixed. After finding the optimum in one direction, it moves to the best solution found and switches to another direction. If the solution does not change after going through all the directions, a local optimum is found and the algorithm finishes.

In this paper, the multistart version of line search is considered; each launch begins in a different randomly selected initial point.

The efficiency of the line search greatly depends on the procedure employed for the univariate bounded optimization. In this paper, two univariate optimization algorithms are compared: the MATLAB function `fminbnd`, and the STEP algorithm.

2.1 MATLAB `fminbnd` function

The MATLAB `fminbnd` function (revision 1.18.4.11 was used) is based on golden-section search and parabolic interpolation. It is able to identify the optimum of quadratic functions in a few steps. On the other hand, it is a local search technique, it can miss the global optimum (of the 1D function). Since this is a rather standard ready-to-use algorithm, it will not be described here in more detail.

2.2 STEP

The acronym STEP stands for *select the easiest point*. The STEP method [4] is univariate global search algorithm based on interval division. It starts from one interval initialized with x_l and x_u , lower and upper bound of the interval, with both points evaluated. In each iteration it selects one interval and divides it to halves by sampling and evaluating the point in the middle.

The STEP algorithm selects the next interval to sample from by evaluating its difficulty, i.e. by its belief how difficult it would be to improve the best-so-far solution by sampling from the respective interval. The measure of interval difficulty chosen in STEP is the value of the coefficient a from a quadratic function $f(x) = ax^2 + bx + c$ which goes through both boundary points and somewhere on the interval reaches the value of $f_{\text{BSF}} - \epsilon$ (ϵ is a small positive number, typically from 10^{-3} to 10^{-8} , meaning an improvement of f_{BSF} by a nontrivial amount). Example of a few STEP iterations can be seen in Fig. 1.

Note that this method is also based on parabolic inter-

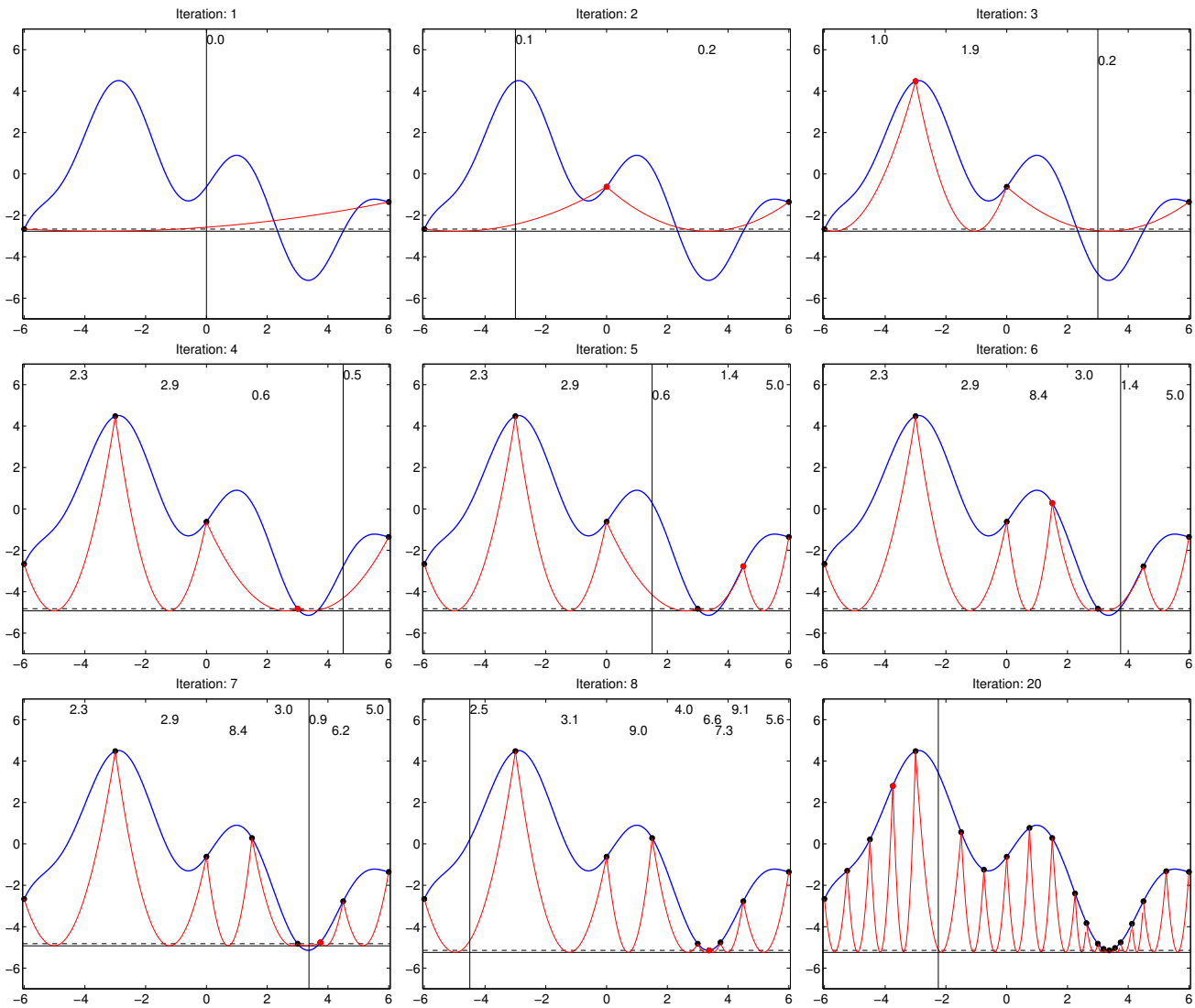


Figure 1: Demonstration of the behavior of the STEP algorithm. Blue line: the objective function. Black dots: previously sampled data points, interval boundaries. Dashed black horizontal line: f_{BSF} level. Solid black horizontal line: $f_{\text{BSF}} - \epsilon$ level. Red line: parabolic functions going through the interval boundaries and touching the objective level $f_{\text{BSF}} - \epsilon$. Red dot: last sampled data point. Vertical line: the place where next point will be sampled. Numbers: the difficulty indices of the respective intervals—the coefficients a of the respective parabolas.

polation (similarly to `fminbnd`), however, applied in a completely different manner than in `fminbnd`. In `fminbnd` it is assumed, that the objective function really has a parabolic shape and the parabola interpolates 3 points on the function. STEP does not assume that the objective function is quadratic; parabola was chosen as the measure of interval difficulty since it is the smoothest function (its curvature does not change, it is a constant value $2a$). It interpolates only 2 points on the function and is required to reach the value $f_{\text{BSF}} - \epsilon$ between the two respective points.

2.3 Parameter Settings

The `fminbnd` function does not have any parameters (except the search space bounds and termination criteria, both

are described later). The STEP algorithm has 2 parameters: the Jones factor $\epsilon = 10^{-8}$ and the maximum interval difficulty set to 10^7 (value determined by experimenting with the Rastrigin function).

2.4 Box Constraints? Yes

Both algorithms were ordered to find the optimum in the hypercube $\langle -6, 6 \rangle^D$. Why this choice if the benchmark functions have the global optimum in the interval $\langle -5, 5 \rangle^D$? Because of the behaviour of the `fminbnd` function. As written in the MATLAB documentation, the `fminbnd` function almost never samples the boundaries of the interval which would make it extremely difficult for the algorithm to find the solution e.g. for the Linear slope function.

2.5 The Crafting Effort

For both algorithms the same setting is used for all problems, so that $\text{CrE} = 0$.

2.6 Invariance Properties

The algorithm searches in axis-parallel directions and thus is *not invariant with respect to rotation*. Both algorithms, `fminbnd` and STEP, make use of the fitness values when deciding where to sample a new point, thus are *not invariant with respect to order-preserving transformations of the fitness function*.

3. EXPERIMENTAL PROCEDURE

The standard experimental procedure of BBOB was used: the algorithms were run on 24 benchmark functions, 5 instances each, 3 runs per instance. Each multistart run was finished

- after finding a solution with fitness difference $\Delta f \leq 10^{-8}$, or
- after performing more than $10^4 \times D$ function evaluations.

Each individual launch of the basic line search algorithm was interrupted (and possibly restarted)

- if any of the above mentioned criteria was satisfied, or
- if two consecutive cycles over all directions ended up with solutions which are closer than 10^{-10} .

The individual univariate searches were stopped

- in case of `fminbnd`, when the target fitness value was reached, or the maximum allowed number of function evaluations was reached (100), or when the boundary points of the interval got closer than 10^{-10} , and
- in case of STEP, when the target fitness value was reached, or when maximum allowed number of function evaluations was reached (1000); moreover an interval was not used for further sampling if its boundary points were closer than 10^{-10} , or when the difficulty of the interval was higher than 10^7 .

4. RESULTS

Results from experiments according to [2] on the benchmark functions given in [1, 3] are presented in Figures 2, 3 and 4 and in Tables 1 and 2.

In Fig. 2, we show only those functions which reveal some interesting differences between the two methods. On functions which are not shown, both functions performed similarly (similarly badly).

5. CPU TIMING EXPERIMENT

The multistart algorithm was run with the maximal number of evaluations set to 10^5 , the basic algorithm was restarted for at least 30 seconds. The results for `fminbnd` were between $3.7 \cdot 10^{-4}$ and $3.9 \cdot 10^{-4}$ seconds per function evaluation for all the tested search space dimensionalities (2–40). For STEP line search, the results were between $5.1 \cdot 10^{-4}$ and $5.9 \cdot 10^{-4}$. The experiment was conducted on Intel Core 2 CPU, T5600, 1.83 GHz, 1 GB RAM with Windows XP SP3 in MATLAB R2007b.

6. CONCLUSIONS

The results for line search method are expectable: line search is unable to optimize non-separable functions. Line search method with the `fminbnd` univariate optimizer is very fast when solving quadratic separable functions (Sphere, Ellipsoid), but is unable to solve the Rastrigin function (due to multimodality the `fminbnd` gets trapped in local optimum). Line search with STEP is effective method for optimization of separable functions. It is slower (about 10 times) than `fminbnd` when optimizing univariate functions (it spends some evaluations on global search), on the other hand it is able to solve the Rastrigin and the Skew Rastrigin-Buche function (since it is a univariate global search method).

On the other hand, the line search method discovered solutions of the two Gallagher functions (or at least solutions of some of their instances) which might be an indication that these functions may be easier for algorithms that exploit separability of the problem.

Acknowledgements

The project was supported by the Ministry of Education, Youth and Sport of the Czech Republic with the grant No. MSM6840770012 entitled “Transdisciplinary Research in Biomedical Engineering II”.

7. 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.
- [2] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2009: Experimental setup. Technical Report RR-6828, INRIA, 2009.
- [3] 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.
- [4] S. Swarzberg, G. Seront, and H. Bersini. S.T.E.P.: The Easiest Way to Optimize a Function. In *IEEE World Congress on Computational Intelligence., Proceedings of the First IEEE Conference on Evolutionary Computation*, pages 519–524 vol.1, 1994.

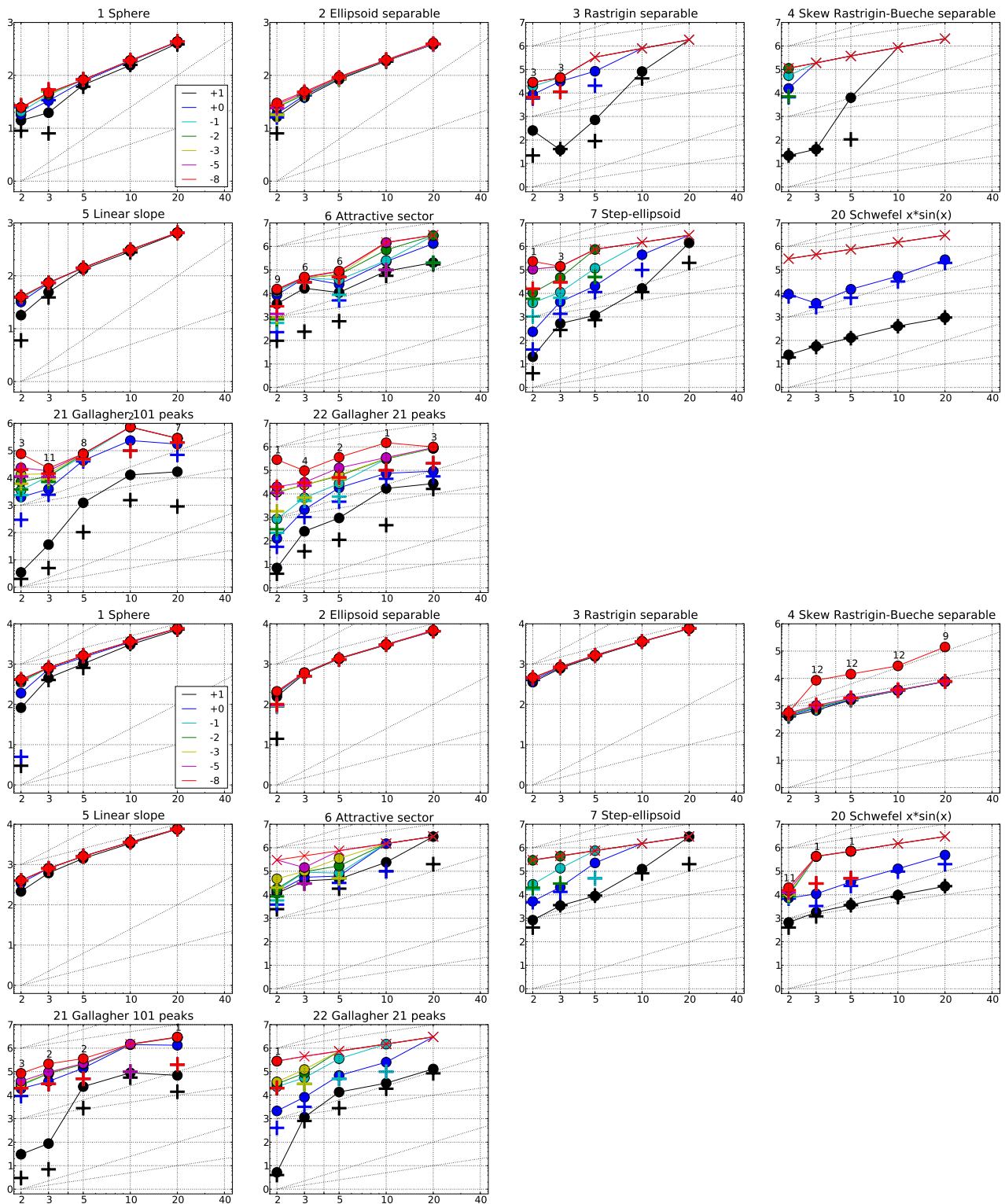


Figure 2: MATLAB `fminbnd` (upper 10 graphs) and STEP (lower 10 graphs): Expected Running Time (ERT, ●) to reach $f_{\text{opt}} + \Delta f$ and median number of function evaluations of successful trials (+), shown for $\Delta f = 10, 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-5}, 10^{-8}$ (the exponent is given in the legend of f_1 and f_{24}) versus dimension in log-log presentation. The $\text{ERT}(\Delta f)$ equals to $\#FEs(\Delta f)$ divided by the number of successful trials, where a trial is successful if $f_{\text{opt}} + \Delta f$ was surpassed during the trial. The $\#FEs(\Delta f)$ are the total number of function evaluations while $f_{\text{opt}} + \Delta f$ was not surpassed during the trial from all respective trials (successful and unsuccessful), and f_{opt} denotes the optimal function value. Crosses (×) indicate the total number of function evaluations $\#FEs(-\infty)$. Numbers above ERT-symbols indicate the number of successful trials. Annotated numbers on the ordinate are decimal logarithms. Additional grid lines show linear and quadratic scaling.

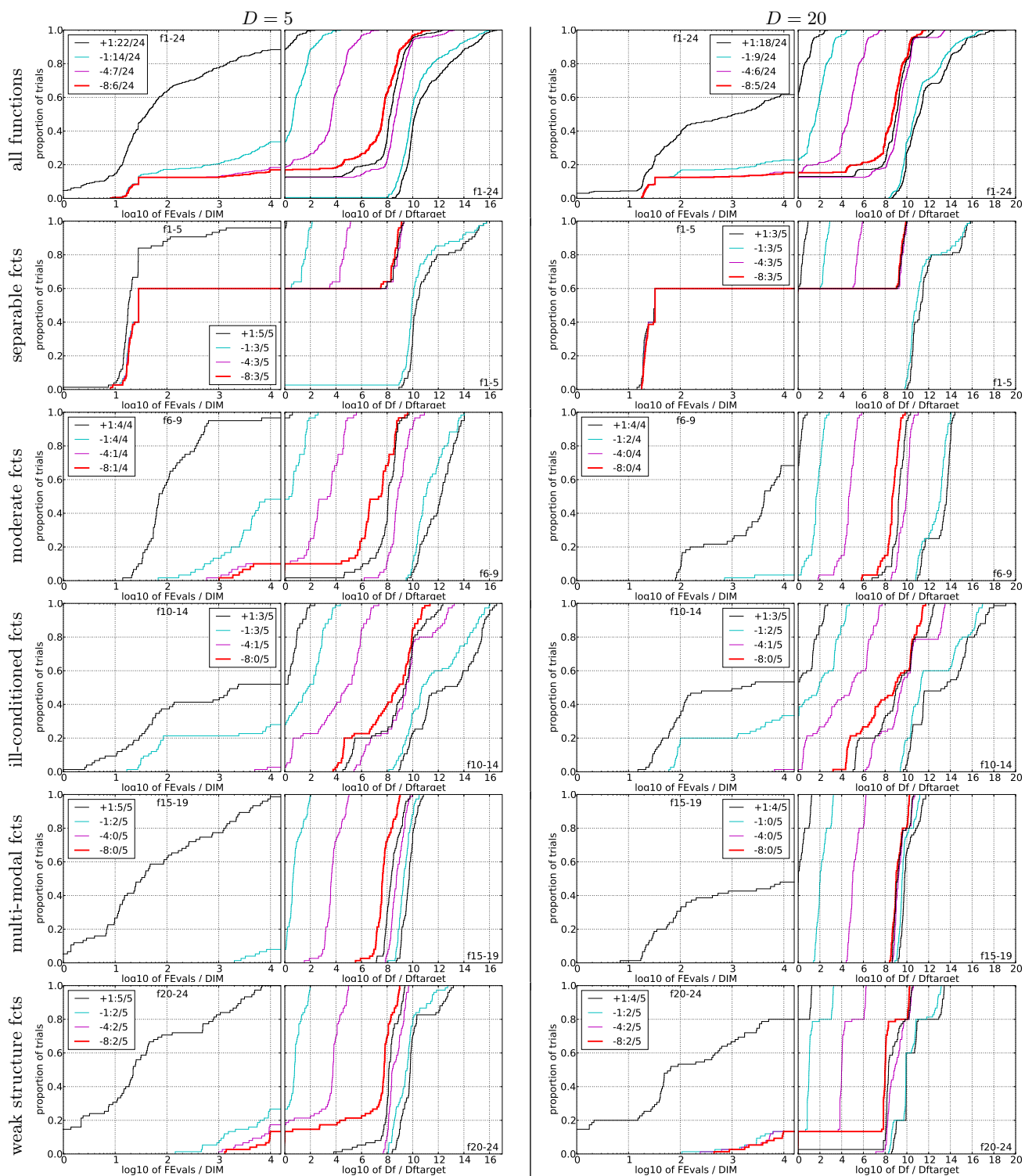


Figure 3: fminbnd Line Search: Empirical cumulative distribution functions (ECDFs), plotting the fraction of trials versus running time (left) or Δf . Left subplots: ECDF of the running time (number of function evaluations), divided by search space dimension D , to fall below $f_{\text{opt}} + \Delta f$ 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 $D, 10D, 100D \dots$ function evaluations (from right to left cycling black-cyan-magenta). Top row: all results from all functions; second row: separable functions; third row: misc. moderate functions; fourth row: ill-conditioned functions; fifth row: multi-modal functions with adequate structure; sixth row: multi-modal functions with weak structure. 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.

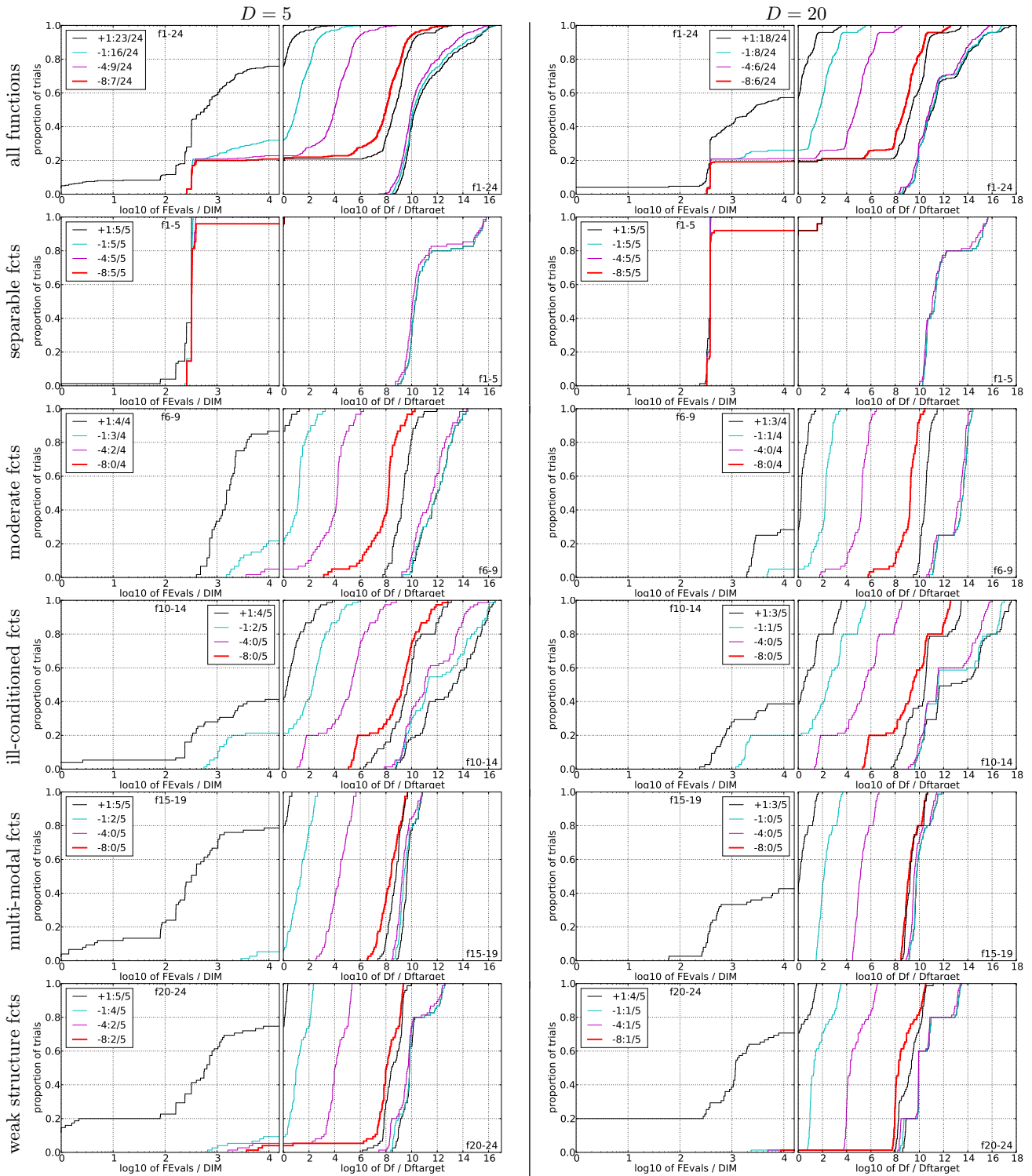


Figure 4: STEP Line Search: Empirical cumulative distribution functions (ECDFs), plotting the fraction of trials versus running time (left) or Δf . Left subplots: ECDF of the running time (number of function evaluations), divided by search space dimension D , to fall below $f_{\text{opt}} + \Delta f$ 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 $D, 10D, 100D \dots$ function evaluations (from right to left cycling black-cyan-magenta). Top row: all results from all functions; second row: separable functions; third row: misc. moderate functions; fourth row: ill-conditioned functions; fifth row: multi-modal functions with adequate structure; sixth row: multi-modal functions with weak structure. 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.