Exploring Superior Structural Materials Using Multi-Objective Optimization and Formal Techniques

(Invited Paper)

Rolf Drechsler*† Stephan Eggersglüß*† Nils Ellendt‡§ Sebastian Huhn*† Lutz Mädler‡§

*Institute of Computer Science
University of Bremen
28359 Bremen, Germany
{drechsle,segg,huhn}
@informatik.uni-bremen.de

†Cyber-Physical Systems
DFKI GmbH
28359 Bremen, Germany

‡Institute of
Materials Science
University of Bremen
28359 Bremen, Germany
ellendt@iwt.uni-bremen.de

§Department of
Production Engineering
University of Bremen
28359 Bremen, Germany
lmaedler@iwt.uni-bremen.de

***Abstract*—The continuous progress in engineering requires increasingly powerful materials being utilized in complex constructions. Commonly used methodologies are hardly capable to explore materials with improved properties, i.e., fulfilling a certain performance profile. An alternative high throughput screening approach exists which allows to process a high number of samples while operating on micro scale. For structural materials, it is not possible to project material properties from the micro to the macro scale. Due to the complex nexuses between the process parameters (applied for sample synthesis), the high-dimensional space of the screening data and the uncertainties concerning material properties while scaling, conventional algorithms are not capable to cope with these data. Thus, this work proposes a general data flow which orchestrates multi-objective optimization approaches as well as formal techniques to handle this challenging computational task. In the end, a framework is drawn that determines the resulting material properties on a macro level by using screening data of micro samples and compares them against a performance profile. For the case that only a slight correlation exists, the framework proposes a batch of alternate process parameters which are, highly probable, leading to superior structural material.**

I. Introduction

Within the last decades, a major engineering progress was achieved that influences all fields of our daily life. In particular, a progress in fields such as safety, mobility, infrastructure and avionics requires steadily improved high-performance structural materials. In these domains, the materials have to exhibit specific material properties, e.g., the elastic limit or the elongation which are specified by a performance profile. The investigated material can only be utilized in a complex construction if a certain performance profile is fulfilled. The exploration of such powerful materials is a challenging, time consuming and highly expensive procedure. Due to this high consumption of resources, only a limited number of newly synthesized materials can be investigated such that established conventional search methods generally do not lead to breakthroughs but rather to slightly improved materials.

An alternative, more promising approach was proposed in [1] that addresses the exploration of new structural materials by a high throughput screening technique. This approach is applied on micro samples that are made with a specific alloy and, subsequently, treated in a certain way. In fact, this heat treatment develops different microstructures leading to specific material properties. The specification of the alloy and the treatment are specified by the namely process parameters. The approach of [1] is operating on a micro level, i.e., only micro samples are considered which enables the high throughput character. The screening data include various measurements, e.g., concerning electrochemical characteristics, X-ray diffraction and novel mechanics tests of the probes that are repeated multiple times for each micro sample to be examined. Due to the high throughput character, these multidimensional measurements result in a huge data set that has to be analyzed to identify promising process parameters which should finally lead to a material candidate that fulfills the desired properties with a high probability. Eventually, the identified parameters are used to produce these improved materials on a macro level such that they can be utilized for practical applications.

For this high throughput approach, four computational aspects have to be considered:

1) No direct correlation between the data which are measured by the screening process and the resulting material properties exist, hence, a method has to be developed which allows to predict these material properties out of the screening data.
2) The screening is applied on micro samples that enables the high throughput character. In contrast to other materials, e.g., functional materials [2], it is not possible to scale the properties of a structural material from the micro level to the macro level. Thus, a mechanism is required which

This work is part of the *Collaborative Research Center* SFB1232: Farbige Zustände (high throughput exploration for evolutionary structural materials). It has started on July 1st (2016) and is funded by the German Research Foundation. The project is planned to run for 12 years, with evaluations after 4 and 8 years.

978-1-5090-2541-1/16/$31.00

allows to project the properties of the micro to the macro level.

3) The actual objective is to find structural materials fulfilling a certain performance profile. Consequently, the methodology has to establish a multidimensional comparison between a material – that is currently under investigation – against a given performance profile. Thus, a global function is realized that indicates the match between the given material properties and the one of the performance profile.
4) In the case that the properties of the investigated micro sample and the ones that are specified by the performance profile are not well matching, an algorithm has to be implemented which identifies process parameters leading to a micro sample with a higher matching rate of the material properties with respect to the performance profile.

The proposed approach addresses all these mentioned shortcomings by proposing a methodology that

- predicts resulting material properties from processed screening data,
- derives knowledge between the measured data on micro level and the resulting material properties on macro level,
- implements high-dimensional relational operators to compare the screening data against each other by using *Multi-Objective Optimization* (MOO) techniques,
- determines promising process parameters for improved materials by invoking a formal optimization-based search procedure,
- allows to specify a hypothesis system by using a newly invented domain-specific language which allows to embed expert knowledge and validates this by using the collected experimental data.

The structure is as follows: Section II gives a brief introduction to the high throughput screening approach. Followed by Section III which provides an overview of the proposed data flow. The main components namely the predictive function as well as the search procedure are described in Section IV. Finally, some conclusions are drawn in Section V.

II. High Throughput Approach

Generally, high throughput techniques try to address the identification of nexuses in highly complex spheres with multiple uncertainties. For instance, one established field of application concerns the nanoparticle toxicology [3]. Accordingly, in the domain of material exploration, the high throughput approach comprises a high number of samples which are synthesized, treated and, subsequently, applied to various measurements to generate the screening data.

Beside the pure data generation, a general flow has to be developed which processes the data and controls the high throughput system. Figure 1 presents this flow consisting of the following components:

1) Step A demonstrates the synthesis, the treatment as well as the screening of micro samples enabling a highspeed processing that leads to high throughput screening.
2) As depicted, correlated experiments on the micro and the macro level exist which are used as the training data set (shown at the very top) and, hence, construct the predictive function. On demand, selected materials are processed on macro scale which is very time as well as resource consuming and, therefore, should be only done for promising candidates or to address remaining uncertainties in the predictive function.
3) The main iterative loop starts at the determined screening data of specific micro sample [cf. 1)] that is investigated.
4) The predictive function is applied on these screening data to determine the resulting material on macro scale with the highest probability.
5) Multiple relational operators are utilized to compare the material properties against a given performance profile.
6) Based on this qualitative evaluation, several experimental designs are generated which control the micro sample generation. As intended, this feedback will potentially produce micro samples for next iterations [cf. 1) to 6)] which are better matching against the performance profile.
7) If a certain quality threshold is reached, a macro sample is generated to, finally, validate the exploration of a new structural material fulfilling the specific performance profile. This generation is shown in step B and contains, in general, the same process steps as the micro sample generation.

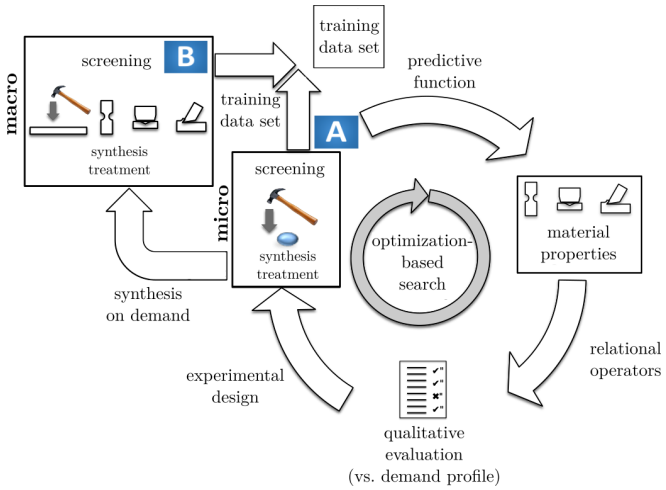

Fig. 1: High throughput screening

III. Data Flow

This section describes the proposed data flow for the framework which allows to determine promising process parameters of improved structural materials, fulfilling a challenging and application-specific performance profile that is defined by the user with respect to its later application.

The set of screening data builds up the experimental database and can contain multiple types of measurements, e.g., the current setting allows eleven different types. Consequently, these measurements have to be projected to various resulting material properties. For the investigated materials with respect to required performance profiles, six different material properties are currently considered. Hereby, a type of measurement does not necessarily affect each material property. For instance, a dilatometry measurement may not be used to infer properties

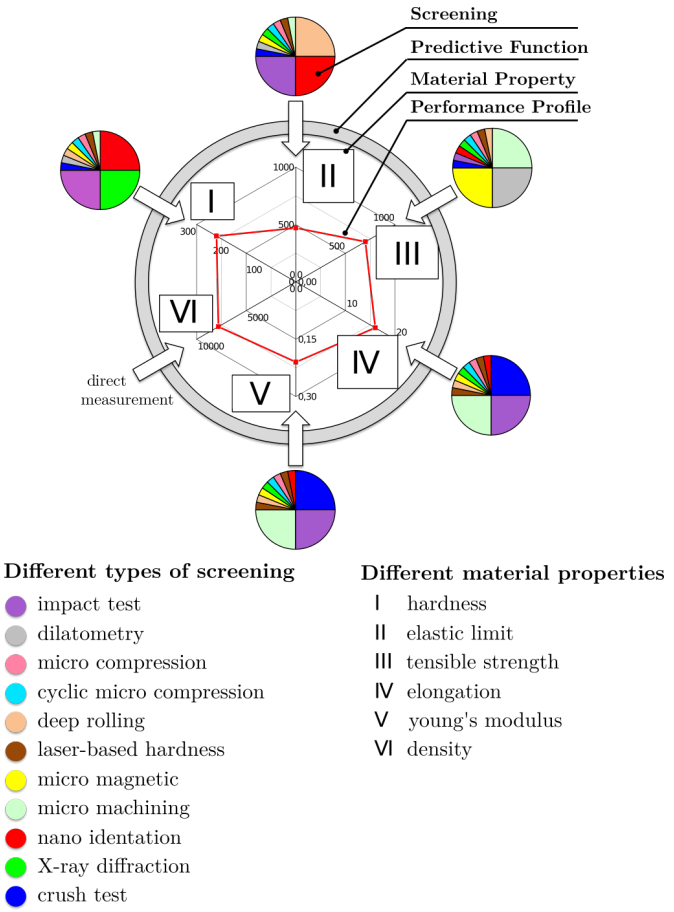

Fig. 2: Possible types of screening data & material properties

regarding the hardness. A possible connection between these data are shown in Figure 2.

The data flow is presented in Figure 3 and includes Steps A to E as described in the following paragraphs.

A: Given are a desired performance profile, which should be fulfilled by the material on macro scale, and an initially constructed *predictive function* Ψ which projects screening data to resulting material properties.

B: Different micro samples have been produced by specific alloys and have been treated as defined by the process parameters. Subsequently, these micro samples are applied to the screening process. All measurements are executed multiple times such that possible process variations can be compensated and, finally, the screening data are stored in the experimental database.

C: The predictive function Ψ receives the performance profile as well as the experimental database and projects the measured screening data of a micro sample to the resulting material property of this material on macro scale.

D: An optimization-based search process is invoked which aims to determine a good candidate, i.e., a set of screening data. Subsequently, Ψ is applied on this candidate, i.e., a promising set of screening data, to predict the resulting material properties that coincide with the performance profile. The optimization finds the best candidate with respect to certain search limits, thus, the quality of a single candidate is evaluated. For this evaluation different, newly developed relational operators are utilized. In fact, the comparison for each material property is implemented by an individual operator.

E: Due to the high dimension of screening data and the material property, the possible search space is tremendous. Consequently, a domain-specific language will be developed which allows experts to formulate so-called hypotheses, e.g., assertions and assumptions about the screening data or material properties, respectively. The concept of a domain-specific language is frequently used to address complex scenarios necessitating user interactions, e.g., as demonstrated in [4], [5]. The user-defined hypotheses are condensed into a system of hypothesis conducting the search procedure by expert knowledge towards a promising candidate.

Listing 1 gives some exemplary hypotheses (as described in step E). The variables D_i refer to the screening data – namely descriptors – for screening type i and the variables M_j refer to a certain material property j, respectively. This example shows the declaration of two assumptions and three assertions concerning material properties as well as a descriptor. Furthermore, it demonstrates a few supported operations concerning the peak calculation and an expected linear correlations between a descriptor D_1 and a material property M_5.

Listing 1: Expert knowledge encoded in hypotheses

```
ASSUME:  D1 > 15
ASSUME:  M5 > 30

ASSERT:  linear_correlation( D1, M5 )
ASSERT:  max( M3, M4, M5 )
ASSERT:  min( M1, M2, M3 )
```

Eventually, if a screening data set is identified that matches best against a certain performance profile, it is possible to derive the process parameters for producing a material leading to such screening data. For compensating possible process variations and handling the existing uncertainties, this method determines instead of a single set of process parameters, a batch of candidates. These candidates are calculated by using well-known approaches related to the statistical experimental design.

IV. Algorithmic Approaches

This section describes the algorithmic approaches to predict resulting material properties as well as the top-level procedure which implements the actual search for improved materials.

A. Predictive Function

The predictive function Ψ is one of the main mechanisms within the proposed data flow and determines the resulting material properties on macro level by a given set of screening data for a micro sample.

As assumed in step A, Ψ is initially constructed. For this construction, a training data set is given which consists of correlated experiments on micro as well as macro scale with known material properties. Within this first construction phase,

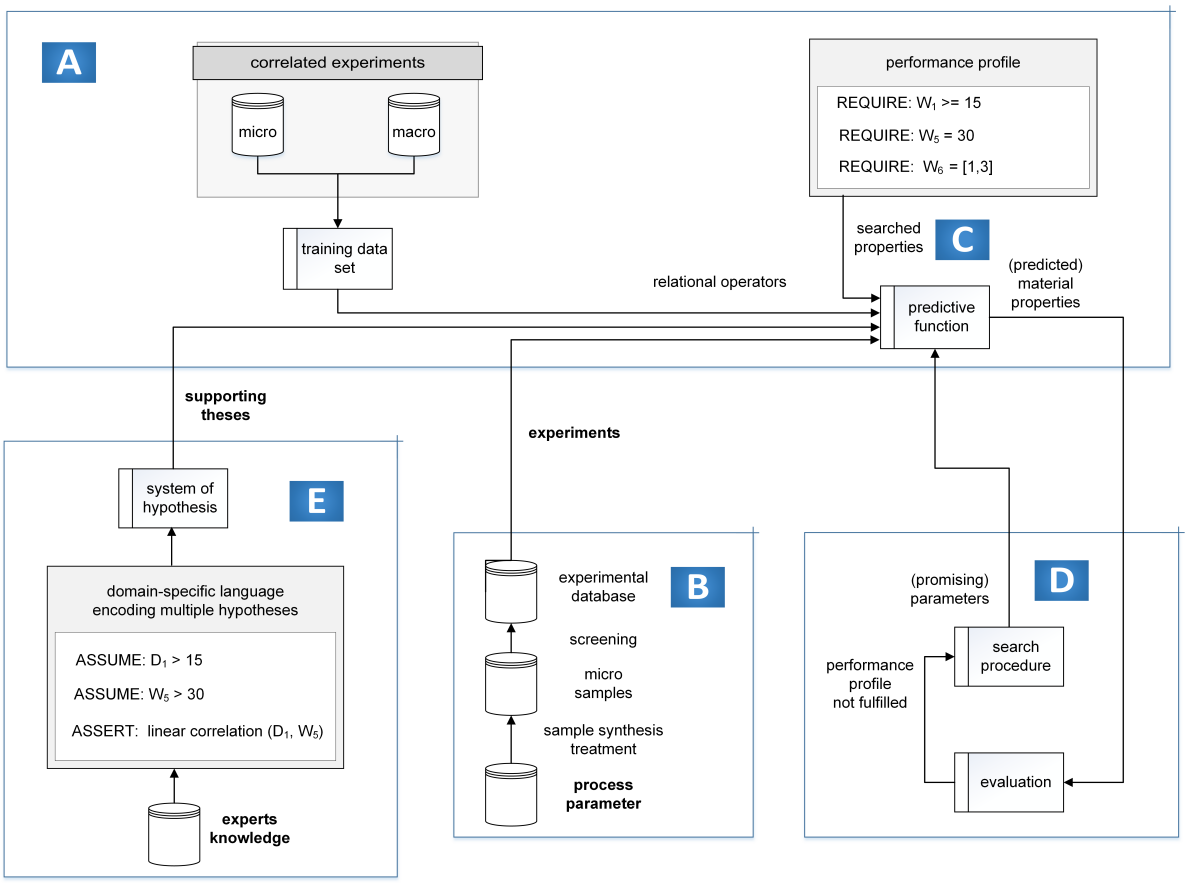

Fig. 3: General data flow

newly developed relational operators are utilized for the comparison between high dimensional material properties. These relational operators are implemented by taking advantage of MOO techniques, e.g., the ϵ-operator proposed in [6], and are crucial for comparison as already stated in step D. Due to the specific characteristics of different material properties, an individual relational operator is developed for each of these properties. After determining this set of relational operators, MOO is also applied on this data set to, finally, generate a parameterized operator. The parameterized operator considers the implicitly given weight of the individual property, e.g., as stated in work [7], and is directly utilized by the predictive function. Figure 4 shows the initial construction procedure as well as the later application of the predictive function where it is applied on screening data to predict resulting material properties.

Beside this, it is possible to utilize further expert knowledge encoded within a system of hypothesis as described in step E. For evaluation and validation of the system of hypotheses, formal techniques are considered, e.g., Bounded Model Checking [8], which have been frequently used for verification of integrated circuits [9], [10] or for the automatic test pattern generation as proposed in work [11].

After the first training phase of Ψ has finished, a further training is possible to achieve a continuous improvement of the prediction quality, thus, Ψ can be improved by processing more correlated experiments.

B. Optimization-based Search Procedure

Beside the predictive function Ψ, one further important component is the optimization-based search procedure. This procedure uses Ψ internally and models the actual iterative loop within the data flow as shown in Figure 3 and, finally, leads to the exploration of new materials with specific properties.

As depicted, Ψ is capable to determine the resulting material properties for a material whose screening data are given. For the exploration of new high-performance materials, a set of screening data have to be identified which result in these material properties – as predicted by Ψ and specified by the performance profile.

The search process determines a new data set and evaluates its quality, i.e., all resulting material properties of this candidate are compared with the corresponding one of the performance profile. For all considered material properties (as shown by Figure 2), an individual relational operator is used that implements a suitable comparator. Subsequently, a formal optimization procedure starts which aims to maximize the accumulated quality of the sample. The degree of freedom for this optimization is tremendous: Generally, various types of screening data can be altered, leading to a high dimensional search space. This hard computational task can be addressed by powerful probabilistic approaches such as evolutionary algorithms. Following this scheme of applying a heuristic search procedure, e.g., an evolutionary algorithm, for optimizing certain parameters as proposed in [12], the data

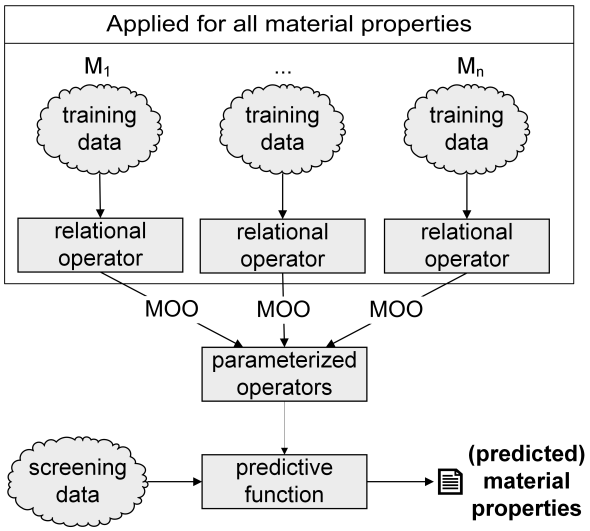

Fig. 4: Creation and application of predictive function

have to be interpreted as following: The overall search space is spanned by the screening data and the search operation consists of altering the measurements. Furthermore, the fitness function is implemented by a distance operation between deviation of the currently predicted material property and the ones specified by the performance profile. As proposed in [13], it is possible to prioritize single objectives for the optimization, i.e., some properties are more important than others. For instance, while optimizing the *promising* parameters for synthesizing an improved material, the resulting tensible strength could be more important than the hardness of the resulting material. Consequently, algorithms are required which support multiple-objectives for the optimization [14], e.g., algorithms that have been successfully applied in [15], [16].

This utilization of algorithmic approaches such as evolutionary algorithms establishes a data flow which is able to cope with the high dimensional screening data and, furthermore, optimize them against an objective target, determined by the resulting material properties and the ones of the performance profile.

V. Conclusions

This work addresses the challenging task of exploring new high-performance structural materials by utilizing the high throughput screening approach proposed in [1]. Following this scheme, a framework is drawn which is capable to address the hard computational task: On the one hand a huge set of multidimensional screen data have to be processed and on the other hand various uncertainties have to be considered.

The Collaborative Research Center SFB1232 aims to orchestrate multi-objective optimization approaches effectively such that a predictive function is implemented that predicts resulting material properties on macro scale by processing screening data of micro samples. Additionally, a set of relational operators is developed which are capable to compare certain material properties. Beside this, formal techniques are combined with these developed functions which leads to an optimization-based search procedure. This functional composition determines a candidate of screening data best matching against a specific performance profile. Eventually, experimental designs will be derived by these identified screening data to, finally, explore and produce high-performance structural materials.

VI. Acknowledgement

Financial support of subproject P01 'Predictive function' and P02 'Experimental design' of the Collaborative Research Center SFB1232 by the German Research Foundation (DFG) is gratefully acknowledged.

References

[1] L. Mädler, "Is high-throughput screening for structural materials/metals possible?" in *International Conference on Nanomanufacturing*, 2014.

[2] A. Zakutayev, F. J. Luciano, V. P. Bollinger, A. K. Sigdel, P. F. Ndione, J. D. Perkins, J. J. Berry, P. A. Parilla, and D. S. Ginley, "Development and application of an instrument for spatially resolved seebeck coefficient measurements," *Review of Scientific Instruments*, vol. 84, no. 5, 2013.

[3] H. Zhang, Z. Ji, T. Xia, H. Meng, C. Low-Kam, R. Liu, S. Pokhrel, S. Lin, X. Wang, Y.-P. Liao, M. Wang, L. Li, R. Rallo, R. Damoiseaux, D. Telesca, L. Mädler, Y. Cohen, J. I. Zink, and A. E. Nel, "Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation," in *ACS Nano*, vol. 6, no. 5, 2012, pp. 4349–4368.

[4] M. Soeken and R. Drechsler, "Grammar-based program generation based on model finding," in *IEEE Design and Test Symposium*, 2013, pp. 1–5.

[5] R. I. McKay, N. X. Hoai, P. A. Whigham, Y. Shan, and M. O'Neill, "Grammar-based genetic programming: a survey," *Genetic Programming and Evolvable Machines*, vol. 11, no. 3, pp. 365–396, 2010.

[6] N. Drechsler, A. Sülflow, and R. Drechsler, "Incorporating user preferences in many-objective optimization using relation ϵ-preferred," in *Natural Computing*, vol. 14, no. 3, 2015, pp. 469–483.

[7] F. di Pierro, S. T. Khu, and D. A. Savic, "An investigation on preference order ranking scheme for multiobjective evolutionary optimization," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 1, pp. 17–45, 2007.

[8] A. Biere, A. Cimatti, E. Clarke, and Y. Zhu, *Symbolic Model Checking without BDDs*. Springer, 1999, pp. 193–207.

[9] G. Fey, A. Sulflow, S. Frehse, and R. Drechsler, "Effective robustness analysis using bounded model checking techniques," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 8, pp. 1239–1252, 2011.

[10] E. Clarke, A. Biere, R. Raimi, and Y. Zhu, "Bounded model checking using satisfiability solving," *Formal Methods in System Design*, vol. 19, no. 1, pp. 7–34, 2001.

[11] S. Eggersglüß, R. Wille, and R. Drechsler, "Improved SAT-based ATPG: More constraints, better compaction," in *International Conference on CAD*, 2013, pp. 85–90.

[12] H.-P. S. T. Bäck, "An overview of evolutionary algorithms for parameter optimization," *Evolutionary Computation*, vol. 1, no. 1, pp. 1–23, 1993.

[13] F. Schmiedle, N. Drechsler, D. Große, and R. Drechsler, "Priorities in multi-objective optimization for genetic programming," in *Conference on Genetic and Evolutionary Computation*. Morgan Kaufmann Publishers Inc., 2001, pp. 129–136.

[14] J. Branke, T. Kaußler, and H. Schmeck, "Guidance in evolutionary multi-objective optimization," *Advances in Engineering Software*, vol. 32, no. 6, pp. 499 – 507, 2001.

[15] A. Sülflow, N. Drechsler, and R. Drechsler, "Robust multi-objective optimization in high dimensional spaces," in *International Conference on Evolutionary Multi-Criterion Optimization*. Springer, 2007, pp. 715–726.

[16] W. Lei and W. Qidi, "Ant system algorithm for optimization in continuous space," in *IEEE International Conference on Control Applications*, 2001, pp. 395–400.