

MEMPSODE: An Empirical Assessment of Local Search Algorithm Impact on a Memetic Algorithm Using Noiseless Testbed

Costas Voglis^{*}
Department of
Computer Science
University of Ioannina
GR-45110 Ioannina, Greece
voglis@cs.uoi.gr

Grigoris S. Piperagkas
Department of
Computer Science
University of Ioannina
GR-45110 Ioannina, Greece
gpiperag@cs.uoi.gr

Konstantinos E. Parsopoulos
Department of
Computer Science
University of Ioannina
GR-45110 Ioannina, Greece
kostasp@cs.uoi.gr

Dimitris G. Papageorgiou
Department of Materials Science
and Engineering
University of Ioannina
GR-45110 Ioannina, Greece
dpapageo@cc.uoi.gr

Isaac E. Lagaris
Department of
Computer Science
University of Ioannina
GR-45110 Ioannina, Greece
lagaris@cs.uoi.gr

ABSTRACT

Memetic algorithms are hybrid schemes that usually integrate metaheuristics with classical local search techniques, in order to attain more balanced intensification/diversification trade-off in the search procedure. MEMPSODE is a recently published software that implements such memetic schemes, based on the Particle Swarm Optimization and Differential Evolution algorithms, as well as on the Merlin optimization environment that offers a variety of local search methods. The present study aims at investigating the impact of the selected local search algorithm in the memetic schemes produced by MEMPSODE. Our interest was focused on gradient-free local search methods. We applied the derived memetic schemes on the noiseless testbed of the Black-Box Optimization Benchmarking 2012 workshop. The obtained results can offer significant insight to optimization practitioners with respect to the most promising approaches.

Categories and Subject Descriptors

G.1.6 [Numerical Analysis]: Optimization—*global optimization*; G.4 [Mathematical Software]

Keywords

Global optimization, memetic algorithms, hybrid algorithms, Black-box optimization, local search

^{*}Corresponding author.

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'12 Companion, July 7–11, 2012, Philadelphia, PA, USA.
Copyright 2012 ACM 978-1-4503-1178-6/12/07 ...\$10.00.

1. INTRODUCTION

Memetic Algorithms (MAs) are the outcome of integrating metaheuristics with local search (LS) procedures. The combination of the two approaches is accompanied by increased effectiveness and accuracy for locating solutions of global optimization problems [5, 11].

The most common MA scheme assumes an evolving population and periodically applies LS to some or all its members. The dynamic of the MA is dictated by the choices of the user with respect to the following issues, also referred to as the *fundamental memetic questions* [11]:

- (I) *Where* will the LS procedures be applied. The user shall define the individuals that will constitute initial points for the LS.
- (II) *When* will the LS procedures be applied. The application frequency of LS shall be determined in order to attain a balanced exploitation of the available computational budget.
- (III) *How much* of the computational budget will be consumed at each application of LS. The user must specify the fraction of the available computational budget that will be devoted to LS.

These issues were studied in [12], under a memetic strategy based on the Particle Swarm Optimization (PSO) algorithm.

MEMPSODE (MEMetic Particle Swarm Optimization and Differential Evolution) is a recently introduced [16] software package that implements closely the PSO-based memetic approaches of [12] and extends them also to the Differential Evolution (DE) framework [14]. More specifically, MEMPSODE contains implementations of the Unified PSO (UPSO) approach [10], which generalizes plain PSO by harnessing the strengths of its standard local and global variants, as well as the five basic DE operators. Also, MEMPSODE borrows LS methods from the Merlin optimization environment [9], which offers some of the most fundamental gradient-based and gradient-free LS methods.

Evidently, the choice of the LS scheme in an MA is expected to have an impact on its performance. Attempts to quantify this impact in hybrid evolutionary algorithms, were made in previous works. For instance, in [6] a hill-climbing method [13] was compared to the nonlinear simplex method [7] and CMA-ES [4]. Moreover, in [2] a hybrid PSO was studied under the influence of the nonlinear simplex method [7] and Powell’s direction set method [8].

The present paper constitutes a first attempt to investigate the impact of the LS method on the MAs implemented in MEMPSODE. For this reason, we fixed the parameters of all algorithms to specific values and alternated the LS procedures combined with PSO and DE. The resulting memetic schemes were tested on the noiseless testbed of the Black-Box Optimization Benchmarking 2012 (BBOB’12) workshop.

The rest of the paper is organized as follows: the employed algorithms are given in Section 2, while Section 3 describes the experimental setting. The obtained experimental results are reported in Section 5, and the paper concludes in Section 6.

2. EMPLOYED ALGORITHMS

The design of the MA schemes presented in [12], were based on PSO. In this context, the following three schemes were proposed:

- Scheme 1:** LS is applied only on the overall best position, p_g , of the swarm.
- Scheme 2:** LS is applied on each locally best position, p_i , $i = 1, 2, \dots, N$, with a prescribed fixed probability, $\rho \in (0, 1]$.
- Scheme 3:** LS is applied both on the best position, p_g , as well as on some randomly selected locally best positions, p_i , $i \in \{1, 2, \dots, N\}$.

These schemes can be applied either at each iteration or whenever a specific number of consecutive iterations has been completed.

In MEMPSODE [16] we extended this strategy to the DE [14] framework and incorporated the Merlin optimization environment for the LS procedures. A detailed presentation of the MEMPSODE software is provided in [16] as well as in the accompanying manual of the software distribution.

A short description of the LS algorithms employed in the present work, are given in the following paragraphs. The notation used for the algorithms’ names, follows the corresponding MEMPSODE (Merlin) standard.

BFGS Method

BFGS belongs to the category of quasi-Newton methods with line search [8]. At the start of k -th iteration a point $\mathbf{x}^{(k)}$, the gradient (or an finite-difference approximation) $\mathbf{g}^{(k)}$ and an approximation $\mathbf{B}^{(k)}$ of the Hessian matrix are available. The method continues by performing a line search to a direction $\mathbf{s}^{(k)}$ defined as $\mathbf{B}^{(k)}\mathbf{s}^{(k)} = -\mathbf{g}^{(k)}$, and an update from $\mathbf{B}^{(k)}$ to $\mathbf{B}^{(k+1)}$, using the BFGS update formula [1]. A careful selection of line search, in addition to the powerful local properties of the Hessian approximation, results in one of the most robust and effective optimization algorithms.

TRUST Method

The TRUST method is a quasi-Newton algorithm that uses also BFGS updates to maintain an approximation of the Hessian, but for the step of the line search is properly selected to guarantee convergence, following the trust region strategy [8]. In this strategy, a quadratic model is being build around the current point using the approximation $\mathbf{B}^{(k)}$ and the gradient vector $\mathbf{g}^{(k)}$. The algorithm *trusts* only the quadratic model in a restricted region around the current iterate. The restriction is usually imposed by a radius that defines a hypersphere.

SIMPLEX Method

This method belongs to the class of direct search methods for nonlinear optimization. It was designed by Nelder and Mead [7]. The algorithm is based on the concept of *simplex* (or polytope) in \mathbb{R}^n , which is a volume element defined by $(n + 1)$ vertices. The input of the algorithm is an initial simplex. The SIMPLEX algorithm then moves this initial simplex towards the minimum by adapting its geometry and finally shrinks it to a small volume element around the minimizer. The procedure is derivative-free and proceeds towards the minimum using a set of $n + 1$ points. Thus, it is expected to be tolerant to ill-conditioned cases. The transformation rules include *contraction*, *expansion* and *shrinkage* of the initial simplex.

ROLL Method

This method belongs to the class of pattern search methods. It proceeds by taking proper steps along each coordinate direction, in turn. Then, the method performs an one-dimensional search on a properly formed direction, in order to tackle possible correlations among the variables. Its input consists of the initial point and a user-defined exploration factor.

AUTO Method

AUTO is a procedure that tries to automatically select the best LS algorithm. The methods BFGS, ROLL, SIMPLEX and TRUST are invoked one after the other. For each one, a rate is calculated by dividing the relative achieved reduction of the function’s value by the number of function calls spent. The method with the highest rate is then invoked again and the procedure is repeated. If all rates assume vanishing values, then all the method tolerances are set to zero and the methods are applied in the following order: ROLL, TRUST, BFGS, SIMPLEX.

3. EXPERIMENTAL SETUP

We used the default restart mechanism provided by the noiseless testbed for a maximum number of $10^5 \times n$ function evaluations. The memetic Scheme 3 was used with LS probability $\rho_i = 0.05$. We selected the default DE operator as the main metaheuristic of the MA, with population size $N = 25$ and parameter values $F = 0.5$ and $CR = 0.7$.

Each LS application was restricted to 2000 function evaluations. Whenever first order derivatives were needed (e.g., in BFGS) we applied an $O(h)$ finite-difference formula where h is an adaptable step size [15]. The experiments were conducted on an Intel I7-2600 3.4 GHz processor machine with 8GB RAM.

| Algorithm | Dimension | | | | | |
|-----------|------------|------------|------------|------------|------------|------------|
| | 2 | 3 | 5 | 10 | 20 | 40 |
| BFGS | 3.1^{-4} | 2.0^{-4} | 1.2^{-4} | 5.0^{-5} | 2.7^{-5} | 2.2^{-5} |
| ROLL | 4.6^{-5} | 5.4^{-5} | 5.7^{-5} | 5.6^{-5} | 4.5^{-5} | 3.8^{-5} |
| SIMPLEX | 2.1^{-4} | 1.5^{-4} | 9.2^{-5} | 5.1^{-5} | 5.5^{-5} | 5.7^{-5} |
| AUTO | 1.5^{-4} | 1.1^{-4} | 6.5^{-5} | 1.5^{-5} | 1.3^{-5} | 1.1^{-5} |

Table 1: Summary of timing results in seconds per function evaluation.

4. CPU TIMING EXPERIMENT

According to [3], for the timing experiment the experimental setting described above was run on f8 with at most 10^3 function evaluations for each call of MEMPSODE, and restarted until at least 30 seconds had passed. The timing experiment was performed on the same platform as the experimental procedure. The timing results using the four LS algorithms are reported in Table 1.

5. EXPERIMENTAL RESULTS

We compared the performance of the memetic DE scheme with the four LS strategies implemented in MEMPSODE. The results are reported in Tables 2 and 3 and graphically illustrated in Figs. 1–3. The AUTO method proved to be superior in robustness, solving in total 1480 of 2160 experiments up to the desired accuracy, 10^{-8} . BFGS was capable of approaching the minimizer using a small number of function evaluations, although it was successful in a smaller number of experiments (1307 out of 2160).

Figures 2 and 3 reveal that in the 5–dimensional experiments, especially for the separable, ill–conditioned and moderate functions, BFGS and AUTO had the best performance in terms of function evaluations. The same holds also for the 20–dimensional case, favoring the aforementioned methods. Considering the SIMPLEX method, it exhibited improved performance only for the small–dimensional cases.

A closer examination of Fig. 1 and Tables 2 and 3, with respect to the ERT values, suggests that the ROLL method performs nicely in the separable functions, although it cannot dominate in all cases. The BFGS performance is improved in f8–f12 while, in the rest of the cases, the results are not conclusive in terms of ERT. Overall, the results suggest that the algorithms implemented in MEMPSODE can be very competitive.

6. CONCLUSION

We presented a comparison among the LS schemes implemented in the MEMPSODE software, within a memetic DE framework. Among the four tested LS methods, BFGS proved to be more efficient in terms of accuracy and function evaluations. On the other hand the AUTO versatile strategy was very robust in solving most of the moderate and ill–conditioned problems to the prescribed accuracy. Since AUTO is a combination of local search strategies (including BFGS), we may conclude that a portfolio of LS approaches may be a valuable addition to hybrid memetic schemes. Further research will consider also the PSO metaheuristic offered by MEMPSODE, as well as different parameter settings.

7. REFERENCES

- [1] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, 1970.
- [2] J. Gimmler, T. Stützle, and T. Exner. Hybrid particle swarm optimization: An examination of the influence of iterative improvement algorithms on performance. *Ant Colony Optimization and Swarm Intelligence*, pages 436–443, 2006.
- [3] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2012: Experimental setup. Technical report, INRIA, 2012.
- [4] N. Hansen and A. Ostermeier. Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*, pages 312–317. IEEE, 1996.
- [5] J. Kennedy and R. C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.
- [6] D. Molina, M. Lozano, C. García-Martínez, and F. Herrera. Memetic algorithms for continuous optimisation based on local search chains. *Evolutionary Computation*, 18(1):27–63, 2010.
- [7] J. Nelder and R. Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.
- [8] J. Nocedal and S. Wright. *Numerical optimization*. Springer verlag, 1999.
- [9] D. Papageorgiou, I. Demetropoulos, and I. Lagaris. MERLIN-3.1. 1. A new version of the Merlin optimization environment. *Computer Physics Communications*, 159(1):70–71, 2004.
- [10] K. E. Parsopoulos and M. N. Vrahatis. Parameter selection and adaptation in unified particle swarm optimization. *Mathematical and Computer Modelling*, 46(1–2):198–213, 2007.
- [11] K. E. Parsopoulos and M. N. Vrahatis. *Particle Swarm Optimization and Intelligence: Advances and Applications*. Information Science Publishing (IGI Global), 2010.
- [12] Y. G. Petalas, K. E. Parsopoulos, and M. N. Vrahatis. Memetic particle swarm optimization. *Annals of Operations Research*, 156(1):99–127, 2007.
- [13] F. Solis. Minimization by random search techniques. *Mathematics of operations research*, pages 19–30, 1981.
- [14] R. Storn and K. Price. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optimization*, 11:341–359, 1997.
- [15] C. Voglis, P. Hadjidoukas, I. Lagaris, and D. Papageorgiou. A numerical differentiation library exploiting parallel architectures. *Computer Physics Communications*, 180(8):1404–1415, 2009.
- [16] C. Voglis, K. Parsopoulos, D. Papageorgiou, I. Lagaris, and M. Vrahatis. Mempsode: A global optimization software based on hybridization of population-based algorithms and local searches. *Computer Physics Communications*, 183(5):1139–1154, 2012.

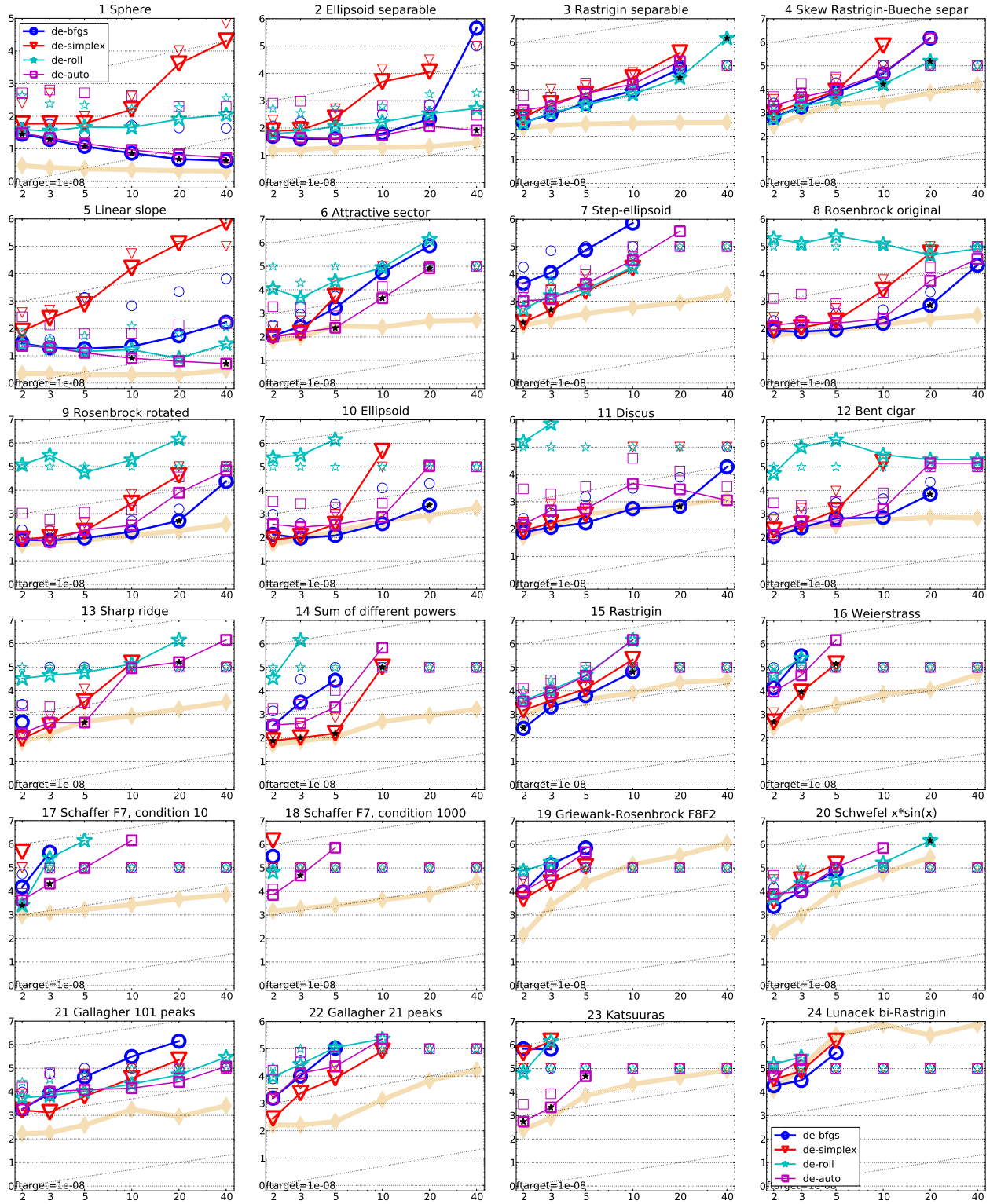


Figure 1: Expected running time (ERT in number of f -evaluations) divided by dimension for target function value 10^{-8} as \log_{10} values versus dimension. Different symbols correspond to different algorithms given in the legend of f_1 and f_{24} . Light symbols give the maximum number of function evaluations from the longest trial divided by dimension. Horizontal lines give linear scaling, slanted dotted lines give quadratic scaling. Black stars indicate statistically better performance compared to all other algorithms with $p < 0.01$ and Bonferroni correction number of dimensions (six). Legend: \circ : bfgs, ∇ : simplex, \star : roll, \square : auto.

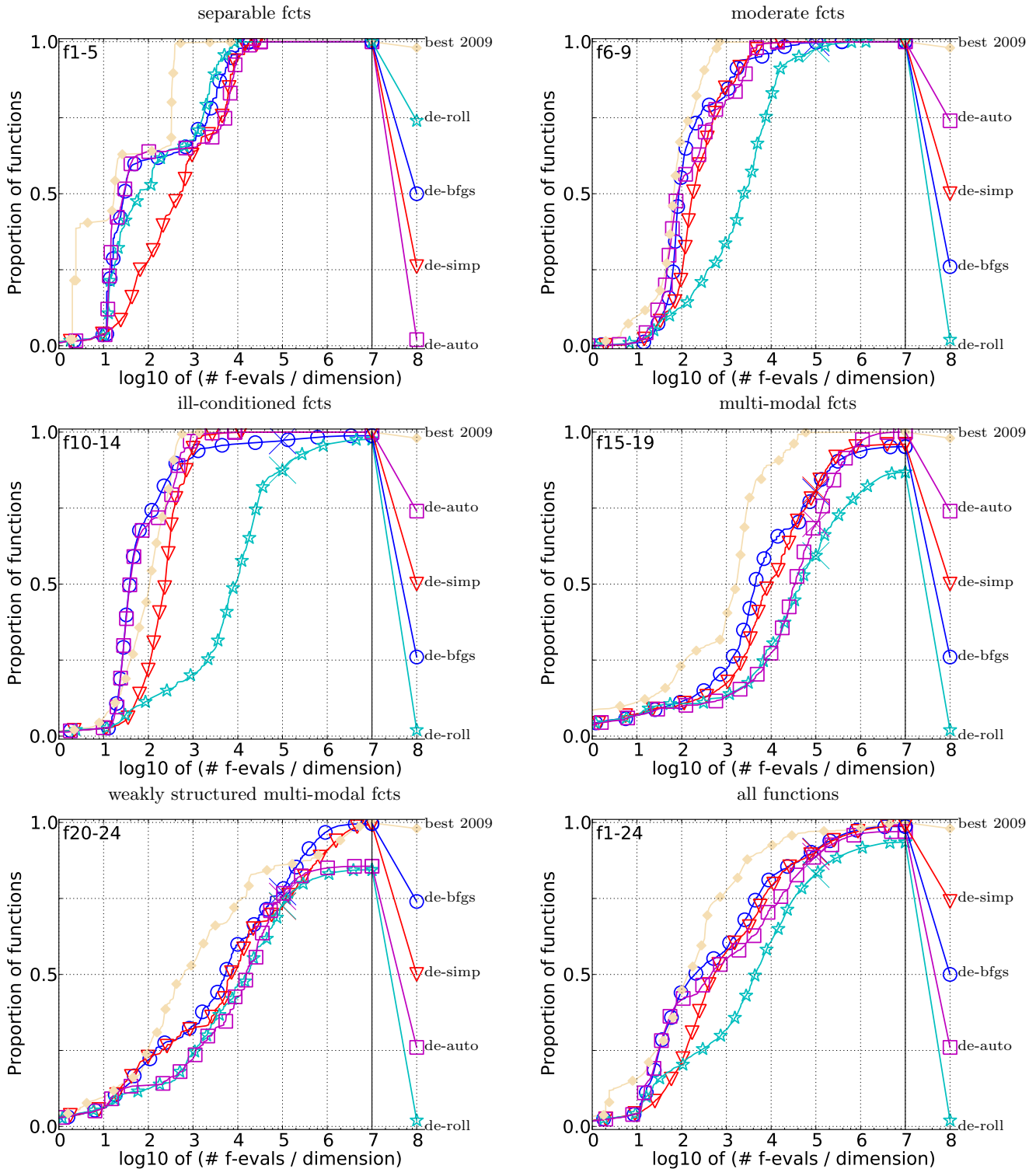


Figure 2: Bootstrapped empirical cumulative distribution of the number of objective function evaluations divided by dimension (FEvals/D) for 50 targets in $10^{[-8..2]}$ for all functions and subgroups in 5-D. The “best 2009” line corresponds to the best ERT observed during BBOB 2009 for each single target.

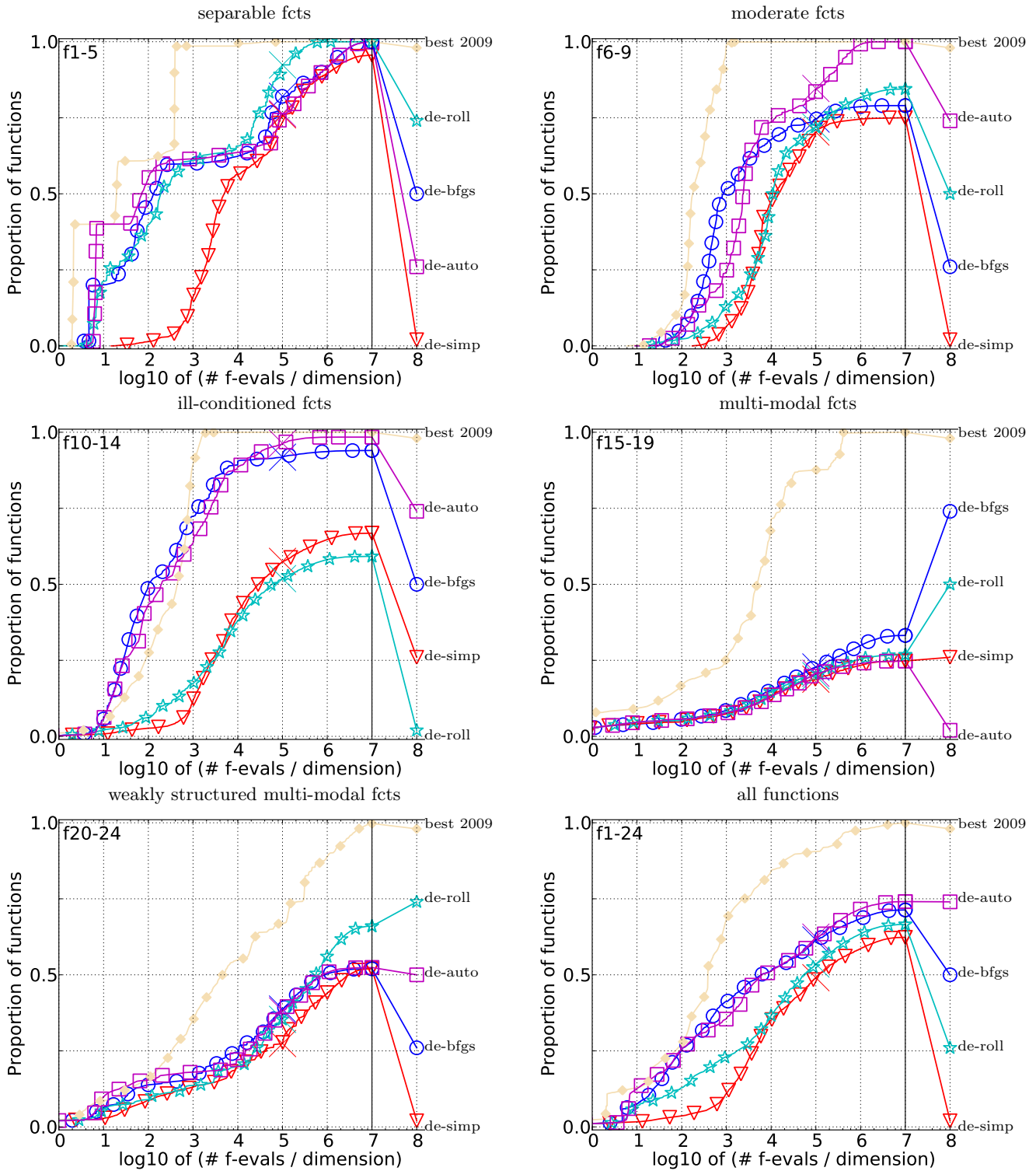


Figure 3: Bootstrapped empirical cumulative distribution of the number of objective function evaluations divided by dimension (FEvals/D) for 50 targets in $10^{[-8..2]}$ for all functions and subgroups in 20-D. The “best 2009” line corresponds to the best ERT observed during BBOB 2009 for each single target.

