

MEMPSODE: Comparing Particle Swarm Optimization and Differential Evolution Within a Hybrid Memetic Global Optimization Framework

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

MEMPSODE is a recently published optimization software that implements memetic Particle Swarm Optimization and Differential Evolution approaches. It combines previously proposed variants of the two algorithms, with the Merlin optimization environment, which includes a variety of established local search methods for continuous optimization. The present study aims at comparing the performance of the memetic variants produced by the two metaheuristics within the framework of MEMPSODE. The algorithms are assessed on the noiseless testbed of the Black-Box Optimization Benchmarking 2012 workshop, providing useful insight regarding their relative efficiency and effectiveness.

Categories and Subject Descriptors

G.1.6 [Numerical Analysis]: Optimization—*global optimization, memetic algorithms, particle swarm optimization, differential evolution, local search*; G.4 [Mathematical Software]

Keywords

Memetic global optimization, Particle swarm optimization, Differential evolution, Black-box optimization, Merlin optimization environment

^{*}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

Particle Swarm Optimization (PSO) and Differential Evolution (DE) have been established as promising optimization tools for solving continuous global optimization problems [2, 8, 13]. They are essentially based on models that draw inspiration from physical systems, exhibiting remarkable capability of locating global solutions in various optimization tasks. However, in many cases, the exploration capability of these algorithms is not accompanied by proportional exploitation properties, i.e., their solution fine-tuning capability is inferior than their global search quality.

This weakness motivated the development of a closely related category of algorithms, namely *Memetic Algorithms* (MAs). MAs constitute a class of hybrid metaheuristics that combine population-based optimization algorithms with established local search procedures [9]. Recently, the MEMPSODE (MEMetic Particle Swarm Optimization and Differential Evolution) software was published [17]. MEMPSODE implements memetic schemes of a generalized PSO variant, namely Unified PSO (UPSO) [12], as well as DE. It combines the memetic schemes proposed in [14] originally for PSO, with the algorithmic local search artillery of the established Merlin optimization environment [10].

The present paper aims at comparing the performance of memetic variants implemented within the framework of MEMPSODE, on the noiseless testbed of the Black-Box Optimization Benchmarking 2012 (BBOB'12) workshop. The primary target is to determine the impact of the choice between UPSO and DE, within the memetic framework presented in [14]. No extensive benchmarking of all the available MEMPSODE variants was possible. Instead, we used a default set of parameters for both UPSO and DE, as well as the same local search procedure, in order to gain insight regarding the most promising uncalibrated approaches offered by MEMPSODE.

The rest of the paper is organized as follows: Section 2 offers brief descriptions of all the employed algorithms, while

Section 3 explains the experimental setting. The obtained experimental results are reported in Section 5, and the paper concludes in Section 6.

2. EMPLOYED ALGORITHMS

As previously mentioned, the MEMPSODE software follows closely the PSO-based memetic schemes proposed in [14], and extends them also to the DE framework. In the following paragraphs, we provide brief descriptions of the algorithms integrated in MEMPSODE, along with a pseudocode that summarizes the corresponding memetic framework.

2.1 Unified Particle Swarm Optimization

The original PSO algorithm was first introduced by Eberhart and Kennedy [1]. UPSO was introduced later as a generalization of PSO that harnesses its local and global variant [12]. If the n -dimensional continuous optimization problem:

$$\min_{x \in X \subset \mathbb{R}^n} f(x), \quad (1)$$

is under investigation, with the search space X being defined as $X \equiv [l_1, r_1] \times [l_2, r_2] \times \dots \times [l_n, r_n]$, then PSO employs a set, called a *swarm*, of N search agents, called the *particles*, $S = \{x_1, x_2, \dots, x_N\}$, to probe X . The i -th particle is defined as $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^\top \in X$, $i \in I = \{1, 2, \dots, N\}$, and moves in X by assuming an adaptable *velocity* (position shift), $v_i = (v_{i1}, v_{i2}, \dots, v_{in})^\top$, $i \in I$, while retaining in memory the *best position* it has ever visited, $p_i = (p_{i1}, p_{i2}, \dots, p_{in})^\top \in X$, $i \in I$.

Also, each particle assumes a *neighborhood* of particles that share information with it. The neighborhood, $\mathcal{N}_i$, is usually defined as a set of indices of the communicating particles, while the shared information is the best position they have ever detected. Henceforth, $g_i \in \mathcal{N}_i$ will denote the index of the particle with the best detected position in the neighborhood of x_i .

If $\mathcal{N}_i \equiv I$, then the whole swarm is considered as the neighborhood of each particle, defining the *global* (also denoted as *gbest*) PSO variant, while $\mathcal{N}_i \subset I$ defines *local* (also denoted as *lbest*) variants. Obviously, in the gbest model it holds that $g_i = g$, $\forall i \in I$, where g is the index of the particle with the overall best position ever detected.

The classical PSO model can be generalized in the UPSO scheme [12]. If we assume that $G_i^{(t+1)}$ and $L_i^{(t+1)}$ denote the velocity update of x_i in the gbest and lbest PSO model, respectively, and t denotes the iteration counter, it holds that [12]:

$$\begin{aligned} G_{ij}^{(t+1)} &= \chi \left[v_{ij}^{(t)} + c_1 \mathcal{R}_1 \left(p_{ij}^{(t)} - x_{ij}^{(t)} \right) + c_2 \mathcal{R}_2 \left(p_{gj}^{(t)} - x_{ij}^{(t)} \right) \right], \\ L_{ij}^{(t+1)} &= \chi \left[v_{ij}^{(t)} + c_1 \mathcal{R}_1 \left(p_{ij}^{(t)} - x_{ij}^{(t)} \right) + c_2 \mathcal{R}_2 \left(p_{gij}^{(t)} - x_{ij}^{(t)} \right) \right], \end{aligned}$$

where $i \in I$; $j = 1, 2, \dots, n$; g is the index of the overall best particle; and $\mathcal{R}_1, \mathcal{R}_2$, are random variables uniformly distributed in $[0, 1]$. Then, the particles are updated as follows [11]:

$$v_{ij}^{(t+1)} = u G_{ij}^{(t+1)} + (1 - u) L_{ij}^{(t+1)}, \quad (2)$$

$$x_{ij}^{(t+1)} = x_{ij}^{(t)} + v_{ij}^{(t+1)}. \quad (3)$$

The parameter $u \in [0, 1]$ is called the *unification factor* and balances the influence (trade-off) of the gbest and lbest velocity update. Obviously, the lbest PSO model is retrieved

for $u = 0$, while for $u = 1$ the gbest PSO model is obtained. All intermediate values produce combinations with diverse convergence properties. The presented UPSO variant is implemented in the MEMPSODE software [17] that was employed in our study.

2.2 Differential Evolution

DE was introduced by Storn and Price [15] as a population-based stochastic optimization algorithm for numerical optimization problems. DE is formulated similarly to PSO. A population, $P = \{x_1, x_2, \dots, x_N\}$, of N individuals is utilized to probe the search space, $X \subset \mathbb{R}^n$. The population is randomly initialized, usually following a uniform distribution within the search space.

Each individual is an n -dimensional vector, $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^\top \in X$, $i \in I = \{1, 2, \dots, N\}$, serving as a candidate solution of the problem at hand. The population is iteratively evolved by applying two operators, *mutation* and *recombination*, on each individual to produce new candidate solutions. Then, a selection takes place, using both new and old individuals, to construct the new population consisting of the N best individuals. The procedure continues in the same manner until a termination criterion is satisfied.

The mutation operator produces a new vector, v_i , for each individual, x_i , $i \in I$, by combining some of the rest of the individuals. There are various DE mutation operators. The most common one is defined as follows:

$$v_i^{(t+1)} = x_g^{(t)} + F \left(x_{r_1}^{(t)} - x_{r_2}^{(t)} \right), \quad (4)$$

where t denotes the iteration counter; $F \in (0, 1]$ is a fixed user-defined parameter; g denotes the index of the best individual in the population, i.e., the one with the lowest function value; and $r_1, r_2 \in I$, are mutually different randomly selected indices that differ also from the index i .

After the mutation, a recombination operator is applied producing a trial vector:

$$u_i = (u_{i1}, u_{i2}, \dots, u_{in}), \quad i \in I,$$

for each individual. This vector is defined as follows:

$$u_{ij}^{(t+1)} = \begin{cases} v_{ij}^{(t+1)}, & \text{if } R_j \leq CR \text{ or } j = \text{RI}(i), \\ x_{ij}^{(t)}, & \text{if } R_j > CR \text{ and } j \neq \text{RI}(i), \end{cases} \quad (5)$$

where $j = 1, 2, \dots, n$; R_j is a random variable uniformly distributed in the range $[0, 1]$; $CR \in [0, 1]$ is a user-defined crossover constant; and $\text{RI}(i) \in \{1, 2, \dots, n\}$, is a randomly selected index.

Finally, the trial vector is compared against the corresponding individual and the best one is selected to be inherited to the next generation, i.e.:

$$x_i^{(t+1)} = \begin{cases} u_i^{(t+1)}, & \text{if } f(u_i^{(t+1)}) < f(x_i^{(t)}), \\ x_i^{(t)}, & \text{otherwise.} \end{cases} \quad (6)$$

The presented DE operator, along with the rest of the basic DE operators, are implemented in the MEMPSODE software [17].

2.3 Local Search

A local solution to an optimization problem can be obtained by applying local search (LS) methods. The hybrid schemes implemented in MEMPSODE employ deterministic

LS procedures. These procedures require a starting point, $x^{(0)}$, and generate a sequence of points, $x^{(t)}$, $t = 1, 2, \dots$, which approximates a minimizer within a prescribed accuracy.

The generation of a new point, $x^{(t+1)}$, in the sequence is based on information collected for the current iterate, $x^{(t)}$. Typically, this information includes the function value at $x^{(t)}$, as well as the first- and probably second-order derivatives of $f(x)$ at $x^{(t)}$. In all cases, the aim is to find a new iterate with lower function value than the current one.

The *Merlin* optimization environment [10] is an efficient and robust general purpose optimization package. It is designed to solve multi-dimensional optimization problems. Merlin offers a variety of well established gradient-based and gradient-free optimization algorithms. Gradient-based algorithms include three methods from the conjugate gradient family, the method of Levenberg-Marquardt, the DFP and several variations of the BFGS algorithms (BFGS) [4]. The gradient-free algorithms include a pattern search and the nonlinear Simplex method.

2.4 Memetic Framework

The design of Memetic PSO (MPSO) in [14] was based on three fundamental schemes, called the *memetic strategies*:

- 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.

Of course, many other memetic strategies can be considered. For instance, a simple one would be the application of LS on every particle. However, such an approach would be costly in terms of function evaluations. In practice, only a small number of particles are considered as starting points for LS, as pointed out in [7]. The memetic strategies proposed in [14] are also implemented in MEMPSODE for both PSO- and DE-based MAs. The general procedure of the memetic algorithms discussed in our study, is given with the pseudocode of Algorithm 1. Notice that the pseudocode uses the same vector notation for both UPSO and DE.

3. EXPERIMENTAL SETUP

We applied the memetic algorithms for a maximum of $10^5 \times n$ function evaluations per run. Whenever MEMPSODE terminated before reaching the global minimum (e.g., when LS resulted in a local minimizer), we performed an independent restart until the maximum number of function evaluations was reached.

The memetic Scheme 3 was applied with probability of local search set to $\rho_i = 0.05$. Both UPSO and DE used a swarm size $N = 25$ particles. In UPSO the unification factor u was set to 1 (gbest) and the initial velocity vector was restraint by a factor of 0.01. For the DE experiments we used the values $F = 0.5$ and $CR = 0.7$. Each local search assumed a maximum number of 2000 function evaluations.

Algorithm 1: Pseudocode of the implemented memetic algorithms.

```

Input: Objective function,  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ; algorithm: algo
(PSO/DE); swarm size:  $N$ ; unification factor:  $UF$ ;
probability for local search:  $\rho$ 
Output: Best detected solution:  $x^*$ ,  $f(x^*)$ .

// Initialization
1 for  $i = 1, 2, \dots, N$  do
2   Initialize  $x_i$  and  $u_i$ 
3   Set  $p_i \leftarrow x_i$  // Initialize best position
4    $f_i \leftarrow f(x_i)$  // Evaluate particle
5    $f_i^p \leftarrow f_i$  // Best position value
6 end

// Main Iteration Loop
7 Set  $t \leftarrow 0$ 
8 while (termination criterion) do
9   Calculate global best index  $g_1$  and local best index  $g_2$ 
10  // Update Swarm/Population
11  if algo = 'PSO' then
12    for  $i = 1, 2, \dots, N$  do
13      Calculate lbest velocity update,  $L_i$ , using  $g_2$ 
14      Calculate gbest velocity update,  $G_i$ , using  $g_1$ 
15       $u_i \leftarrow UF L_i + (1 - UF) G_i$  // Unified PSO
16       $x_i = x_i + u_i$  // Update particle's position
17    end
18  else if algo = 'DE' then
19    for  $i = 1, 2, \dots, N$  do
20       $x_i \leftarrow p_i$  // Replicate best positions
21    end
22    for  $i = 1, 2, \dots, N$  do
23      Calculate  $v_i$  using Eq. (4) // Mutation
24    end
25    for  $i = 1, 2, \dots, N$  do
26      Calculate  $u_i$  using Eq. (5) // Recombination
27    end
28    for  $i = 1, 2, \dots, N$  do
29      Calculate  $x_i$  using Eq. (6)
30    end
31  // Evaluate Swarm/Population
32  for  $i = 1, 2, \dots, N$  do
33     $f_i \leftarrow f(x_i)$ 
34  end
35  // Update Best Positions/Individuals
36  for  $i = 1, 2, \dots, N$  do
37    if  $f_i < f(p_i)$  then
38       $p_i \leftarrow x_i$ 
39       $f_i^p \leftarrow f_i$ 
40    end
41  end
42  // Apply Memetic Strategy
43  Apply one of the memetic strategies with probability  $\rho$ 
44 end

```

The employed LS algorithm was the BFGS method using the default parameters provided by Merlin. For the first order derivatives we applied an $O(h)$ finite differences formula where h was an adaptable step size (see [16] for details). The experiments were conducted on an Intel I7-2600 3.4 GHz machine with 8GB RAM using GNU compiler suite v.4.4.3.

4. CPU TIMING EXPERIMENT

For the timing experiment, according to [5], the experimental procedure described above was run on f8 with at most 1000 function evaluations in each call to MEMPSODE and restarted until at least 30 seconds had passed. The timing experiment was performed on the same platform as the experimental procedure. The results for the UPSO variant were 2.8, 1.8, 1.2, 0.5, 0.26, and 0.24 times 10^{-4} seconds per function evaluation and for the DE variant 3.1, 2.0, 1.2, 0.5,

0.27, and 0.22 times 10^{-4} seconds per function evaluation, for dimensions 2, 3, 5, 10, 20, and 40, respectively.

5. EXPERIMENTAL RESULTS

Results from experiments according to [5] on the benchmark functions given in [3, 6] are presented in Figs. 1, 2 and 3 and in Table 1. The *expected running time* (ERT), reported in figures and tables, depends on a given target function value, $f_t = f_{\text{opt}} + \Delta f$, and it 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 [5].

A direct comparison between the memetic gbest variant of UPSO and the DE variant implemented in MEMPSODE, can be deduced by observing the scatter plots in Fig. 2 and the starred records in Table 1. In the separable case, DE variant outperforms gbest PSO, especially as dimensionality increases. For the moderate category, DE seems to outperform the gbest PSO variant only on function 7 (step ellipsoid) and scores marginally better in all other cases. The same behaviour is repeated for the ill-conditioned cases where DE variant is slightly superior. Finally, DE variant outperforms gbest PSO in all multimodal functions but the opposite is observed for the weak structured cases.

The DE variant's superiority against gbest PSO is also obvious by inspecting Fig. 3. In almost all cases (except the weak structured functions) the ECDF of DE variant lies higher than the corresponding ECDF of PSO variant and this pattern is repeated in all levels of accuracy.

It is also worth mentioning that the DE variant scored the best recorded ERT for some accuracy levels in the case of ill-conditioned functions (functions 10–14). From the corresponding lines of Table 1, we can see that for relatively low levels of accuracy the achieved ERT scores are quite competitive. Overall, the results suggest that the algorithms implemented in MEMPSODE can be very competitive.

6. CONCLUSIONS

We presented an empirical evaluation between two implementations of the memetic algorithm introduced in [14]. The original one uses gbest PSO for search space exploration, and the variation applies DE. Both methods are included in the MEMPSODE software that provides a versatile environment for global optimization. The results from the comparison on the noiseless testbed indicate that the DE variant has a relative advantage against the gbest PSO model. Both methods attain competitive performance, scoring way above the average, against the majority of algorithms on the same noiseless testbed, as suggested by the results.

7. REFERENCES

- [1] R. C. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. In *Proceedings Sixth Symposium on Micro Machine and Human Science*, pages 39–43, Piscataway, NJ, 1995. IEEE Service Center.
- [2] A. P. Engelbrecht. *Fundamentals of Computational Swarm Intelligence*. Wiley, 2006.
- [3] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Presentation of the noiseless functions.
- [4] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, 1970.
- [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] W. E. Hart. *Adaptive Global Optimization with Local Search*. PhD thesis, University of California, San Diego, USA, 1994.
- [8] J. Kennedy and R. C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.
- [9] P. Moscato. Memetic algorithms: A short introduction. In D. Corne, M. Dorigo, and F. Glover, editors, *New Ideas in Optimization*, pages 219–235. McGraw-Hill, London, 1999.
- [10] 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.
- [11] K. E. Parsopoulos and M. N. Vrahatis. UPSO: A unified particle swarm optimization scheme. In *Lecture Series on Computer and Computational Sciences, Vol. 1, Proceedings of the International Conference of Computational Methods in Sciences and Engineering (ICCMSE 2004)*, pages 868–873. VSP International Science Publishers, Zeist, The Netherlands, 2004.
- [12] 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.
- [13] K. E. Parsopoulos and M. N. Vrahatis. *Particle Swarm Optimization and Intelligence: Advances and Applications*. Information Science Publishing (IGI Global), 2010.
- [14] Y. G. Petalas, K. E. Parsopoulos, and M. N. Vrahatis. Memetic particle swarm optimization. *Annals of Operations Research*, 156(1):99–127, 2007.
- [15] 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.
- [16] C. Voglis, P. Hadjidoukas, I. Lagaris, and D. Papageorgiou. A numerical differentiation library exploiting parallel architectures. *Computer Physics Communications*, 180(8):1404–1415, 2009.
- [17] 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.

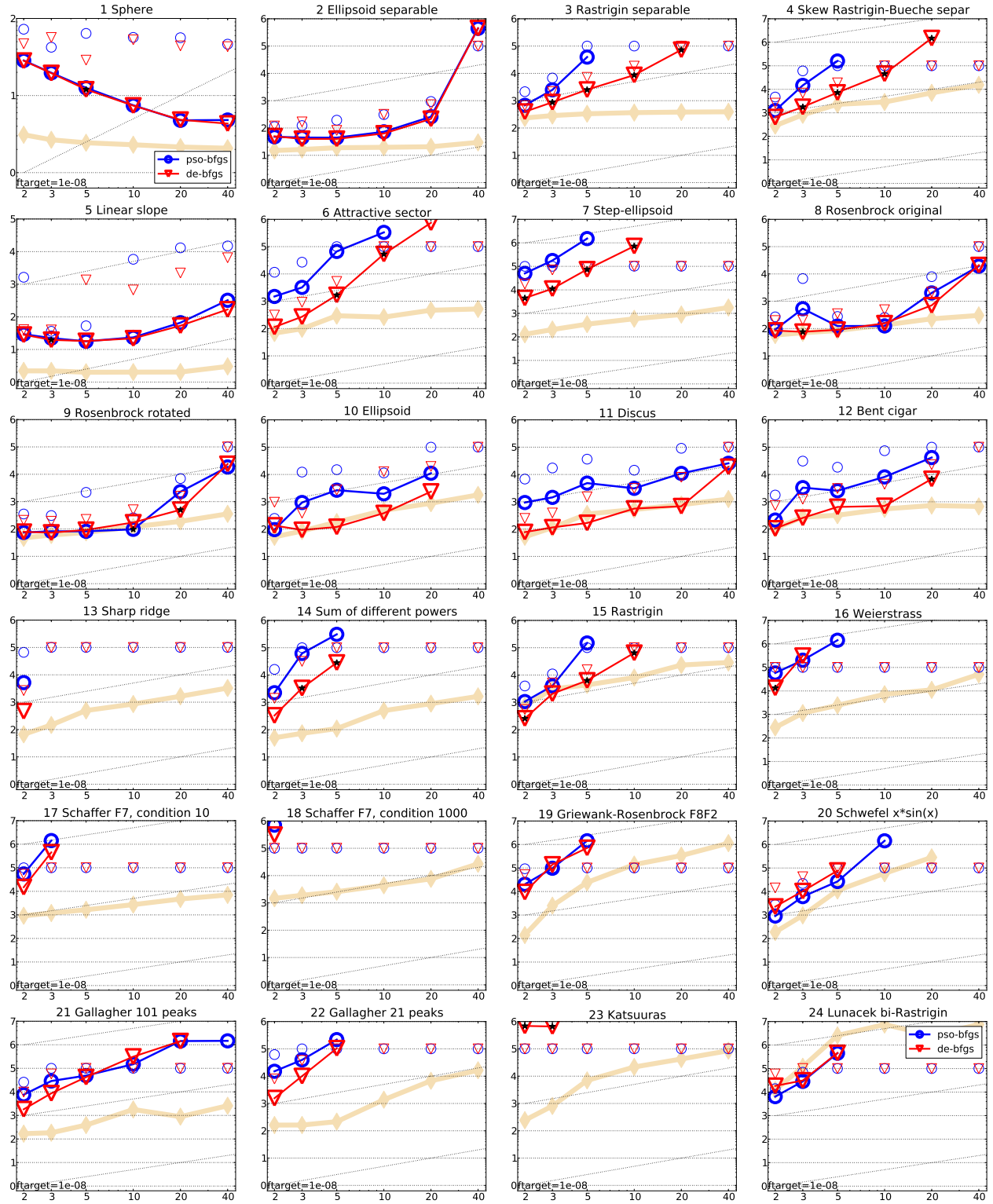


Figure 1: Expected running time (ERT in number of f -evaluations) divided by dimension for target function value 10^{-8} as $\log_{10}$ values versus dimension. Different symbols correspond to different algorithms given in the legend of f_1 and f_{24} . Light symbols give the maximum number of function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

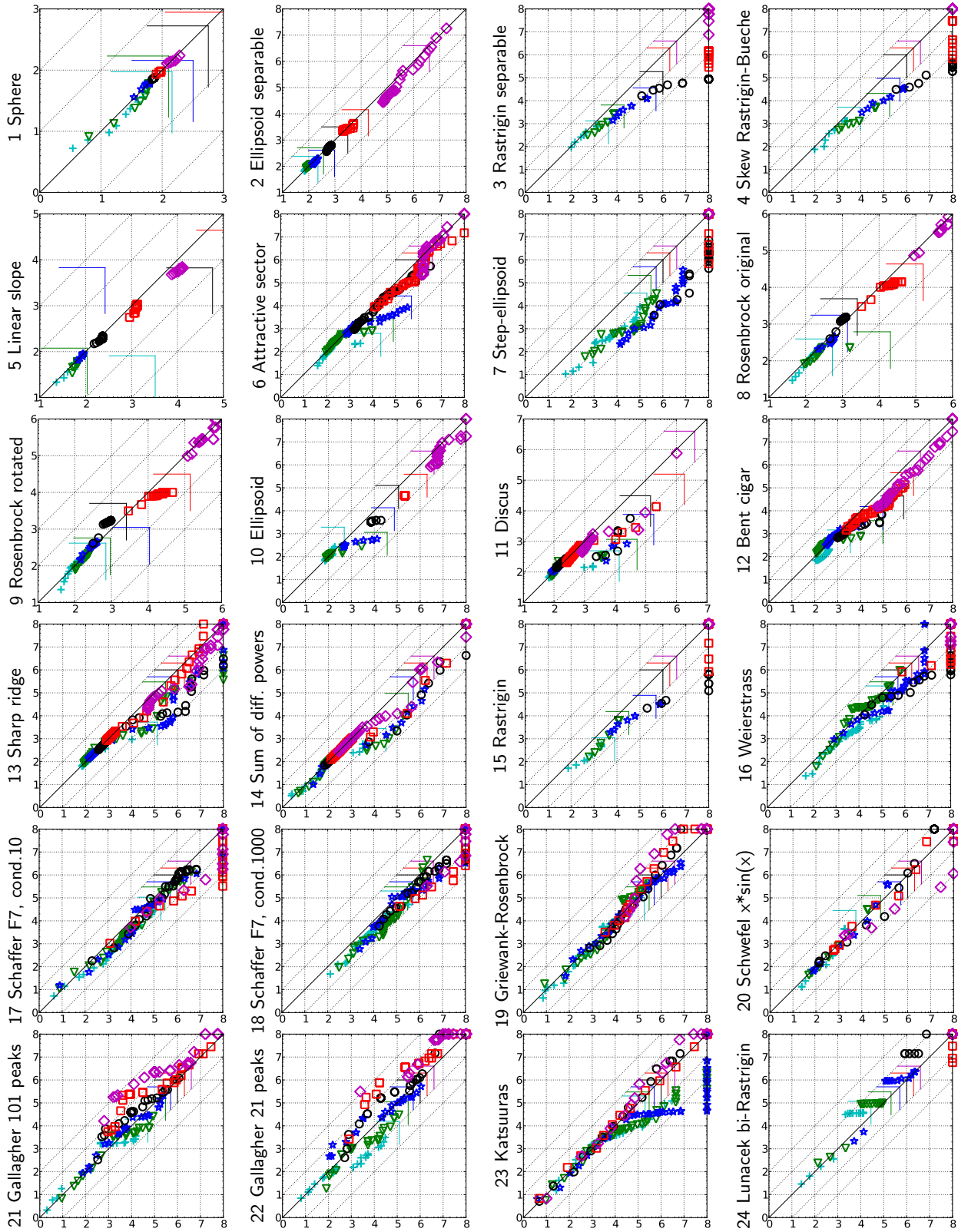


Figure 2: Expected running time (ERT in $\log_{10}$ of number of function evaluations) of pso-bfgs (x -axis) versus de-bfgs (y -axis) for 46 target values $\Delta f \in [10^{-8}, 10]$ in each dimension on functions f_1 – f_{24} . Markers on the upper or right edge indicate that the target value was never reached. Markers represent dimension: 2:+, 3:∇, 5:*, 10:○, 20:□, 40:◇.

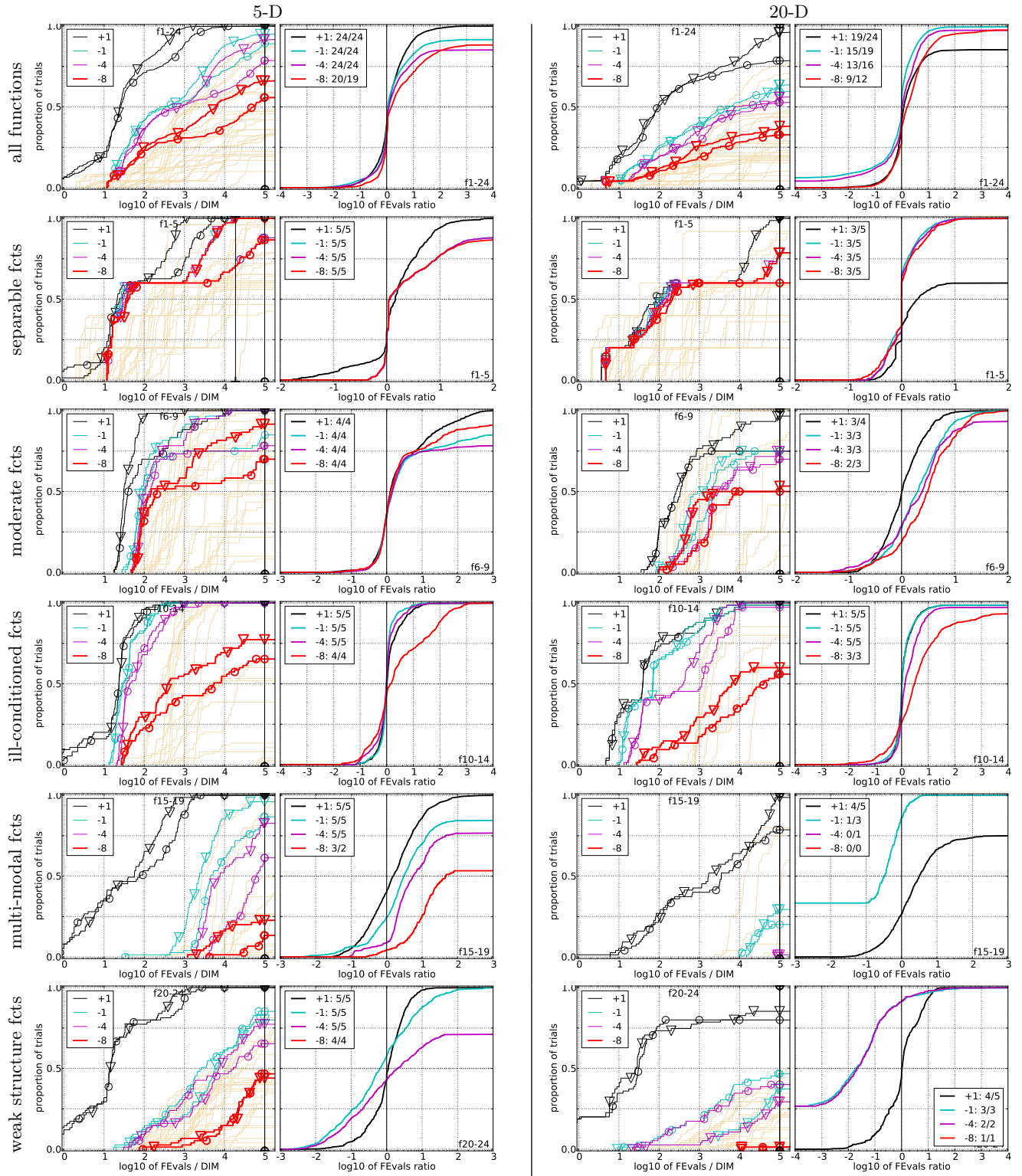


Figure 3: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/ D) to reach a target value $f_{\text{opt}} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ is given by the first value in the legend, for pso-bfgs ($\circ$) and de-bfgs (∇). Light beige lines show the ECDF of FEvals for target value $\Delta f = 10^{-8}$ of all algorithms benchmarked during BBOB-2009. Right sub-columns: ECDF of FEval ratios of pso-bfgs divided by de-bfgs, 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 (pso-bfgs first).

