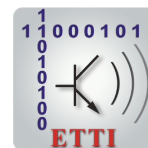
POLITEHNICA UNIVERSITY OF BUCHAREST



Doctoral School of Electronics, Telecommunications and Information Technology

Decision No. 1062 **from** 10-07-2023

Ph.D. THESIS SUMMARY

Georgian NICOLAE

APLICAȚII DE ÎNVĂȚARE AUTOMATĂ ÎN PROIECTAREA TRANZISTOARELOR MOSFET DE PUTERE

MACHINE LEARNING APPLICATIONS IN POWER MOSFET DESIGN

THESIS COMMITTEE

Prof. Dr. Ing. Mihai CIUC Politehnica Univ. of Bucharest	President
Prof. Dr. Ing. Corneliu BURILEANU Politehnica Univ. of Bucharest	PhD Supervisor
Prof. Dr. rer. nat. Georg PELZ Infineon Technologies AG & Univ. Duisburg-Essen	Referee
Prof. Dr. Ing. Marina ȚOPA Technical Univ. of Cluj-Napoca	Referee
Prof. Dr. Ing. Cristian RAVARIU Politehnica Univ. of Bucharest	Referee

BUCHAREST 2023

Table of contents

1 Introduction **1**

1.1 Designing a new generation of power MOSFETs 1
1.2 Scope of the Research 2
1.3 Motivation 2
1.4 Thesis Structure 2

2 Related Work and Theoretical Fundamentals **3**

2.1 Finite Element Analysis of Power MOSFET 3
2.1.1 Finite Element Simulation 3
2.1.2 Electro-thermal limitations of power MOSFETs 3
2.1.3 Practical considerations finite element simulation 3
2.2 General Automatic Parameter Tuning and Algorithm Configuration . . . 4
2.3 Machine Learning Models in Semiconductors 4
2.3.1 Application of Machine Learning Techniques 4
2.3.2 Applications of Regression Models 4
2.4 Automatic Design Optimization in Semiconductors 4

3 Advancing Time-Efficient Simulation of Power MOSFETs **5**

3.1 Introduction 5
3.2 Reducing Overall Simulation Time by Automation 5
3.2.1 Generic Power MOSFET Design 6
3.2.2 Aiding Tool for Simulation Parameter Tuning 6
3.2.3 Automated Simulation Setup 8
3.3 Metamodel-based Prediction of On Resistance 8
3.3.1 Overview 8
3.3.2 Simulation Dataset 8
3.3.3 Fitting Machine Learning Metamodel for R_{on} Prediction 8
3.3.4 Approaching Discrete Changes in Power MOSFET Design . . . 9
3.3.5 Metamodel Performance Evaluation 9
3.4 Experimental Results 9
3.4.1 Automatic Parameter Tuning – Experimental Scenario 9
3.4.2 Metamodel Use-case 1: *Technology A* Devices 10

3.4.3 Metamodel Use-case 2: *Technology B* Devices 11
3.5 Summary and Conclusions 12

4 Metamodel Applications for Power MOSFETs 13
4.1 Introduction 13
4.2 Sensitivity Analysis 13
4.3 Design of experiment analysis 14
4.4 Improving Parametric Yield 16
4.4.1 Metamodel-Based Yield Estimation Challenges 16
4.4.2 Proposed Method for Parametric Yield Estimation 16
4.4.3 Experimental Scenario 16
4.5 Summary and Conclusions 17

5 On Cost-aware Design of Power MOSFETs 18
5.1 Introduction 18
5.2 Proposed Automatic Design Optimization Method 18
5.2.1 Overview 18
5.2.2 Considerations on Automatic Design Optimization using Metamodels 19
5.2.3 Optimization Problem Formulation 20
5.2.4 Design Optimization for Continuous R_{on} Function 21
5.2.5 Approaching Discontinuities 22
5.2.6 Validation of the Optimal Design 22
5.3 Experimental Results 23
5.3.1 Optimization use-case 1: *Technology A* 23
5.3.2 Optimization use-case 2: *Technology B* 23
5.4 Summary and Conclusions 26

6 General Conclusions 27
6.1 Objectives and Results 27
6.2 Original contributions 28
6.3 List of Original Publications 29
6.4 List of Technical Reports 30
6.5 Future Work 30

References 31

Chapter 1

Introduction

The purpose of this thesis is to conduct an investigation into the design-related techniques that are currently in use and to provide solutions for improving those techniques in order to maximize device performance while simultaneously reducing associated costs. First and foremost, the objective is to reduce the amount of simulation time required for advanced power MOSFET analyses. A second objective is to determine a trade-off between the cost and performance of any new power MOSFET devices that are designed.

1.1 Designing a new generation of power MOSFETs

Finite element simulation is a powerful tool in the evaluation of semiconductor devices. A reliable simulation covers the physical processes that influence the performance of the MOSFET. With the physics understood, the simulator can be used to model new physical effects in the MOSFET structure. Determining the performance parameters may require further processing of the simulation results. To accomplish this, the industry uses both commercial and proprietary simulation tools.

With a simulator ready, the next step is to evaluate the performance of the novel device. A MOSFET design featuring structural improvements is developed. The analysis of the proposed structure continues with the defining limitations from a manufacturing point of view. Design rules are defined during this step. Using both theoretical and manufacturing limitations, the performance is evaluated with regard to the requirements of the applications.

At this moment the proposed design of devices is finally approved to be employed. Design and simulation engineers work in close collaboration for designing new devices based on the provided set of performance requirements. Starting from the template, multiple design iterations are performed until an optimal design is achieved.

The final stage is to make all the necessary adjustments to the new device design in order to ensure its effective manufacturing. To achieve this, an initial estimate of the yield is required.

1.2 Scope of the Research

The objective is to automate the design process for future generations of power MOSFETs. In the preceding section, a concise summary of this topic was provided. Through analysis of this, we are able to identify the challenges with the overall methodology: *(1)* Setting up the simulation environment requires considerable effort and cost; *(2)* In order to conduct simulations, the MOSFET design needs to be manually adjusted numerous times; *(3)* The finite element simulations, which are required in order to determine the performance of a MOSFET design, take manual and computational effort and are time-consuming; *(4)* The initial phases of the process of designing the new MOSFETs are carefully addressed by the design engineer based on the modeling of the physical processes; *(5)* The initial yield estimation for new MOSFETs influences the number of test chip series.

1.3 Motivation

As mentioned in the preceding sections, the process of designing new types of MOSFETs has various challenges. The design process is largely manual, even if the simulation itself is automated because design variations are manual so far. These can be overcome by the application of machine learning concepts. Techniques based on machine learning have already been successfully employed in a wide variety of different semiconductor-related tasks. The following premises serve as the basis for the proposed methods in this thesis: *A.* Data volume limitation; *B.* Data consistency; *C.* Routine tasks; *D.* Modelling alternatives.

1.4 Thesis Structure

Chapter 2 presents the fundamentals of the power MOSFET and its modelling using finite element simulation, a literature review with regard to automatic parameter tuning methods, general machine learning concepts, and a review of automatic design optimization in semiconductors.

The methods proposed in Chapter 3 aim to improve the overall simulation time necessary for an advanced analysis of power MOSFETs using metamodels while Chapter 4 presents immediate applications of the metamodels.

The focus of Chapter 5 is on the automatic design optimization of power MOSFETs with regard to the performance specifications. Hence, the proposed methods achieve a cost-effective transistor design.

Chapter 6 draws general conclusions on the proposed methods, describes the impact of the current topic and illustrates the main contributions of the author.

Chapter 2

Related Work and Theoretical Fundamentals

2.1 Finite Element Analysis of Power MOSFET

The task of numerical modeling the thermal and electromagnetic phenomena in semiconductor devices is very challenging because it requires finding a solutions to several numerical problems. Additionally, the non-linearities of the models increase level of difficulty even more.

2.1.1 Finite Element Simulation

Electro-thermal simulators [Rin01, PJS08, BIMR18] are tools that have been specifically created for determining the temperature of a device at a certain point in time while it is operating. Estimating the device's maximum internal temperature is a goal for figuring out the safe operating area, but there are additional considerations that can also reveal the thermal stress placed on the relevant structures [Nic18].

2.1.2 Electro-thermal limitations of power MOSFETs

The Safe Operating Area (SOA) refers to a section of the VDS – ID characteristic curve that denotes the range of operating voltages and currents at which a power MOSFET may be used without the risk of causing irreparable damage. Short-term safe operating area and long-term safe operating area are the two categories of SOA [HP05].

2.1.3 Practical considerations finite element simulation

Finite element simulation determines a numerical approximation of the real solution. In general, there will be an approximation error that varies in function of the mesh structure because determining the simulation result is equivalent to solving a system of equations.

2.2 General Automatic Parameter Tuning and Algorithm Configuration

A generic overview of the methods employed in *parameter tuning* is presented in [Hoo12]. The authors describe that most of the approaches fall into three classes of algorithms: *Racing Procedures*, *ParamILS*, and *Sequential Model-Based Optimization*. Then other approaches are presented with regard to application-specific tasks.

2.3 Machine Learning Models in Semiconductors

Machine Learning approaches for increasing the efficiency of the processes involved in semiconductor industry, as well as related data analysis, have gained a lot of research interest in both industry and academia [ICFQ93][Wan17]. This is because of the strong potential for both of these areas of research. In recent years, several machine learning techniques have been proposed for solving semiconductor related tasks [HHL21].

2.3.1 Application of Machine Learning Techniques

In this work, the proposed methods fall into the category *Supervised learning*, *Regression*, according to [Mur12]. Regression techniques suppose determining a function for mapping the input variables to a continuous output variable. The first requirement is to build a collection of input-output pairs structured in a dataset.

2.3.2 Applications of Regression Models

Model approximation, or metamodelling, represents a technique for reducing the runtime of these computation-intensive processes. Moreover, complex tasks which require a huge number of computationally expensive simulations are enabled by employing metamodelling techniques.

2.4 Automatic Design Optimization in Semiconductors

The methods proposed in the literature are usually of the derivative-free optimization type. Compared to deterministic optimization, this is preferred by the authors because it is less difficult to configure. Techniques can be classified based on whether a surrogate model is employed. Optimization using a surrogate model is preferred in the literature. The surrogate model in the presented works is either based on Gaussian Processes or Neural Networks.

Chapter 3

Advancing Time-Efficient Simulation of Power MOSFETs

3.1 Introduction

The required performance of the power MOSFETs is constantly rising due to their usage in a wide spectrum of modern applications. Therefore, designing the next generation of power MOSFETs has become much more challenging. Simulation is a key component, as the design process begins with modelling the transistor using finite element analysis. Ensuring high accuracy in simulation is a requirement for achieving high performance devices. Then, the trade-off between accuracy and simulation time needs to be carefully addressed.

3.2 Reducing Overall Simulation Time by Automation

Designing discrete semiconductors, such as MOSFETs, involves using high-end Computer Aided Design (CAD) tools. Among CAD tools, finite element simulators are of high importance in designing new devices and technologies. The configuration of the simulation software for the specific task needs to be addressed before starting to collect simulation results. This process is defined as simulation parameter selection or tuning.

Parameter selection, as a part of simulation, consists of an iterative manual process performed by the simulation engineer. The previous experience of the human expert and its correlation to the current task are the main factors in determining the duration of this task. An experienced simulation engineer will be capable of optimizing this manual trial and error process by empirically choosing a good starting set of parameters.

The approach for performing a simulation-based task needs to be discussed in order to assess the overall simulation time. Therefore, a simulation-based analysis consists of four main steps, as follows: *(1)* Simulation environment setup; *(2)* Design of Experiments; *(3)* Running simulation; *(4)* Post-processing results.

3.2.1 Generic Power MOSFET Design

Before setting up the simulation environment, it is important to present the power MOSFET structure. Fig. 3.1 illustrates a vertical cross-section of a generic 3D structure of a power MOSFET package, reduced to 8 main components (layers).

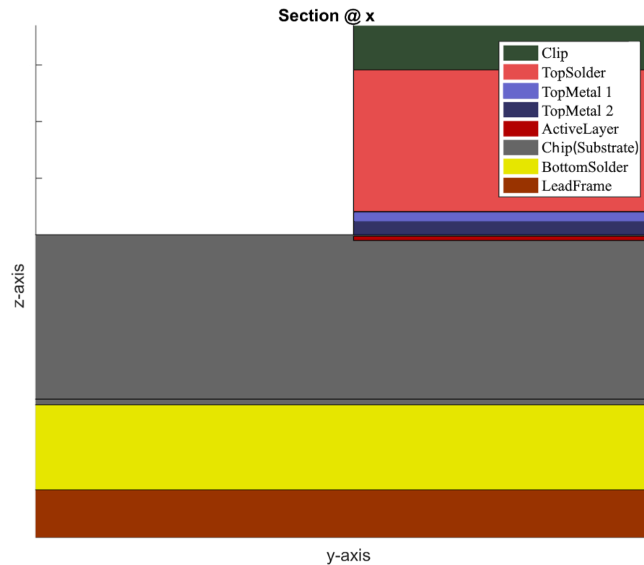


Fig. 3.1 Device subsection on x-axis [NBC+22]

Each layer of the MOSFET package is defined by length, width, thickness, and material properties. Each layer is supposed to be homogeneous for simplification purposes. Table 3.1 presents the design parameters available to the engineer. Length and width parameters are defined for the chip and the clip while thickness is defined for each layer.

Table 3.1 Design Parameters [NBF+21]

No.	Parameter name	No.	Parameter name
1	Clip Width	6	Clip Thickness
2	Clip Length	7	Top Metal 1 Thickness
3	Chip Width	8	Top Metal 2 Thickness
4	Chip Length	9	Top Solder Thickness
5	Chip Thickness	10	Bottom Solder Thickness

3.2.2 Aiding Tool for Simulation Parameter Tuning

With regard to the current case, the simulator has two parameters that need to be configured. Those two parameters are mesh settings for the finite element simulation. One parameter is responsible for the density of the mesh, while the other defines the maximum volume of a finite element. The size of the mesh is very important as it represents a trade-off between the results accuracy and simulation time.

An overview of the proposed method is illustrated in Fig. 3.5. There are two main functional blocks: the Experiment Planner and the Simulator Framework.

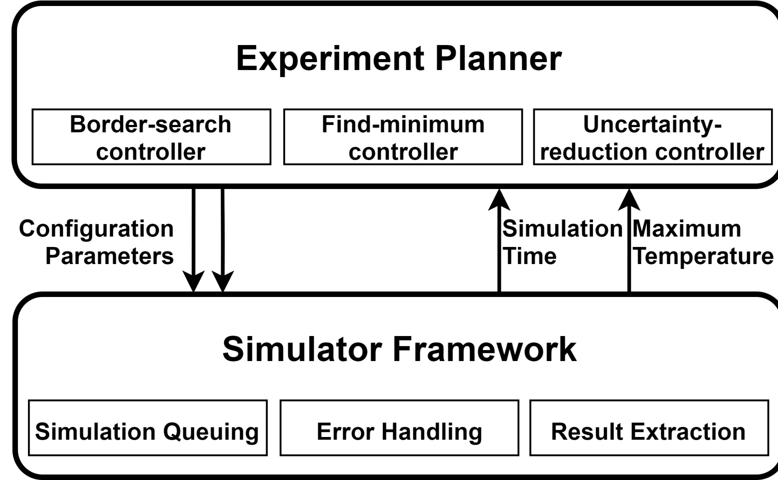


Fig. 3.5 Proposed Method Diagram [NBB+20]

There are three main objectives which need to be accomplished by *Experiment Planner*: *(i)* determine the convergence area as a function of configuration parameters; *(ii)* determine global or local minima of simulation time in the convergence area; *(iii)* describe results behavior in the convergence area. Before describing the algorithm, a couple of factors must be defined: $\boldsymbol{N}$ is the number of parallel simulations, $\boldsymbol{k}$ is the stop condition, $\boldsymbol{n}_{bs}$ is the number of simulations allocated to Border-Search Controller, $\boldsymbol{n}_{fm}$ is the number of simulations allocated to Find-minimum Controller, $\boldsymbol{n}_{ur}$ is the number of simulations allocated to Uncertainty-reduction Controller.

Algorithm 1: Automatic Parameter Tuning [NBB+20]

Input : Configuration space $\mathscr{C}$
Output : Convergence region, Temperature, Simulation Time

1 Initialize $\boldsymbol{N}$, $\boldsymbol{k}$, $\boldsymbol{n}_{bs} = 0$, $\boldsymbol{n}_{fm} = 0$, $\boldsymbol{n}_{ur} = \boldsymbol{N}$ and SVM model from Border-search controller;
2 Each controller generates samples according to $\boldsymbol{n}_{bs}$, $\boldsymbol{n}_{fm}$, $\boldsymbol{n}_{ur}$;
3 **while** *k is not met* **do**
4   The Simulator Framework runs the simulations and returns Simulation Time, Temperature and Convergence Information;
5   Use SVM from Border-search controller to predict results for current set of samples and compare with simulation values;
6   Count all misclassified samples: $\boldsymbol{h}$;
7   Update number of simulations allocated: $\boldsymbol{n}_{bs} = \boldsymbol{h}$, $\boldsymbol{n}_{fm} = \text{floor}(\frac{\boldsymbol{N}-\boldsymbol{h}}{2})$, $\boldsymbol{n}_{ur} = \text{floor}(\frac{\boldsymbol{N}-\boldsymbol{h}}{2})$;
8   All data is used for training in all controllers;
9   Each controller generates samples according to $\boldsymbol{n}_{bs}$, $\boldsymbol{n}_{fm}$, $\boldsymbol{n}_{ur}$;
10 **end**

3.2.3 Automated Simulation Setup

The functionality of the Automated Simulation Setup is presented as follows. The first block, 2D Chip Layout, generates the 2D layout based on its inputs (chip width, chip length, clip width, and clip length).

The second block has the role of translating the 2D layout into the 3D layout, generating the 3D structure (package). To achieve this, the first step is defining the planar size of all the existing layers. The next step is building the 3D geometry by including the thickness parameters of the layers and adjusting the shapes according to the technology parameters. In the end, electro-thermal material properties are assigned to the 3D elements, and the structure is now ready for simulation.

3.3 Metamodel-based Prediction of On Resistance

The formal description of our target metamodel is represented as the function $f(X) = R_{on}$, where $f : \mathbb{R}^N \to \mathbb{R}$ and X represents the vector of N design parameters. The main advantage of the metamodels is their low inference time. This is significantly lower than finite element simulation time by several orders of magnitude based on our results [NCB+23]. Another advantage is that, once the metamodel is fitted, it can be used for multiple analysis tasks. These tasks usually require a large number of simulations.

3.3.1 Overview

There are three main functional blocks in the proposed method: *(1)* Design of Experiments – starts from design parameters ranges and design rule constraints and has as outcome the simulation dataset. *(2)* Metamodel Fitting – takes the simulation dataset from the previous block and outputs the metamodel featuring the highest prediction performance. *(3)* Metamodel Prediction – is the final stage, where the best metamodel is employed for R_{on} prediction based on design parameters.

3.3.2 Simulation Dataset

Not every combination of design parameters leads to an actual device that can be manufactured. However, simulation of these invalid combinations of design parameters is still possible. Therefore, all experiments present in the simulation dataset need to be double checked with regard to the design rule constraints.

3.3.3 Fitting Machine Learning Metamodel for R_{on} Prediction

In the current perspective, approximating the R_{on} function of design parameters takes on a statistical dimension. The estimation function needs to have sufficient accuracy

compared to simulation with regard to the application of the metamodel. The aim is to fit a metamodel which has a low probability $\mathbb{P}$ of predicting inaccurate R_{on} in the hyperspace of valid design parameters, D. Four main machine learning model types have been evaluated in this work: *Linear regression*, *Support Vector Machine (SVM)*, *Gaussian Process*, and *Neural networks*.

3.3.4 Approaching Discrete Changes in Power MOSFET Design

This thesis takes on the challenge of fitting a metamodel when there are discontinuous changes in the transistor design and includes the following particularities along with its examination of the topic. Our approach consists in dividing the valid design parameter space into multiple sub-spaces, based on the continuity of the response, R_{on}.

3.3.5 Metamodel Performance Evaluation

The metamodel evaluation metric used in this work is the maximum relative error (MRE), illustrated in (3.3). Hence, the prediction results can be expressed together with the confidence, $R_{on} \pm MRE\%$. This is very important from a practical point of view, as the simulation and design engineers are already familiar with this format.

$$MRE[\%] = \max_{X \in D} \left(\frac{|R(X) - \widehat{R}(X)|}{R(X)} \times 100 \right) \tag{3.3}$$

where: X is the n-dimensional vector of design parameters, $D \subset \mathbb{R}^n$ is the valid parameter space, $R(X)$ is the R_{on} simulation result, $\widehat{R}(X)$ is the R_{on} metamodel estimate.

3.4 Experimental Results

An experimental scenario illustrating the reduction of simulation time is presented. For fitting the R_{on} prediction metamodels of the devices, there are two use-cases, each featuring a specific transistor design technology.

3.4.1 Automatic Parameter Tuning – Experimental Scenario

The following experiments are performed on a transistor structure featuring a lower chip area. We start by illustrating the experiments with regard to the first objective: define the convergence area of the configuration parameters. By using the Border-search controller in the Experiment Planner to determine this area, a large number of non-convergent simulations are avoided. In Fig. 3.16 we show an example of using Border-search controller for determining the convergence area. The input of the Border-search controller is binary; the simulation was either convergent or not.

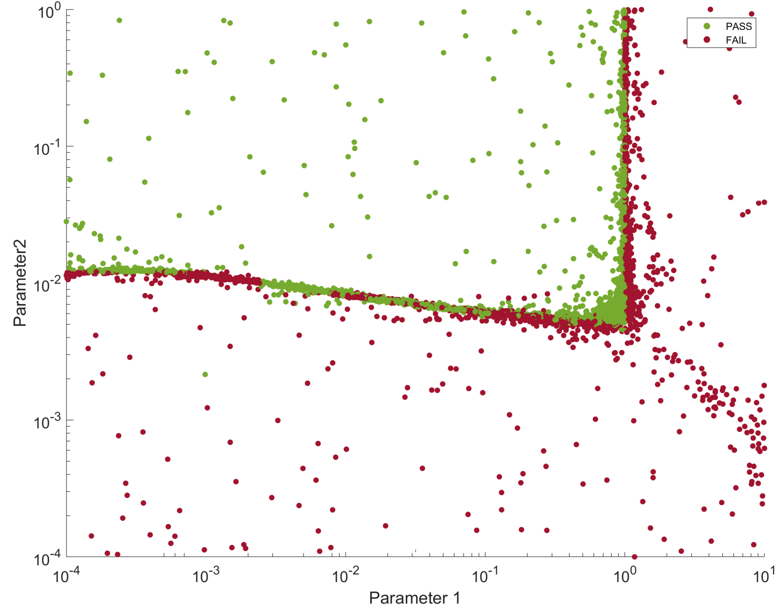


Fig. 3.16 Convergence region function of configuration parameters[NBB+20]

Table 3.3 shows that around 90% of the configurations generated by the other two controllers would not converge. In the absence of the Border-search controller, only 10.56% of the experiments performed by the Uncertainty-reduction controller are convergent. In the same conditions, the Find-minimum controller has similar performance. By employing the Border-search controller, the Experiment Planner has a convergence rate of 81%.

Table 3.3 Convergence rate for separate controllers [NBB+20]

Controller	Convergence rate
Uncertainty-reduction controller	10.65%
Find-minimum controller	12.25%
Experiment Planner	81%

3.4.2 Metamodel Use-case 1: *Technology A* Devices

Technology A involves a generic power MOSFET. This inherits all the characteristics described in Section 3.2.1. As a specific feature, this technology has the property of being parameterized in a continuous manner.

To cover the parameter space, a number of 1300 simulations were performed using an automated simulation setup. The dataset is split into three categories of data for neural network training: 70% training, 15% validation, and 15% evaluation. For the other machine learning regressors, we use a 10 k-fold validation [NBF+21].

The results of our experiments are presented in Table 3.4. As can be seen, the extent of the Fully Connected Layer, denoted by the variable *LayerSize*, has been adjusted in a

wide range, beginning with 4 nodes, and progressing as high as 50 nodes. The best result is obtained with a Fully Connected Layer of size 40 and training performed by *Bayesian regularization*. For this metamodel, the MAE is 0.00026 *mOhms* and the maximum relative error is 0.14%.

Table 3.4 Neural Network Results: Maximum absolute value of error is presented for each trained model [NCB+23]

No.	LayerSize	Training algorithm		
		Levenberg–Marquardt	Bayesian regularization	Scaled Conjugate Gradient
1	4	0.01432	0.01328	0.06964
2	6	0.00803	0.00409	0.03714
3	8	0.00257	0.00265	0.04503
4	10	0.00256	0.00291	0.09026
5	15	0.00163	0.00078	0.05323
6	20	0.00157	0.00045	0.04566
7	25	0.00058	0.00039	0.07499
8	30	0.00074	0.00043	0.06108
9	35	0.00083	0.00048	0.04214
10	**40**	0.00054	**0.00026**	0.11500
11	45	0.00051	0.00064	0.07066
12	50	0.00045	0.00042	0.19870

Metamodel fitting results are presented in Table 3.5. As can be observed, the best results are achieved by Neural Networks, but the Gaussian Process shows promising results.

Table 3.5 *Technology A* Metamodel Evaluation [NCB+23]

Algorithm	MRE [%]
Neural Networks	0.14
GPR Matern 5/2	0.70
GPR Rational Quadratic	1.13
SVM Cubic Regression	5.58
Linear Regression	74.02

3.4.3 Metamodel Use-case 2: *Technology B* Devices

For *Technology B*, the process of design involves two techniques: (i) geometry scaling, (ii) adjusting the number of gate fingers. Hence, the performance result R_{on} can be considered a function that is discontinuous at the points where the second process is involved. Also, being a advanced technology, the simulation time is longer compared to *Technology A*.

The experiments for *Technology B* feature a reduced set of design parameters to illustrate the relevant application differences at a minimum simulation time cost. Parameters 1–5 from Table 3.1 are varied, while the rest are kept constant. Then, in order to achieve an accurate metamodel, our proposal is to train separate machine learning models for each continuous region of R_{on}. The R_{on} prediction consists in a piece-wise function, defined on two distinct domains.

The simulation dataset is composed of 500 samples, where sampling design space parameters needs to cover the continuity regions of the R_{on} function. In the presented case, the R_{on} response space is divided into two continuous regions. Therefore, 100 samples cover a reduced region, and the rest of the 400 samples cover the other region. For neural networks, the percentage of data used for training, validation, and testing is the same as that used for *Technology A* and 10 k-fold validation for the rest.

Two types of approaches have been tried: (i) training a single metamodel (referred to as *Single Model*) for the entire design parameter space even though it has a discontinuity hyperplane, and (ii) training separate metamodels for each continuous region, referred to as *M1* for the small subspace and *M2* for the largest subspace. The results of metamodel fitting are presented in Table 3.7 as the maximum relative error expressed in percent. Training a single metamodel for the entire parameter space does not yield the best results.

Table 3.7 *Technology B* Metamodel Evaluation using MRE[%] [NCB+23]

Algorithm	Single metamodel	Combined metamodel		
		M1	M2	Combined
Neural Networks	1.81	0.59	**0.27**	0.59
GPR Matern 5/2	2.40	**0.50**	0.42	0.50
GPR Rational Quadratic	2.91	0.70	0.60	0.70
SVM Cubic Regression	14.21	6.15	4.79	6.15
Linear Regression	82.03	41.23	25.41	41.23

3.5 Summary and Conclusions

This chapter presented a method to reduce the simulation time of power MOSFETs, followed by a machine learning technique to estimate R_{on} based on simulation results. In this way, we achieve our main goal of reducing the overall simulation time.

The proposed method for parameter tuning that provides an analysis is of great importance, especially in industry, because it allows the user to choose the parameters that fit the requirements best. The introduction of machine learning techniques to estimate R_{on} of a MOSFET shows promising results. Metamodels represent reliable alternatives to simulation-only approaches and have a significant advantage when a large number of simulations is mandatory.

Chapter 4

Metamodel Applications for Power MOSFETs

4.1 Introduction

The aim of this chapter is to demonstrate the importance of metamodels through a series of applications. One important application is sensitivity analysis. A second application is based on whether we can use already available metamodels to make the process of fitting even more efficient. However, the most advanced application concerns yield estimation, an aspect of high importance for the newly designed devices.

4.2 Sensitivity Analysis

An immediate application of the metamodel is sensitivity analysis. This is one of the tasks that requires a large number of simulations In this way, we effectively demonstrate the benefits of metamodels. Chip size and clip size are, by definition, the two most important influencing factors for R_{on}. Hence, the analysis is performed on the rest of the parameters presented in Table 3.1. A sensitivity analysis is performed using 10,000 samples from the valid design parameter space. These methods are briefly described, as the aim is to highlight the speed of analysis.

For correlation, a value close to 0 is the least influential, while a value close to 1 or -1 represents a high influence. For the rest of the methods, a higher value indicates a greater influence. Table 4.1 illustrates the results of the sensitivity analysis. It can be observed that silicon and clip thicknesses are generally the most influential. However, the aim is to illustrate that the low inference time enables this type of analysis. Compared to a single simulation of R_{on} performed using standard commercial tools, the inference time is millions of times lower, considering the same computing hardware. On the computing node used by us, R_{on} simulation takes over 10 minutes while single inference takes $100\mu s$. The inference time was computed as an average over 10,000 inferences.

Table 4.1 Sensitivity Analysis of Design Parameters

Parameter	Method					
	Correlation	Entropy Pair	Entropy Simple	Brownian Distance	EFAST	Jansen
Silicon Thickness	0.29	1.00	0.08	0.02	0.00	1.11
Clip Thickness	-0.79	1.00	1.00	0.79	0.09	0.11
Top Metal 1 Thickness	-0.17	0.36	0.04	0.02	0.01	0.01
Top Metal 2 Thickness	-0.02	0.01	0.00	0.03	0.01	0.01
Top Solder Thickness	0.19	0.50	0.05	0.02	0.04	0.03
Bottom Solder Thickness	-0.01	0.00	0.00	0.02	0.00	0.00

4.3 Design of experiment analysis

In this experiment, chip length and width are variables, while the rest of the design parameters are fixed to an arbitrary value. This way, we illustrate in Fig. 4.1 the valid design parameter space in the current setup.

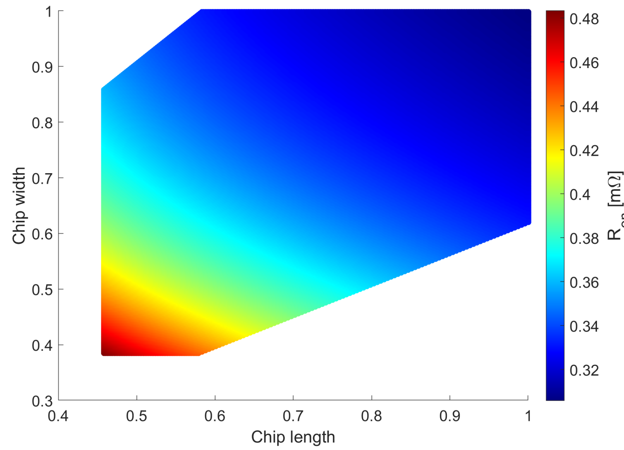


Fig. 4.1 Valid design parameter space in the current setup
Chip width and length normed to $[0, 1]$

We further use this artificially generated dataset for an investigation on design of experiment techniques. In the previous chapter, we presented our approach for designing the dataset used for metamodel fitting. The first step in designing the dataset is to determine the corners of the design parameter space. Hence, we perform an experiment to investigate how the method used to design the experiments in the dataset impacts the overall performance of the metamodel.

We propose two methods for deciding the coordinates of the next samples that will be included in the dataset. *HS* – next sample is positioned at the coordinates of the maximum absolute error. In this way, we estimate that each new sample will add a significant contribution to the metamodel's prediction performance. *RS* – random sampling.

To show the error distribution in the valid design parameter space, we further detail the evolution of the error from 8 up to 11 samples in Fig. 4.4. The plots placed on the left side illustrate the error for *HS* sampling while on the right side are RS1 plots.

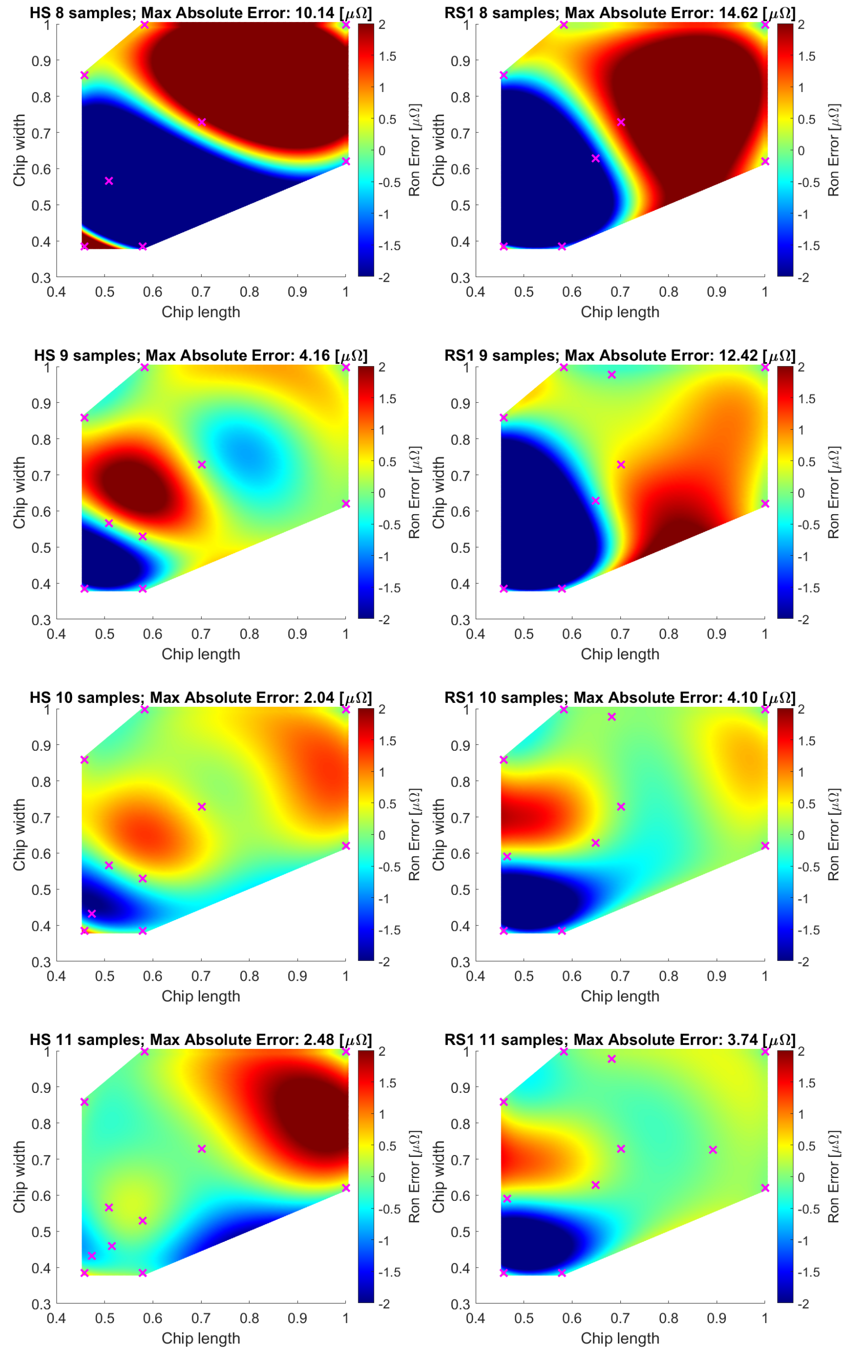


Fig. 4.4 HS vs RS1 – 7 to 11 samples
Chip width and length normed to $[0, 1]$

4.4 Improving Parametric Yield

A subcategory of yield known as parametric yield loss deals with functional devices which do not meet the required performance standards. Variations in one or more design parameters during the production process are the root cause of parametric failures. As a result, the distribution of parameter variance is crucial since it might result in certain devices not meeting their requirements.

4.4.1 Metamodel-Based Yield Estimation Challenges

There are many proposed methods of sampling and modeling of circuit characteristics and performance, as well as different yield estimation techniques for analog and mixed-signal circuits. These can be classified into three main types of approaches: *(1)* straightforward Monte Carlo (MC) [GYH11, PP02];*(2)* statistical MC-based [LLGP04, LZP08, LZW$^+$12, FYZL14, GTZZ20]; *(3)* Non-MC [GR08, GLY$^+$12, GR21];

4.4.2 Proposed Method for Parametric Yield Estimation

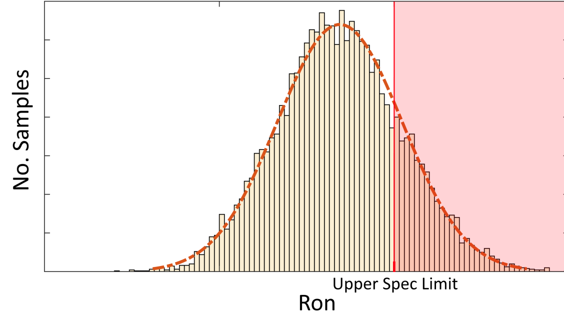
The hypothesis is that a type of device is usually manufactured in a single fabrication plant. Therefore, after a new device of the same type is designed, it will be sent for manufacturing to that plant. The set of manufacturing machines will be the same as for the previous devices. Fabrication data of these devices is available, and we aim to process this data and use it for estimating the device parameter variance of a new device design. The method consists of two functional blocks *(1)* Block *A*– Manufacturing Variation Modelling; *(2)* Block *B*– Technology Modelling.

Block *A* is in charge of modeling the behavior of the semiconductor fabrication machinery with respect to changes in design parameters. Block *B* has the objective of fast R_{on} estimation as a function of design parameters for a certain technology.

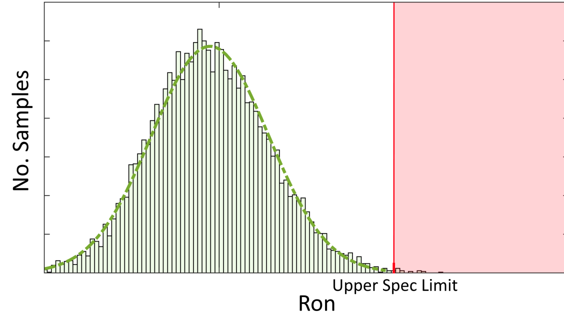
4.4.3 Experimental Scenario

Fig. 4.9a shows a dispersion estimate of R_{on} throughout the manufacturing process for a design. We can observe a number of devices, placed within the red zone, which do not meet the R_{on} requirement. The red probability density function curve illustrates an estimate of how R_{on} will be impacted by the manufacturing process of this newly designed device. The vertical red line, *Upper Spec Limit*, represents the R_{on} specification. If this product design progresses into the production phase, around 20% of the manufactured devices will fail to meet the specified performance, and the yield component for parametric failures will be around 80.66%.

The proposed method represents a design aiding tool for determining the best parametric yield without sacrificing cost or performance. From this point of view, the



(a) Initial design yield: 80.66%



(b) Optimal design yield: 99.82%

Fig. 4.9 R_{on} distributions estimation in the manufacturing process [NBC+22]

reliability of the optimal design is strongly correlated to the design parameters and the precision degree of their characterization. As shown in Fig. 4.9, shifting the distribution of the design parameters results in a 19% increase in the parametric yield. The main advantage of the proposed technique consists in designing products with improved performance at reduced manufacturing costs in every instance.

4.5 Summary and Conclusions

The first important application is sensitivity analysis. For this, we have performed a number of 10,000 inferences, and we have determined an average CPU time of $100\mu s$ for a single prediction of Ron, compared to 10 minutes using conventional simulation on the same computing hardware. A second application proposed was a design of experiment analysis. To investigate the importance of sample position, we have performed several experiments, using over 50,000 metamodel inferences. The results indicated the positions of the most important samples, which we can further include in the dataset for the next metamodels.

Finally, we proposed a method for parametric yield improvement by estimating this yield in pre-silicon. This is based on modelling manufacturing data using a multivariate distribution in the first step. Then the yield is estimated using metamodels to predict R_{on} of samples generated from the multivariate distribution. In this way, the parametric yield can be quickly estimated, offering a tool for yield improvement.

Chapter 5

On Cost-aware Design of Power MOSFETs

5.1 Introduction

The focus of this chapter is on the automatic design optimization of power MOSFETs with regard to their performance specifications. The proposed methods achieve a cost-aware design.

5.2 Proposed Automatic Design Optimization Method

The design process speed needs to be improved to meet the high demand for application-specific power devices. The impact of reducing design time is high because it enables producing devices that are optimal with regard to their requirements for specific applications. An optimal set of design parameters needs to be determined to design high-performance devices while saving silicon area. The decision on whether a candidate MOSFET design meets the specification is taken by the design engineers. Although this flow ensures the performance of the MOSFET device, it does not guarantee the minimum silicon area. From this point of view, the characteristics of the final MOSFET design, validated by the experts, need to be the best design that can be accomplished. Considering this, the proposed method includes an optimization algorithm capable of ensuring both meeting the requirements and minimizing silicon area.

5.2.1 Overview

The schematic representation of the proposed method is presented in Fig. 5.2. The industry application of the proposed method is to develop a MOSFET design aiding tool for engineers. This tool is intended to increase the efficiency of the design process and relieve design experts from repetitive manual work.

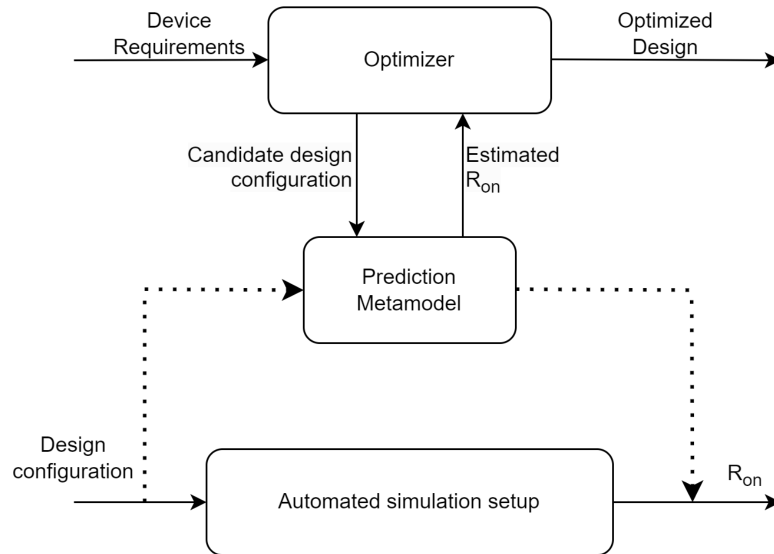


Fig. 5.2 Design optimization overview [NCB+23]

5.2.2 Considerations on Automatic Design Optimization using Metamodels

A practical example is shown in Fig. 5.3. This plot illustrates a section of parameter space consisting of two variables: the length and width of the chip. The rest of the design parameters are fixed to an arbitrary value. The samples have been evaluated using the R_{on} metamodel for *Technology A*, described in the previous chapter. The surface delimited by a green border resembles the valid combinations of length and width of the chip. Design constraints are not respected beyond the green border. In those regions exists at least one design rule violation, hence, design parameter configurations from these areas represent devices which cannot be manufactured.

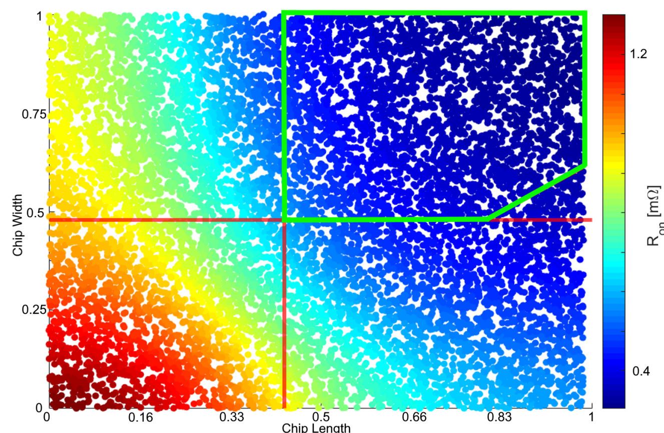


Fig. 5.3 Section of the design parameter space. Chip Length and Width are normed to [0,1]

5.2.3 Optimization Problem Formulation

Finding the optimal trade-off between cost and performance in a MOSFET device design implies identification of the three components of an optimization problem: variables, constraints, and objective function.

Variables

1) $ChipLength$, 2) $ChipWidth$, 3) $ClipLength$, 4) $ClipWidth$

Constraints

(a) Minimum chip length $\leq ChipLength \leq$ Maximum chip length

(b) Minimum chip width $\leq ChipWidth \leq$ Maximum chip width

(c) Minimum clip length $\leq ClipLength \leq$ Maximum clip length

(d) Minimum clip width $\leq ClipWidth \leq$ Maximum clip width

(e) Minimum chip aspect ratio $\leq \frac{ChipWidth}{ChipLength} \leq$ Maximum chip aspect ratio

(f) Minimum clip aspect ratio $\leq \frac{ClipWidth}{ClipLength} \leq$ Maximum clip aspect ratio

(g) $\frac{ChipLength - ClipLength}{2} \geq \delta_{\text{length}}$

(h) $\frac{ChipWidth - ClipWidth}{2} \geq \delta_{\text{width}}$

(i) $R_{on} \leq R_{on_specification}$

where:

- δ_{length} is the minimum distance between chip and clip margin on the longer side
- δ_{width} is the minimum distance between chip and clip margin on the shorter side

Objective function

- $F(X) = ChipLength \times ChipWidth$

where $X = \begin{pmatrix} ChipLength & ChipWidth & ClipLength & ClipWidth \end{pmatrix}^T$ is the input vector which contains the design variables.

With the variables, constraints, and objective function defined, finding the optimal trade-off between cost and performance (R_{on}) for a MOSFET device is formulated as follows.

$$\begin{aligned} \text{Minimize} \quad & F(X) \\ \text{subjected to:} \quad & \\ & \bullet \ \text{linear constraints (a) – (h);} \\ & \bullet \ \text{non-linear constraint (i);} \end{aligned} \tag{5.1}$$

5.2.4 Design Optimization for Continuous R_{on} Function

In the present case, the problem formulation is adapted to fit on design optimization of power MOSFETs. This is one of our main contributions and is presented below.

$$\text{Minimize } F(X) \text{ such that} \begin{cases} c(X) \leq 0 \\ A \cdot X \leq b \\ lb \leq x \leq ub \end{cases} \tag{5.3}$$

where:

- $c(X) = R(X)$ - $R_{on_specification}$; (constraint (i))
- lb & ub are the lower and upper boundary vectors for design parameters; (constraints (a) – (d))
- $F(X)$ is the objective function

Constraints (e) and (f) are defined in (5.4) using the middle inequality from (5.3).

$$\begin{pmatrix} ChAR_{\min} & -1 & 0 & 0 \\ -ChAR_{\max} & 1 & 0 & 0 \\ 0 & 0 & ClAR_{\min} & -1 \\ 0 & 0 & -ClAR_{\max} & 1 \end{pmatrix} \times \begin{pmatrix} ChipLength \\ ChipWidth \\ ClipLength \\ ClipWidth \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{5.4}$$

where:

- $ChAR_{min}$ & $ChAR_{max}$ represent minimum and maximum of chip aspect ratio
- $ClAR_{min}$ & $ClAR_{max}$ represent minimum and maximum of clip aspect ratio
- all aspect ratios are defined as $\dfrac{width}{length}$

Constraints (g) and (h) are defined in (5.5) using the middle inequality from (5.3).

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} ChipLength \\ ChipWidth \\ ClipLength \\ ClipWidth \end{pmatrix} \leq \begin{pmatrix} 2\delta_{\text{length}} \\ 2\delta_{\text{width}} \end{pmatrix} \tag{5.5}$$

Finally, the middle inequality from (5.3) is expressed as (5.6).

$$\begin{pmatrix} ChAR_{\min} & -1 & 0 & 0 \\ -ChAR_{\max} & 1 & 0 & 0 \\ 0 & 0 & ClAR_{\min} & -1 \\ 0 & 0 & -ClAR_{\max} & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix} \times \begin{pmatrix} ChipLength \\ ChipWidth \\ ClipLength \\ ClipWidth \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2\delta_{\text{length}} \\ 2\delta_{\text{width}} \end{pmatrix} \tag{5.6}$$

5.2.5 Approaching Discontinuities

A discontinuity has a great impact in optimization as it creates a three-dimensional discontinuity hyperplane in the domain of the R_{on} function. Therefore, the optimizer needs to be robust to this kind of discontinuity. To overcome this, we propose employing differential evolution, a meta-heuristics optimization algorithm.

The problem formulation for the differential evolution optimizer is illustrated in (5.7).

$$\text{Minimize } f(X) \text{ such that } lb \leq X \leq ub \tag{5.7}$$

where lb & ub are lower and upper boundary vectors for design parameters (constraints (a) – (d)) and $f(X)$ is the objective function .

From (5.7) we observe that differential evolution does not provide mechanisms for other types of constraints. To overcome this, the proposed approach consists in adjusting the objective function to include constraints (e) – (i). Therefore, we propose that the objective function be the sum of three factors, as illustrated in (5.8). By expressing this in the form of a sum, we can intuitively handle the mathematical characteristics of the objective function.

$$F_{DE}(X) = F(X) + F_{cost_DR}(X) + F_{cost_Ron}(X) \tag{5.8}$$

where $F(X)$ represents chip area as in (5.3); $F_{cost_DR}(X)$ represents constraints (e) – (h) as in (5.9); $F_{cost_Ron}(X)$ represents constraint (i) as in (5.10).

The first term, $F(X)$, has already been presented in the previous subsection. $F_{cost_DR}(X)$ and $F_{cost_Ron}(X)$ are designed to add a penalty to $F_{DE}(X)$ if the required constraints are not respected, while subtracting a bonification if the candidate is within the valid design parameter space.

$$F_{cost_DR}(X) = \begin{cases} -DR_{Scale} & \text{,if } \Delta_{DR}(X) \leq 0 \\ DR_{Scale} \cdot \Delta_{DR}(X) & \text{,if } \Delta_{DR}(X) > 0 \end{cases} \tag{5.9}$$

where $\Delta_{DR}(X)$ is a cumulative metric of design rule violations, DR_{Scale} is a scaling coefficient for $\Delta_{DR}(X)$.

$$F_{cost_Ron}(X) = \begin{cases} -R_{on_Scale} & \text{,if } \Delta_{R_{on}}(X) \leq 0 \\ R_{on_Scale} \cdot \Delta_{R_{on}}(X) & \text{,if } \Delta_{R_{on}}(X) > 0 \end{cases} \tag{5.10}$$

where $\Delta_{R_{on}}(X) = R(X) - R_{on_specification}$, R_{on_Scale} is a scaling coefficient for $\Delta_{R_{on}}(X)$.

5.2.6 Validation of the Optimal Design

We propose a procedure for optimal solution validation structured into two stages as follows. First stage aims to demonstrate that the optimal solution is not dependent on the starting point of the optimization algorithms. The second stage consists of comparing the

solution that resulted from the automatic optimization method with the optimal solutions achieved by designers using the standard methodology.

5.3 Experimental Results

For automatic design optimization of the devices, there are two use-cases, each featuring a specific transistor technology. *Technology A* employs a set of standard design techniques consisting of layer resizing; therefore, R_{on} function is continuous. *Technology B* is characterized by discrete geometrical elements in the design, which depend on the size of the transistor.

5.3.1 Optimization use-case 1: *Technology A*

In the present case, R_{on} is estimated using the machine learning metamodel of *Technology A*. An interior-point algorithm can be employed for optimization. Fig. 5.9 illustrates that regardless of the starting point, the optimal design determined by the algorithm is the same.

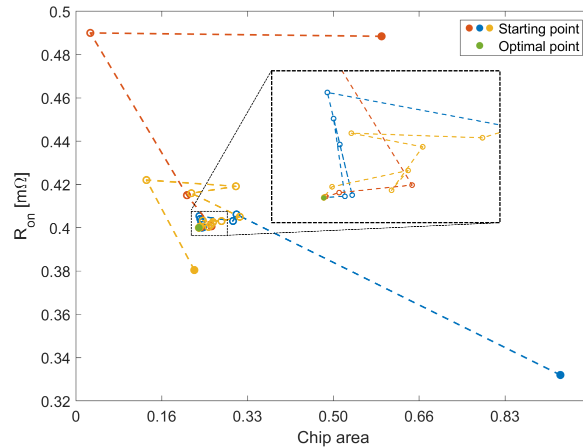


Fig. 5.9 Chip area optimization for *Technology A* starting from three different design configurations [NCB+23]. Chip area is normed to [0,1]

5.3.2 Optimization use-case 2: *Technology B*

Fig. 5.11 illustrates the discontinuity of the *Technology B* metamodel. Increasing the chip size, beyond a certain threshold, results in categorical changes in the device design. In *Technology B*, the number of gate fingers is dependent on chip size in order to achieve higher device performance. To illustrate these categorical changes, chip length and clip dimensions were fixed to arbitrary values while chip width was varied within the valid range. R_{on} was estimated using the metamodel for *Technology B*. The observed discontinuity is defined in (5.12).

$$
\begin{gathered}
\lim_{ChW \to Th^-} R(X) \neq \lim_{ChW \to Th^+} R(X) \\
\lim_{ChW \to Th^-} R(X) = R(X_{Th})
\end{gathered}
\tag{5.12}
$$

where $R(X)$ is the R_{on} estimation function; ChW is the chip width of X; Th is the threshold value of chip width ; $X_{Th} = \left(ChipLength \quad Th \quad ClipLength \quad ClipWidth \right)^T$.

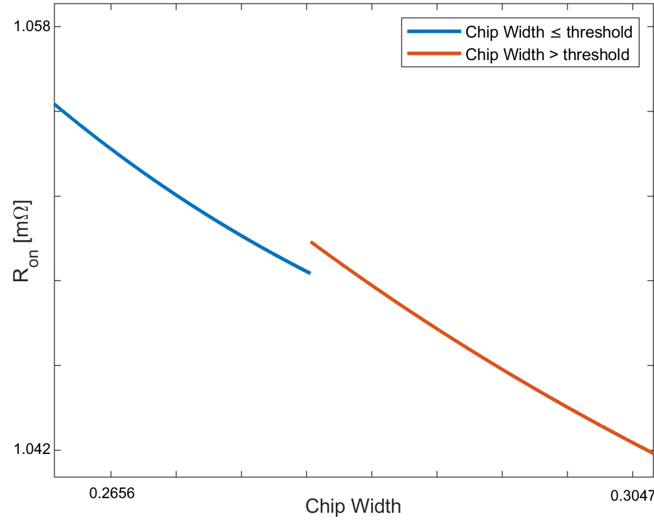


Fig. 5.11 Discontinuity point in R_{on} function. Chip width normed to [0,1]

With F_{cost_DR} and F_{cost_Ron} designed, the next step is to determine the objective function F_{DE} as in (5.8). Fig. 5.19 illustrates an example of F, F_{cost_DR}, and F_{cost_Ron} together with their influence on F_{DE}. For a better representation, each plot consists of 100,000 samples uniformly distributed in the design parameter space. Clip dimensions are fixed at an arbitrary value and the range of clip dimensions is the same in all four plots. The aim is to illustrate an example of how the designed F_{DE} objective function leads to achieving the optimal. The four plots of Fig. 5.19 are the following: *(1)* represents the function F, the first term of F_{DE}, the area of the chip; *(2)* represents the function F_{cost_DR}, the second term of F_{DE}, design rules cost; *(3)* represents function F_{cost_Ron}, the third term of F_{DE}, R_{on} cost; *(4)* represents the objective function $F_{DE} = F + F_{cost_DR} + F_{cost_Ron}$.

Towards a statistical validation of the solution, Fig. 5.21 presents a comparison between the results determined by differential evolution and the results found by interior point algorithm for *Technology B*. As R_{on} is not continuous, using the interior point optimizer does not reach the optimal for all starting points. For generating this plot, 1,000 uniform samples of the design space are used as starting points for optimization. The orange dots illustrate the optimal values found by the interior point optimizer. There are five solutions discovered in function of the starting point. As expected, the discontinuity is not properly handled by the interior point optimizer. However, the solutions obtained

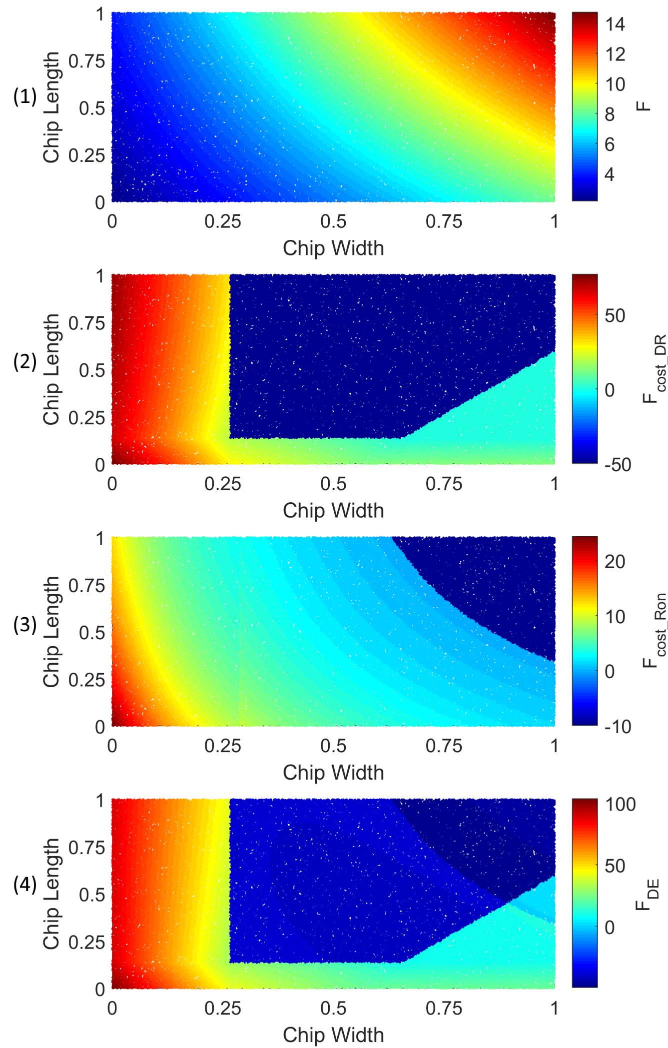


Fig. 5.19 Example for the objective function and its components.
Dimensions are normed to [0,1]

by differential evolution are placed in a relatively small area, compared to the orange dots, very close to the R_{on} specification limit. Therefore, these are closer to the real optimum. For a numerical comparison, the worst solutions are considered for both the differential evolution and the interior point. In this specific case, the experimental results show a chip area 3.11% larger for the interior-point optimizer compared to differential evolution.

Evaluation of the physical devices resulting from the optimization method is yet to be determined because manufacturing a new device involves many other aspects with regard to production (e.g. manufacturing process flow). Therefore, the evaluation of these optimal designs has been performed by design engineering experts only at the simulation level. However, previous real product measurements and simulation data were used in the evaluation of devices resulting from automatic design optimization. Although design experts confirmed the improvements of the method, a special mention is that this evaluation is confidential and therefore cannot be disclosed.

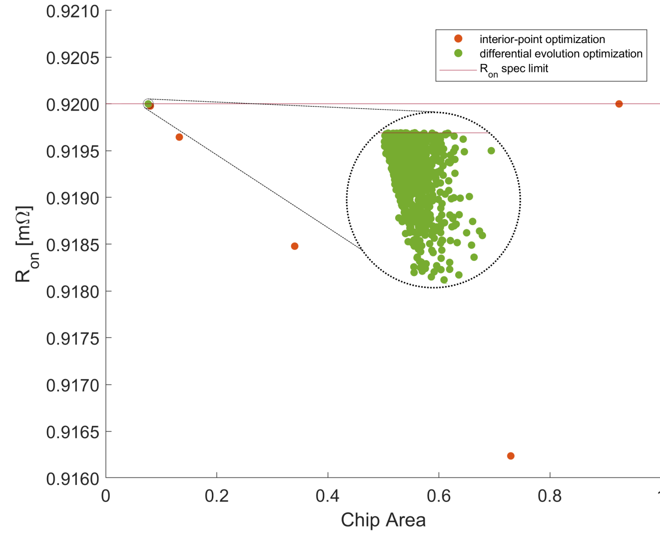


Fig. 5.21 Chip Area optimization for *Technology B* [NCB+23]. Chip Area is normed to [0,1]

5.4 Summary and Conclusions

This chapter presented an automatic design optimization technique for power MOSFETs. In this way, the approach is focused on the trade-off between cost and performance.

We proposed a flow consisting of three blocks to accomplish automatic optimization: an automated simulation setup, a prediction metamodel, and a design optimizer. In the preceding chapter, technical challenges, and strategies for overcoming them were described for the first two. Hence, the focus in this chapter was oriented toward the remaining block. The preliminary steps towards optimization consist of defining design variables, design rules, R_{on} constraints, and the objective function. First, design optimization with a continuous R_{on} function was presented. Then, special attention was paid to discontinuities in the R_{on} estimation function and the challenges raised by this. Optimal designs are highly accurate with regard to statistical evaluation. Insignificant differences, lower than 10^{-5}, of R_{on} and chip area with respect to the starting point are caused by numerical noise.

Chapter 6

General Conclusions

This research aims to improve the design methodology for future power MOSFET devices. To meet the higher requirements for power applications, the current manual methods have been enhanced. All of these advancements rely on machine learning techniques.

6.1 Objectives and Results

There are two major objectives that contribute to the present design methodology for new power MOSFETs. Each is related to a critical component of the design flow. These are presented as follows.

O1. Speeding up power MOSFETs simulation analysis techniques
The necessary time for analysis is an important factor in the final performance of the device. For accomplishing this objective there are three specific objectives.

O1.1. Reducing the time for simulation configuration
The deliverable of this objective is an automatic framework for parameter tuning in finite element simulation of power MOSFETs. This technique is presented in Chapter 3, Section 3.2.2.

O1.2. Reducing the time for simulation based analysis
This objective is accomplished by developing an automatic simulation environment. The simulation results for a set of defined experiments are obtained automatically using this flow. More details have been illustrated in Chapter 3, Section 3.2.3.

O1.3. Reducing the number of simulations in difficult tasks
This is achieved by fitting prediction metamodels based on simulation results. A limited number of simulations is required for metamodel training and evaluation. Then prediction can be used to determine the results. This was detailed in Chapter 3, Section 3.3.

O2. Ensuring cost-awareness and production yield in power MOSFETs design
Manufacturing high-performance devices at the lowest cost requires an optimal design with regard to both material loss and yield.

O2.1. Defining the automatic optimization problem
The outcome of this objective is a mathematical formulation for achieving a cost-effective MOSFET design. This has been presented in Chapter 5, Section 5.2.3.

O2.2. Applying the automatic design optimization
An automatic optimization framework has been developed considering the practical aspects with regard to transistor technologies. Sections 5.2.4 and 5.2.5 of Chapter 5 include the important steps towards the accomplishment of this objective.

O2.3. Improving Parametric Yield
Reliability of the design in production is ensured by employing a method for an initial estimation of the parametric yield based on previous manufacturing data. This has been presented in Chapter 4, Section 4.4.2.

6.2 Original contributions

The main contributions regarding power MOSFETs analysis, presented in Chapter 3, are the following:

- A method for speeding up parameter tuning in finite element simulation. This has the role of helping the engineers reduce the risk of obtaining erroneous simulation results and find a good trade-off between simulation time and accuracy. It also contributes to reducing the overall simulation time [NBB+20].
- A method for fitting R_{on} metamodels. With an initial set of finite element simulations for building the dataset, the metamodels are capable of R_{on} estimation. The length of the dataset, which is necessary for this technique to function properly, is dependent not only on the number of input parameters but also on how those parameters influence the final result [NBF+21].
- A method for handling discontinuities in fitting R_{on} metamodels. This is an extension of the previous method, which ensures high accuracy when modelling advanced technologies [NCB+23].

The main contributions with regard to power MOSFETs metamodel applications, presented in Chapter 4, are the following:

- A study on the most influential design parameters of the power MOSFET with emphasis on the requirements for performing this kind of sensitivity analysis.

- A study comparing two design of experiment methods for building the fitting dataset demonstrating that previous sampling knowledge has a significant impact on the final performance of the metamodel.
- A method for yield improvement by an initial parametric yield estimate in pre-silicon. This method uses the available production data to model the parameter variation in production. Then, using metamodels, a distribution of the device's performance in manufacturing is obtained [NBC+22].

The main contributions with regard to power MOSFETs design techniques, presented in Chapter 5, are the following:

- A design optimization method for power MOSFETs based on deterministic algorithms [NCB+23]. This employs machine learning metamodels instead of finite element simulation to speed up the optimization process. Deterministic optimization is appropriate for reaching an optimal design in simple transistor technologies. Results have demonstrated high performance in determining the minimum area of the chip for a required R_{on}.
- A design optimization method for power MOSFETs based on meta-heuristic algorithms [NCB+23]. With a proper configuration of the optimization method, the optimal can be determined in any circumstance.

6.3 List of Original Publications

1. [NBB+20] G. Nicolae, C. Boianceanu, A. Buzo, C.V. Diaconu, H. Cucu, G. Pelz, and C. Burileanu. Automatic parameter tuning in finite element analysis of semiconductor packages. In *2020 International Semiconductor Conference (CAS)*, pages 41–44, Virtual Conference, 2020, ISI WOS:000637264600009
2. [NBF+21] G. Nicolae, A. Buzo, C. Feuerbaum, C.V. Diaconu, H. Cucu, G. Pelz, and C. Burileanu. Metamodel-based prediction of on resistance for microelectronic power switches. In *2021 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)*, pages 1–3, Virtual Conference, 2021, ISI WOS:000927194000010
3. [NBC+22] G. Nicolae, A. Buzo, H. Cucu, C. Burileanu, and G. Pelz. Manufacturing variation estimation of on resistance in power semiconductors. In *2022 18th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, pages 1–4, Villasimius, Sardinia, Italy, 2022
4. [NCB+23] G. Nicolae, H. Cucu, C. Burileanu, A. Buzo, C. Feuerbaum, and G. Pelz. Automatic design optimization of microelectronic power switches. *UNIVERSITY POLITEHNICA OF BUCHAREST SCIENTIFIC BULLETIN SERIES*

C-ELECTRICAL ENGINEERING AND COMPUTER SCIENCE, 85(1):377–388, 2023, ISI WOS:000957721700001

5. [GND+20] Andrei Gaita, Georgian Nicolae, Emilian C. David, Andi Buzo, Corneliu Burileanu, and Georg Pelz. A SIFT-based Waveform Clustering Method for aiding analog/mixed-signal IC Verification. In *2020 IEEE EUROPEAN TEST SYMPOSIUM (ETS 2020)*, Proceedings of the European Test Symposium, Virtual Conference, 2020. 25th IEEE European Test Symposium (ETS), ELECTR NETWORK, MAY 25-29, 2020, ISI WOS:000615974000037

6.4 List of Technical Reports

1. [Nic19b] G. Nicolae. Optimization of power electronics packaging, fabrication process and device simulation. *Technical Report No. 1, University Politehnica of Bucharest*, June 2019

2. [Nic19b] G. Nicolae. A machine learning approach for mesh optimization in thermal analysis of semiconductor packages. *Technical Report No. 2, University Politehnica of Bucharest*, December 2019

3. [Nic20] G. Nicolae. Multi-objective optimization of electronic packages. *Technical Report No. 3, University Politehnica of Bucharest*, June 2020

6.5 Future Work

Improving the parameter tuning method for finite element simulation for more advanced MOSFET structures and defining a metric for determining the optimal combination of simulation parameters.

Reducing the size of the simulation dataset for fitting metamodels by sharing knowledge across technologies. State of the art transfer learning techniques may be a solution.

Improving the metamodels by including more R_{on} influencing factors in the metamodels to cover more difficult situations. Technology CAD parameters, which are now fixed in simulation, represent a good example.

Extending the proposed metamodeling technique for estimation of multiple performance parameters of power MOSFETs. For example, Z_{thJC} and safe operating area are important indicators of transistor performance.

Improving the metamodeling fitting technique for a better handling of categorical changes in the transistor design. One way to do this is to use machine learning classification techniques to define the edges of continuous spaces while regression estimates the response in those spaces.

References

[BIMR18] Giovanni Breglio, Andrea Irace, Luca Maresca, and Michele Riccio. Electro-thermal simulations of power semiconductor devices during high stress events. In *2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)*, pages 1–4, 2018.

[FYZL14] Chenlei Fang, Fan Yang, Xuan Zeng, and Xin Li. Bmf-bd: Bayesian model fusion on bernoulli distribution for efficient yield estimation of integrated circuits. In *2014 51st ACM/EDAC/IEEE Design Automation Conference (DAC)*, pages 1–6, 2014.

[GLY+12] Fang Gong, Xuexin Liu, Hao Yu, Sheldon X. D. Tan, Junyan Ren, and Lei He. A fast non-monte-carlo yield analysis and optimization by stochastic orthogonal polynomials. *ACM Trans. Des. Autom. Electron. Syst.*, 17(1), jan 2012.

[GND+20] Andrei Gaita, Georgian Nicolae, Emilian C. David, Andi Buzo, Corneliu Burileanu, and Georg Pelz. A SIFT-based Waveform Clustering Method for aiding analog/mixed-signal IC Verification. In *2020 IEEE EUROPEAN TEST SYMPOSIUM (ETS 2020)*, Proceedings of the European Test Symposium, Virtual Conference, 2020. 25th IEEE European Test Symposium (ETS), ELECTR NETWORK, MAY 25-29, 2020.

[GR08] Chenjie Gu and Jaijeet Roychowdhury. An efficient, fully nonlinear, variability-aware non-monte-carlo yield estimation procedure with applications to sram cells and ring oscillators. In *2008 Asia and South Pacific Design Automation Conference*, pages 754–761, 2008.

[GR21] Zhengqi Gao and Ron Rohrer. Efficient non-monte-carlo yield estimation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pages 1–1, 2021.

[GTZZ20] Zhengqi Gao, Jun Tao, Dian Zhou, and Xuan Zeng. Efficient parametric yield estimation over multiple process corners via bayesian inference based on bernoulli distribution. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 39(10):3144–3148, 2020.

[GYH11] Fang Gong, Hao Yu, and Lei He. Stochastic analog circuit behavior modeling by point estimation method. In *Proceedings of the 2011 International Symposium on Physical Design*, ISPD ’11, page 175–182, New York, NY, USA, 2011. Association for Computing Machinery.

[HHL21] Chen He, Hanbin Hu, and Peng Li. Applications for machine learning in semiconductor manufacturing and test (invited paper). In *2021 5th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)*, pages 1–3, 2021.

[Hoo12] Holger H. Hoos. Automated algorithm configuration and parameter tuning. In Youssef Hamadi, Eric Monfroy, and Frédéric Saubion, editors, *Autonomous Search*, pages 37–71. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[HP05] P.L. Hower and S. Pendharkar. Short and long-term safe operating area considerations in ldmos transistors. In *2005 IEEE International Reliability Physics Symposium, 2005. Proceedings. 43rd Annual.*, pages 545–550, 2005.

[ICFQ93] K.B. Irani, J. Cheng, U.M. Fayyad, and Z. Qian. Applying machine learning to semiconductor manufacturing. *IEEE Expert*, 8(1):41–47, 1993.

[LLGP04] Xin Li, Jiayong Le, P. Gopalakrishnan, and L.T. Pileggi. Asymptotic probability extraction for non-normal distributions of circuit performance. In *IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004.*, pages 2–9, 2004.

[LZP08] Xin Li, Yaping Zhan, and Lawrence T. Pileggi. Quadratic statistical *max* approximation for parametric yield estimation of analog/rf integrated circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 27(5):831–843, 2008.

[LZW+12] Xin Li, Wangyang Zhang, Fa Wang, Shupeng Sun, and Chenjie Gu. Efficient parametric yield estimation of analog/mixed-signal circuits via bayesian model fusion. In *2012 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 627–634, 2012.

[Mur12] Kevin P Murphy. *Machine learning: a probabilistic perspective*. Cambridge, MA, 2012.

[NBB+20] G. Nicolae, C. Boianceanu, A. Buzo, C.V. Diaconu, H. Cucu, G. Pelz, and C. Burileanu. Automatic parameter tuning in finite element analysis of semiconductor packages. In *2020 International Semiconductor Conference (CAS)*, pages 41–44, Virtual Conference, 2020.

[NBC+22] G. Nicolae, A. Buzo, H. Cucu, C. Burileanu, and G. Pelz. Manufacturing variation estimation of on resistance in power semiconductors. In *2022 18th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD)*, pages 1–4, Villasimius, Sardinia, Italy, 2022.

[NBF+21] G. Nicolae, A. Buzo, C. Feuerbaum, C.V. Diaconu, H. Cucu, G. Pelz, and C. Burileanu. Metamodel-based prediction of on resistance for microelectronic power switches. In *2021 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)*, pages 1–3, Virtual Conference, 2021.

[NCB+23] G. Nicolae, H. Cucu, C. Burileanu, A. Buzo, C. Feuerbaum, and G. Pelz. Automatic design optimization of microelectronic power switches. *UNIVERSITY POLITEHNICA OF BUCHAREST SCIENTIFIC BULLETIN SERIES C-ELECTRICAL ENGINEERING AND COMPUTER SCIENCE*, 85(1):377–388, 2023.

[Nic18] G. Nicolae. Sensitivity analysis on electro-thermal limit of safe operating area in power mosfets. Master's thesis, University Politehnica of Bucharest, 2018.

[Nic19a] G. Nicolae. A machine learning approach for mesh optimization in thermal analysis of semiconductor packages. *Technical Report No. 2, University Politehnica of Bucharest*, December 2019.

[Nic19b] G. Nicolae. Optimization of power electronics packaging, fabrication process and device simulation. *Technical Report No. 1, University Politehnica of Bucharest*, June 2019.

[Nic20] G. Nicolae. Multi-objective optimization of electronic packages. *Technical Report No. 3, University Politehnica of Bucharest*, June 2020.

[PJS08] Martin Pfost, Joachim Joos, and Matthias Stecher. Measurement and simulation of self-heating in dmos transistors up to very high temperatures. In *2008 20th International Symposium on Power Semiconductor Devices and IC's*, pages 209–212, 2008.

[PP02] A. Papoulis and S.U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill series in electrical engineering: Communications and signal processing. McGraw-Hill, 2002.

[Rin01] N. Rinaldi. On the modeling of the transient thermal behavior of semiconductor devices. *IEEE Transactions on Electron Devices*, 48(12):2796–2802, 2001.

[Wan17] Li-C. Wang. Experience of data analytics in eda and test—principles, promises, and challenges. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 36(6):885–898, 2017.