Data Efficient Black-Box Optimization using Tree-structured Parzen Estimator

Vipul Chalotra[1, 2] **Inneke Van Nieuwenhuyse**[1, 2] **Sasan Amini**[1, 2]

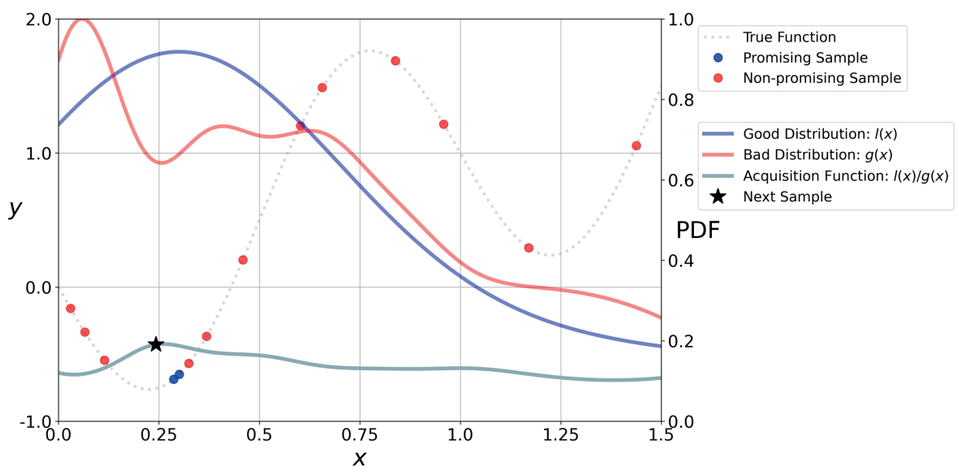
1. Computational Mathematics Research Group & Data Science Institute, Hasselt University, Belgium

2. FlandersMake@UHasselt

Have you ever felt like you're trying to tune a radio with multiple knobs?

Many of us face black-box problems, situations where we don't know the exact mathematical formulation of our problem but we can evaluate the problem. **Bayesian Optimization (BO) can help you solve such problems in a data-efficient way**. One such BO approach is the Tree-structured Parzen Estimator (TPE).

TPE is a density-based surrogate that models p(x|y) instead of p(y|x). It separates promising from non-promising regions using two density estimators and scales efficiently to **high-dimensional, mixed, and conditional search spaces.**

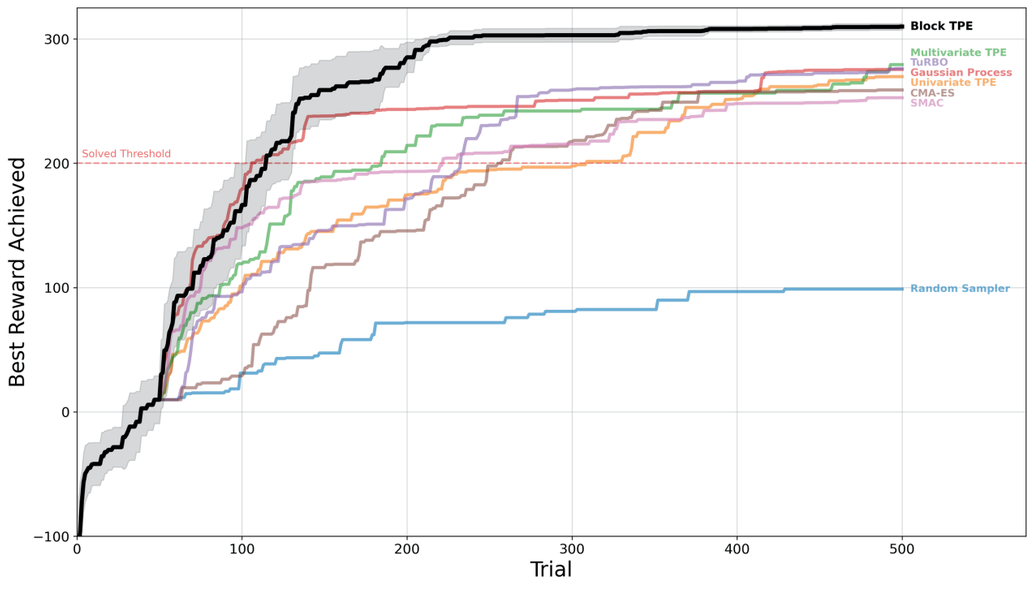


TPE is widely used for hyperparameter optimization (HPO) due to its versatility. We introduce **Block TPE,** an extension of the state-of-the-art TPE implementation in Optuna, by enhancing its efficiency in high-dimensional spaces through supervised latent space transformations to focus the density estimation directly on the most influential parameter directions. In particular, Block TPE improves the TPE algorithm by:

1. Grouping input dimensions together in blocks → solve multiple TPEs in lower dimensions
2. Transforming blocks from input space X to latent space Z → rotate the search space for better alignment with the objective

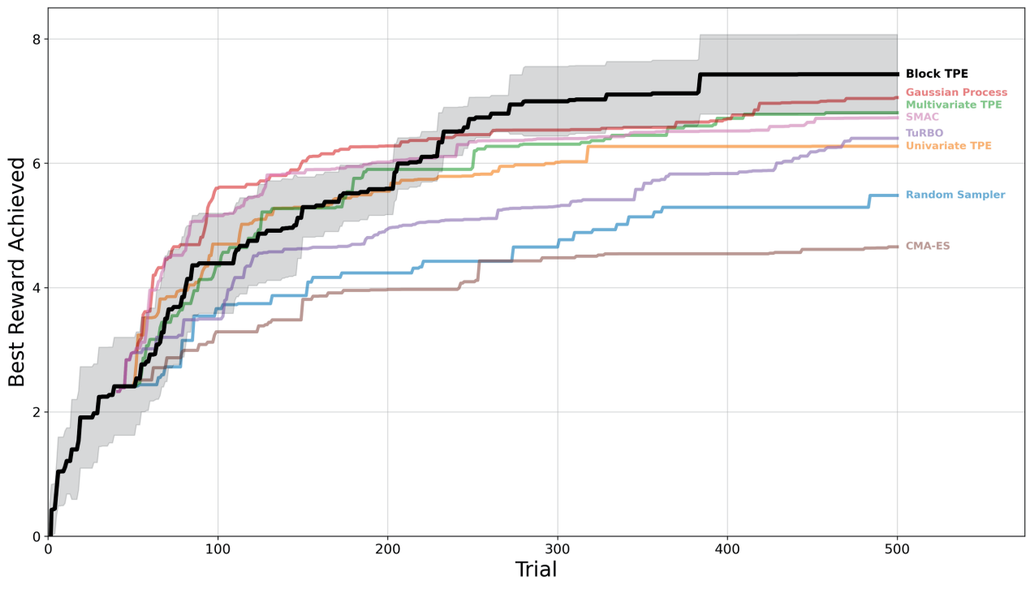
12D Benchmark: Lunar Lander - Small Budget Policy Optimization

Block TPE beats multiple BO methods across 10 runs of 500 trials



14D Benchmark: Robot Pushing - Control Optimization Problem

Outperforms Gaussian Process, SMAC, TuRBO, and more



Block TPE starts slowly but significantly outperforms Random Search, Univariate TPE, & Multivariate TPE in higher dimensions.

Curves that rise to 100% earlier and more steeply indicate better and more consistent optimization performance.

The empirical cumulative distribution functions (ECDFs) are aggregated over 6 COCO BBOB functions, 15 instances, and 3 seeds per dimension.