Delft University of Technology

Continuous surrogate-based optimization algorithms are well-suited for expensive discrete problems

Rickard Karlsson, Laurens Bliek, Sicco Verwer and Mathijs de Weerdt

Faculty of Electrical Engineering, Mathematics & Computer Science, Delft University of Technology

Abstract

One method to solve expensive black-box optimization problems is to use a surrogate model that approximates the objective based on previous observed evaluations. The surrogate, which is cheaper to evaluate, is optimized instead to find an approximate solution to the original problem.

In the case of discrete problems, recent research has revolved around surrogate models that are specifically constructed to deal with with discrete structures. A main motivation is that literature considers continuous methods, such as Bayesian optimization with Gaussian processes as the surrogate, to be sub-optimal (especially in higher dimensions) because they ignore the discrete structure by e.g. rounding off real-valued solutions to integers.

In this paper, we claim that this is not true. In fact, we present empirical evidence showing that the use of continuous surrogate models displays competitive performance on a set of high-dimensional discrete benchmark problems, including a real-life application, against state-of-the-art discrete surrogate-based methods. Our experiments on different discrete structures and time constraints also give more insight into which algorithms work well on which type of problem.

Problem description

We consider a d-dimensional **discrete black-box optimization problem**

$$\min_x f(x)$$

$$\text{s.t. } l_i \leq x_i \leq u_i, i = 1 \ldots, d$$

where $x \in \mathbb{Z}^d$ and $f : \mathbb{Z}^d \longrightarrow \mathbb{R}$. Assume that

- function f is a black-box, i.e. no information is known about it.
- function f is expensive to evaluate.
- output has additive noise $y = f(x) + \epsilon$.

Surrogate-based optimization

Approximate f with a surrogate model M and solve $\min_x M(x)$ instead. The main benefit is that the surrogate model is cheaper to evaluate and can be fitted on a history of evaluations from f.

Surrogate-based optimization is an **iterative procedure** as follows:

1. Evaluate $y = f(x) + \epsilon$
2. Fit surrogate model M on the previous history of evaluations
3. Propose a new x^{next} to evaluate on next with the help of an acquisition function
4. Go back to step 1, unless evaluation budget is reached.

Benchmark with limited evaluation budget

Experimental setup

Strict evaluation budget up until 500 evaluations of the objective function; large problem search spaces in the order of $10^{40} - 10^{50}$; each experiment repeated five times; all algorithms start with five uniform samples in the search space before the surrogate model guides the search.

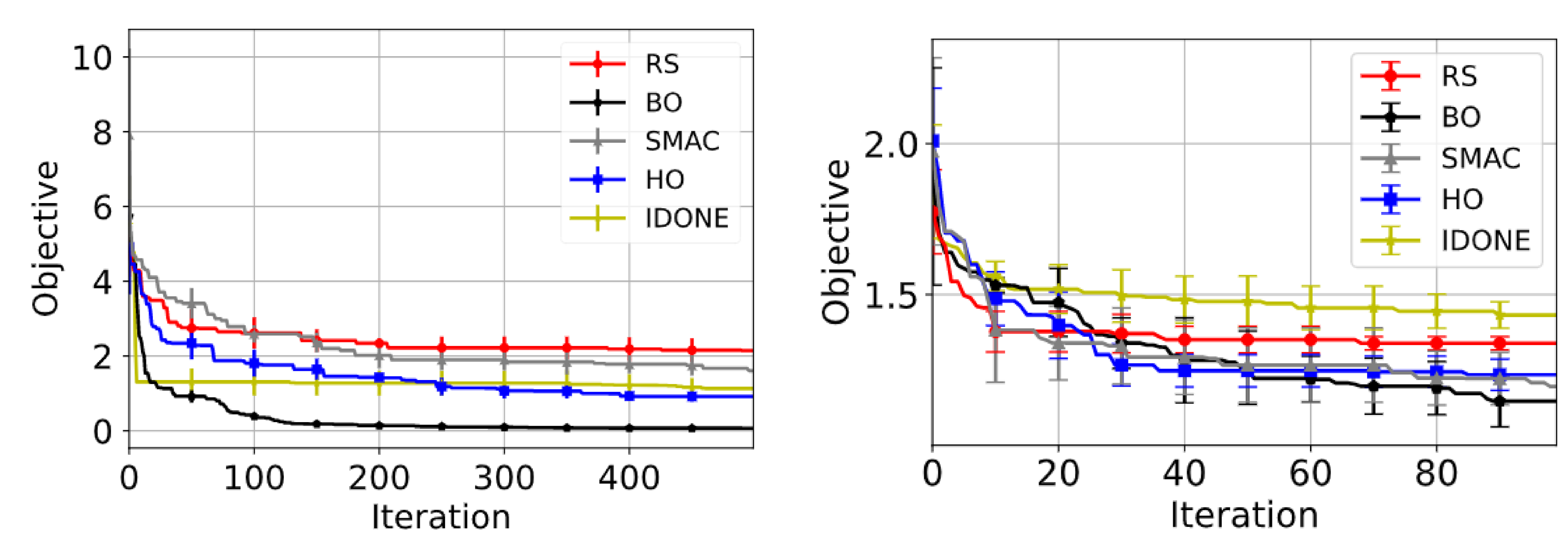
Algorithms

Continuous: **Bayesian optimization** [1] (BO) with Gaussian processes as surrogate, **IDONE** [2] with piece-wise linear surrogate model that has guaranteed integer solutions as the minimum.

Discrete: **SMAC** [3] with random forest as surrogate, **HyperOpt** [4] with tree of Parzen estimators as surrogate.

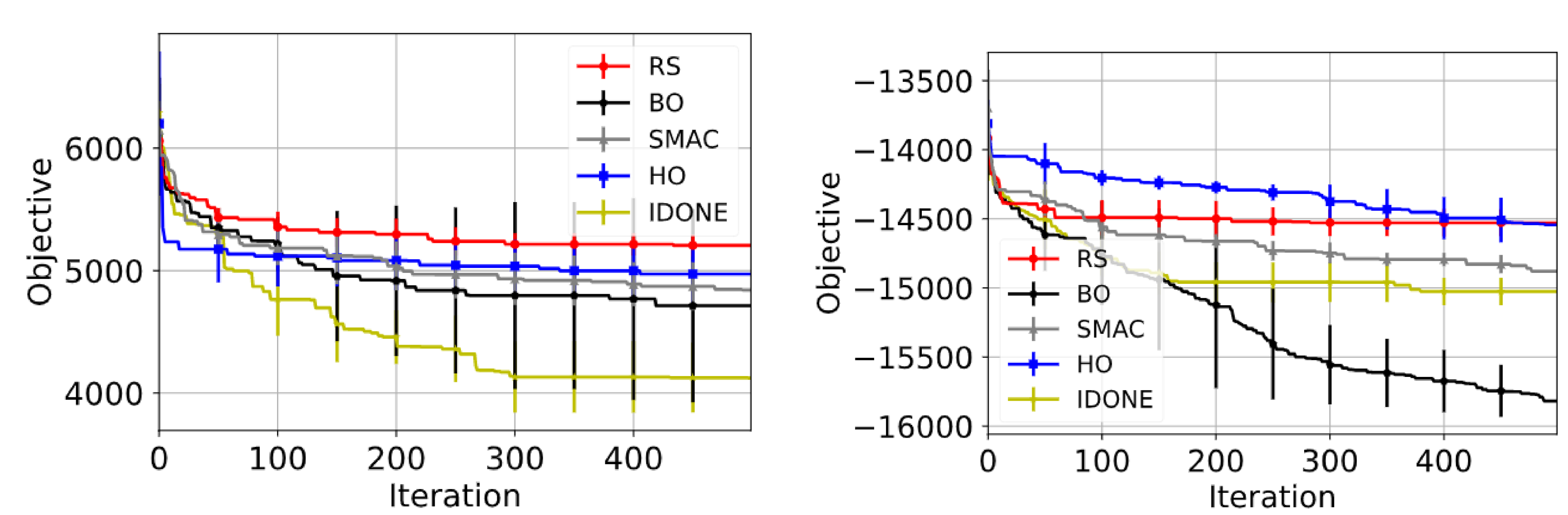
Experiments

Bayesian optimization perform the best on three out of four discrete benchmark problems (Rosenbrock, Electrostatic Recipitator, Weighted Max-Cut), while IDONE performs the best on the fourth (Traveling salesman problem). Both use continuous surrogate models.



(a) (b)

Figure: Left: Discrete Rosenbrock Problem, right: The Electrostatic Precipitator Problem



(a) (b)

Figure: Left: Traveling salesman problem, right: Weighted Max-Cut problem

Benchmark under time constraints

If we limit the time budget instead of the evaluation budget and vary the evaluation time for each problem, then the results appears different. Continuous surrogates shows to appear well on all of the problems with certain constraints, but the discrete surrogate-based algorithms show more promise in this scenario.

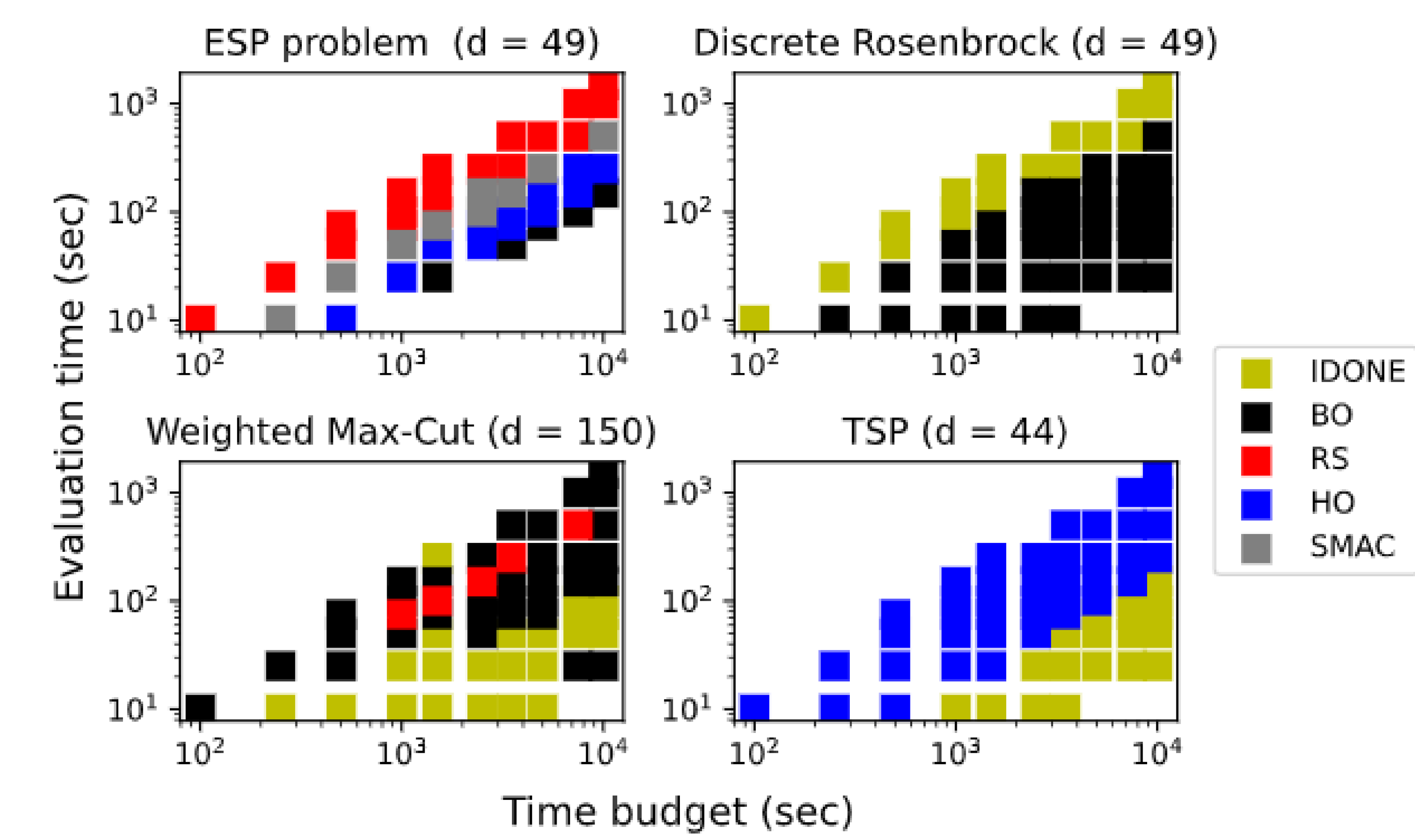


Figure: Each box represents the best-performing algorithm when restricting the time budget according to the x-axis and varying the evaluation time of the objective according to the y-axis.

Conclusion

- Continuous surrogates applied to discrete black-box optimization problems should get more attention.
- Restricting the time budget gives additional insight, continuous and discrete surrogates work good with different time constraints.

References

[1] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical Bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951--2959, 2012.

[2] Laurens Bliek, Sicco Verwer, and Mathijs de Weerdt. Black-box combinatorial optimization using models with integer-valued minima. *Annals of Mathematics and Artificial Intelligence*, pages 1--15, 2020.

[3] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *International conference on learning and intelligent optimization*, pages 507--523. Springer, 2011.

[4] James Bergstra, Dan Yamins, and David D Cox. Hyperopt: A Python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, pages 13--20, 2013.