Robust Resistive Open Defect Identification Using Machine Learning with Efficient Feature Selection

Zahra Paria Najafi-Haghi, Florian Klemme, Hanieh Jafarzadeh, Hussam Amrouch, Hans-Joachim Wunderlich Institute of Computer Architecture and Computer Engineering, University of Stuttgart, Germany Email: {najafi-haghi, klemme, jafarzadeh, amrouch, wu}@informatik.uni-stuttgart.de

Abstract—Resistive open defects in FinFET circuits are reliability threats and should be ruled out before deployment. The performance variations due to these defects are similar to the effect of process variations which are mostly benign. In order not to sacrifice yield for reliability the effect of defects should be distinguished from process variations. It has been shown that machine learning (ML) schemes are able to classify defective circuits with high accuracy based on the maximum frequencies F_{max} obtained under multiple supply voltages $V_{dd} \in V_{op}$. The paper at hand presents a method to minimize the number of required measurements. Each supply voltage V_{dd} defines a feature $F_{max}(V_{dd})$. A feature selection technique is presented, which uses also the already available F_{max} measurements. It is shown that ML-based techniques can work efficiently and accurately with this reduced number of $F_{max}(V_{dd})$ measurements.

Index Terms—Resistive open defects, process variations. machine learning, feature selection

I. INTRODUCTION

Resistive open defects represent themselves as Small Delay Faults (SDFs) and are reliability threats even if the device behaves within its specification [1] [2] [3]. An imperfect connection is the physical source of a resistive open defect. It has a great potential to degrade further in the field and should be ruled out before deployment.

Pure random process-induced variability is a substantial concern in the chip production. But they are mostly benign and the affected device can be used safely. It is a demanding challenge to distinguish performance variations due to maleficent defects from the effects due to benign process-induced variability, even if both introduce a similar additional delay within the specification of the device.

Machine learning-based classification has been used in literature [4] [5] [6] [7] to identify defects from variations for logic cells and in [8] for interconnects. Recently, it has been shown that machine learning-based classification techniques using the maximum frequencies $F_{max}^c(V_{dd})$ of a circuit $c \in C$ under multiple voltages V_{dd} can distinguish defects on critical paths from variations with a high accuracy [6]. C is the population of all produced chips. Let V_{op} be the set of applied supply voltages. Vectors $M_c = (F_{max}^c(V_{dd}) | V_{dd} \in V_{op}), c \in C$ contain frequencies $F_{max}^c(V_{dd})$ under different supply voltages $V_{dd} \in V_{op}$ for each circuit c.

The advent of the FinFET technology reduced the effects of variations to some extent but did not remove it completely [9] [10]. Adaptive Voltage Frequency Scaling (AVFS) is used as means to overcome the effects of variations and provides the necessary infrastructure to measure the vectors M_c . However,

the measurement for numerous devices under multiple voltages is costly. On the other hand, speed-binning reports of Adaptive Voltage Frequency Scaling (AVFS) systems contain some fraction of these vectors. Each speed-binning procedure has a specific range and resolution for the applied supply voltages, and $F_{max}(V_{dd})$ is generated for specific voltages V_{dd} . To use the speed-binning report for the aim of defect identification the missing data has to be measured separately.

This paper proposes an efficient feature selection technique to identify a minimized set of features $F_{max}(V_{dd})$, which takes the already measured timing information from the speedbinning report into account for an accurate classification.

II. EFFICIENT FEATURE SELECTION

The starting point of the proposed technique is a set of vectors $\{M_c \mid c \in C\}$, where the set of voltages V_{op} sweeps from the minimum voltages V_{min} over the nominal voltage V_{nom} to the maximum voltage V_{max} in a certain step width Δ , leading to $\lceil \frac{V_{max}-V_{min}}{\Delta} \rceil + 1$ different values in V_{op} . The vectors M_c can be obtained by a Monte Carlo technique using the process variability parameters for timing-aware simulation, Static Timing Analysis (STA), or even physical measurements for a set of real devices. The vectors have to include results of both defective and defect-free devices, and are used as a dataset for supervised learning as described in [6].

Our proposed feature selection method is a variant of the methods in textbooks such as [11] [12]. It takes this dataset to determine a minimized subset $\tilde{V}_{op} \subset V_{op}$ of voltages and hence features $F_{max}(V_{dd})$, $V_{dd} \in \tilde{V}_{op}$, which contains already available measurements $V'_{op} \subset \tilde{V}_{op}$. The new shorter vectors $\tilde{M}_c = (F_{max}(V_{dd}) \mid V_{dd} \in \tilde{V}_{op})$ are used to train a classifier for the circuits $c \in C$. The reduced set $\tilde{V}_{op} \subset V_{op}$ is not only used for reducing the measurement costs, but also for reducing the noise in the dataset. As a result, the classification precision with the selected smaller feature set can be even higher than the precision obtained by using the complete set V_{op} . It also proves the robustness of the machine learning-based defect identification against reduced or missing data.

The special format of our problem allows us to aim for an optimal, exhaustive solution. The original set of features determined by V_{op} is rather small and an exhaustive search of the $2^{|V_{op}|}-1-|V_{op}|$ possible subsets is practicable. In addition, so-called *wrapper* techniques can increase the efficiency [13].

The detailed steps of the proposed feature selection method can be seen in Algorithm 1. N is defined by the user and given

to the algorithm as the maximum number of features. ACC is a 2D array with two rows and *i* columns. The first row stores the classification accuracy for each round of *i* subset selection, and the second row presents the corresponding subsets of voltages. The ordered subsets of voltages according to their obtained accuracy is called the Importance List in this work, in a way that the first element in this list is the most important voltage set (V_{op}) . *i* is the number of possible subset selections with the size of $N - |V'_{op}|$ from the remaining voltages (V_{rem}), which is determined by the binomial coefficients. For instance the number of possible 2-tuple voltages from the set of $|V_{rem}| =$ 12 voltages is $C_{12}^2 = 66$.

Algorithm 1 Algorithm for Selection of the Most Efficient Features in ML-based Defect Identification

Input: Original Dataset $\{M_c \mid c \in C\}$, Operating V Set V_{op} , Preselected V Set V'_{op} , Max Number of Features N

Output: Optimum Set V_{op}

- Initialisation :
- Remaining V Set V_{rem} := V_{op} \ V'_{op};
 Classification Accuracy ACC = [0..1][0..i − 1] Loop Process for each Selected Subset:
- 3: for i = 1 to $\binom{|V_{rem}|}{N |V'_{op}|}$ do
- Select A New Voltage Set V_{op}^{i} in which 4: $V_{in}^i \subset V_{rem} \& |V_{in}^i| = N - |V_{in}'|$

5: Vector
$$M_c^i := (F_{max}(V_{dd}) \mid V_{dd} \in V_{op}^i \bigcup V_{op}');$$

6: Dataset
$$DS^i := \{M_c^i \mid c \in C\};$$

- Create RF Classifier with DS^i : 7:
- $ACC[0][i] := Acc^i;$ 8.
- $ACC[1][i] := V_{op}^i;$ 9.
- 10: end for
- 11: Sort ACC[[][] Descending based on ACC[0][];
- 12: Importance List := ACC[1][];
- 13: $\tilde{V}_{op} := ACC[1][0];$
- 14: return V_{op}

III. CLASSIFICATION RESULTS

The approach has been applied on datasets generated for 3 multiplier circuits with different operand sizes as case studies. The circuits are synthesized by a commercial synthesis tool under tight timing constraints, which results in circuits with 440, 1987, and 7444 cell instances for Multipliers with 8, 16, and 32-bit operands respectively. For a fair comparison, the 2-input gate equivalents are 31.5, 45.5, and 56 on the longest paths for each of the multipliers.

The original datasets for each feature selection contain $F_{max}(V_{dd})$ for 13 different supply voltages between 0.4V and 1.0V, with the step of 0.05: $V_{op} = \{0.4, 0.45, .., 0.95, 1.0\}$. The details of the circuits and datasets can be found in [6].

To select the most effective feature combinations two scenarios for the preselected voltage conditions are considered:

- Cond.1: One feature is preselected corresponding to the $F_{max}(V_{dd})$ measurement at the nominal voltage $V_{dd} = 0.7 V$. One Single or two 2-Tuple voltages are selected from the V_{rem} .

- Cond.2: Three features are preselected corresponding to the maximum, minimum, and nominal voltages 0.4 V, 0.7 V, 1.0 V, and the feature selection method selects one Single voltage from the V_{rem} .

The feature selection algorithm reports the importance list for each of the scenarios and creates the classifiers for each V_{op} . The most relevant metric for our problem is **precision**, which corresponds to the possible yield loss, while the fault coverage mainly depends on the design and the test program. The classification precision by having the most important voltage in the V_{op} and when all the voltages are available is presented in Tab. I.

TABLE I CLASSIFICATION PRECISION FOR THE TOP SELECTED VOLTAGES FROM THE IMPORTANCE LIST (Prec.sel) VS. FOR ALL VOLTAGES (Prec.all)

Cond.	Circuits	Single/2-Tuple	Sel. V	$Prec{sel}$	Prec.all
	Mul.8	Single	[1.0]	0.881	0.916
		2-Tuple	[1.0,0.4]	0.914	0.916
1	Mul.16	Single	[1.0]	0.741	0.784
		2-Tuple	[1.0,0.6]	0.782	0.784
	Mul.32	Single	[1.0]	0.674	0.719
		2-Tuple	[1.0,0.8]	0.718	0.719
	Mul.8	Single	[0.5]	0.921	0.916
2	Mul.16	Single	[0.9]	0.801	0.784
	Mul.32	Single	[0.9]	0.721	0.719

It can be seen in Tab. I that with the selection of only one or two additional features the RF-based defect identification can already obtain a precision close to the one when all 13 features are available or even higher, with at least 77% savings in computational or measurement effort.

ACKNOWLEDGEMENTS

This work is funded by the DFG under grant WU 245/22-1 (ACCROSS), partially supported by Advantest as part of the Graduate School "Intelligent Methods for Test and Reliability" (GS-IMTR) at the University of Stuttgart.

References

- S. Hellebrand, T. Indlekofer, M. Kampmann, M. Kochte, C. Liu, and H. Wunderlich, "FAST-BIST: Faster-than-at-Speed BIST Targeting Hidden Delay Defects," in *Proc. IEEE Int'l Test Conf.*, 2014, pp. 1–8.
 C. Liu, E. Schneider, M. Kampmann, S. Hellebrand, and H. Wunder-lich, "Extending Aging Monitors for Early Life and Wear-Out Failure Prevention," in *Proc. IEEE Asian Test Symp.*, 2018, pp. 92–97.
 M. Tabranizor K. Paeng and K. Chakrabarty. *Test and diagnosis for*
- M. Tehranipoor, K. Peng, and K. Chakrabarty, *Test and diagnosis for small-delay defects*. Springer, 2011.
 Z. Najafi-Haghi, M. Hashemipour Nazari, and H. Wunderlich, "Variation-Aware Defect Characterization at Cell Level," in *Proc. IEEE Trans. Test*.
- *European Test Symp.*, 2020, pp. 1–6. Z. P. Najafi-Haghi and H.-J. Wunderlich, "Resistive Open Defect Classification of Embedded Cells under Variations," in *Proc. IEEE Latin*-[5]
- American Test Symp., 2021, pp. 1–6.
 [6] Z. P. Najafi-Haghi, F. Klemme, H. Amrouch, and H.-J. Wunderlich, "On extracting reliability information from speed binning," in *Proc. of the*

- [6] Z. P. Najah-Haghi, F. Klemme, H. Ahnouen, and n.-s. wandernen, S. extracting reliability information from speed binning," in *Proc. of the IEEE European Test Symp.*, 2022.
 [7] Y. Liao, Z. P. Najah-Haghi, H.-J. Wunderlich, and B. Yang, "Efficient and robust resistive open defect detection based on unsupervised deep learning," in *Proc. of IEEE Int'l Test Conf.*, 2022, pp. 185–193.
 [8] A. Sprenger, S. Sadeghi-Kohan, J. D. Reimer, and S. Hellebrand, "Variation-aware test for logic interconnects using neural networks a case study," in *Proc. IEEE Int'l Test Conf.*, 2022, pp. 100–103.
 [9] H. Amrouch, G. Pahwa, A. D. Gaidhane, C. K. Dabhi, F. Klemme, O. Prakash, and Y. S. Chauhan, "Impact of variability on processor performance in negative capacitance finfet technology," *IEEE Trans. on Circuits and Systems*, vol. 67, no. 9, pp. 3127–3137, 2020.
 [10] S. Natarajan *et al.*, "A 14nm logic technology featuring 2nd-generation finfet, air-gapped interconnects, self-aligned double patterning and a 0.0588 µm2 sram cell size," in *IEEE Intl. Electron Devices Meeting*.
 [11] M. Kuhn and K. Johnson, *Feature engineering and selection: A practical approach for predictive models*. CRC Press, 2019.
 [12] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research*, pp. 1157–1182, 2003.
 [13] R. Kohavi and G. H. John, "Wrappers for Feature Subset Selection," *Artificial Intelligence*, vol. 97, no. 1, pp. 273–324, 1997.