Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

BREAKING THE MONOTONY: PROMOTING DIVERSITY IN HIGH-DIMENSIONAL BATCH SURROGATE OPTIMIZATION

Nazanin Nezami
Department of Mechanical and Industrial Engineering
University of Illinois at Chicago
842 W Taylor St
Chicago, IL 60607, USA

ABSTRACT

In the realm of high-dimensional batch surrogate optimization, the challenge of fostering diversity while pursuing optimal solutions is paramount. Traditional approaches often result in monotonous exploration patterns, hindering the discovery of promising solutions and reducing efficiency. This thesis introduces innovative strategies, prioritizing diversity and exploration to break free from the monotony inherent in such tasks. Additionally, the thesis explores the impact of algorithmic hyperparameters on the exploration-exploitation trade-off to establish a robust framework. The "Elevating Exploration" strategies prioritize diverse candidate batch generation through adaptive sampling techniques, infusing vitality into the optimization process and effectively exploring uncharted regions of the search space. Empirical validation on optimization problems confirms their effectiveness in navigating complex landscapes. Beyond theoretical advancements and empirical validation, this thesis lays the groundwork for a paradigm shift, empowering practitioners to approach complex optimization challenges with renewed vigor and precision by promoting diversity and elevated exploration.

1 RESEARCH PROPOSAL

The optimization of resource-intensive black-box functions represents a critical challenge encountered in a wide array of real-world applications, ranging from engineering design and drug discovery to algorithm design. In this thesis, we delve into the complexities of this task, emphasizing the need for novel techniques to enhance the efficiency and effectiveness of Surrogate Optimization (SO) algorithms. SO plays a pivotal role in addressing black-box optimization (BBO) problems (Anahideh et al. 2022), where the underlying analytical form of the target function is either unknown or inaccessible, all while complying with a limited evaluation budget. SO revolves around the creation of a cost-effective surrogate model capable of approximating the intricate behavior of the black-box function. This surrogate model then guides the selection of promising regions within the input space for subsequent sampling and evaluations.

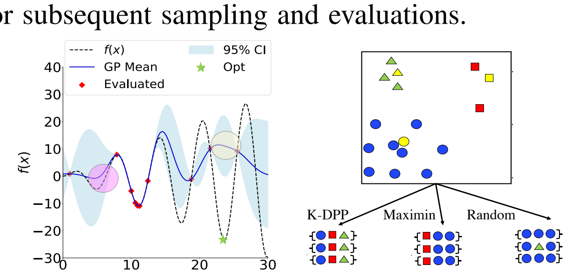

Figure 1: Diverse Sampling Adversarial Examples.

One of the key challenges in SO, particularly in scenarios involving high-dimensional and complex functions, lies in striking a balance between exploration and exploitation while ensuring diverse and promising candidate selections. The practice of concurrently assessing multiple data points in each iteration of SO, referred to as batch sampling, serves to reduce the overhead cost linked with individual evaluations and explore various areas at once for faster convergence. However, existing SO algorithms face several challenges in effectively balancing the exploration-exploitation trade-off, ensuring diversity in the selected batches, and maintaining computational efficiency. To address these challenges, we advocate the adoption of new methodologies that emphasize diversity and efficient exploration. In pursuit of a superior exploration-exploitation balance, we shed light on the limitations of conventional exploration techniques such as Maximin distance and surrogate's uncertainty (variance) through adversarial examples as shown in Figure 1. While Maximin promotes exploration by maximizing the spacing between sampled points, it may not guarantee a diverse set of solutions, especially in high-dimensional spaces, as it may struggle to distribute points effectively in such spaces. Moreover, SO approaches using variance tend to favor areas with high uncertainty for exploration. This might cause the optimization algorithm to converge prematurely to suboptimal solutions and it may not guarantee the exploration of poorly explored regions due to lack of control over the diversity of the selected points.

To tackle the intricacies of batch sampling, we introduce the concept of diversity through Determinantal Point Processes (DPPs) (Bıyık et al. 2019) in our proposed approach, termed DPPSO. Diversity here refers to the degree of variation or heterogeneity present in a set of observations. DPPSO leverages DPP-based candidate generation and DPP-based sample selection techniques to promote diversity within the selected batch. While DPPSO shows promise in solving low-dimensional BBO problems such as DNA Binding in Figure 2, its application to high-dimensional settings poses computational challenges due to expensive matrix calculations.

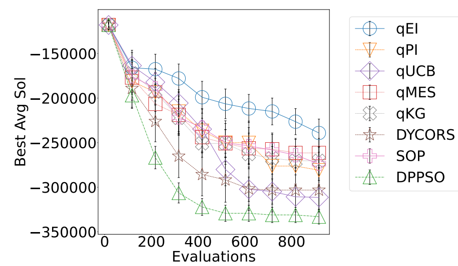

Figure 2: DNA Binding Optimization Problem ($d = 8$).

To address the inefficiency issue, we propose various approaches including the incorporation of dynamic candidate generation methods and parallel computation. Our dynamic candidate generation approach, DEEPA (Nezami and Anahideh 2023), enhances exploration by perturbing dimensions based on their importance derived from prior data. DEEPA employs a Pareto batch sampling strategy, although it does not inherently guarantee diversity within the batch and may face scalability issues with large batch sizes as many of the existing SO approaches. On the other hand, the parallelization of DPP not only enhances the computational efficiency of DPP-based strategies but also empowers them to contend with the computational demands posed by high-dimensional optimization tasks. This parallelization holds the potential to revolutionize the practicality and scalability of DPP-based Surrogate Optimization (DPPSO), ushering in a new era of efficiency and competitiveness in tackling multifaceted optimization challenges.

Moreover, similar to other SO algorithms, our proposed strategies include different algorithmic hyperparameters including perturbation radius and surrogate models' hyperparameters, which require initialization or manual tuning. The choice of hyperparameter configuration further complicates the accessibility of any SO algorithms for practitioners. Therefore, we propose a self-adjusting framework that enables the adjustment of critical hyperparameters within the SO procedure. Lastly, we seek adaptable methodologies to accommodate both continuous and categorical dimensions in real-world problems, thereby advancing the landscape of surrogate-based optimization.

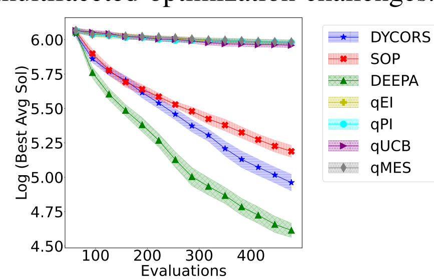

Figure 3: Rastrigin Global Optimization Problem ($d = 30$).

REFERENCES

Anahideh, H., J. Rosenberger, and V. Chen. 2022. "High-dimensional Black-box Optimization Under Uncertainty". *Computers & Operations Research* 137:105444.

Bıyık, E., K. Wang, N. Anari, and D. Sadigh. 2019. "Batch Active Learning Using Determinantal Point Processes". arXiv preprint arXiv:1906.07975, Accessed August 30th 2023.

Nezami, N., and H. Anahideh. 2023. "Dynamic Exploration-Exploitation Pareto Approach for High-Dimensional Expensive Black-Box Optimization". Available at SSRN 4382756, Accessed August 30th 2023.