Proceedings of the 2022 Winter Simulation Conference
B. Feng, G. Pedrielli, Y. Peng, S. Shashaani, E. Song, C.G. Corlu, L.H. Lee, E.P. Chew, T. Roeder, and P. Lendermann, eds.

THE INTERPLAY OF OPTIMIZATION AND STATISTICS TO SOLVE LARGE-SCALE BLACK-BOX NOISY FUNCTIONS

Pariyakorn Maneekul

Industrial and System Engineering
University of Washington
Seattle WA ,USA

ABSTRACT

The main research focus is to develop optimization methodologies that integrate with statistics for solving black-box noisy functions with discrete-time performance analysis. The proposed research will improve the efficiency of a partition-based optimization method, Probabilistic Branch and Bound (PBnB) for level set approximation. Probabilistic Branch and Bound approximates a level set by branching subregions and statistically classifying them as maintained (inside the level set) or pruned (no intersection with the level set). We propose the multi-level PBnB that uses importance sampling to identify promising subregions for classification compared with uniform sampling in the original PBnB. We also incorporate Gaussian processes as a surrogate model to guide the importance sampling and aid in classifying subregions. We present the proposed algorithms with a finite-time performance analysis in terms of incorrect pruning and maintaining of subregions of the solution space. Numerical experiments are presented and compared to the analytical results for demonstration purposes.

1 BLACK-BOX NOISY FUNCTIONS OPTIMIZATION

Black-box optimization requires the design and analysis of algorithms that can be applied to objective functions and the constraints where the structure of the functions is unknown with no explicit mathematical formulation. In addition, the objective function and constraints cannot be evaluated directly, but must be estimated often by running a simulation. The task of black-box noisy optimization depends on acquiring or estimating the function value given a design vector to gain information and create a search path to the optimum. Black-box optimization often has a limited budget for function evaluations due to time-consuming function evaluations (e.g., a simulation) and limit on computation expense. The design of an algorithm for black-box optimization raises a challenge of trading off exploration, exploitation, and estimation to investigate regions without compromising efficiency and quality of the solution. When the objective function is computationally expensive to evaluate, recent advances in statistical models and machine learning algorithms offer the opportunity to leverage analytical tools to enhance the prediction of a global optimum with probabilistic bounds on the quality of the solution.

2 PROBABILISTIC BRANCH AND BOUND (PBNB)

The main research focus is to develop an optimization methodology that integrates statistics with global optimization for large-scale black-box noisy functions with discrete-time performance analysis. The proposed research will first improve the efficiency of a black-box partition-based optimization method, Probabilistic Branch and Bound (PBnB), for level set approximation. The goal for level set approximation is to find a set of solutions that achieves a target quantile (e.g., best 10%). The problem of estimating level sets of black-box noisy objective functions arises in a wide range of applications, and is an active area of research.

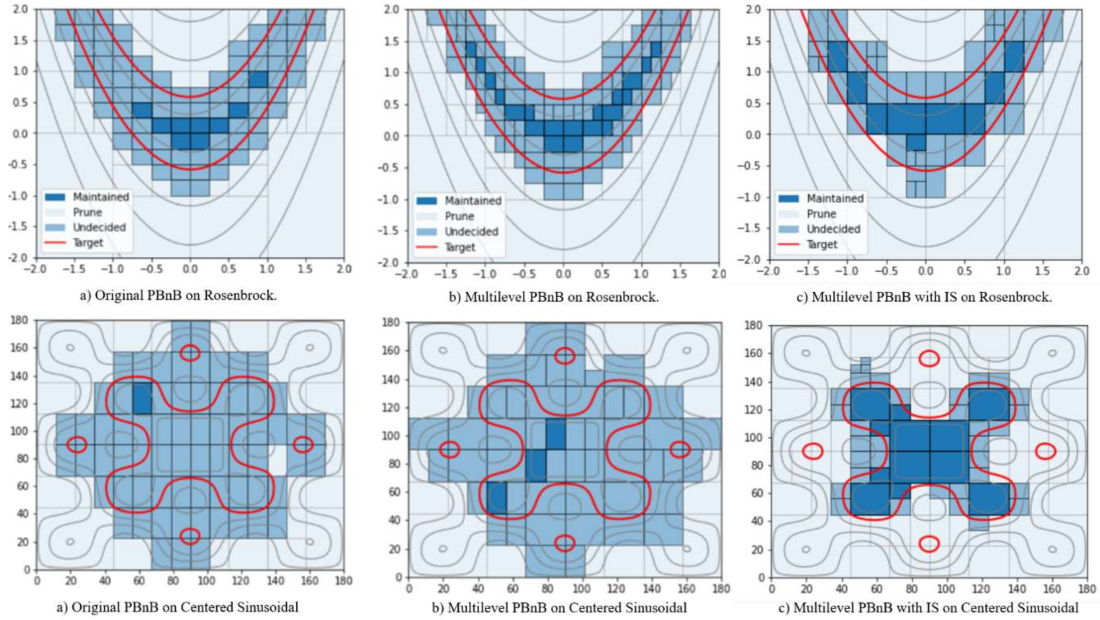

Figure 1: Approximating the 0.2 quantile level set (contour shown in red) on the tenth iteration of Original, Multilevel PBnB, and Multilevel PBnB with Importance sampling (IS) on Rosenbrock (Top row) and Centered Sinusoidal (Bottom row) in two dimensions.

The main computational challenge of Original PBnB stems from making too many function evaluations and branching too many subregions. This motivates the concept of Multilevel PBnB, where only the most promising subregions are branched. Also, Original PBnB samples uniformly on the current subregions, which is a conservative sampling distribution. We consider a form of importance sampling that biases samples towards more promising subregions, called Multilevel PBnB with IS. Computational results, as shown in Figure 1 (and presented in the WSC 2022 paper), demonstrate improved efficiency by only branching on the most promising subregions, and that adding importance sampling to focus the sampling density on the subregions with better observed function values further improves performance. Notice that Multilevel PBnB improves on Original PBnB by maintain more subregions, while Multilevel PBnB with IS improves further by maintaining more subregions of larger size in 10 iterations. More results are presented in the WSC 2022 paper.

The proposed research will also incorporate surrogate modeling into PBnB to provide statistical predictions to inform branching, guide the importance sampling distribution, and aid in classifying subregions. A proposed Multilevel PBnB with Gaussian processes will build on machine learning methods to bound the probabilities of misclassifying subregions and streamline the dynamic update of branching and sampling distributions. The proposed research will build on the previous finite-time analysis of PBnB that provides probability bounds on incorrect classification of subregions, and will present a finite-time performance analysis in terms of incorrect classification of subregions of the solution space. A challenge to extending the analysis is to account for dependencies in the adaptive sampling. The use of order statistics and stochastic processes has led to preliminary results of a finite-time statistical analysis of Multilevel PBnB with IS and with Gaussian processes to provide probability bounds on correctly identifying near-optimal solutions. We anticipate that this analysis may also be applied to other adaptive random search algorithms. Numerical experiments on benchmark problems are presented. In addition, a computational experiment is compared to the analytical results for demonstration purposes.