



**HAL**  
open science

# Optimizing With Low Budgets: A Comparison On the Black-Box Optimization Benchmarking Suite and OpenAI Gym

Elena Raponi, Nathanaël Carraz Rakotonirina, Jérémy Rapin, Carola Doerr,  
Olivier Teytaud

► **To cite this version:**

Elena Raponi, Nathanaël Carraz Rakotonirina, Jérémy Rapin, Carola Doerr, Olivier Teytaud. Optimizing With Low Budgets: A Comparison On the Black-Box Optimization Benchmarking Suite and OpenAI Gym. IEEE Transactions on Evolutionary Computation, In press, pp.1-1. 10.1109/TEVC.2023.3346788 . hal-04370224

**HAL Id: hal-04370224**

**<https://hal.sorbonne-universite.fr/hal-04370224>**

Submitted on 2 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimizing with Low Budgets: a Comparison on the Black-box Optimization Benchmarking Suite and OpenAI Gym

Elena Raponi, Nathanaël Carraz Rakotonirina, Jérémy Rapin, Carola Doerr, Olivier Teytaud

**Abstract**—The growing ubiquity of machine learning (ML) has led it to enter various areas of computer science, including black-box optimization (BBO). Recent research is particularly concerned with Bayesian optimization (BO). BO-based algorithms are popular in the ML community, as they are used for hyperparameter optimization and more generally for algorithm configuration. However, their efficiency decreases as the dimensionality of the problem and the budget of evaluations increase. Meanwhile, derivative-free optimization methods have evolved independently in the optimization community. Therefore, we urge to understand whether cross-fertilization is possible between the two communities, ML and BBO, i.e., whether algorithms that are heavily used in ML also work well in BBO and vice versa. Comparative experiments often involve rather small benchmarks and show visible problems in the experimental setup, such as poor initialization of baselines, overfitting due to problem-specific setting of hyperparameters, and low statistical significance.

With this paper, we update and extend a comparative study presented by Hutter et al. in 2013. We compare BBO tools for ML with more classical heuristics, first on the well-known BBOB benchmark suite from the COCO environment and then on Direct Policy Search for OpenAI Gym, a reinforcement learning benchmark. Our results confirm that BO-based optimizers perform well on both benchmarks when budgets are limited, albeit with a higher computational cost, while they are often outperformed by algorithms from other families when the evaluation budget becomes larger. We also show that some algorithms from the BBO community perform surprisingly well on ML tasks.

**Index Terms**—Benchmarking, Black-box optimization, BBOB, OpenAI Gym, Bayesian Optimization, Reinforcement Learning.

## I. INTRODUCTION

Black-Box Optimization (BBO) is an affirmed and rapidly growing field of optimization and a topic of critical importance in many application areas including complex systems engineering, energy and environment, materials design, drug discovery, chemical process synthesis, and computational biology [1]. As in other classical optimization contexts, BBO assumes that we are facing an objective function  $f$  for which we aim to provide a solution  $x$  with  $f(x)$  as good as possible

using as little computational effort as possible. The key distinguishing property of BBO is that the algorithms learn about the problem instance  $f$  only by querying the quality  $f(x)$  of possible solution candidates  $x$ . This may be because we indeed lack an explicit representation of  $f$  (e.g., if simulations or experiments are required for assessing the quality of a possible solution) or when we lack efficient approaches to make use of the instance information (e.g., many real-world scheduling and routing problems are solved with heuristic approaches). Simplifying the complexity measurement, in BBO we typically only account for the number of function evaluations, and hence aim for identifying high-quality solutions  $x$  using as few function evaluations as possible.

Because of its high practical relevance, people with many different backgrounds are drawn into BBO, leading to a multitude of approaches in the area, ranging from simple heuristics such as local search to local/global modeling approaches. To understand the strengths and weaknesses of these different methods, fair performance comparisons are needed. Several platforms and benchmark suites (i.e., collections of benchmark problems) address this empirical comparison. Some examples are the Black-Box Optimization Benchmarking (BBOB) collection of the COCO environment [3], Large-Scale Global Optimization (LSGO) [4], Nevergrad [5], Pseudo-Boolean Optimization (PBO) [6], and Machine Learning and Data Analysis (MLDA) [7]. Apart from these more general benchmarking suites, there are also problem collections for evaluating and comparing algorithms for specific BBO tasks such as algorithm configuration and selection [8], [9], neural architecture search [10], [11], and expensive global optimization [12]; see [13, Section 3] for a more exhaustive summary.

An approach commonly used for expensive optimization problems (for which the available number of function evaluations can be very small) uses surrogates to approximate the problem instance  $f$ , with the idea that the approximation  $\hat{f}$  can be used to locate interesting solution candidates without requiring evaluations of the true problem  $f$ . Recently, Machine Learning (ML) has gone down this road and proposed several tools for BBO along these lines. In particular, a large field of research is Bayesian Optimization (BO), which is based on the Efficient Global Optimization (EGO) algorithm [14]. Despite its success, BO is stated to be limited to less than 15 parameters [15], [16] and a few thousand evaluations [17] according to the literature. To overcome this issue, recent research started to explore space partitioning and local modeling. In fact, learning a classifier that locates the samples on a promising

E. Raponi (the corresponding author) is with LIACS, Leiden University, Leiden, The Netherlands (e-mail: e.raponi@liacs.leidenuniv.nl). At the time of writing this article, she was initially affiliated with Meta AI Research in Paris, France, and later with the Technical University of Munich in Munich, Germany, and LIP6, CNRS, Sorbonne Université in Paris, France. N. C. Rakotonirina is with Universitat Pompeu Fabra, Barcelona, Spain (e-mail: nathanael.rakotonirina@upf.edu). At the time of writing this article, he was affiliated with LIMA, Université d’Antananarivo, Madagascar. J. Rapin and O. Teytaud are with Meta AI Research, Paris, France (e-mail: jrapin, oteytaud@fb.com). C. Doerr is with LIP6, CNRS, Sorbonne Université, Paris, France (e-mail: carola.doerr@lip6.fr).

subregion of the domain with high probability might be more effective than learning a regressor on the whole domain. Among other partitioning strategies, the Latent Action Monte Carlo Tree Search (LA-MCTS) [18] recursively learns space partition in a hierarchical manner using Monte Carlo Tree Search (MCTS) [19]. Therefore, in addition to BO, MCTS has also been adapted from control and games to BBO [18], [20]. However, both the BO and MCTS tools for BBO have rarely been compared to existing BBO methods in a systematic and satisfactory manner. For example, the comparisons in the LA-MCTS paper [18] depend heavily on poor initialization of competitors, and the baselines used in the paper are not made available in the provided code, while the comparisons in [21] consider only the simple (1+1) sampling method and not the more sophisticated (and often better performing) BBO methods provided by the Nevergrad platform, although they refer to the comparison as ‘Nevergrad’. However, we note some efforts to make neutral and comprehensive comparisons. Extensive comparisons in Nevergrad [5] tend to favor more classical methods such as tools from mathematical programming like Cobyta [22] and others [23], [24] or evolution strategies [25] like CMA-ES [26], possibly equipped with surrogate models [27]–[29]. Hutter et al. [30] ran SMAC-BBOB, a well-known BO framework, on the BBOB benchmark suite and got positive results for BO when the budget of admissible function evaluations is fairly small. However, their investigation on expensive black-box functions only assesses SMAC against CMA-ES, yielding a rather limited benchmarking study.

Overall, there is widespread utilization of black-box optimization algorithms in the ML field, yet there exists a deficiency in conducting comprehensive comparisons between the algorithms favored in ML domains and those favored by evolutionary computation researchers. Recent papers [2] raised doubts on the reproducibility of some results in the ML community. In this work, we update and extend the comparison made in [30] by adding several state-of-the-art solvers and by comparing not only on the BBOB benchmark suite but also – as it is closer to ML – on direct policy search for OpenAI Gym problems [31]. We evaluate the performance of various BBO algorithms, considering solvers commonly associated with the ML community as well as more classical heuristics, and testing over a range of budgets that allow us to draw fair and unbiased conclusions.

## II. BLACK-BOX OPTIMIZATION ALGORITHMS

We briefly summarize the algorithms included in our comparison, along with a brief description of two main classes of interest, BO and LA-MCTS.

### A. Bayesian Optimization

BO [14], [32] is a sequential design strategy targeting global optimization of black-box functions that do not assume any functional forms. It is particularly advantageous for problems where the objective function is difficult to evaluate, is a black box with some known structure, relies upon less than 20 dimensions, and where no information about sensitivity and derivatives is available. Since the objective function does not

have an explicit mathematical formulation, BO treats it as a random function and places a prior over it. A Kriging model, also known as *Gaussian Process Regression* (GPR), can be used as a prior probability distribution over functions in BO.

BO starts with sampling an initial Design of Experiments (DoE) of size  $n_0$ :  $\mathbf{X} = [\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n_0)}]^\top$ , where the sample  $i$  is denoted as  $\mathbf{x}^{(i)} = (x_1^{(i)}, \dots, x_D^{(i)}) \in \mathcal{X} \subset \mathbb{R}^D$  [33]. The corresponding objective function values are denoted as  $\mathbf{y} = (f(\mathbf{x}^{(1)}), f(\mathbf{x}^{(2)}), \dots, f(\mathbf{x}^{(n_0)}))^\top$ . Conventionally, a centered Gaussian process prior is assumed on the objective function:  $f \sim gp(0, K(\cdot, \cdot))$ , where  $K: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  is a positive definite function – also known as *kernel function* – which computes the autocovariance of the process.

Often, a Gaussian likelihood is taken, leading to a conjugate posterior process [34], i.e.,  $f | \mathbf{y} \sim gp(\hat{f}(\cdot), K'(\cdot, \cdot))$ , where  $\hat{f}$  and  $K'$  are the posterior mean and covariance function, respectively.

On an unknown point  $\mathbf{x}$ ,  $\hat{f}(\mathbf{x})$  yields the maximum *a posteriori* estimate of  $f(\mathbf{x})$  whereas  $\hat{s}^2(\mathbf{x}) := K'(\mathbf{x}, \mathbf{x})$  quantifies the uncertainty of this estimation. The posterior process is, again, a GPR. Based on the posterior process, promising points are identified via the so-called *infill-criterion*, i.e., by optimizing an *acquisition function* that balances  $\hat{f}$  with  $\hat{s}^2$  (exploitation vs. exploration). A variety of infill-criteria has been proposed in the literature, e.g., Probability of Improvement [32], [35], Expected Improvement [35], and the Upper Confidence Bound [36], [37]. When a new candidate point is selected by the infill criterion, it is evaluated and added to the BO data set, which is used to update the GPR posterior. This process is repeated until a stopping condition is met, i.e., a good enough result is located or the computational budget is exhausted.

### B. Monte Carlo Tree Search

Monte Carlo Tree Search (MCTS) [19] is a solver that migrated from trees of bandits for games and control, including alpha-zero [38], [39], to applications in BBO [18], [40]. Its evolved version, LA-MCTS [18], progressively learns and generalizes promising regions in the problem space by recursively partitioning so that solvers like BO can access these regions to improve their performance. At any iteration  $t$  of the algorithm, a training dataset  $\mathbf{D}_t = (\mathbf{X}, \mathbf{Y})$  of all the points evaluated so far is available. A tree node  $A$  represents a subregion of the search space  $\Omega_A$ . Therefore,  $\mathbf{D}_t \cap \Omega_A$  is the set of all the samples falling in the subregion  $\Omega_A$ . MCTS uses the Monte Carlo simulation to accumulate value estimates that lead to highly rewarding trajectories in the search tree. In other words, MCTS pays more attention to promising nodes (i.e., subregions of the search space), in order to avoid the need to brute force all possibilities. This is done by using an Upper Confidence Bound (UCB) policy. More specifically, each node has an associated UCB value and, during selection, the child node with the highest UCB value is considered. The statistics used to compute the UCB are (1)  $n_A$ , which is the number of samples in  $\mathbf{D}_t \cap \Omega_A$ , and (2) the node value  $v_A := (\sum_{\mathbf{x}_i \in \mathbf{D}_t \cap \Omega_A} f(\mathbf{x}_i)) / n_A$ .

Therefore, in LA-MCTS, which is the MCTS-based optimizer that we include in the comparisons of this study,

promising regions are found by recursively partitioning the search space based on latent actions. In one iteration, LA-MCTS starts building the tree by partitioning and then selects a region based on UCB. Finally, sampling is performed in the selected region using BO. In this way, BO avoids over-exploring the search space, and its performance improves, especially for high-dimensional problems.

### C. Selection of BO-based Algorithms

From the large collection of existing BO-based solvers, we have selected the following ones for our empirical comparison:

- BO: the Bayesian Optimization algorithm [41] implemented in Nevergrad. The python class is a wrapper over the `bayes_opt` package [42].
- LA-MCTS [18]: as described above, LA-MCTS is an MCTS-based derivative-free meta-solver that recursively learns space partition in a hierarchical manner. Sampling is then performed in the selected region using BO.
- Turbo [43]: a trust-region-inspired algorithm using Thompson sampling rather than the optimization of an acquisition function to find new candidate solutions in each subregion. Turbo20 denotes the multi-trust-regions counterpart of Turbo.
- AX [44]: a modular BO framework that uses BoTorch primitives for optimization over continuous spaces. It automates the selection of optimization routines, reducing the amount of fine-tuning required.
- SMAC [45]: a sequential model-based algorithm for algorithm configuration to optimize the parameters of arbitrary algorithms. It scales well to high dimensions and is particularly suitable for hyperparameter optimization of ML algorithms. The main core consists of BO. SMAC2 refers to SMAC-HPO, i.e., SMAC with hyperparameter optimization.
- HyperOpt [46]: library for serial and parallel hyperparameter optimization, designed to accommodate BO algorithms based on Gaussian processes and regression trees. We use the version based on Parzen estimates.
- Optuna [47]: automatic hyperparameter optimization software framework which uses state-of-the-art algorithms for sampling hyperparameters and pruning unpromising trials. By default, Optuna implements a BO algorithm (Tree-structured Parzen Estimator).

### D. Classical Black-box Optimization Algorithms

As baselines commonly used in the broader BBO context we consider **CMA**, which stands for CMA-ES, a well-known evolution strategy [26], **Cobyla**, a tool from mathematical programming [22], particle swarm optimization (PSO [48]), and **NGOpt** from Nevergrad [5], a wizard that combines many classical algorithms in various ways [49]. In the use cases considered in this work (sequential, low-dimensional, noise-free problems), NGOpt mainly uses CMA, Cobyla, and (1+1)-type sampling equipped with metamodels.

We also include DefaultCMA [26], a version of CMA without the BBOB-specific initialization used for the experiments of CMA on BBOB [50]. We do this to show that

ad hoc initialization of CMA-ES does not lead to significant improvement over DefaultCMA.

## III. BENCHMARK PROBLEMS

Extensive comparisons have already been proposed in Nevergrad [49], focusing on reproducibility, real-world, and different problem sizes. We propose here additional experiments in the well-known BBOB framework [50], chosen for its simplicity/canonicity, and for OpenAI Gym<sup>1</sup>, commonly used in the reinforcement learning (RL) environment [51]–[59].

**BBOB:** The BBOB collection contains 24 functions, with known difficulty (e.g., non-separability, ill-conditioning, different levels of multimodality, adequate or weak global structure, etc.). Although all functions are defined and can be evaluated over  $\mathbb{R}^D$ , the default search domain is  $[-5, 5]^D$ ; that is, in contrast to Nevergrad’s experiments in [5], [60], the BBOB suite has a focus on bounded domains.

BBOB offers a possibility to randomize both the position of the optimal solution (that is, test functions are randomly shifted in the domain) and the function value of the optimum (test functions are randomly shifted in the co-domain). To reduce bias with respect to problem encoding, we run the algorithms on 15 randomly chosen instances per test function and dimension. We consider six different dimensions (2, 3, 5, 10, 20, and 40).

As suggested in [30], we focus on a relatively small budget of  $10D$  or  $100D$  function evaluations (the default setting for BBOB experiments is  $1000D$ ). When a method crashes, we rerun it with the remaining budget.

For the BBOB experiments, our key performance measure is the empirical cumulative distribution function (ECDF) of the runtimes needed to reach the optimal objective value with a given precision  $\Delta t$ , i.e., the runtime depends on a given target function value ( $f_t = f_{\text{opt}} + \Delta t$ ) and is computed across all runs of an optimizer on a function, as the total number of function evaluations before reaching  $f_t$  divided by the number of trials that actually reached  $f_t$ . Informally, the ECDF shows the proportion of problems solved within a given budget, with the budget indicated on the x-axis.

**OpenAI Gym with Direct Policy Search:** Firstly, we work on a multi-deterministic Open AI Gym with tiny neural nets. Indeed, we analyze a small version (small number of neurons, low budget) to fit the low-budget context of the present work. *Multi-deterministic* here means that for a given algorithm, a different seed is drawn at random for each repetition, and this list of seeds is used for all algorithms. There are different seeds so that we avoid overfitting, and the list of seeds is the same over the different algorithms so that we have statistical pairing. This means that parameters are not tuned based on specific landscapes (e.g., bounded/unbounded domains, prescribed optimal regions, and multimodality of the objective function). This is somewhat analogous to random perturbation of the optimum in classical BBO benchmarks in the sense that the objective function is deterministic but drawn randomly.

<sup>1</sup><https://www.gymnasium.dev/>



Within the gym library, we select a few environments to compare algorithms based on (1) average loss (i.e. average negated reward) and (2) average winning rates (i.e., the frequencies at which one method outperforms another): MountainCar-v0 ( $D = 12$ ), Pendulum-v0 ( $D = 15$ ), GuessingGame-v0 ( $D = 24$ ), HotterColer-v0 ( $D = 24$ ), CartPole-v0 ( $D = 28$ ), CartPole-v1 ( $D = 28$ ), NChain-v0 ( $D = 40$ ), and MountainCarContinuous-v0 ( $D = 8$ ). We chose these problems with  $D < 50$  because they are sufficiently challenging and are not too hard in our context of tiny networks and minimal budget, i.e., not all algorithms perform equally. Here we use a *neural factor* (i.e., the scaling coefficient used in the benchmarking suite to choose the size of neural networks) of 1. The dimension, i.e., the total number of weights, is a consequence of the number of neurons, which in turn is based on the scaling factor (the number of neurons per hidden layer is the neural factor multiplied by the input dimension).

In a second set of experiments we include problems with dimensions  $D \leq 264$  for larger networks defined by setting the neural factor to 3, which parameterizes the size of the hidden layer. This bound on the dimension and the different neural factor lead to a different subset of OpenAI Gym problems: LunarLanderContinuous-v2, Blackjack-v0, Pendulum-v0, HotterColder-v0, MountainCarContinuous-v0, CartPole-v0, Acrobot-v1, NChain-v0, GuessingGame-v0, CartPole-v1.

Winning percentage rates are evaluated at fixed budgets of 25, 50, 100, 200, and 400 function evaluations for both settings (neural factor 1 and 3), while bigger budgets are also considered for the bigger networks (see Sections IV and V in the Supplementary Material).

#### IV. RESULTS OF THE EMPIRICAL COMPARISON

**Results for BBOB:** Figs. 1 and 2 present results with budget equal to  $10D$  and  $100D$ , respectively, in dimension  $D$ . The x-axis is the budget divided by the dimension, in logarithmic scale, while the y-axis shows the frequency of problems solved, i.e., the higher, the better. The plots are built by using the COCO/BBOB post-processing tool. It aggregates problems with different target precision values and displays the runtime distributions with simulated restarts [3], [61]. The default target precision values are 51 evenly log-spaced values between  $10^2$  and  $10^{-8}$ . The complete data set is archived on Zenodo [62] and allows further comparison with other BBOB data sets through the original COCO platform [3] or the IOHanalyzer web-interface [63].

Our experiments for a budget of  $10D$  reproduce the results in [30], where SMAC outperforms CMA. However, SMAC performance decreases with increasing dimension. Similar to all BO-based algorithms, its computational complexity increases significantly with increasing dimension and with increasing budget.

Although we tested fewer optimizers for the  $D = 20$  and  $D = 40$  cases due to the high computational cost, it is interesting to note that Cobyla, which is simply based on linear interpolation, often performs better than all other solvers. Although there is no tuning for our present results from Cobyla

on BBOB, it outperformed all BO-based algorithms. Cobyla is also the algorithm selected by the NGOpt wizard for the test cases considered, which explains their similar performance.

Fig. 2 shows that both versions of CMA – defaultcma and cmafm2 (and NGOpt16, which uses CMA as a component) – perform better than all other solvers when a larger budget of function evaluations is available.

We conclude by remarking that, in the BBOB benchmark with our specific setting using a low budget, there are elements that have a significant impact on the overall result:

- The LinearSlope function, which has optima in the corners, is difficult for methods that assume that the optimum is supposed to be inside.
- The precision parameter, ranging from  $1e - 8$  to 100 by default, has a significant impact on results. Since our budget is much smaller than the default setting, the percentage of the problems that are successfully solved is smaller than the typically reported ones.

In an additional set of experiments, we verified that removing LinearSlope or changing the precision parameter does not change the overall picture of our results.

Table I and Table II list the total execution time in seconds for 15 runs on the 24 BBOB functions for budget  $10D$  and  $100D$ , respectively. Note that NGOpt sometimes, in particular when the budget is sufficiently large compared to the dimension, spends a significant time learning a meta-model and checking in cross-validation whether this meta-model could be applied. NGOpt can hence be unexpectedly more expensive in lower dimensions, which in our experiments leads to the non-monotonic behavior of the running time with respect to the dimension. DefaultCMA, Turbo20, and SMAC2 have comparable execution times to CMAFmin2, Turbo, and SMAC, respectively. AX, BO, LAMCTS, and SMAC showed a total time of more than 20 000 seconds in dimension 10. As their computational cost became unmanageable for dimension 20, we stopped the runs after 3-day wall-clock time. We therefore do not report results for these algorithms for the 20- and 40-dimensional BBOB functions with  $10D$  budget. The same algorithms, with the exception of SMAC, plus Optuna, are excluded from the experiments with a higher total evaluation budget ( $100D$ ) due to their excessive computational cost for that budget.

#### **Results for the Direct Policy Search on OpenAI Gym:**

In Fig. 3, we provide an independent comparison of black-box solvers applied to the Ng-Full-Gym benchmark, which is Nevergrad’s direct policy search applied to OpenAI Gym. More precisely, is the optimization of a neural network as a controller for OpenAI Gym problems. We plot the unscaled loss, defined as the opposite of the reward the agent accumulates over time. Here, the lower the curve, the better the performance. The non-monotonic trend of the curves with respect to the elapsed budget is due to the fact that, for a given problem and algorithm, we perform new runs each time the total budget of the evaluations is updated. As a consequence, the algorithm may choose a different parametrization depending on the total budget or the ratio between the budget and the dimension of the specific benchmark, leading to a statistically significant difference in performance. In addition,

TABLE I  
TOTAL COMPUTATIONAL TIME IN SECONDS FOR THE BBOB BENCHMARK  
BUDGET = 10D

Dimension	2	3	5	10	20	40
CmaFmin2	30	31	33	38	50	90
AX	32 267	65 044	63 808	291 845		
Turbo	116	1 118	3 572	11 535	58 670	48 700
Cobyla	16	24	32	51	86	202
NGOpt	90	204	303	563	167	336
BO	1087	2 129	5 486	23 274		
Optuna	48	121	409	1 733	7 388	37 431
Lamcts	389	1 941	9 804	47 218		
SMAC	5 978	9 258	26 295	66 733		
HyperOpt	34	117	420	1 969	7 716	21 835
PSO	15	20	31	70	183	587

TABLE II  
TOTAL COMPUTATIONAL TIME IN SECONDS FOR THE BBOB BENCHMARK  
BUDGET = 100D

Dimension	2	3	5	10
CmaFmin2	36	41	53	89
Turbo	13 241	22 161	38 903	141 852
Cobyla	197	297	474	845
NGOpt	625	1 254	3 070	333
SMAC	224 031	205 209	351 508	807 572
HyperOpt	1 038	1 655	4 260	17 836
PSO	60	101	179	458

since we perform independent runs for different budgets and average over the repetitions, some variability in performance is to be expected.

We add the following algorithms to our comparison: QO-RandomSearch [64], MetaTuneRecentering [65], and MetaRecentering [66] are variants of random search that are fully parallel and reduce redundancies compared to random search. We also add NGOpt16RL and SpecialRL from Nevergrad: NGOpt16RL is Nevergrad’s wizard NGOpt optimized specifically for RL problems; SpecialRL runs the base NGOpt16RL for half of the budget and then uses test-based population size adaptation (TBPSA) [67] in the second half. Moreover, since the OpenAI Gym suite contains problems that are unbounded with optima at a small scale, we add algorithms that can adjust the scale of the search. GeneticDE and RotatedTwoPointsDE, for example, are designed to extrapolate the scale from some variables to others and find the right scale before running the classic DE [68] algorithm. MetaTuneRecentering performs a DoE entirely designed to choose the right scale depending on the relationship between budget and dimension. MultiCMA is a tentative robustification of CMA that runs three copies of CMA. We also consider a (1+1) evolutionary algorithm [69], which is a simple and fast state-of-the-art heuristic for adjusting the scale, Diagonal CMA [26], which uses a diagonal covariance matrix, and its scaled version, Scaled Diagonal CMA [5]. Finally, we add to our portfolio Lamcts-Turbo [18], an improved version of vanilla LA-MCTS coupled with Turbo to draw new samples in the selected node of the Monte

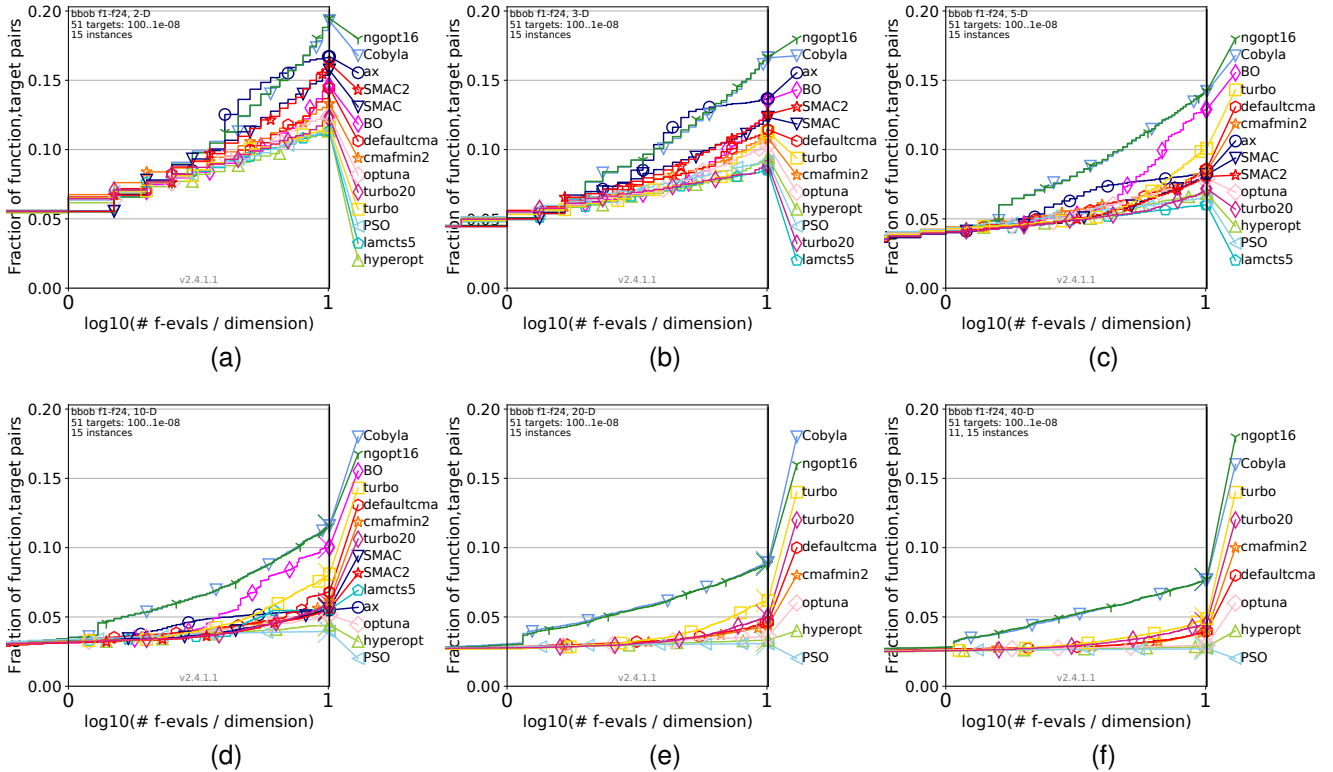


Fig. 1. Illustration of the ECDF of runtimes on the BBOB functions f1–f24, using 51 targets uniformly spaced on a log scale between  $1e-8$  and 100. Plots are shown for dimension  $D$  equal to (a) 2, (b) 3, (c) 5, (d) 10, (e) 20, and (f) 40 and budget  $10D$ . X-axis: budget/dimension in log-scale. Y-axis: frequency of solving at the requested precision. Overall, Cobyla and NGOpt16 (which heavily relies on Cobyla) perform best in these examples.

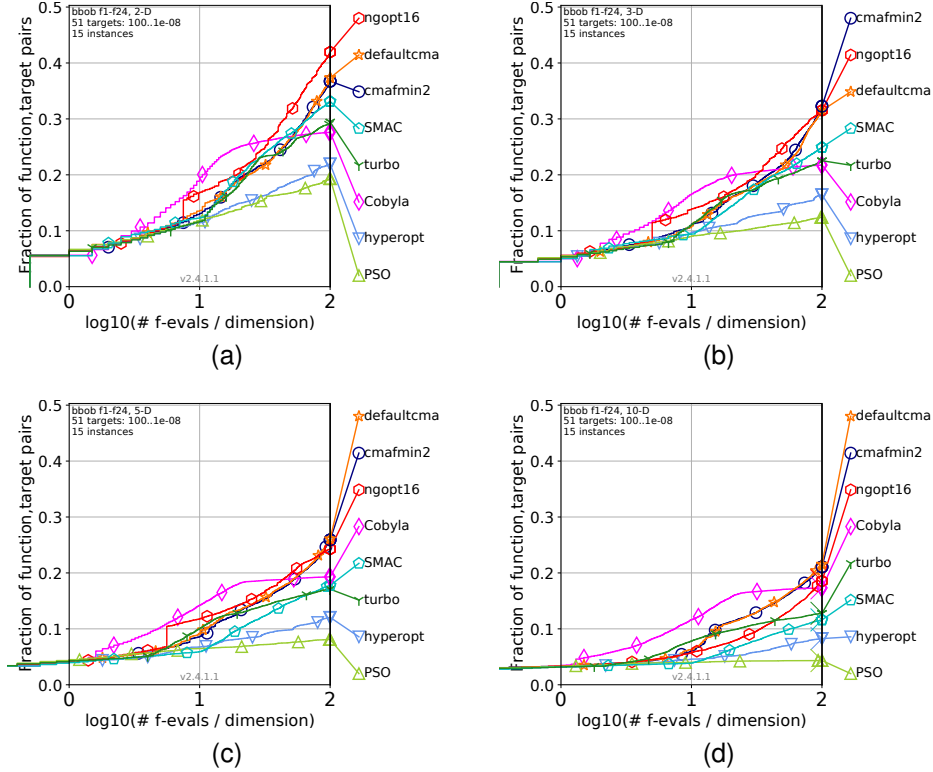


Fig. 2. As Fig. 1 but with a 100D budget and only for dimensions (a) 2, (b) 3, (c) 5, and (d) 10. Some methods, which are too slow, are removed from the analysis. Overall, CMA algorithms and NGOpt16 perform best here.

Carlo search tree at each iteration. For a complete list of all algorithms compared, see Section VI in the Supplementary Material.

Although we cannot derive comprehensive recommendations from Fig. 3 because of the high variability of results across different benchmarks and dimensions, some conclusions can be drawn from the observation of similar patterns. First, we note that Fig. 3b, 3c, and 3d corresponding to problems of intermediate dimension ( $D = 15$  and  $D = 24$ ) show no significant differences in solver performance. On the contrary, for the low-dimensional test cases in Fig. 3a and 3g, we observe a very clear superiority of the BO-based solvers (BO and AX), HyperOpt, PSO, and the wizard NgOpt16RL, all consistently belonging to the 5-best group. On the other hand, Fig. 3e, 3f, and 3h show that the quality of the performance of the 5-best group deteriorates as the dimension of the problem increases. Indeed, we observe the lowest loss values at the end of the total evaluation budget for population-based algorithms like DE and CMA, which are known to be powerful heuristics to address difficult black-box problems when a large number of function evaluations are available. However, for lower budgets of up to 100 function evaluations, the best-performing algorithms are still BO, AX, HyperOpt, PSO, and NgOpt16RL. On the contrary, Lamcts-Turbo consistently performs poorly on OpenAI Gym benchmarks, regardless of problem budget and dimension, in sharp contrast to the results shown in [18], where LA-MCTS was reported to perform well on some OpenAI Gym problems. We suspect that this discrepancy is caused by a poor initialization

of the competitors in [18].

For an in-depth comparison, we also present aggregate plots based on average winning rates (e.g., Fig. 4), where we observe good results for the BO-based methods: BO is the most powerful solver for the lowest budget, while AX, NGOpt16RL and PSO turn out to be the best algorithms (within the confidence intervals) when aggregating results across all budgets (Fig. 4f). This means that vanilla BO is actually good for fast low-precision approximations on difficult problems, whereas AX, NGOpt16RL and PSO can be recommended when optimizing low-dimensional problems with low to medium budgets. Furthermore, Fig. 4 confirms, on the one hand, the good performance of NGOpt16RL, PSO, and HyperOpt for benchmarks with tiny neural networks, and, on the other hand, the poor search capabilities of Lamcts-Turbo, regardless of the available budget. For higher dimensions, Fig. 5 shows that BO-based solvers become less competitive, although they still perform among the best for the smallest budgets (25 and 50 evaluations). For larger budgets, we see that CMA and DE start becoming more competitive, which is in line with [49]. However, Fig. 5f shows a supremacy of HyperOpt, which is designed for large-scale optimization for models with hundreds of parameters, and PSO, which performs surprisingly well considering that it is a classical heuristic originating from the BBO field and rarely used in ML applications. It can also be noted that Cobyla does not perform as well as for BBOB, while PSO performs constantly well for intermediate budgets from 50 to 400 total evaluations.

Fig. 1 in the Supplementary Material extends the present

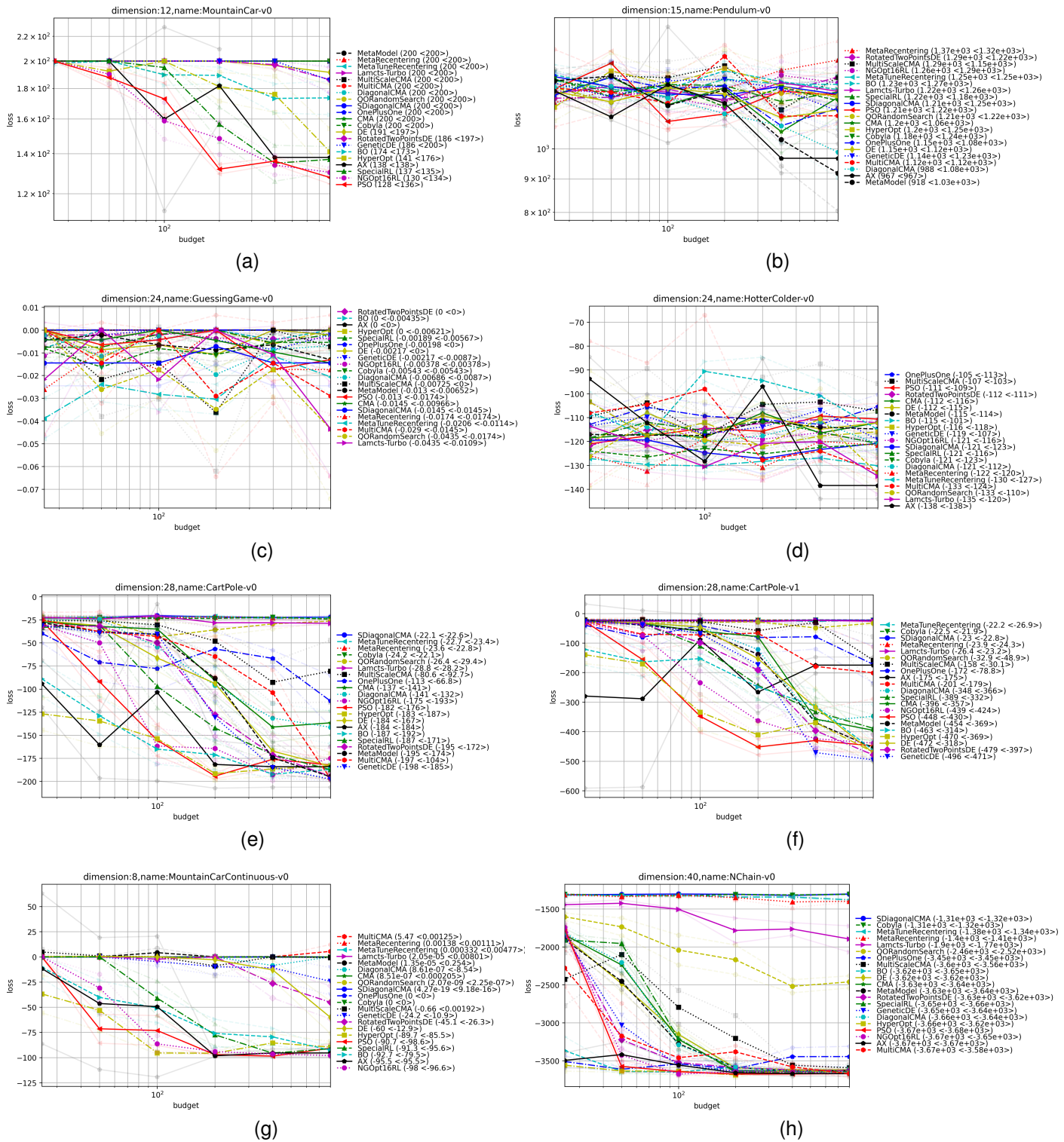


Fig. 3. Multi-deterministic Open AI Gym with tiny neural net: a random seed is randomly drawn for each optimization run so that overfitting is more difficult. See Fig. 4 for an aggregated view. The legend lists all the compared algorithms with two numbers in parentheses: These are the performance values for the last and second-to-last budgets on the x-axis, respectively. The first number is also used to sort the algorithms by performance.

results to budgets 800, 1600, and 3200 for both tiny and big neural nets. Here, as the total budget increases, a smaller set of algorithms are compared: AX is affordable for runs on tiny neural nets up to a budget of 800 evaluations, and BO and Lamcts-Turbo become too expensive when exceeding

a total budget of 1600 evaluations for big neural nets. The RL-specific algorithms, SpecialRL, and NGOpt16RL achieve the best performance for the largest budgets, followed by the various CMA and DE versions, leaving PSO and HyperOpt behind.



All in all, our results on OpenAI Gym show the competitiveness of BO-based methods for small dimensions and evaluation budgets (up to 100 evaluations), comparably to other solvers, while solvers from other families perform better as the budget increases. We note that this is different behavior from the BBOB benchmarks, where BO-based methods never rank first and show only average performance.

## V. CONCLUSION

Our results provide insight into the comparison between different state-of-the-art BO methods, commonly used in ML, and more classical BBO heuristics. We compared the solvers on the BBOB benchmark suite from the COCO environment, which is well-known in the BBO community, and on OpenAI Gym problems, which are from the RL domain. On the BBOB comparison, we noted that Cobyla and the Nevergrad wizard perform best for any tested dimensionality from 2D to 40D with a total budget of 10D. They were consistently better than SMAC and Turbo, which was highlighted as a strong BO method in [21]. It is worth noting that Turbo and HyperOpt

have the advantage of being computationally cheaper than other BO methods. Moreover, HyperOpt performed among the best on OpenAI Gym, with the exception of the largest considered budget for big neural nets. Good performance was also observed for PSO, which is commonly considered a classical heuristic and hardly ever used for ML tasks. For the larger budget of 100D, CMA or the Nevergrad wizard perform best regardless of the considered dimension of the problem.

In general, we found that the OpenAI Gym benchmark is very sensitive to variable scaling. While BBOB focuses on translations of optima and might therefore favor algorithms tuned for this setting, direct policy search for OpenAI Gym has an unbounded domain and depends differently on the initialization scaling and the ability of the algorithm to change scaling as needed. We optimized the scaling to Bayesian Optimization methods for obtaining results in Fig. 4, and then we ran the experiments to get the (possibly more neutral) results in Fig. 5 without any change.

Overall, tools based on heavy use of machine learning are computationally more expensive and perform roughly

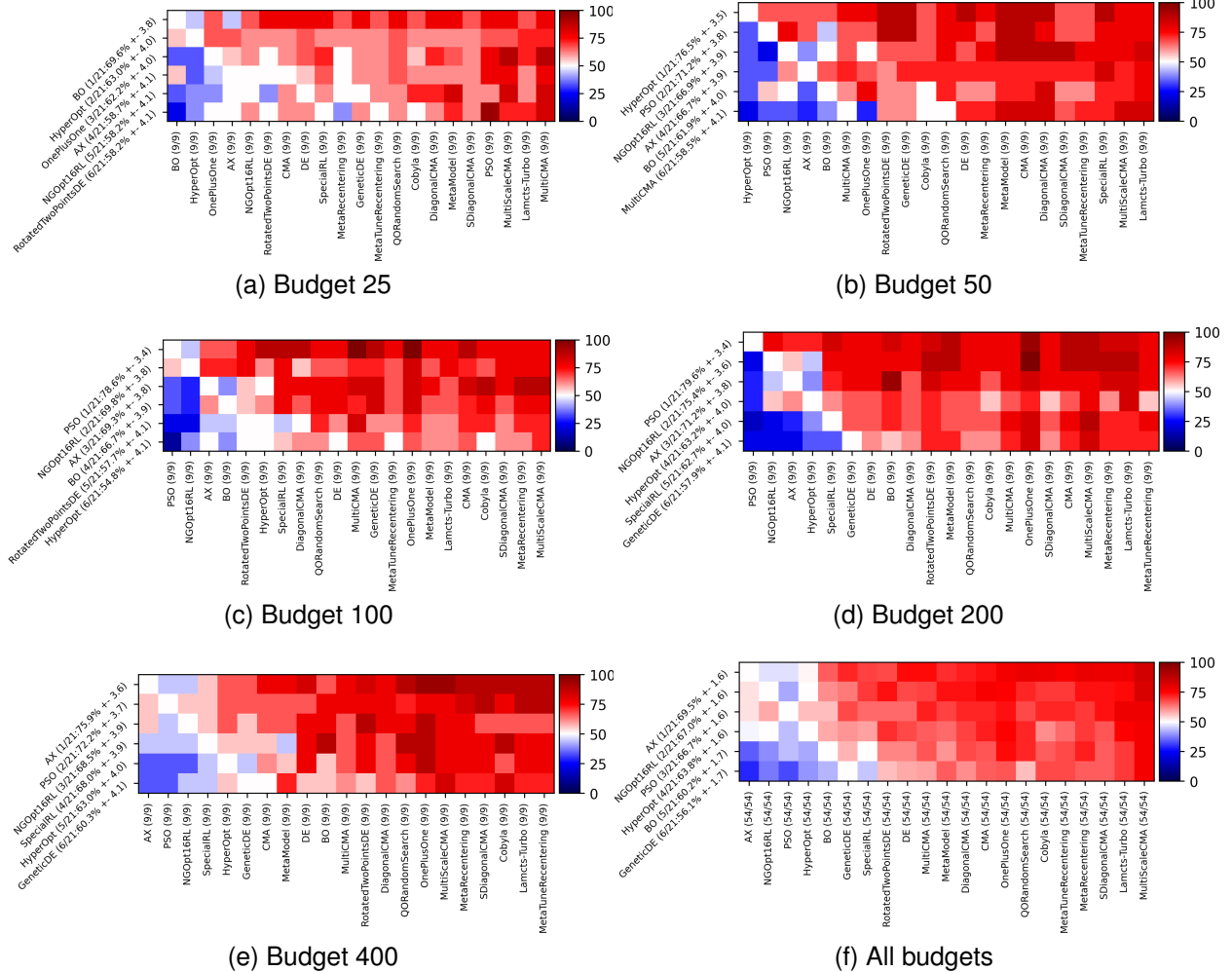


Fig. 4. Same results as reported in Fig. 3, but aggregated comparison as provided by Nevergrad [5]. Row A column B shows the frequency at which method A outperformed method B for the given budget. 9 distinct problems per budget. We include only problems for which the dimension is  $D < 50$ . Methods are ranked per average winning rate; best methods are listed first. Note that winning rates are all very close to each other: only PSO is significantly better. Fig. 5 presents similar experiments but with bigger neural nets. Fig. 1 in the Supplementary Material extends the present results to budgets 800, 1600, and 3200.

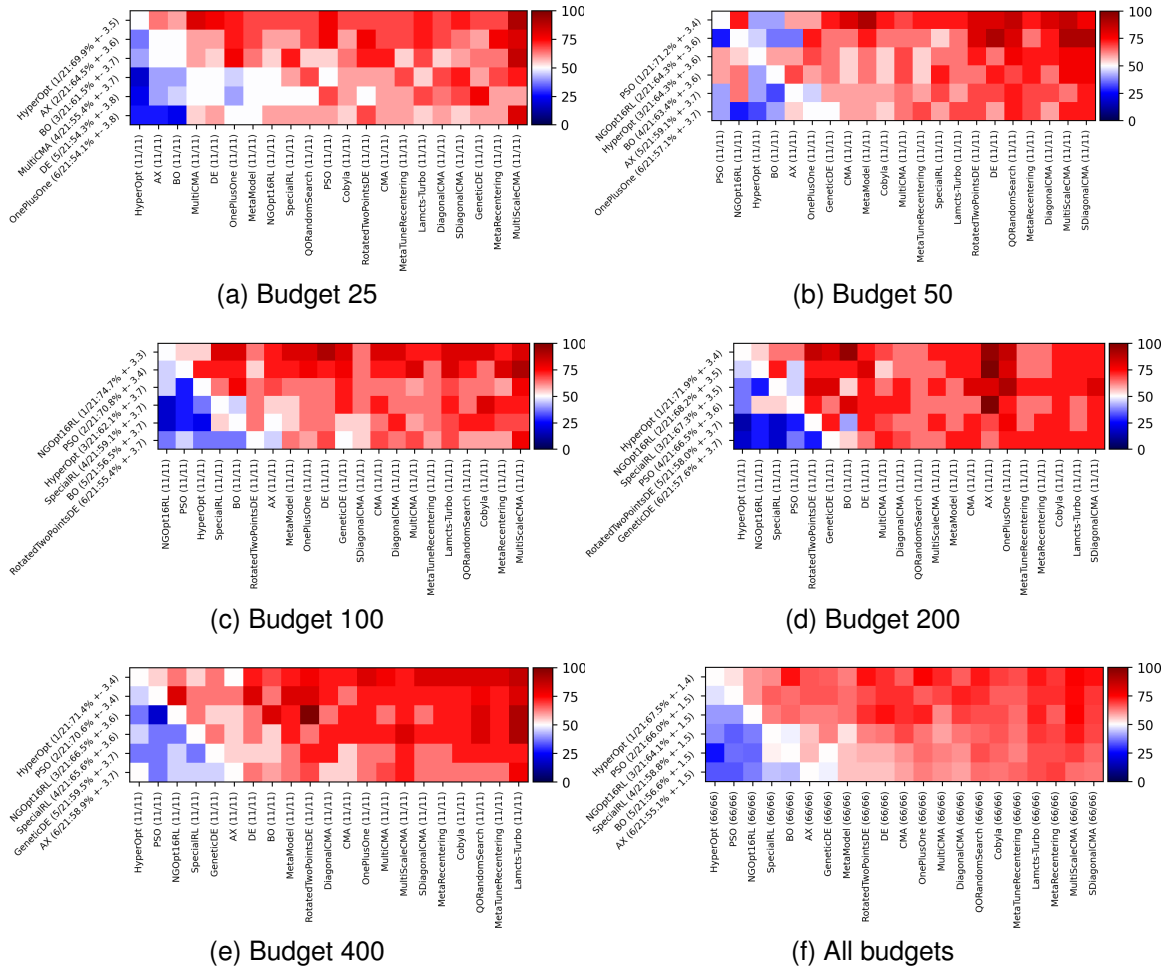


Fig. 5. Same as Fig. 4, but with bigger nets (neural factor 3 in Nevergrad’s benchmark scaling). 11 distinct problems per budget. We truncated at dimension  $\leq 264$ . Dimension ranges from 24 to 264 instead of 8 to 40 in Fig. 4. Due to the computational cost, it was not possible to finish the runs for SMAC. Fig. 1 in the Supplementary Material extends the present results to budgets 800, 1600, and 3200.

equivalently to mathematical programming or evolutionary techniques. In future work, we plan to compare BO and other methods in discrete settings as well. Our results also indicate a high relevance of initializing the solvers with the right scaling. Identifying suitable methods for a dynamic control policy is therefore another topic that we aim to investigate in future works. Of course, it remains interesting to periodically update our comparisons with new state-of-the-art solvers. Since all our experiments are performed with Nevergrad, such an ongoing benchmarking is largely facilitated: users can simply add their favorite method and compare their results to the ones reported above.

#### ACKNOWLEDGMENTS

Our work is supported by ANR, project ANR-22-ERCS-0003-01, by the CNRS INS2I emergence project RandSearch, and by the PRIME programme of the German Academic Exchange Service (DAAD) with funds from the German Federal Ministry of Education and Research (BMBF).

#### REFERENCES

[1] I. Bajaj, A. Arora, and M. M. F. Hasan, “Black-Box Optimization: Methods and Applications,” in *Black Box Optimization, Machine Learn-*

*ing, and No-Free Lunch Theorems*, ser. Springer Optimization and Its Applications, P. M. Pardalos, V. Rasskazova, and M. N. Vrahatis, Eds. Cham: Springer International Publishing, 2021, pp. 35–65.

[2] I. L. Markov. The false dawn: Reevaluating Google’s reinforcement learning for chip macro placement. *CoRR*, abs/2306.09633, 2023.

[3] N. Hansen, A. Auger, R. Ros, O. Mersmann, T. Tušar, and D. Brockhoff, “COCO: A platform for comparing continuous optimizers in a black-box setting,” *Optimization Methods and Software*, vol. 36, no. 1, pp. 114–144, Jan. 2021.

[4] K. Tang, X. Li, P. N. Suganthan, Z. Yang, and T. Weise, “Benchmark Functions for the CEC’2010 Special Session and Competition on Large-Scale Global Optimization,” University of Science and Technology of China, Tech. Rep., 2010.

[5] J. Rapin and O. Teytaud, “Nevergrad - A gradient-free optimization platform,” <https://GitHub.com/FacebookResearch/Nevergrad>, 2018.

[6] C. Doerr, H. Wang, F. Ye, S. van Rijn, and T. Bäck, “IOHprofiler: A Benchmarking and Profiling Tool for Iterative Optimization Heuristics,” *arXiv e-prints:1810.05281*, Oct. 2018. [Online]. Available: <https://arxiv.org/abs/1810.05281>

[7] J. Rapin, M. Gallagher, P. Kerschke, M. Preuss, and O. Teytaud, “Exploring the MLDA benchmark on the nevergrad platform,” in *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, ser. GECCO ’19. New York, NY, USA: Association for Computing Machinery, Jul. 2019, pp. 1888–1896.

[8] B. Bischl, P. Kerschke, L. Kotthoff, M. Lindauer, Y. Malitsky, A. Fréchet, H. H. Hoos, F. Hutter, K. Leyton-Brown, K. Tierney, and J. Vanschoren, “ASlib: A benchmark library for algorithm selection,” *Artif. Intell.*, vol. 237, pp. 41–58, 2016. [Online]. Available: <https://doi.org/10.1016/j.artint.2016.04.003>

- [9] F. Hutter, M. López-Ibáñez, C. Fawcett, M. Lindauer, H. H. Hoos, K. Leyton-Brown, and T. Stützle, “AClib: A benchmark library for algorithm configuration,” in *Proc. of Learning and Intelligent Optimization (LION'14)*, ser. LNCS, vol. 8426. Springer, 2014, pp. 36–40. [Online]. Available: [https://doi.org/10.1007/978-3-319-09584-4\\_4](https://doi.org/10.1007/978-3-319-09584-4_4)
- [10] Y. Mehta, C. White, A. Zela, A. Krishnakumar, G. Zaberjia, S. Moradian, M. Safari, K. Yu, and F. Hutter, “NAS-benchmark: NAS evaluation is (now) surprisingly easy,” in *The Tenth International Conference on Learning Representations, ICLR 2022*. OpenReview.net, 2022. [Online]. Available: <https://openreview.net/forum?id=0DLwqQLmqV>
- [11] A. Zela, J. Siems, and F. Hutter, “NAS-Bench-1Shot1: Benchmarking and Dissecting One-shot Neural Architecture Search,” in *Proc. of 8th International Conference on Learning Representations (ICLR'20)*. OpenReview.net, 2020. [Online]. Available: <https://openreview.net/forum?id=SJx9ngStPH>
- [12] L. Bliet, A. Guijt, R. Karlsson, S. Verwer, and M. de Weerd, “Expobench: Benchmarking surrogate-based optimisation algorithms on expensive black-box functions,” *CoRR*, vol. abs/2106.04618, 2021. [Online]. Available: <https://arxiv.org/abs/2106.04618>
- [13] T. Bartz-Beielstein, C. Doerr, J. Bossek, S. Chandrasekaran, T. Eftimov, A. Fischbach, P. Kerschke, M. López-Ibáñez, K. M. Malan, J. H. Moore, B. Naujoks, P. Orzechowski, V. Volz, M. Wagner, and T. Weise, “Benchmarking in optimization: Best practice and open issues,” *CoRR*, vol. abs/2007.03488, 2020. [Online]. Available: <https://arxiv.org/abs/2007.03488>
- [14] D. R. Jones, M. Schonlau, and W. J. Welch, “Efficient Global Optimization of Expensive Black-Box Functions,” *Journal of Global Optimization*, vol. 13, no. 4, pp. 455–492, Dec. 1998.
- [15] A. Nayebi, A. Munteanu, and M. Poloczek, “A Framework for Bayesian Optimization in Embedded Subspaces,” in *Proceedings of the 36th International Conference on Machine Learning*. PMLR, May 2019, pp. 4752–4761.
- [16] Z. Wang, F. Hutter, M. Zoghi, D. Matheson, and N. de Freitas, “Bayesian Optimization in a Billion Dimensions via Random Embeddings,” *arXiv:1301.1942 [cs, stat]*, Jan. 2016.
- [17] Z. Wang, C. Gehring, P. Kohli, and S. Jegelka, “Batched Large-scale Bayesian Optimization in High-dimensional Spaces,” *arXiv:1706.01445 [cs, math, stat]*, May 2018.
- [18] L. Wang, R. Fonseca, and Y. Tian, “Learning search space partition for black-box optimization using monte carlo tree search,” in *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, 2020. [Online]. Available: <https://proceedings.neurips.cc/paper/2020/hash/e2ce14e81dba66dbff9cbe35ecfdb704-Abstract.html>
- [19] R. Coulom, “Efficient selectivity and backup operators in Monte-Carlo tree search,” in *Proceedings of the 5th International Conference on Computers and Games*, ser. CG'06. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 72–83.
- [20] R. Munos, “Optimistic optimization of a deterministic function without the knowledge of its smoothness,” in *Proceedings of the 24th International Conference on Neural Information Processing Systems*, ser. NIPS'11. Red Hook, NY, USA: Curran Associates Inc., 2011, p. 783–791.
- [21] R. Turner, D. Eriksson, M. McCourt, J. Kiili, E. Laaksonen, Z. Xu, and I. Guyon, “Bayesian optimization is superior to random search for machine learning hyperparameter tuning: Analysis of the black-box optimization challenge 2020,” in *NeurIPS 2020 Competition and Demonstration Track*, vol. 133. PMLR, 2020, pp. 3–26. [Online]. Available: <http://proceedings.mlr.press/v133/turner21a.html>
- [22] M. J. Powell, *A Direct Search Optimization Method That Models the Objective and Constraint Functions by Linear Interpolation*. Springer Netherlands, 1994, pp. 51–67.
- [23] C. Cartis, J. Fiala, B. Marteau, and L. Roberts, “Improving the flexibility and robustness of model-based derivative-free optimization solvers,” 2018.
- [24] M. J. Powell, “An efficient method for finding the minimum of a function of several variables without calculating derivatives,” *The Computer Journal*, vol. 7, no. 2, pp. 155–162, 1964.
- [25] H.-G. Beyer, *The Theory of Evolution Strategies*, ser. Natural Computing Series. Springer, Heidelberg, 2001.
- [26] N. Hansen and A. Ostermeier, “Completely derandomized self-adaptation in evolution strategies,” *Evolutionary Computation*, vol. 11, no. 1, 2003.
- [27] A. Auger, M. Schoenauer, and O. Teytaud, “Local and global order 3/2 convergence of a surrogate evolutionary algorithm,” in *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO '05. New York, NY, USA: Association for Computing Machinery, Jun. 2005, pp. 857–864.
- [28] K. Khowaja, M. Shcherbatyy, and W. K. Härdle, “Surrogate models for optimization of dynamical systems,” 2021.
- [29] Z. Luksic, J. Tanevski, S. Dzeroski, and L. Todorovski, “Meta-model framework for surrogate-based parameter estimation in dynamical systems,” *IEEE Access*, vol. 7, pp. 181 829–181 841, 2019.
- [30] F. Hutter, H. Hoos, and K. Leyton-Brown, “An evaluation of sequential model-based optimization for expensive blackbox functions,” in *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation*, ser. GECCO '13 Companion. New York, NY, USA: Association for Computing Machinery, 2013, p. 1209–1216. [Online]. Available: <https://doi.org/10.1145/2464576.2501592>
- [31] G. Brockman, V. Cheung, L. Pettersson, J. Schneider, J. Schulman, J. Tang, and W. Zaremba, “OpenAI Gym,” (arXiv, 2016) Available: <https://arxiv.org/pdf/1606.01540>
- [32] J. Mockus, *Bayesian Approach to Global Optimization: Theory and Applications*. Springer Science & Business Media, Dec. 2012.
- [33] T. J. Santner, B. J. Williams, and W. I. Notz, *The Design and Analysis of Computer Experiments*, ser. Springer Series in Statistics. New York: Springer-Verlag, 2003.
- [34] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, ser. Adaptive Computation and Machine Learning. Cambridge, Mass: MIT Press, 2006.
- [35] A. I. J. Forrester, A. Söbester, and A. J. Keane, *Engineering Design via Surrogate Modelling - A Practical Guide*. John Wiley & Sons Ltd., 2008.
- [36] T. L. Lai and H. Robbins, “Asymptotically efficient adaptive allocation rules,” *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.
- [37] N. Srinivas, A. Krause, S. M. Kakade, and M. Seeger, “Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design,” *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012.
- [38] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, jan 2016.
- [39] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. P. Lillicrap, K. Simonyan, and D. Hassabis, “Mastering chess and shogi by self-play with a general reinforcement learning algorithm,” *CoRR*, vol. abs/1712.01815, 2017. [Online]. Available: <http://arxiv.org/abs/1712.01815>
- [40] R. Munos, “From Bandits to Monte-Carlo Tree Search: The Optimistic Principle Applied to Optimization and Planning,” Tech. Rep., 2014.
- [41] J. Snoek, H. Larochelle, and R. P. Adams, “Practical Bayesian optimization of machine learning algorithms,” in *Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 2951–2959.
- [42] F. Nogueira, “Bayesian Optimization: Open source constrained global optimization tool for Python,” 2014. [Online]. Available: <https://github.com/fmfn/BayesianOptimization>
- [43] D. Eriksson, M. Pearce, J. Gardner, R. D. Turner, and M. Poloczek, “Scalable global optimization via local Bayesian optimization,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32. Curran Associates, Inc., 2019. [Online]. Available: <https://proceedings.neurips.cc/paper/2019/file/6c990b7aca7bc7058f5e98ea909e924b-Paper.pdf>
- [44] FacebookResearch, “Ax - adaptive experimentation,” ax.dev, 2020.
- [45] F. Hutter, H. H. Hoos, and K. Leyton-Brown, “Sequential model-based optimization for general algorithm configuration,” in *LION*, ser. Lecture Notes in Computer Science, C. A. C. Coello, Ed., vol. 6683. Springer, 2011, pp. 507–523. [Online]. Available: <http://dblp.uni-trier.de/db/conf/lion/lion2011.html#HutterHL11>
- [46] J. Bergstra, B. Komer, C. Eliasmith, D. Yamins, and D. D. Cox, “Hyperopt: a python library for model selection and hyperparameter optimization,” *Computational Science & Discovery*, vol. 8, no. 1, p. 014008, 2015.
- [47] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A next-generation hyperparameter optimization framework,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019*. ACM, 2019, pp. 2623–2631. [Online]. Available: <https://doi.org/10.1145/3292500.3330701>

- [48] J. Kennedy, R. Eberhart, "Particle Swarm Optimization," in *Proceedings of IEEE International Conference on Neural Networks*. IEEE, 1995, pp. 1942–1948. [Online]. Available: doi:10.1109/ICNN.1995.488968
- [49] L. Meunier, H. Rakotoarison, P. Wong, B. Rozière, J. Rapin, O. Teytaud, A. Moreau, and C. Doerr, "Black-box optimization revisited: Improving algorithm selection wizards through massive benchmarking," *IEEE Trans. Evol. Comput.*, vol. 26, no. 3, pp. 490–500, 2022. [Online]. Available: <https://doi.org/10.1109/TEVC.2021.3108185>
- [50] N. Hansen, A. Auger, S. Finck, and R. Ros, "Real-parameter black-box optimization benchmarking 2009: Experimental setup," INRIA, France, Tech. Rep. RR-6828, 2009.
- [51] M. Thoma, "RL-agents," <https://martin-thoma.com/rl-agents/>.
- [52] M. Thoma, "Q-Learning," <https://martin-thoma.com/q-learning/>.
- [53] A. Manukyan, M. A. Olivares-Mendez, M. Geist, and H. Voos, "Deep reinforcement learning-based continuous control for multicopter systems," in *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2019, pp. 1876–1881.
- [54] W. C. Lewis-II, M. Moll, and L. E. Kavraki, "How much do unstated problem constraints limit deep robotic reinforcement learning?" *CoRR*, vol. abs/1909.09282, 2019. [Online]. Available: <http://arxiv.org/abs/1909.09282>
- [55] R. Henry and D. Ernst, "Gym-ANM: Reinforcement learning environments for active network management tasks in electricity distribution systems," 2021.
- [56] A. Zubow, S. Rösler, P. Gawłowicz, and F. Dressler, "GrGym: When GNU Radio Goes to (AI) Gym," in *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, ser. HotMobile '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 8–14. [Online]. Available: <https://doi.org/10.1145/3446382.3448358>
- [57] S. Green, C. M. Vineyard, and Ç. K. Koç, "Impacts of Mathematical Optimizations on Reinforcement Learning Policy Performance," in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–8.
- [58] S. Sinha, H. Bharadhwaj, A. Srinivas, and A. Garg, "D2rl: Deep dense architectures in reinforcement learning," 2020.
- [59] F. Rezazadeh, H. Chergui, L. Alonso, and C. Verikoukis, "Continuous multi-objective zero-touch network slicing via twin delayed ddpg and openai gym," 2021.
- [60] J. Rapin and O. Teytaud, "Dashboard of results for Nevergrad platform," <https://dl.fbaipublicfiles.com/nevergrad/allxps/list.html>, 2020.
- [61] N. Hansen, A. Auger, D. Brockhoff, and T. Tusar, "Anytime performance assessment in blackbox optimization benchmarking," *IEEE Trans. Evol. Comput.*, vol. 26, no. 6, pp. 1293–1305, 2022. [Online]. Available: <https://doi.org/10.1109/TEVC.2022.3210897>
- [62] E. Raponi, N. R. Carraz, J. Rapin, C. Doerr, and O. Teytaud, "Comparison of Low-budget Black-box Optimization Algorithms on BBOB," in *Zenodo*, Sep. 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.8375417>
- [63] H. Wang, D. Vermetten, F. Ye, C. Doerr, and T. Bäck, "IOHanalyzer: Detailed Performance Analyses for Iterative Optimization Heuristics," in *ACM Transactions on Evolutionary Learning and Optimization*, vol. 2, no. 3, pp. 1–29, 2022. [Online]. Available: <https://doi.org/10.1145/3510426>
- [64] S. Rahnemayan, H. R. Tizhoosh, and M. M. A. Salama, "Quasi-oppositional differential evolution," in *2007 IEEE Congress on Evolutionary Computation*, Sep. 2007, pp. 2229–2236.
- [65] L. Meunier, C. Doerr, J. Rapin, and O. Teytaud, "Variance reduction for better sampling in continuous domains," in *Parallel Problem Solving from Nature - PPSN XVI - 16th International Conference, PPSN 2020, Leiden, The Netherlands, September 5-9, 2020, Proceedings, Part I*, ser. Lecture Notes in Computer Science, vol. 12269. Springer, 2020, pp. 154–168.
- [66] M. Cauwet, C. Couprie, J. Dehos, P. Luc, J. Rapin, M. Rivière, F. Teytaud, O. Teytaud, and N. Usunier, "Fully parallel hyperparameter search: Reshaped space-filling," in *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*, vol. 119. PMLR, 2020, pp. 1338–1348. [Online]. Available: <http://proceedings.mlr.press/v119/cauwet20a.html>
- [67] V. Khalidov, M. Oquab, J. Rapin, and O. Teytaud, "Consistent population control: generate plenty of points, but with a bit of resampling," Proceedings of the 15th ACM/SIGEVO Conference on Foundations of Genetic Algorithms, FOGA 2019, 116–123, Potsdam, Germany, August 27-29, 2019. Available: <https://doi.org/10.1145/3299904.3340312>
- [68] R. Storm and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, pp. 341–359. 1997. <https://doi.org/10.1023/A:1008202821328>
- [69] M. A. Schumer and K. Steiglitz, "Adaptive Step Size Random Search," *IEEE Trans. Automat.*, vol. AC-13, no. 3, pp. 270–276, 1968.