Rapid MOSFET Contact Resistance Extraction from Circuit using SPICE Augmented Machine Learning without Feature Extraction

Journal:	*Transactions on Electron Devices*
Manuscript ID	TED-2021-07-1764-R
Manuscript Type:	Regular
Date Submitted by the Author:	19-Jul-2021
Complete List of Authors:	Lu, Thomas; Stanford Online High School, ; San Jose State University, Electrical Engineering Kanchi, Varada; San Jose State University Mehta, Kashyap; San Jose State University Oza, Sagar ; San Jose State University Ho, Tin; San Jose State University Wong, Hiu-Yung; San Jose State University, Electrical Engineering
Area of Expertise:	CMOSFET switches, CMOSFET logic devices, Yield optimization, Reliability, Machine Learning, Circuit simulation

Rapid MOSFET Contact Resistance Extraction from Circuit using SPICE Augmented Machine Learning without Feature Extraction

Thomas Lu, Varada Kanchi, Kashyap Mehta, Sagar Oza, Tin Ho, and
Hiu Yung Wong, *Senior Member, IEEE*

***Abstract*—It is desirable to monitor the degradation of integrated circuits (IC's) or perform their failure analysis through their electrical characteristics (such as the voltage transfer characteristic, *VTC*, of an inverter). Such a method is non-destructive, low-cost, and can be applied to a large number of samples. Machine learning is naturally an excellent tool to perform this task. However, it is very expensive, in terms of time and cost, to generate enough experimental data with well-controlled defects to train a reliable machine. Moreover, IC defect signatures and features are usually embedded in the hyper-space of their electrical characteristics and are difficult to extract. In this paper, we propose to use dimensionality reduction to extract the defect signature from the IC electrical characteristics by using data generated through simulations. A CMOS inverter is used for demonstration. The drain contact resistances, which can increase due to defect or degradation, of the nMOSFET and pMOSFET in an inverter are extracted using a machine based on Autoeconder (AE). The machine is trained using data generated from SPICE simulation. The machine is then tested using experimental data and high accuracy is obtained ($R^2 > 0.9$). In particular, for the first time, through the analysis of the hidden variables, we demonstrate that the machine has effectively extracted the features automatically which obviates the cumbersome feature extraction process.**

***Index Terms*—Autoencoder, CMOS Inverter, Contact Resistance, Defects, Machine Learning, Reverse Engineering, SPICE simulation**

I. Introduction

DEFECT identification [1] and reverse engineering [2] are the daily routines in semiconductor manufacturing and fabrication processes. They rely heavily on destructive and expensive, in terms of cost and time, failure analysis (FA) tools such as SEM [3] and TEM [4]. As Moore's law reaching its end [5], fabrication processes become more complicated by using emerging 3D structures such as FinFET, stacked nanosheet, and complimentary FET [6][7][8], and the demand for rapid FA increases. Moreover, since the parasitic resistance and capacitance start dominating in highly scaled devices in the Design-Technology Co-Optimization (DTCO) era [9], the understand and analysis of the parasitic components and their degradation also become more important.

To meet the increasing demand of FA, it is desirable to find a low-cost and high-throughput methodology to analyze, narrow down or even pinpoint the defect properties qualitatively (e.g. identify the type of defects) and quantitatively (e.g. identify the contact location and its resistance value). One attractive proposal is to use the electrical characteristics to deduce the defect qualitatively and quantitively.

However, the electrical characteristics of devices (such as the Current-Voltage, *I-V*, and Capacitance-Voltage, *C-V*) and circuits (such as the *VTC* of an inverter or the butterfly curves of an SRAM) usually are not explicitly correlated to the defect properties. Therefore, machine learning is proposed in [10]-[18] to learn the correlations. To use machine learning successfully, enough training data, i.e. electrical characteristics of devices/circuits with well-controlled defects, is required. This cannot be obtained easily and accurately through the experiment at a low cost. In [10], it is proposed to use Technology Computer Added Design (TCAD) to generate *I-V* curves of *p-i-n* diodes with various epitaxial thicknesses and doping and a machine is trained to deduce the thicknesses and doping based on the measured *I-V*. This is called TCAD-augmented machine learning. Note that no feature extraction is performed in [10] and this obviates domain expertise and the resources required in feature extractions. However, such an approach can lead to overfitting and cannot be used in experimental data due to the noise in the experimental *I-V*. Principal component analysis [11], noise technique [12], and Autoencoder (AE) [13][14] are proposed to solve the overfitting issue when domain expertise and feature extraction are not used. Among them, AE is the most promising in that it gives similar results as with feature extraction when applied to experimental data [14]. Moreover, it is shown that by using the same AE architecture, one can apply to both *I-V* and *C-V* of FinFET without the need of performing separate feature

Manuscript received July 18, 2021. *(Corresponding author: Hiu Yung Wong).* This material is based upon work supported by the National Science Foundation under Grant No. 2046220.

Thomas Lu is an intern in M-PAC Lab, Electrical Engineering Department, San Jose State University, San Jose, CA, USA and is also affiliated with Stanford Online High School, CA, USA

Varada Kanchi, Kashyap Mehta, Sagar Oza, Tin Ho, and Hiu Yung Wong are with the M-PAC Lab, Electrical Engineering Department, San Jose State University, San Jose, CA, USA (E-mail: hiuyung.wong@sjsu.edu).

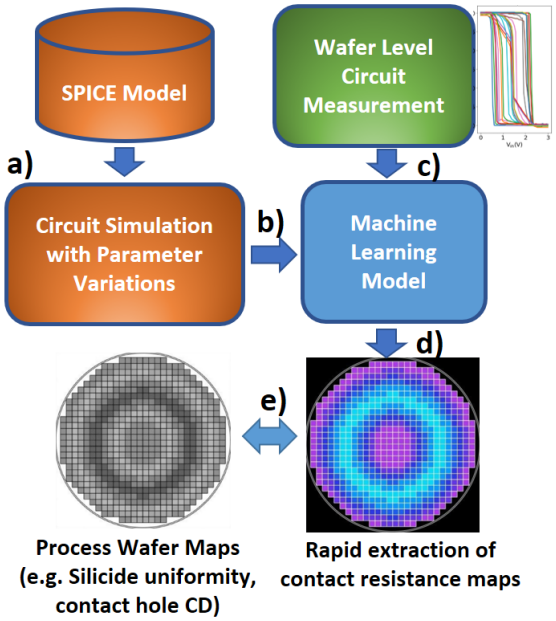

Fig. 1. Overview of the application of this study. SPICE circuit simulations, a), are performed to generate data to build a machine, b), without feature extraction. Wafer level measurement, e.g. *VTC* of an inverter, are fed into the machine, c), which can deduce the contact resistances distribution on the wafer level, d), and can be used to compare to process wafer maps to identify the defect root cause, e).

extractions for the *I-V* and *C-V* curves [15]. Note that only [11] and [14] have the results verified with experimental data.

Based on the successful experience in using AE in TCAD-augmented machine learning, in this paper, we applied a similar framework to extract the drain contact resistances of the NMOS and PMOS in an inverter based on its *VTC*. The drain contact resistance can be considered as a result of manufacturing defects [20] or degradation (e.g. due to radiation damage [21]). The training data is generated using SPICE simulation and the result is verified experimentally. This work has four major achievements. Firstly, it shows that the concept of TCAD-augmented machine learning can be extended to other types of simulations (in this case, SPICE) directly. Secondly, it shows AE-based machine is applicable also to circuit electrical characteristics without the need for architectural change. Thirdly, this is the first time to verify a SPICE-augmented machine learning on a circuit using experimental data. Lastly, through hidden node analysis, we show that the machine has successfully performed accurate feature extraction automatically. It should also be noted that feature extraction in [14] for TCAD-augmented machine learning is relatively easy. However, in the *VTC* case, the extraction of features based on *VTC* is far from trivial.

It is worth noting that there were also efforts in the literature to use TCAD trained data to deduce SOI BV properties [16], nanowire properties [17], and SRAM defects locations [18][19]. However, feature extractions are used and they have not been verified using experimental data (i.e. the trained machines are tested using unseen TCAD data only).

II. Overview

The idea of this paper can be used in the common scenario in a fab as illustrated in Fig. 1. The goal is to use SPICE simulation to rapidly generate circuit electrical characteristics (e.g. *VTC* of an inverter) with variations of the parameters in interest (e.g. drain contact resistances) to train a machine without feature extraction. Circuit measurements on the wafer level are then fed into the machine to deduce the distribution of the parameters, which is then compared to any existing process wafer maps (e.g. contact hole CD) for correlations to understand the source of defects. To demonstrate this, inverter and drain contact resistance are the circuit and parameters in interest, respectively. Inverters with various contact resistances are first measured experimentally (Section III). SPICE simulations with contact resistance variations are conducted to generate training data (Section IV) and a machine is trained (Section V) to deduce the contact resistances for any given measured *VTC*.

III. Experiment Setup

The inverter is constructed using commercial-off-the-self (COTS) PMOS and NMOS in ALD1103 by Advanced Linear Devices, Inc. fabricated using enhanced ACMOS silicon gate CMOS process [22]. Discrete resistors with accuracy better than 10% are used to model the *extra* drain contact resistance of PMOS and NMOS due to defect or degradation. They are labeled as R_0 and R_1, respectively. Fig. 2 shows the schematic of the setup. A USB-powered data acquisition module, ADALM2000 by Analog Devices [23] is used for biasing, input sweeping, and data acquisition. Since the input impedance of the sensing channel of ADALM2000 is only 1MΩ, a unity gain buffer is added at the inverter output. V_{DD} is set to be 3V.

The PMOS and NMOS are characterized by setting $|V_{GS}| = |V_{DS}| = V_{DD}/2 = 1.5V$. The drain current, $|I_{DS}|$, is measured and their effective resistance, R_p for PMOS and R_n for NMOS, is found by $1.5V/|I_{DS}|$. It is found that $R_p = 2243\Omega$ and $R_n = 744\Omega$.

Three sets of inverter experiments are then conducted. Each set uses 1 NMOS and NP PMOS in parallel, where NP = 1, 2, or 3. This represents 3 different types of inverter circuits. For example, in the set with NP = 3, it has 3 PMOS connected in parallel to represent an inverter with a stronger PMOS, i.e. the width is tripled. V_{in} is swept from 0V to 3V and the V_{out} is measured to construct the *VTC*. Different values of R_0 and R_1, between 0 to 1MΩ, are used with 0 representing no extra

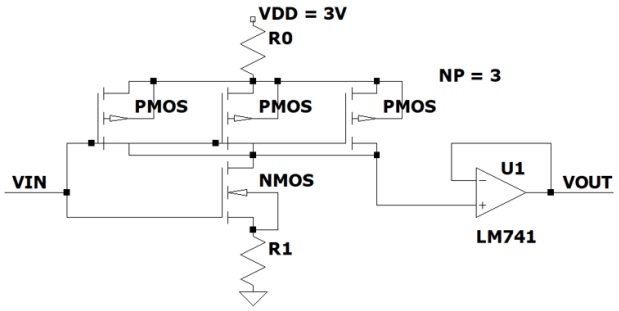

Fig. 2. Schematic circuit of the experiment. NP = 3 case is shown.

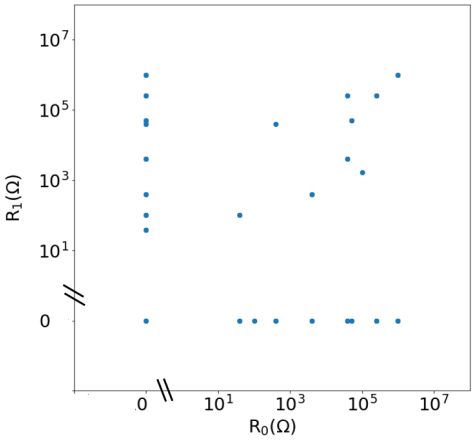

Fig. 3. Distribution of R_0 and R_1 pairs used to construct the inverter circuits for each NP group.

contact resistance. There are 25 circuits in each set of experiments. Fig. 3 shows the distribution of R_0 and R_1 used in the experiments. Fig. 4 shows the measured *VTC* of the circuits.

IV. Training Data Generation

The SPICE models of the PMOS/NMOS in ALD1103 are not provided by the vendor. To model the PMOS/NMOS in ALD1103, a SPICE LEVEL 3 model is created so that the threshold voltage and R_n and R_p match the experimental values. In any modern IC fabrication process, an accurate SPICE model is usually available and thus it is not necessary to perform this extra calibration. However, if it is a new process under development, the SPICE model might not be complete. It is thus important to show that using a primitive SPICE model (such as in our case) to generate simulation data is still useful in simulation augmented machine learning.

About 10,000 VTC curves are simulated using Cadence Spectre [24] by varying R_0 and R_1 for each NP value. The resistance values vary from 10Ω to 10MΩ logarithmically. Fig. 5 shows 75 randomly selected simulation curves from each NP group.

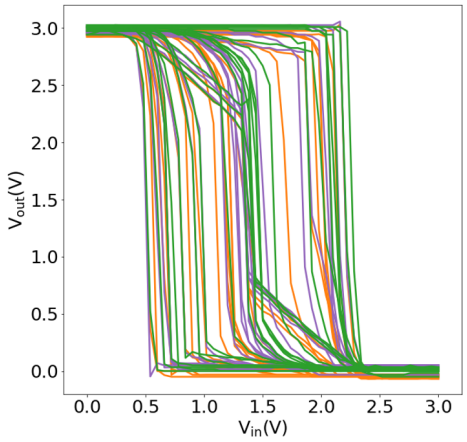

Fig. 4. *VTC's* of experimental inverter circuits with various R_0 and R_1. 3 groups of circuits are identified. Orange: NP = 1. Purple: NP = 2, Green: NP = 3.

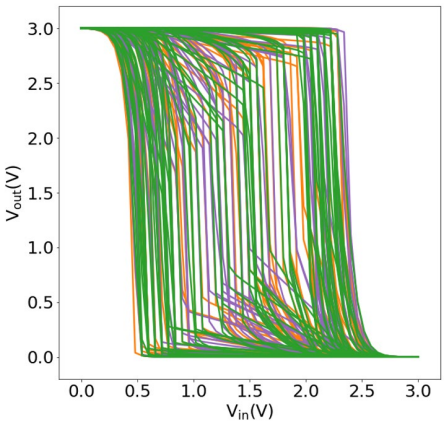

Fig. 5. Simulation VTC's. 3 groups of inverter circuits are identified. Orange: NP = 1. Purple: NP = 2, Green: NP = 3. Only 75 curves from each group are shown for clarity.

V. Machine Learning

As shown in Fig. 4 and Fig. 5, it is not trivial to extract the features that are associated with R_0 and R_1. One possibility is to extract the V_M (i.e. $V_M = V_{in}|_{V_{in}=V_{out}}$) and slopes of the regions before and after V_M which are affected not only by R_0 and R_1 but also R_n and R_p. As a result, for each set of inverters, one might need to re-optimize the feature extractions. Moreover, for small R_0 and R_1, the *VTC* is smoother but becomes more piecewise linear when R_0 and R_1 are large. Therefore, feature extractions require a lot of domain expertise (knowledge of how the R_0, R_1, R_n, and R_p affect the shape of the *VTC*).

To obviate the cumbersome feature extraction and re-extraction when different inverters are used, an autoencoder-based machine is used in this study. Autoencoder [25][26], a type of manifold learning algorithm, is capable of performing non-linear dimensionality reduction of hyperspace data into latent variables. The *VTC* curves are the hyperspace representing the inverter configuration. Since in the simulation

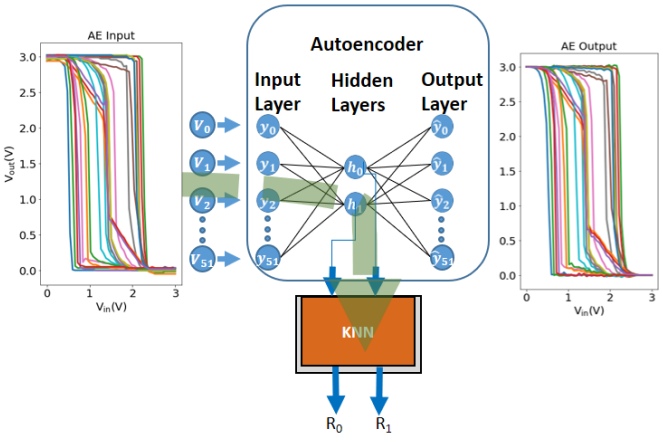

Fig. 6. The autoencoder-based machine used in this study. The machine has 5 hidden layers, each has 80, 50, 2, 50, and 80 nodes respectively. For clarity, only the middle layer is shown. Each *VTC* is discretized into 51 points and encoded as 2 latent variables, h_0 and h_1. The machine is trained such that the output values equal the input values in the autoencoder and h_0 and h_1 are correlated to R_0 and R_1 using KNN with k=3.

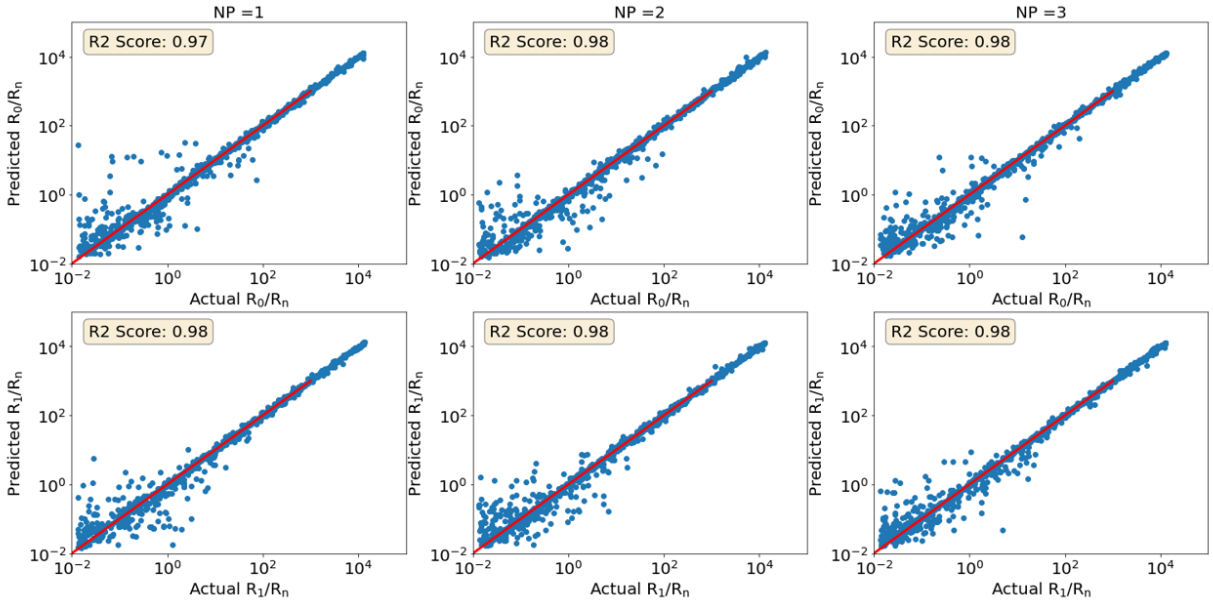

Fig. 7. Validation results of NP = 1, 2, 3 cases. R_0 and R_1 are normalized to R_n.

and experiment, only R_0 and R_1 are varied in each set of inverters, the *VTC* can be represented by a set of latent variables with much reduced dimensions. The latent variables can then be correlated to R_0 and R_1.

Fig. 6 shows the algorithm used and for clarity, only the middle hidden layer is shown. Tensorflow platform is used [27]. Each *VTC* is discretized into 51 points (y_0 to y_{50}). An AE with 5 hidden layers is used. 90% of the simulation curves are used as the training data and the rest are used for validation. The model is trained to have less than 5×10^{-5} mean squared error loss by comparing the AE input (y_0 to y_{50}) and output ($\hat{y}_0$ to $\hat{y}_{50}$). A decaying learning rate scheme is used, in which the learning rate is gradually reduced from 0.001 before 50 epochs to 0.0002 after 1700 epochs. This is found to help avoid local minima in the optimization process. The number of nodes in each hidden layer is 80, 50, 2, 50, and 80, respectively, and node counts are manually adjusted to minimize loss.

The latent variables in the middle layer, h_0 and h_1, are correlated to R_0 and R_1 through the k-th nearest neighbors (KNN) algorithm with k = 3. Overfitting is monitored and avoided by tracking the validation loss and evaluating KNN performance once every 50 epochs. No overfitting occurs and 2000 epochs are used in the final model.

Fig. 7 shows that the model performs very well and the coefficients of determination [28], R^2, are very close to 1 for both validation and training (not shown) data.

After the machine is trained, experimental *VTC's* from Fig. 4 are then fed into the machine to deduce the R_0 and R_1 of the

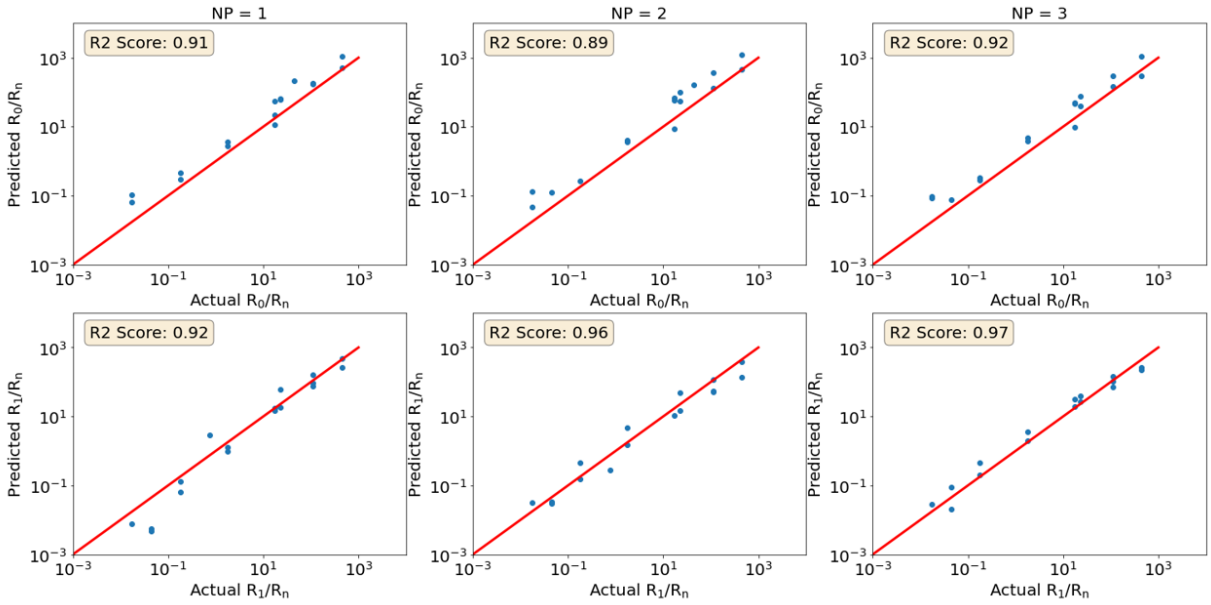

Fig. 8. Prediction of experimental R_0 and R_1. R_0 and R_1 are normalized to R_n.

inverters, without any manual feature extraction. Fig. 8 shows that it can predict R_0 and R_1 very well with R^2 larger than 0.9.

VI. Discussion

We demonstrated experimentally that, without feature extraction, by using AE, the drain contact resistances in an inverter can be deduced based on measured *VTC's*. Statistically, it deduces the extra resistance location (at PMOS, i.e. R_0, or at NMOS, i.e. R_1) and extra resistance values very well as shown in Fig. 8. In certain experiments, it can predict the abnormal resistance with values of only a few percentages of R_n well (e.g. R_1 in NP = 2 case in Fig. 8 even when $R_1/R_n <$ 10%), which is impossible in other FA methods.

Although discrete components are used in the experiment, the transistors are made in a traditional CMOS IC process. Therefore, the experimental circuits constructed represent an integrated circuit inverter well. Such an approach also obviates the need to fabricate a test chip which might have a limitation on the number of probe pads. On the other hand, such a setup introduces more noises due to parasitic capacitance and resistance than a fully integrated circuit. Since DC measurement is performed, parasitic capacitance has minimal impacts. Parasitic resistance is also negligible as it is much lower than the MOSFET resistance which is further confirmed with the accuracy in the contact resistance prediction result. Such extra noise and non-ideality are expected to make it more difficult to apply simulation augmented machine learning to the experimental data than if a fully integrated inverter were used. Since good results are obtained with the current experimental setup, it is expected that such a methodology will perform even better on fully integrated circuits.

To further quantify the ability of the AE on performing accurate feature extraction, Fig. 9 shows the input and output of the AE. Note that the AE encodes the 51-dimensional input data (y_0 to y_{50}) to a 2-dimensional space spanned by h_0 and h_1, which is then used to reconstruct the original curve as the output ($\hat{y}_0$ to $\hat{y}_{50}$). As shown, the output resembles the input very well which means that the AE has successfully automatically extracted 2 features to represent the *VTC*.

Fig. 10 and Fig. 11 show the relationship between the latent

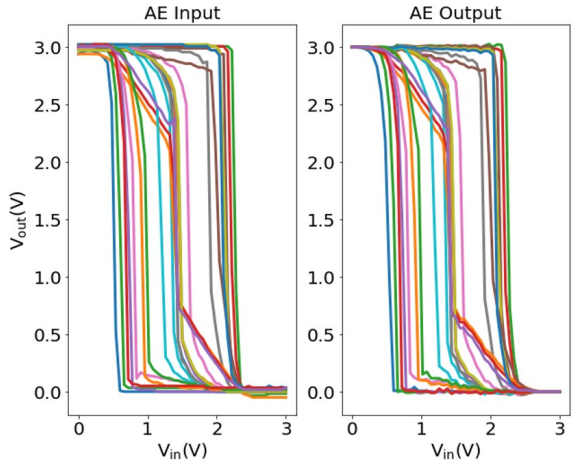

Fig. 9. Experimental *VTC's* input to the machine (Left) and the *VTC's* output from the machine (Right). NP = 3 is used.

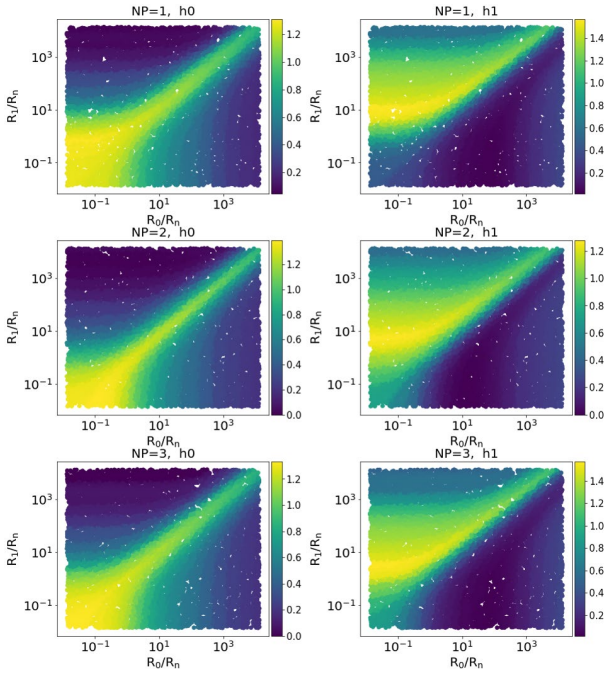

Fig. 10: h_0 and h_1 as the functions of R_0 and R_1 for NP = 1, 2, and 3.

variables and R_0 and R_1 for all 3 sets of inverters. It can be seen that all 3 different sets of inverters (NP = 1, 2, and 3), which have different V_M, show the same qualitative and very similar quantitative relationships between the contact resistances and the hidden variables. Therefore, the AE is repeatable in all 3 cases and has automatically extracted the features that

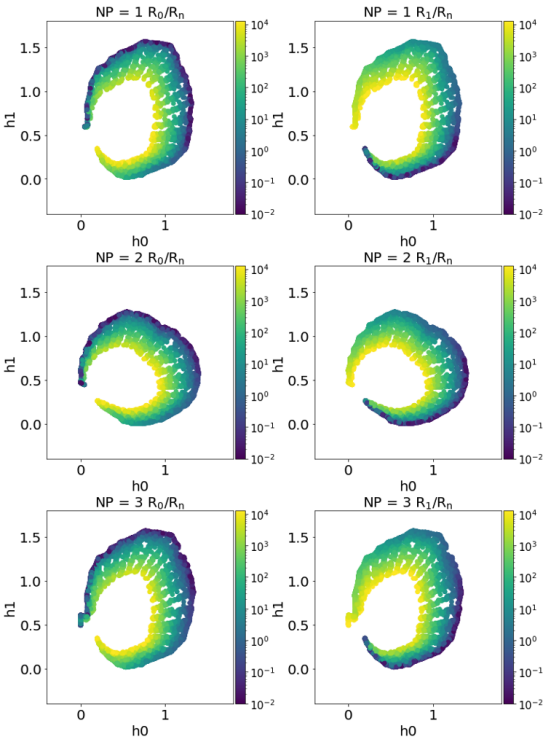

Fig. 11: R_0 and R_1 as the functions of h_0 and h_1 for NP = 1, 2, and 3.

faithfully represent the parameters, i.e. R_0 and R_1, that we are interested in.

The time required to perform 10,000 SPICE simulations for training data generation is less than 3 hours with 1 Spectre license and the time required to train the machine is only 20 minutes using Intel i7-8705G CPU with 8 logical cores. The SPICE simulation time can be further reduced to 16 minutes if only 1,000 curves are used. The R^2 of predicting the experimental data is found to be still better than 0.85 when only 1000 data are used for training. Therefore, this approach is very rapid and can be performed almost on the fly for semiconductor defect debugging and yield enhancement without the need for expensive FA.

VII. Conclusion

In this paper, we demonstrated a method to deduce the contact resistances of NMOS and PMOS in an inverter based on its *VTC* by using simulation augmented machine learning without the need for feature extraction. An autoencoder-based machine is used, which can reduce the hyper-dimensional *VTC* to a 2-dimensional latent subspace representing the contact resistances. Such a method is validated experimentally and can achieve accuracy with $R^2 > 0.9$. It is clearly shown that the latent subspace has faithfully represented the parameters in interest. Such an approach is not inverter-specific and no feature extraction is required. It is thus expected to be applicable to other circuits. Since circuit simulations are fast, this method is expected to be very useful in improving the defect debugging and yield enhancement processes.

References

[1] G. A. Susto, M. T. and A. Beghi, “Anomaly Detection Approaches for Semiconductor Manufacturing”, Procedia Manufacturing 11 (2017) 2018 – 2024.

[2] R. Torrance and D. James, "Reverse Engineering in the Semiconductor Industry," 2007 IEEE Custom Integrated Circuits Conference, San Jose, CA, 2007, pp. 429-436. doi: 10.1109/CICC.2007.4405767

[3] S. J. Pennycook and L. A. Boatner, “Chemically sensitive structure-imaging with a scanning transmission electron microscope,” Nature, vol. 336, no. 6199, pp. 565–567, Dec. 1988, doi: 10.1038/336565a0.

[4] M. J. Williamson, R. M. Tromp, P. M. Vereecken, R. Hull, and F. M. Ross, “Dynamic microscopy of nanoscale cluster growth at the solid–liquid interface,” Nat. Mater., vol. 2, no. 8, pp. 532–536, Aug. 2003, doi: 10.1038/nmat944.

[5] (2020). International Roadmap for Devices and Systems - More Moore. [Online]. Available: https://irds.ieee.org/images/files/pdf/2020/2020IRDS_MM.pdf

[6] S. -W. Chang et al., "First Demonstration of CMOS Inverter and 6T-SRAM Based on GAA CFETs Structure for 3D-IC Applications," 2019 IEEE International Electron Devices Meeting (IEDM), 2019, pp. 11.7.1-11.7.4, doi: 10.1109/IEDM19573.2019.8993525.

[7] A. Guler and N. K. Jha, "Three-Dimensional Monolithic FinFET-Based 8T SRAM Cell Design for Enhanced Read Time and Low Leakage," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 27, no. 4, pp. 899-912, April 2019, doi: 10.1109/TVLSI.2018.2883525.

[8] A. Elwailly, J. Saltin, M. J. Gadlage and H. Y. Wong, "Radiation Hardness Study of LG = 20 nm FinFET and Nanowire SRAM Through TCAD Simulation," in IEEE Transactions on Electron Devices, vol. 68, no. 5, pp. 2289-2294, May 2021, doi: 10.1109/TED.2021.3067855.

[9] V. Moroz, "Technology Inflection Points: Planar to FinFET to Nanowire", 2016 International Symposium on Physical Design, p. 55.

[10] Y. S. Bankapalli and H. Y. Wong, "TCAD Augmented Machine Learning for Semiconductor Device Failure Troubleshooting and Reverse Engineering," *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Udine, Italy, 2019, pp. 1-4, doi: 10.1109/SISPAD.2019.8870467.

[11] Hiu Yung Wong, Ming Xiao, Boyan Wang, Yan Ka Chiu, Xiaodong Yan, Jiahui Ma, Kohei Sasaki, Han Wang, and Yuhao Zhang, "TCAD-Machine Learning Framework for Device Variation and Operating Temperature Analysis With Experimental Demonstration," in *IEEE Journal of the Electron Devices Society*, vol. 8, pp. 992-1000, 2020, doi: 10.1109/JEDS.2020.3024669.

[12] S. S. Raju, B. Wang, K. Mehta, M. Xiao, Y. Zhang and H. -Y. Wong, "Application of Noise to Avoid Overfitting in TCAD Augmented Machine Learning," 2020 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Kobe, Japan, 2020, pp. 351-354, doi: 10.23919/SISPAD49475.2020.9241654.

[13] K. Mehta, S. S. Raju, M. Xiao, B. Wang, Y. Zhang and H. Y. Wong, "Improvement of TCAD Augmented Machine Learning Using Autoencoder for Semiconductor Variation Identification and Inverse Design," in *IEEE Access*, vol. 8, pp. 143519-143529, 2020, doi: 10.1109/ACCESS.2020.3014470.

[14] Harsaroop Dhillon, Kashyap Mehta, Ming Xiao, Boyan Wang, Yuhao Zhang, and Hiu Yung Wong, "TCAD-Augmented Machine Learning with and without Domain Expertise," in IEEE Transactions on Electron Devices, doi: 10.1109/TED.2021.3073378.

[15] K. Mehta and H. -Y. Wong, "Prediction of FinFET Current-Voltage and Capacitance-Voltage Curves Using Machine Learning With Autoencoder," in IEEE Electron Device Letters, vol. 42, no. 2, pp. 136-139, Feb. 2021, doi: 10.1109/LED.2020.3045064.

[16] J. Chen et al., "Powernet: SOI Lateral Power Device Breakdown Prediction With Deep Neural Networks," in *IEEE Access*, vol. 8, pp. 25372-25382, 2020, doi: 10.1109/ACCESS.2020.2970966.

[17] H. Carrillo-Nuñez, N. Dimitrova, A. Asenov and V. Georgiev, "Machine Learning Approach for Predicting the Effect of Statistical Variability in Si Junctionless Nanowire Transistors," in *IEEE Electron Device Letters*, vol. 40, no. 9, pp. 1366-1369, Sept. 2019. doi: 10.1109/LED.2019.2931839.

[18] C. Teo, K. L. Low, V. Narang and A. V. Thean, "TCAD-Enabled Machine Learning Defect Prediction to Accelerate Advanced Semiconductor Device Failure Analysis," *2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Udine, Italy, 2019, pp. 1-4, doi: 10.1109/SISPAD.2019.8870440.

[19] J. Pan, et al.,"Transfer Learning-Based Artificial Intelligence-Integrated Physical Modeling to Enable Failure Analysis for 3 Nanometer and Smaller Silicon-Based CMOS Transistors," Early Access, https://doi.org/10.1021/acsanm.1c00960.

[20] Y. Cui, J. Y. Jeong, Y. Gao, and S. G. Pyo, “Effect of Contact Plug Deposition Conditions on Junction Leakage and Contact Resistance in Multilevel CMOS Logic Interconnection Device,” *Micromachines*, vol. 11, no. 2, p. 170, Feb. 2020. doi.org/10.3390/mi11020170.

[21] C. Fares and F. Ren, “"Effect of alpha-particle irradiation dose on SiNx/AlGaN/GaN metal–insulator semiconductor high electron mobility transistors", Journal of Vacuum Science & Technology B 36, 041203 (2018) https://doi.org/10.1116/1.5042261.

[22] https://www.aldinc.com/pdf/ALD1103.pdf

[23] https://www.analog.com/en/design-center/evaluation-hardware-and-software/evaluation-boards-kits/adalm2000.html#eb-overview

[24] https://www.cadence.com/en_US/home/tools/custom-ic-analog-rf-design/circuit-simulation/spectre-simulation-platform.html

[25] G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," Science, Vol. 313, Issue 5786, pp. 504-507. DOI: 10.1126/science.1127647

[26] M. A. Kramer, "Nonlinear principal component analysis using autoassociative neural networks" AIChE Journal. 37 (2): 233–243. doi:10.1002/aic.690370209.

[27] https://www.tensorflow.org/

[28] Coefficient of Determination. In: The Concise Encyclopedia of Statistics. Springer, New York, NY., 2008. doi.org/10.1007/978-0-387-32833-1_62