Ain Shams University
Faculty of Engineering
Electronics and Communications Department

Analog Integrated Design Automation

A thesis
submitted in partial fulfillment of the requirements of the doctoral degree
in Electrical Engineering

Submitted by
Reem ElAdawi Shaban
M.Sc. of Electrical Engineering
Electronics and Communications Department
Ain Shams University, 2001

Supervised by
Prof. Dr. Hani Fikry
Prof. Dr. Mohamed Dessouky
Cairo, 2018

Analog Integrated Design Automation: Detecting Rare-Events using Machine Learning

Reem ElAdawi Shaban

PhD dissertation
Electronics and Communications Department
Faculty of Engineering - Ain Shams University

ABSTRACT

Circuit reliability in the presence of statistical process variation is becoming a challenging domain, as transistors sizes get smaller and smaller in the nanoscale era. It is not enough to add safety margins to designs or depend on simple assumptions about "worst case" corners. Circuits that are replicated in millions of instances on large designs such as SRAMs and flip flops face the problem that a rare event in a circuit cell may have a large impact on the system yield. Monte Carlo simulation is an ideal tool for estimation of the yield of such high replication circuit (HRCs) but not a practical one because of the high yield requirements for individual cells. Numerous techniques have been proposed to overcome this disadvantage in Monte Carlo approach. Some approaches are analytical or semi analytical that model the behavior of SRAM cell. Other statistical techniques aim to reduce the number of Monte Carlo simulations. Statistical blockade algorithm aims at detecting rare events in the tail through the classification of tail points and the fitting of GPD distribution to the tail points.

In this thesis the statistical blockade method is investigated and a couple of problems are tackled and solution proposed to enhance the performance of the algorithm. The complexity problem is addressed and a subset selection algorithm is proposed. The classification method used is studied and compared to other classification methods to overcome the drawback of SVM classifier.

Key words: Monte Carlo Techniques, Yield Estimation, Statistical Blockade, Classification, Random Forests, Support Vector Machines, Random Forests, Deep Neural Networks

AIN SHAMS UNIVERSITY
FACULTY OF ENGINEERING
CAIRO-EGYPT
Electronics and Communication Department

Analog Integrated Design Automation

Reem ElAdawi Shaban

Supervised by
Prof. Dr. Hani Fikry

Prof. Dr. Mohamed Amin Dessouky

Thesis Summary

Circuit reliability in the presence of statistical process variation is becoming a challenging domain, as transistors sizes get smaller and smaller in the nanoscale era. It is not enough to add safety margins to designs or depend on simple assumptions about “worst case” corners. Circuits that are replicated in millions of instances on large designs such as SRAMs and flip flops face the problem that a rare event in a circuit cell may have a large impact on the system yield. Monte Carlo simulation is an ideal tool for estimation of the yield of such high replication circuit (HRCs) but not a practical one because of the high yield requirements for individual cells. Numerous techniques have been proposed to overcome this disadvantage in Monte Carlo approach. Some approaches are analytical or semi analytical that model the behavior of SRAM cell. Other statistical techniques aim to reduce the number of Monte Carlo simulations. Statistical blockade algorithm aims at detecting rare events in the tail through the classification of tail points and the fitting of GPD distribution to the tail points. In this thesis the statistical blockade method is investigated

and a couple of problems are tackled and solution proposed to enhance the performance of the algorithm. The complexity problem is addressed and a subset selection algorithm is proposed. The classification method used is studied and compared to other classification methods to overcome the drawback of SVM classifier.

Thesis Outline

The thesis is divided into seven chapters including lists of contents, tables and figures as well as list of references and one appendix.

- Chapter 1: This chapter is an introduction which states the problem being tackled and the motivation for this study.
- Chapter 2: This chapter discusses rare event simulation and the challenges facing current Monte Carlo techniques in calculating yield for memory circuits. State-of-art techniques are presented.
- Chapter 3: In this chapter, an overview of subset selection techniques is discussed and an algorithm for subset selection is presented.
- Chapter 4: In this chapter, different classification techniques are discussed and the metrics used to measure the performance of a classifier are explained
- Chapter 5: This chapter discusses the use of subset selection using Lasso algorithm in reducing the complexity of the problem and the use of Random Forests and Deep Neural Networks to replace SVM classifiers.
- Chapter 6: In this chapter, the results for subset selection are presented. Also the results for replacing SVM classifiers with Random Forests and DNN are presented, compared and discussed.
- Chapter 7: This chapter concludes the work presented in this thesis, and highlights the proposed future work.

Thesis supervisors:
Prof. Dr. Hani Fikry
Prof. Dr. Mohamed Amin Dessouky

ACKNOWLEDGMENT

All gratitude to ALLAH

Many thanks my supervisor Prof. Dr. Mohamed Dessouky. and Prof. Dr. Hani Firky for their insightful thoughts and helpful discussions.

I would like to express my gratitude to Prof. Mohamed Dessouky for his support, guidance and hours of fruitful discussions.

My deep appreciation to Dr. Hazem Eltahawy for his support and encouragement throughout this Ph.D study.

Last but not the least, I would like to thank my family who have patiently supported me all the way to accomplish this work.

Contents

List of Figures **XIII**

List of Tables **XVI**

Nomenclature **XVIII**

1 Introduction **1**
1.1 Motivation 1
1.2 Scope and Contributions 3
1.3 Thesis Organization 3

2 Rare Event Simulation **5**
2.1 Introduction 5
2.2 Monte Carlo (MC) 6
2.3 Variance Reduction Techniques 7
2.3.1 Control Variates 7
2.3.2 Antithetic Variates 8
2.3.3 Importance Sampling 9
2.4 Quasi-Monte Carlo Methods 10
2.5 Computing Probabilistic Hypervolumes 11
2.5.1 Using Simplicial Approximation 12
2.5.2 Worst-Case Distance Approximation 13
2.5.3 Ellipsoidal Approximation 14
2.5.4 Euler-Newton Curve Tracing 14
2.6 Statistical Blockade 15
2.6.1 Calculation of Tail Failure Probability 15
2.6.2 Sample Maximum 16
2.6.3 Statistical Blockade Algorithm Steps 17

2.6.3.1 Simulate Training Data Set 17
2.6.3.2 Label Training Data Set 17
2.6.3.3 Training the Classifier 18
2.6.3.4 Sample Point Generation 18
2.6.3.5 Classification 18
2.6.3.6 Tail distribution estimation 18
2.7 Conclusion 19

3 Subset Selection Techniques **20**
3.1 Introduction 20
3.2 Multiple Linear Regression 20
3.3 Subset Selection 24
3.3.1 Best Subset Selection 25
3.3.2 Stepwise Selection 25
3.3.2.1 Forward Stepwise Selection 25
3.3.2.2 Backward Stepwise Selection 25
3.3.2.3 Hybrid Approaches 26
3.3.3 Shrinkage Methods 26
3.3.3.1 Ridge Regression 26
3.3.3.2 Lasso Regression 27
3.3.4 Principle Component Analysis (PCA) 28
3.3.5 Other Techniques for Subset Selection 29
3.4 Conclusion 30

4 Classification Techniques **31**
4.1 Introduction 31
4.2 Classification Techniques 31
4.2.1 Bayes' Classifier 32
4.2.2 Logistic Regression 34
4.2.3 Linear discriminant Analysis (LDA) 34
4.2.4 K Nearest Neighbors (K-NN) 36
4.2.5 Support Vector Machine 37
4.2.5.1 Maximal Margin Classifier 37
4.2.5.2 Support Vector Classifier 38
4.2.5.3 Support Vector Machines (SVMs) . . . 40
4.2.6 Decision Trees 42
4.2.7 Random Forest 44

4.2.7.1 Bagging 44
4.2.7.2 Random Forest for Classification . . . 45
4.2.8 Neural Networks 46
4.2.8.1 The Human Brain 46
4.2.8.2 The Neuron 47
4.2.8.3 The Perceptron 48
4.2.8.4 Deep Learning Networks 51
4.2.8.5 Supervised Learning - Back-propagation 52
4.3 Cross Validation 53
4.3.1 Leave-One-Out Cross-Validation (LOOCV) . . . 54
4.3.2 k-Fold Cross-Validation 55
4.4 Measuring the Accuracy of the Classifier 55
4.4.1 Confusion Matrix 55
4.4.2 Sensitivity 56
4.4.3 Specificity 56
4.4.4 Accuracy 57
4.4.5 ROC Curves 57
4.5 Conclusion 59

5 Proposed Classification and Subset Selection 60
5.1 Introduction 60
5.2 Subset Selection 61
5.3 Selected Classification Techniques 62
5.4 Imbalanced Data Sets 63
5.5 Solving the Imbalance Problem 65
5.6 Neural Networks 66
5.6.1 Activation Functions 67
5.6.2 Optimization Algorithm 67
5.6.3 Optimizing DNN Parameters 68
5.7 Proposed Flowchart 71
5.7.1 Tools 72
5.7.2 Data Preparation 74
5.7.3 Classifier Training 75
5.7.3.1 SVM Training 76
5.7.3.2 RF Training 78
5.7.3.3 DNN Training 79
5.7.4 Yield Estimation 81

5.8 Conclusion 83

6 Case Study and Discussion **84**
6.1 Introduction 84
6.2 Performance Metrics 84
6.2.1 Write Time 85
6.2.2 Leakage Current 86
6.3 Standard Monte Carlo Simulation 86
6.4 Sampling and Data Preprocessing 87
6.5 Subset Selection Results 88
6.5.1 Statistical blockade with proposed parameter selection 88
6.5.2 Model corresponding to minimum cross-validated error: 91
6.5.3 Model corresponding to one standard error of the minimum: 92
6.5.4 Yield Comparison 93
6.6 Classification Results 94
6.6.1 Data Used 95
6.6.2 Linear SVM Classifier 96
6.6.2.1 Training 96
6.6.2.2 Prediction 97
6.6.2.3 Yield Calculation 98
6.6.3 SVM Radial Classifier 101
6.6.3.1 Training 101
6.6.3.2 Prediction 103
6.6.3.3 Yield Calculation 105
6.6.4 Random Forests Classifier 107
6.6.4.1 Training 107
6.6.4.2 Prediction 110
6.6.4.3 Yield Calculation 111
6.6.5 Deep Neural Networks 114
6.6.5.1 Training 114
6.6.5.2 Prediction 116
6.6.5.3 Yield Calculation 118
6.6.6 Summary 121
6.6.6.1 Time to Train Classifier 121

6.6.6.2 Sensitivity, Specificity and Accuracy . 122
6.6.7 Conclusion 123

7 Conclusion and Future Work **125**
7.1 Conclusion 125
7.2 Future Work 125

Publications **127**

References **129**

Appendix **137**

List of Figures

2.1 (a) Pseudo-random Sequence (b) Sobol Sequence [30] 11
2.2 Yield calculation [28] 12
2.3 Curve tracing using the Euler-Newton method [36] 15
2.4 Calculation of tail distribution 16
2.5 Statistical Blockade flow chart 17

3.1 Linear regression hyperplane to minimize the residual sum of squares [2] 21
3.2 Ridge and Lasso regression coefficient constraints [4] 28

4.1 KNN classification example 37
4.2 Support Vector Classifier with data points on the wrong side of the margin 40
4.3 (a) Non-linearly separable data (b) Using polynomial kernel function of degree 3 (c) Using radial kernel function [2] . . . 42
4.4 Illustration of structure of a Decision Tree 43
4.5 An illustration for Bagging 44
4.6 Random Forest of Decision Trees 45
4.7 Illustration of biological neuron 46
4.8 Neural Network Neurons 47
4.9 Two input Perceptron 49
4.10 Decision Boundary for Two Input Perceptron 50
4.11 (a) Multi-Layer Perceptron with one hidden layer (b) Multi-Layer Perceptron with two hidden layers 52
4.12 Leave-One-Out Cross-Validation [2] 54
4.13 5-fold Cross-Validation illustration [2] 55
4.14 Five discrete classifiers represented on ROC graph [19] . . . 58

4.15 ROC curves with different areas, curve A is the perfect classifier, curve B is a good classifier and curve C has an area of 0.5 59

5.1 Rectified Linear Unit 67
5.2 (a) Grid Search (b) Random Search [44] 69
5.3 Bayesian Optimization 70
5.4 R Studio 72
5.5 Spyder 73
5.6 Flow chart for data preparation 74
5.7 Flow chart for SVM classifier training 76
5.8 Flow chart for RF classifier training 78
5.9 Flow chart for DNN classifier training 80
5.10 Flow chart for estimating the yield by fitting GPD 81

6.1 6T SRAM bit circuit 85
6.2 Probability density distribution for scaled Write Time . . 85
6.3 Probability density distribution for scaled Leakage Current 86
6.4 Normal distribution and 6 sigma percentages 87
6.5 The coefficients included in the model as a function of tuning parameter (a) Write Time (b) Leakage Current . . 89
6.6 Mean squared error versus tuning parameter (a) Write Time (b) Leakage Current 90
6.7 CDF for fitted tail GPD - using 58 parameters 92
6.8 CDF for tail fitted GPD (using 38 parameters) 93
6.9 GPD Fit for Write Time tail points t = 99; parameters for GPD: location = 2.83 , shape = -0.0703 and scale = 0.327 99
6.10 GPD Fit for Leakage Current tail points for t =99 ; parameters for GPD: location = 3.82 , shape = 0.138 and scale = 1.69 . . 100
6.11 Area under the ROC curve versus the values of the tuning parameter for Write Time 102
6.12 Area under the ROC curve versus the values of the tuning parameter for Leakage Current 103
6.13 GPD Fit for Write Time tail points; parameters for GPD: location = 2.83 , shape = -0.0749 and scale = 0.329 106
6.14 Error Rate (OOB) versus number of trees grown for Write Time (a) Using all parameters (b) Using 38 parameters 108

6.15 Error Rate (OOB) versus number of trees grown for Leakage Current (a) Using all parameters (b) Using 21 parameters . . . 109
6.16 GPD Fit for Write Time tail points for t = 99; parameters for GPD: location = 2.83 , shape = -0.0979 and scale = 0.342 . . 112
6.17 GPD Fit for Leakage Current tail points for t = 99; parameters for GPD: location = 3.82 , shape = 0.166 and scale = 1.46 . . 113
6.18 GPD Fit for Write Time tail points t = 99; parameters for GPD: location = 2.83 , shape = -0.116 and scale = 0.362 119
6.19 GPD fit for Leakage Current tail points for t = 99; parameters for GPD: location = 3.82 , shape = 0.145 and scale = 1.66 . . 120