Optimizing solvers for real-world expensive black-box optimization with applications in vehicle design
Long, F.X.

Chapter 1

Introduction

In general, Black-Box Optimization (BBO) problems refer to a class of optimization problems, where the objective and constraint functions are typically available in the form of black boxes [6]. In real-world applications, for instance, the inputs of BBO problems can be evaluated using laboratory experiments and/or virtual simulations to obtain the corresponding outputs, where an analytical understanding of the underlying process is still lacking. For solving BBO problems, various optimization algorithms have been developed over the years, such as the nature-inspired Evolutionary Algorithm (EA) like Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) [45], swarm-based algorithms like Particle Swarm Optimization (PSO) [55], and surrogate-assisted algorithms like Bayesian Optimization (BO) [85]. Given that different algorithms have their own strengths and weaknesses, there is no such universal optimization algorithm that can effectively solve all kinds of BBO problems [156]. Consequently, identifying the most time- and resource-efficient algorithm for specific BBO problems is critical to ensure an optimal optimization performance, which is also commonly known as the Algorithm Selection Problem (ASP) [112]. Nevertheless, selecting the most suitable optimization algorithm can be extremely challenging, especially for practitioners inexperienced in the optimization field and lack of domain knowledge.

One of the ongoing research directions towards automated Algorithm Selection (AS) in the evolutionary computation community focuses on associating the performance of optimization algorithms with the optimization landscape of BBO problems. Essentially, based on some landscape analysis tools, the optimization landscape characteristics of BBO problems can be numerically quantified, such as global structure and multi-modality [80]. The so-called landscape features can be related with the performance of optimization algorithms, e.g., using Machine Learning (ML) approaches [14, 49, 58]. By employing supervised ML models, for instance, the performance of optimization algorithms on the BBO problems can be roughly estimated based on their landscape characteristics. Subsequently, the most suitable optimization algorithm for unseen BBO problems can be conveniently identified, once their optimization landscape characteristics have been computed.

In previous work like [58], landscape-aware AS approaches have shown encouraging success on BBO benchmark suites, such as the Black-Box Optimization Benchmarking (BBOB) functions [44]. When using benchmark suites, a large set of problem instances necessary for the training of reliable ML models can be easily generated and cheaply evaluated. On the other hand, little work has been attempted to investigate landscape-aware AS for BBO in real-world applications with expensive function evaluations. Compared to benchmark suites, the generation of real-world problem instances can be particularly limited, the landscape characteristics of real-world problems are not well understood, and the function evaluations are prohibitively costly and/or time-consuming. Consequently, generating a large training dataset for ML models is infeasible, and thus, typical AS approaches are not practical for solving real-world expensive BBO problems. To the best of our knowledge, exploration of suitable AS approaches for real-world expensive BBO problems is still lacking.

1.1 Thesis Objective

Motivated to fill the gaps, the primary objective of this thesis is to develop an automated optimization approach for an optimal solving of real-world expensive BBO problems w.r.t. real-world constraints, such as allocated time and resources available for optimization. Precisely, our aim is to investigate an automated approach that can optimally fine-tune optimization algorithms for solving expensive BBO problems using a limited function evaluation budget. Eventually, the proposed optimization pipeline can be deployed to assist practitioners unfamiliar with AS in properly fine-tuning optimization algorithms for their applications.

The general workflow of our proposed optimization pipeline is illustrated in Figure 1.1, consisting of altogether four steps. Fundamentally, by analyzing the optimization landscape characteristics of expensive BBO problems, some test functions that are (i) *cheap-to-evaluate* and (ii) belong to the *same optimization problem class* can be identified. Essentially, these test functions, which we refer to as *representative func-*

tions, can be considered as *surrogates* of the BBO problems. Unlike in a ML context, we are not interested in replacing or approximating the true functions of expensive BBO problems using some representative functions. Instead, we aim to establish a preferably large set of test functions with similar landscape characteristics, which can be considered for the fine-tuning of optimization algorithms, since the performance of optimization algorithms should be comparable between the actual BBO problems and their representative functions. By exploiting these cheap representative functions, near-optimal optimization algorithms for the BBO problems can be identified at a significantly lower computational cost, prior to the actual optimization runs using expensive function evaluations. In other words, a much more flexible function evaluation budget can be afforded for the fine-tuning of optimization algorithms based on some representative functions. Eventually, the actual BBO problems can be effectively solved using the identified near-optimal optimization algorithms.

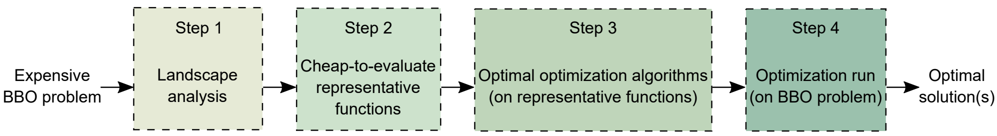

Figure 1.1: Outline of our automated optimization pipeline proposed for solving expensive BBO problems using a limited function evaluation budget. Principally, the optimization landscapes of expensive BBO problems are analyzed (Step 1) to identify cheap-to-evaluate representative functions (Step 2). By exploiting these representative functions, near-optimal optimization algorithms can be cheaply identified (Step 3) and then applied to solve the BBO problems (Step 4).

Within the scope of this thesis, we consider crashworthiness optimization problems in the automotive industry as representative real-world engineering BBO problems with expensive function evaluation. As vehicle design becomes increasingly sophisticated nowadays, solving crashworthiness optimization problems using conventional optimization approaches is getting notoriously challenging and tedious, which start losing their effectiveness. Subsequently, alternatives that can provide better vehicle designs, and thus, accelerate the overall vehicle development progress are gaining more attention in the automotive industry. Compared to conventional optimization approaches for solving automotive crash problems, we aim to achieve a better performance using our proposed optimization pipeline in terms of:

- Using less computational resources w.r.t. expensive function evaluations;
- Being less time-consuming w.r.t. total wall-clock time; and/or

- Finding better solution(s) w.r.t. optimization objective.

Despite automotive crashworthiness optimization problems are being considered as representative real-world applications in this thesis, the potential of our pipeline is not limited to them. In fact, the pipeline is designed and developed with the long-term vision that it can be applied to other expensive BBO problems beyond automotive crashworthiness optimization. Principally, the proposed pipeline could be further extended to other engineering applications with appropriate modifications, such as vehicle control system calibration in the automotive industry [131], ship design in the maritime industry [25], and turbomachinery design in the aerospace industry [103].

Beyond that, another crucial aspect of our proposed optimization pipeline is that optimal optimization algorithms can be identified automatically, requiring minimal input from practitioners. From our perspective, this is essential for real-world applications and can potentially reduce the workload of practitioners, especially for those that are unfamiliar with fine-tuning of optimization algorithms. Interestingly, this is in line with the latest trend, where the so-called *one-click* optimizers are getting introduced in commercial optimization tools [115].

1.2 Research Questions

During the journey to achieve the research objective defined for this thesis, i.e., developing an automated optimization pipeline for optimally solving expensive BBO problems, we focus on addressing the following research questions. In this context, our investigations are primarily based on the widely-used BBOB functions and automotive crashworthiness optimization as a representative example of real-world expensive BBO problems, refer to Section 1.1.

RQ1: Within the feature space defined by optimization landscape features, how are the distributions of real-world expensive BBO problems situated w.r.t. some benchmark functions?

In the landscape-aware AS approach, the optimization landscape characteristics of BBO problems are exploited to identify optimal optimization algorithms. To the best of our knowledge, however, an investigation about the optimization landscape of real-world expensive BBO problems is still lacking. Thus, what are the optimization problem classes of real-world BBO problems? Is there any similarity between real-world BBO problems and benchmark functions in terms of optimization landscape characteristics? To what extent can benchmark functions be considered as representative functions for real-world BBO problems? (Chapter 3)

RQ2: If none of the benchmark functions can sufficiently represent the optimization problem classes of real-world expensive BBO problems, how to augment the benchmark functions with other test functions that are appropriate to serve as representative functions?

Despite the fact that benchmark functions cover a variety of optimization problem classes, a sufficient coverage of the problem classes in real-world applications might still be lacking. For such cases, how can we complement the set of benchmark functions with more test functions to further improve the overall diversity of optimization problem classes? More importantly, are any of these newly included test functions appropriate to be considered as representative functions for the fine-tuning of optimization algorithms? (Chapter 4)

RQ3: How well can we estimate the actual performance of optimization algorithms on real-world expensive BBO problems based on some cheap-to-evaluate representative functions?

In our optimization approach, we propose to optimally fine-tune optimization algorithms for real-world expensive BBO problems, by exploiting some cheap-to-evaluate representative functions from the same optimization problem classes. In this context, how well can we predict the performance of different optimization algorithms on unseen BBO problems based on some representative functions? More critically, can we identify optimal optimization algorithms for BBO problems in this way? Moreover, are all representative functions appropriate for the fine-tuning of optimization algorithms? Otherwise, how can we properly identify those that are appropriate for this task? (Chapter 4)

RQ4: How to specifically generate test functions that belong to the same optimization problem classes of real-world expensive BBO problems?

Since the hand-crafted benchmark function suites can never fully cover all optimization problem classes, it is valuable to have an alternative that can generate test functions belonging to specific optimization problem classes. In this regard, how can we guide the generation of test functions towards a particular optimization problem class, e.g., based on some optimization landscape characteristics? What is the potential of such approach, and if any, the challenges that must be overcome? (Chapter 4)

RQ5: What is the performance of optimal optimization algorithms identified using the proposed optimization approach (Section 1.1), when tested on some benchmark functions? More crucially, can the optimal optimization algorithms outperform some state-of-the-art approaches for solving real-world expensive BBO problems?

In this thesis, we propose to fine-tune optimization algorithms for real-world expensive BBO problems based on some cheap-to-evaluate representative functions. From this perspective, what is the performance of such optimal optimization algorithms in general? In other words, can we identify well-performing optimization algorithms in this way? Eventually, can we optimally solve real-world expensive BBO problems using the fine-tuned optimization algorithms w.r.t. some real-world constraints? (Chapter 5)

RQ6: What is the potential of pre-trained general purpose predictive models in identifying optimal optimization algorithms for real-world expensive BBO problems?

Similar to a landscape-aware ASP context, we are inspired to explore an alternative that can identify optimal optimization algorithms for real-world expensive BBO problems using some predictive models. In this regard, how can we improve the performance of predictive models to preferably generalize well across different applications? Compared to typical landscape-aware ASP, how can we take it one step further, towards selecting optimal algorithms as well as fine-tuning their hyperparameters? (Chapter 5)

1.3 Thesis Outline

In summary, the structure of this thesis is highlighted in the following:

Chapter 2 briefly presents a literature review and additional information that are essential for comprehending and gaining insights into different research topics that have been explored within the framework of this thesis, such as a gentle introduction to BBO, fine-tuning of algorithm configurations, fitness landscape analysis using Exploratory Landscape Analysis (ELA), and automotive crashworthiness optimization.

Chapter 3 first introduces an approach for capturing the optimization landscape characteristics of BBO problems based on some ELA features. Apart from the classical ELA features, we also explore a feature-free alternative that can characterize BBO problems based on some latent representations, e.g., computed using deep Neural Network (NN) models. To have a better understanding, we also take a closer look at a large set of BBOB instances w.r.t. their optimization landscape characteristics and optimization performances, since the BBOB suite is heavily utilized in our investigations. Lastly, we analyze the optimization landscape of several real-world automotive crashworthiness optimization problems in terms of ELA features, based on a comparison against those of the BBOB functions.

Chapter 4 summarizes the generation of cheap-to-evaluate representative functions for real-world expensive BBO problems using a tree-based random function generator. Apart from the fact that such representative functions are similar in terms of optimization landscape, we show that they can be exploited for estimating the actual performance of optimization algorithms on unseen BBO problems. Beyond that, we evaluate the potential of guiding the function generation towards specific optimization problem classes based on some ELA features, by extending the random function generator with Genetic Programming (GP).

Chapter 5 describes in detail the automated optimization pipeline proposed for optimally solving real-world expensive BBO problems. For a comprehensive performance assessment, the proposed approach is first evaluated across a wide range of optimization problem classes using the BBOB suite. More importantly, the approach is applied and tested on a real-world automotive crashworthiness optimization problem, using expensive Finite Element (FE) simulation runs. Furthermore, we investigate the potential of training predictive ML models that can efficiently identify optimal algorithms for BBO problems, potentially improving the overall effectiveness of the proposed approach.

Chapter 6 summarizes significant research findings of this thesis, discusses current limitations, and provides potential improvements for future work.

1.4 Author's Contributions

Within the scope of this thesis, all important scientific findings and results have been previously published and contributed as part of the research community, as summarized in the following:

Journal Publications

1. Fu Xing Long, Bas van Stein, Moritz Frenzel, Peter Krause, Markus Gitterle, and Thomas Bäck. Generating Cheap Representative Functions for Expensive Automotive Crashworthiness Optimization. *ACM Trans. Evol. Learn. Optim.*, 4(2), jun 2024.

2. Fu Xing Long, Niki van Stein, Moritz Frenzel, Peter Krause, Markus Gitterle, and Thomas Bäck. Surrogate-based automated hyperparameter optimization for expensive automotive crashworthiness optimization. *Struct. Multidiscip. Optim.*, 68(4), April 2025.

Peer-reviewed Conference Publications

1. Fu Xing Long, Bas van, Stein, Moritz Frenzel, Peter Krause, Markus Gitterle, and Thomas Bäck. Learning the characteristics of engineering optimization problems with applications in automotive crash. *In Proceedings of the Genetic and Evolutionary Computation Conference, GECCO '22*, page 1227–1236, New York, NY, USA, 2022. Association for Computing Machinery.*

2. Fu Xing Long, Diederick Vermetten, Bas van Stein, and Anna V. Kononova. BBOB Instance Analysis: Landscape Properties and Algorithm Performance Across Problem Instances. *In Applications of Evolutionary Computation: 26th European Conference, EvoApplications 2023*, Held as Part of EvoStar 2023, Brno, Czech Republic, April 12–14, 2023, Proceedings, page 380–395, Berlin, Heidelberg, 2023. Springer-Verlag.†

3. Bas van Stein, Fu Xing Long, Moritz Frenzel, Peter Krause, Markus Gitterle, and Thomas Bäck. DoE2Vec: Deep-Learning Based Features for Exploratory Landscape Analysis. *In Proceedings of the Companion Conference on Genetic and Evolutionary Computation, GECCO '23 Companion*, page 515–518, New York, NY, USA, 2023. Association for Computing Machinery.

4. Fu Xing Long, Diederick Vermetten, Anna V. Kononova, Roman Kalkreuth, Kaifeng Yang, Thomas Bäck, and Niki van Stein. Challenges of ELA-Guided Function Evolution Using Genetic Programming. *In Proceedings of the 15th International Joint Conference on Computational Intelligence - Volume 1: ECTA*, pages 119–130. INSTICC, SciTePress, 2023.

*Best Paper Award
†Outstanding Students of 2023

5. Fu Xing Long, Niki van Stein, Moritz Frenzel, Peter Krause, Markus Gitterle, and Thomas Bäck. Surrogate-Based Algorithm Selection and Hyperparameter Tuning for Automotive Crashworthiness Optimization. *In The 15th World Congress of Structural and Multidisciplinary Optimisation*, 2023.

6. Roy de Winter, Fu Xing Long, Andre Thomaser, Thomas H.W. Bäck, Niki van Stein, and Anna V. Kononova. Landscape analysis based vs. domain-specific optimization for engineering design applications: A clear case. *In 2024 IEEE Conference on Artificial Intelligence (CAI)*, pages 776–781, 2024.

7. Fu Xing Long, Moritz Frenzel, Peter Krause, Markus Gitterle, Thomas Bäck, and Niki van Stein. Landscape-Aware Automated Algorithm Configuration Using Multi-output Mixed Regression and Classification. *In Parallel Problem Solving from Nature – PPSN XVIII*, pages 87–104, Cham, 2024. Springer Nature Switzerland.