Accelerating Reliability Analysis for Aging and Self-heating using Machine Learning

Tarek Mohamed[1] and Hussam Amrouch[1,2]

[1]Semiconductor Test and Reliability, University of Stuttgart, Stuttgart, Germany
[2]Chair of AI Processor Design, Technical University of Munich, Munich, Germany
email: amrouch@tum.de

Abstract—**The relentless drive for advanced technologies, fueled by the demands of AI and safety-critical applications, has intensified the focus on transistor aging—a pivotal concern that undermines both transistor reliability and overall circuit performance. This issue is further exacerbated by advancements in packaging and 3D integration, where elevated operating temperatures accelerate aging mechanisms. As technology nodes scales below 3 nm, transistor self-heating emerges as a fundamental challenge, driven by the thermal constraints of 3D confined structures. Traditional physics-based simulation tools, such as Technology CAD (TCAD), struggle to meet the growing computational demands of these intricate designs, with escalating simulation times that impede comprehensive design exploration and optimization. Here, we present a novel framework leveraging machine learning (ML) to accelerate transistor and circuit reliability analysis. These ML-driven methodologies achieve accurate predictions of self-heating and aging effects, enabling rapid identification of aging-prone transistors while drastically reducing computational overhead. Furthermore, by obviating the need to share proprietary, physics-based models from semiconductor foundries, these techniques preserve data confidentiality, addressing critical industry concerns. Such approach not only enhances the scalability of reliability assessments but also offers a transformative pathway for tackling the multifaceted challenges of next-generation semiconductor technologies.**

Index Terms—**Reliability, aging, self-heating, graph neural network, TCAD, SPICE, self-supervised graph attention network**

I. INTRODUCTION

Reliability of semiconductor devices has become the most important part of the design process. One of the main concerning factors for semiconductor device reliability is the transistor aging [1], which stems from the building material of the transistor breaking down and degrading overtime. The resulting impact on transistor performance can have a wide range of effects on the performance of the circuit from simple errors all the way to complete circuit failure [2]. Transistor aging is thus at the forefront of research as the main contributor to the overall semiconductor reliability. The ongoing reduction of transistor node size has led to accelerated transistor aging due to SHE [3]. SHE has become one of the most significant transistor reliability factors, as the smaller transistor size and more confining geometry leading to trapped heat inside the

transistor channel creating a hotspot [4]. Which accelerates the degradation of transistor material thus a faster aging process cutting down the lifetime of the transistor. The main driving factors behind transistor aging can be traced to two major phenomena. The first is Bias Temperature Instability (BTI), which occurs due to the constant and prolonged exposure to voltage bias and elevated device temperature. The accumulation of the charges on and near the gate leads to a change in the transistor's threshold voltage (V_{th}) [5]. The shift in V_{th} causes slower transistor switching times leading to timing errors [5]. BTI effect is accelerated by SHE as the elevated temperature amplify the charges trapping process, resulting in accelerated aging for the transistor and a shorter lifetime for the circuit [6]. The second main contributor to transistor aging is Hot Carrier Injection (HCI). HCI occurs when high energy carriers collide with the gate oxide or substrate, resulting in the carriers being injected into the gate oxide or substrate [7]. This carrier collision results in defects to the material, generated defects over time accumulate and result in a shift in V_{th} [8]. As temperature of the transistor is one of the main factors influencing HCI, as carriers become more energetic with the higher temperature. Elevated temperature alongside, higher drain-to-source voltages, and technology scaling leading to higher electric field within the transistor are all factors amplifying the effects of HCI. To be able to anticipate and work around the aging process and its complicated impact on transistor performance, highly detailed physics-based simulation tools need to be implemented. As such, using Technology CAD (TCAD) [9] tools, has become the standard for both academic and industrial design-technology co-optimization work. The high accuracy is by virtue of employing detailed physics models and equations. Due to recent advancements in technology scaling, as well as the introduction of complex 3D transistor structures. The time and computational requirements for TCAD tools have increased exponentially to the point where it is no longer feasible, to find an optimal solution to a design problem. Additionally, when designing at the circuit level, the problem is even more pronounced as due to the huge number of transistors only wasteful worst-case predictions are used in order to guarantee functionality. Machine Learning (ML) and smart learning approaches applied to design methods at all levels from transistor to circuit in conjunction with existing simulation-based tools offer the next evolution of the

technology design process. In this paper, we provide a deep dive into four ML methods used to accelerate and enhance the design process from transistor to circuit level. Highlighting different ML approaches and techniques like Graph Neural Network (GNN) [10], Convolutional Neural Network (CNN), and self-supervised learning. These ML models alongside domain knowledge offer a paradigm shift in the time and computational requirements of the design process, without compromising the privacy of manufacturing foundries.

Our key contributions within this work are as follows:

1) We present GNN-based TCAD acceleration model [11]. We employ a self-supervised variation of the Graph Attention Network (SuperGAT), we tested our model on a calibrated $14\,\mathrm{nm}$ FDSOI device TCAD model. Our model achieves an accuracy of $97.1\,\%$ in predicting the transistor transfer characteristics of $I_d - V_g$, while achieving a speedup of more than $100\,000$ x compared to TCAD simulation tools.

2) We present CNN-based thermal TCAD acceleration model. Our model is tested on a calibrated $28\,\mathrm{nm}$ FDSOI device TCAD model from cryogenic to room temperature. We employ a hybrid CNN architecture with dense and 1D convolution layers. Our model achieves an accuracy of $96.66\,\%$ for transistor thermal profile and $97.44\,\%$ for hotspot prediction, while achieving a speedup of more than $13\,000$ x compared to TCAD simulation tools.

3) We present ML-based susceptibility analysis and classification model [12]. We employ a Graph Attention Network (GAT). Our model is tested on $7\,\mathrm{nm}$ FinFET standard cells from the open-source ASAP7 standard cell library. Our classifier achieves a $80.4\,\%$ accuracy in identifying transistors susceptible to aging in a circuit only via examining the topology of the circuit.

4) We present an ML-based workload-based aging prediction tool that predicts the ΔV_{th} after 10 years given a specific workload. We employ a CNN-based model, our model is trained and tested on a large circuit 32-bit MAC and tested on an 8-bit adder to demonstrate the generality of our model. Achieving $94.29\,\%$ and $92.91\,\%$ accuracy for 32-bit MAC and 8-bit adder, respectively.

II. MACHINE LEARNING FOR TCAD ACCELERATION

In this section we focus on the transistor simulation acceleration using ML. In Section-II-A we introduce a self-supervised GNN-based model predicting electrical characteristics of various transistor configurations. In Section-II-B we introduce a CNN-based model for transistor thermal profile prediction.

A. Self-Supervised Graph Neural Network for Transistor Electrical characteristics Prediction

The electric characteristics simulation is one of the most important aspects of TCAD simulation. Our focus in this approach [11] is to capture as much detail as possible from the transistor TCAD mesh structure and translate that into
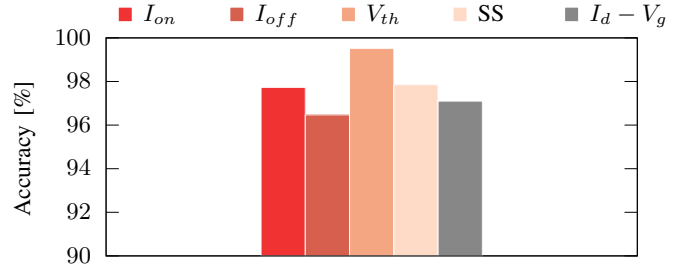

Fig. 1. Accuracy of prediction in in terms of on current (I_{on}), off current (I_{off}), threshold voltage (V_{th}), subthreshold swing (SS), and $I_d - V_g$ accuracy % of our SuperGAT model, compared to calibrated TCAD model.

an ML appropriate format. To be able to capture transistor information including 3D structure, we needed to abandon traditional ML approaches like CNN, which can only deal with flat data structures. Our solution is to make use of GNN which is a graph based ML method that can directly work with non-flat data captured in a graph format. We employ a self-supervised variation of the Graph Attention Network (SuperGAT) [13] architecture implemented via PyTorch Geometric [14] and Nvidia's Cuda library [15] training on an NVIDIA A100 graphics card. The graph interpretation of a transistor TCAD mesh structure is straightforward. Mesh points are converted into graph nodes and the mesh connection into graph edges. We use a calibrated $14\,\mathrm{nm}$ FDSOI device model as the basis for our dataset for training and testing. The device model is calibrated according to experimental data provided by [16]. Our SuperGAT-based model is able to predict important electrical characteristics of never seen before transistor configurations, such as V_{th}, subthreshold swing (SS), on current (I_{on}), and off current (I_{off}), and reproduce the full the $I_d - V_g$ curves. Our SuperGAT model is trained and tested using a calibrated TCAD dataset comprised of 540 $I_d - V_g$ curves, which is then split via the K-fold [17] method into training/validation/testing splits as follows: $80\,\%$ / $10\,\%$ / $10\,\%$. As seen in Fig. 1 the resulting prediction from the SuperGAT model over the testing never seen before dataset has a minimum accuracy value of $96.47\,\%$ for the I_{off} prediction and a maximum accuracy of $99.5\,\%$ for the SS prediction. While I_{on}, V_{th}, and $I_d - V_g$ illustrate high accuracy numbers of $97.73\,\%$, $97.85\,\%$, and $97.1\,\%$ respectively. Our SuperGAT model achieves a minimum R^2 score of 0.992 across all transistor characteristics. In Fig. 2, a comparison between TCAD simulation and prediction of our SuperGAT model in terms of $I_d - V_g$ curve, with our model achieving an accuracy of $98.56\,\%$ for this transistor configuration. Our SuperGAT model achieves a speedup of more than $100,000$x, with an inference time of just 0.007 seconds per transistor, compared to 13 minutes in TCAD simulation.

B. Convolutional Neural Network for Transistor Self-Heating Effect (SHE) Prediction

SHE has become a major reliability and longevity concern, as it accelerates the transistor aging process. SHE is especially prominent in cryogenic temperatures as at low temperatures,
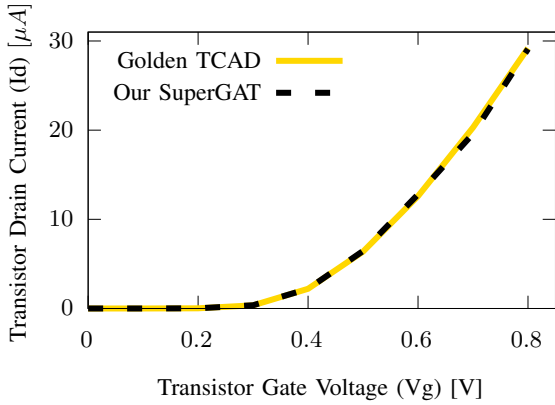
Fig. 2. $I_d - V_g$ curve prediction from our SuperGAT models against calibrated TCAD simulation data. Our SuperGAT model achieves an accuracy % of 98.56 % for this transistor configuration operating at 300K ambient temperature.

heat dissipation is lower, thermal conductivity significantly decreases, and carrier mobility increases, resulting in amplified SHE. The only accurate way to anticipate SHE on the device's internal temperature and performance is to use physics-based simulation tools like TCAD, which although offering very accurate simulation results has two major drawbacks. First is the model calibration effort, which is a long and intensive manual process. Model calibration is also only possible given the existence of experimental data at the desired temperature for the model to be calibrated against and re-calibrated if a different ambient temperature needs to be studied. Secondly, even after the long calibration process simply running the simulation tool is both a time and computationally expensive endeavor limiting the scope of design space that can be explored. Our approach to addressing these drawbacks is to use a CNN-based method that can extrapolate and predict the thermal profile of never seen before ambient temperatures from cryogenic to room temperature. Our CNN model is running via NVIDIA's Cuda package [15] and implemented by TensorFlow [18] training on an NVIDIA A100 graphics card. We employ a 28 nm FDSOI device, which is commonly used in cryogenic applications due to it offering a voltage body bias. We create a dataset based on 28 nm FDSOI device TCAD calibrated model operating at an ambient temperature of 77K, 150K, and 300K. Our device model is calibrated according to experimental data provided by [19]. To generate dataset we do a full operational range sweep of the device 0-0.9 V and 0-1.2 V for V_d and V_g respectively. Resulting in 390 TCAD simulations, with 130 for each ambient temperature. The training dataset is made of 260 TCAD simulation results for both 77K, and 300K, while the 130 TCAD simulations for 150k are used for testing. TCAD thermal data is converted into 2D graph displaying the temperature values staring from the drain across the channel to the source of the device. This thermal map is used as the input to the CNN model alongside 19 input features we have developed to support the learning process of the CNN model. Feature engineering is a huge part of this work, as simply using the voltage biases i.e (V_d and V_g) as
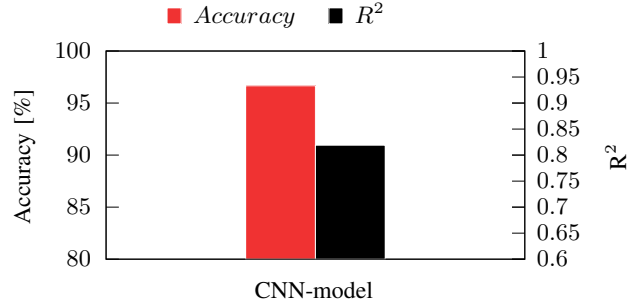

Fig. 3. Accuracy percentage (left-hand axis) illustrated in red and the R^2 score (right-hand axis) illustrated in black, for our CNN-based model tested on transistors operating at a temperature of 150K.

input features showed poor correlation with the device SHE. We provide two examples of the features used as input for our CNN. The correlation coefficients are used to illustrate the relevance of each feature to the prediction target (device SHE resulting temperature). The first engineered input feature Feature-1 is based on the operating voltage of the transistor and the ambient temperature. Eq.1 shows Feature-1, and it achieves a correlation coefficient score of (0.867).

$$\texttt{Feature-1} = -\frac{1 - ((V_g * V_d))}{Ambient.Temp} \qquad (1)$$

Another example of an engineered input feature is Feature-2, which is also based on operating voltage and ambient temperature, but is calculated in a different arithmetic way to give a different correlation aspect to the device temperature. Eq. 2 shows Feature-2 calculation, Feature-2 achieves a correlation coefficient score of (0.764).

$$\texttt{Feature-2} = ((V_g + V_d) * Ambient.Temp) \qquad (2)$$

Our CNN model is made of six dense layers with a hidden dimension of 1024, 512, 256, 64, 8, and a 1 respectively, and a 1D convolution layer with a dimension of 4096 and a kernel size of 3, finally two dropout layers placed after the first and second dense layers. Illustrated on Fig. 3 is the accuracy percentage (96.66 %) and the R^2 score (0.819) for oAmbientur CNN model tested on never seen before operating temperature of 150K. Fig. 4 illustrates the hotspot comparison between our CNN model prediction and TCAD simulation, our model achieves an accuracy of 97.44 % applied to the hotspot at an ambient temperature of 150K across a sweep of the V_g. Our CNN model achieves high accuracy of 96.66 % for transistor thermal profile, 97.44 % for hotspot prediction, and a high R^2 0.819. Our model achieves a significant speedup of more than 13 000 x, while maintaining high accuracy.

III. MACHINE LEARNING FOR CIRCUIT TRANSISTOR SUSCEPTIBILITY ANALYSIS AND CLASSIFICATION

In this work [12], we provide a method for identifying aging susceptible transistors in a circuit via the use of a Graph Attention Network (GAT)-based [20] model. To generate a dataset where we can identify aging susceptible transistors,
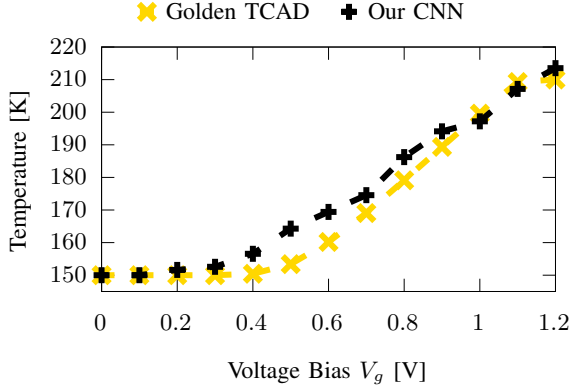
Fig. 4. The voltage bias sweep across V_g on the X-axis, With the V_d fixed at the highest operating value of 0.9 V for operating temperature 150K. The Y-axis shows the highest temperature of the device channel at each voltage bias. The TCAD simulation data is illustrated in the gold, and the CNN predicted data is illustrated in black. With the values points highlighted in a cross.

we needed to simulate circuits with degraded transistors. We represent aging degradation in this work as a shift in the transistor's V_{th} (ΔV_{th}). We selected ΔV_{th} to represent the aging effect as it is the most commonly effect parameter, due to the aging process. We label a transistor as susceptible if after applying ΔV_{th} to represent the aging effect, the circuit delay is impacted negatively i.e (the circuit delay increases with the aged transistor). The ΔV_{th} value selected is $50\,\mathrm{mV}$ as it represents a 10-year aging degradation in an intense operating environment [21]. To identify susceptible transistors a single simulation at a high ΔV_{th} is not sufficient data, as shown in [22] the biggest change in the circuit delay can occur at a lower ΔV_{th}, so a full simulation sweep of different ΔV_{th} ($10\,\mathrm{mV}$, $20\,\mathrm{mV}$, $30\,\mathrm{mV}$, $40\,\mathrm{mV}$ and $50\,\mathrm{mV}$) values had to be done to have accurate transistor classification data. If the aged transistor impacts the circuit delay negatively at any ΔV_{th} value then the transistor is labeled as susceptible. An example of the resulting transistor classification is carried out on a NAND3x2 cell and can be seen in Fig. 5, transistors P1, P2, P3, and N3 are classified as susceptible highlighted in red, and transistors N1, and N2 are classified as non-susceptible highlighted in green. The simulation and labeling process is carried out for 130 cells and 1152 transistors in the ASAP7 Process Design Kit (PDK) [23] standard cell library. After creating the dataset with transistor labels as susceptible or non-susceptible, the netlist of all cells is converted into a format that can be processed by the GAT model, the conversion process is illustrated in Fig. 6. Standard cell netlist is converted into a heterogeneous graph featuring different node types to represent different elements of a netlist such as transistors, supply ports (VDD, GND), and I/O ports. All connections between elements in the netlist are converted into graph edges to convey a topologically accurate depiction of the netlist representing all elements and connections between them. Our dataset is then split via the K-fold [17] method into training/validation/testing sets ($72\,\%$ / $8\,\%$ / $20\,\%$ respectively), with random selection performed five times to generate five folds. Our GAT-based model is implemented using the Pytorch
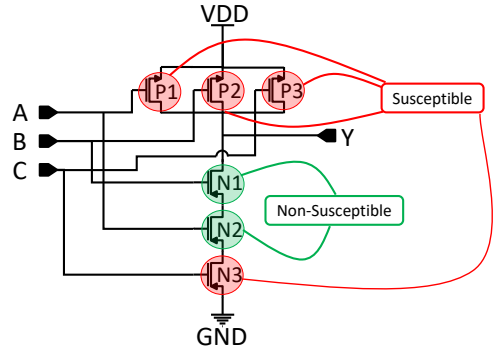

Fig. 5. An illustration of an NAND3x2 cell. The resulting classification of transistors after aged transistor simulation. Transistors P1, P2, P3, and N3 being classified as susceptible are highlighted in red. Transistors N1, and N2 being classified as non-susceptible are highlighted in green.
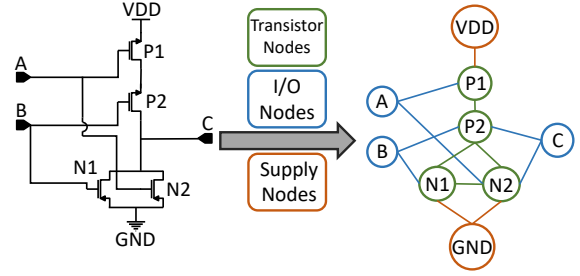

Fig. 6. An illustration of the conversion process of an example NOR two input cell. The different types of nodes are highlighted in different colors transistor nodes in green, I/O nodes in blue, and supply nodes in orange.

Geometric package [14] and running via NVIDIA's Cuda package [15] training. The model consists of two GAT layers each followed by a linear layer, a single dropout layer is utilized to counter overfitting to the training data [24]. Our GAT model achieves an $80.4\,\%$ accuracy identifying aging susceptible transistors on unseen circuits.

IV. MACHINE LEARNING FOR TRANSISTOR WORKLOAD DEPENDENT AGING

The extent of a transistor's aging is dependent on the workload of that specific transistor. This makes it difficult to have an estimate for the guard-bands to protect the transistor and the circuit. As the foundries are not willing to share their physics-based aging models with designers due to the confidential information enclosed in them. Even if the foundries were willing to share their aging models, there is still a problem with the computational and time requirement of their models. Making the worst-case guard-bands estimate the only viable option for the designer, wasting potential performance. In this work. We propose a ML-based solution to resolve the confidentiality issue while offering a much faster and less computationally demanding solution. We generate a dataset made up of two larger circuits, an 8-bit adder and a 32-bit MAC. The dataset generation and SPICE simulation are done with the CARAT framework [25]. The circuits are consisting of 14nm FinFET [26]. The CARAT framework illustrated in Fig. 7 starts by carrying out SPICE simulation for each workload, as since in larger circuits, not every transistor
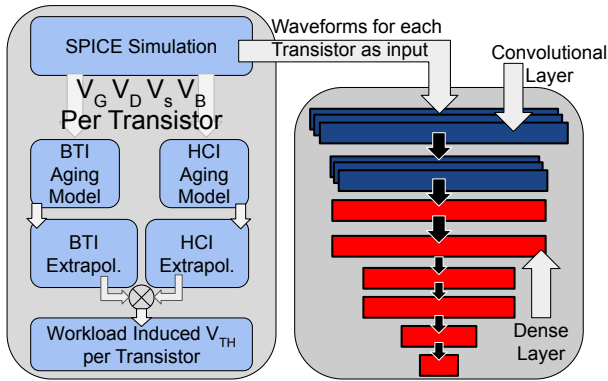
Fig. 7. On the right of the dataset generation flow via the CARAT framework the first step is SPICE simulation then short term physics based models, finally the extrapolation. On the left is an illustration of the CNN with the convolution and dense layers illustrated in blue and red respectively.

is connected directly to the input, the circuit needs to be simulated in order to get the voltage waveforms of each transistor. The generated voltage waveforms are V_D, V_G, V_S, and V_B. The individual transistor voltage waveforms as well as the temperature T are used as input to the short-term physics-based aging model to generate the resulting ΔV_{th} from short-term aging. The CARAT framework employs two physics-based aging models, BTI model (BAT [27]) and HCD model (HEAT [28]). Carrying-out simulation for 10 years lifespan in SPICE is not practical, so the output of the physics based models is extrapolated by following the trajectory of the DC degradation cure to give a degradation cure that is based on the workload of the transistor. The output of the CARAT tool is the ΔV_{th} after 10 years, given the workload of each transistor. This work flow is then repeated for every stimuli input to each circuit to generate the dataset used for training and testing for our ML model. The 8-bit adder is made up of 222 transistors, and the 32-bit MAC is made up of 2790 transistors. To simulate different workloads, random stimuli are generated as input for both circuits. 16 random workloads are generated for the 32-bit MAC circuit, and 200 random workloads are generated for the 8-bit adder. This results in a total of 89040 transistors aging workload data entries in an even separation of 44640 from 32-bit MAC and 44400 from 8-bit adder. The output of the ΔV_{th} is used as the regression target of the CNN, and waveforms (V_D, V_G, V_S, and V_B) are used as input to the CNN. A further two input features are computed. First is V_{gs} for the NMOS transistors and V_{sg} for PMOS transistors, the second input feature is a binary activity indicator showing if the transistor is on/off, with 1 indicating on and 0 indicating off across the time of the voltage waves. Giving the CNN a total of six input features and ΔV_{th} as a regression target for each transistor. Our CNN model is running via NVIDIA's Cuda package [15] and implemented by TensorFlow [18] training on an NVIDIA A100 graphics card. Our CNN illustrated in Fig. 7 consists of two 1D convolution layers made up of 1024 filters each, followed by six dense layers made up of 4096, 4096, 2048, 2048, 512, and 1 units, respectively. Our 32-bit MAC data is split into training and testing, 85 % (37940 transistors) and 15 % (6700 transistors), respectively, and the 8-bit adder is

TABLE I
CNN MODEL TRAINED ON 32-BIT MAC TRAINING DATASET AND TESTED ON 32-BIT MAC TRAINING AND 8-BIT ADDER DATASET

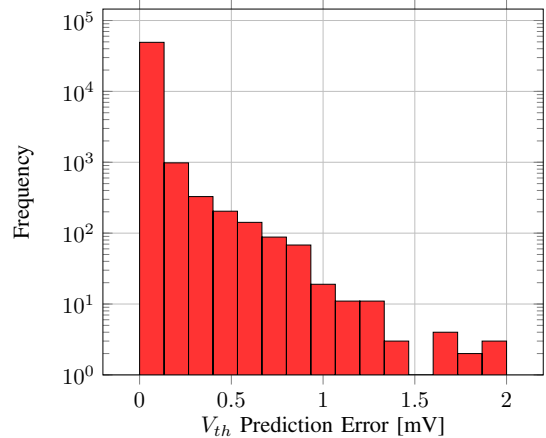| Testing Dataset	Dataset Size/Circuit Size [Transistors]	R^2	Accuracy[%]
32-bit MAC Test dataset	6700 / 2790	0.998	94.29 %
8-bit Adder	44400 / 222	0.997	92.91 %
Accuracy Difference			1.38 %

Fig. 8. Frequency histogram of the error in the ΔV_{th} prediction in mV by our CNN compared to the physics-bases aging model data with logarithmic frequency scale, testing across the total testing dataset of 51,100 transistors.

used entirely as a testing dataset to demonstrate the generality of our model when applied to a new never seen before circuit. The 8-bit adder is data is not used at any point during the training of the CNN and is just used for testing. The CNN is trained for 350 epochs and tested on a never seen before transistors from the 32-bit MAC circuit test dataset as well as the 8-bit adder.

Our CNN archives an accuracy of 94.29 % training and test on 32-bit MAC, and R^2 score of 0.998 and an accuracy of 92.91 % and a R^2 score of 0.997 while testing on 8-bit adder, which demonstrates the generality of our model, since it loses only drops 1.38 % when applied to a never seen before circuit as shown in Table I. Fig. 8 illustrates the frequency of the error value in the ΔV_{th} predicted by our CNN across both the 32-bit MAC test and the 8-bit adder datasets totaling 51100 transistors. Our CNN model shows a maximum error value of less than 2 mV.

V. CONCLUSION

In this special session paper, through four unique ML methods applied to TCAD electrical profile acceleration, TCAD thermal acceleration, transistor susceptibility model, and transistor workload dependent aging. We have illustrated the potential of ML to revolutionize the reliability and design process for both transistors and circuits at large. As ML can not only offer a huge speedup of up to 100,000 x over traditional TCAD simulation, but also conserve the privacy of the foundry as sharing sensitive and confidential models is no longer required to achieve accurate simulation performance.

REFERENCES

[1] A. Schaldenbrand, J. Xie, and H. Elhak, "Recent updates to transistor level reliability analysis," in *2019 IEEE International Reliability Physics Symposium (IRPS)*, 2019, pp. 1–8.

[2] E. Maricau and G. Gielen, "Transistor aging-induced degradation of analog circuits: Impact analysis and design guidelines," in *2011 Proceedings of the ESSCIRC (ESSCIRC)*, 2011, pp. 243–246.

[3] C. Prasad, S. Ramey, and L. Jiang, "Self-heating in advanced cmos technologies," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2017, pp. 6A–4.1–6A–4.7.

[4] V. Negro and L. Pannone, "Self-heating and gate leakage current in a guarded mosfet," *Proceedings of the IEEE*, vol. 60, no. 3, pp. 342–343, 1972.

[5] H. Amrouch, S. Mishra, V. van Santen, S. Mahapatra, and J. Henkel, "Impact of bti on dynamic and static power: From the physical to circuit level," in *2017 IEEE International Reliability Physics Symposium (IRPS)*, 2017, pp. CR–3.1–CR–3.6.

[6] Y. Zhao and Y. Qu, "Impact of self-heating effect on transistor characterization and reliability issues in sub-10 nm technology nodes," *IEEE Journal of the Electron Devices Society*, vol. 7, pp. 829–836, 2019.

[7] S. Kiamehr, F. Firouzi, and M. B. Tahoori, "Input and transistor reordering for nbti and hci reduction in complex cmos gates," in *Proceedings of the Great Lakes Symposium on VLSI*, ser. GLSVLSI '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 201–206. [Online]. Available: https://doi.org/10.1145/2206781.2206829

[8] A. Bhattacharjee and S. N. Pradhan, "Impact of transistor aging on the reliability of the analog circuit," in *2020 International Conference on Computational Performance Evaluation (ComPE)*, 2020, pp. 212–216.

[9] *Synopsys Sentaurus TCAD® v. U-2022.12-SP1 Available: https://www.synopsys.com/manufacturing/tcad.html*.

[10] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.

[11] T. Mohamed and H. Amrouch, "Super-tcad: Self-supervised graph attention networks for accelerated transistor simulations," Sep. 2024. [Online]. Available: http://dx.doi.org/10.36227/techrxiv.172684312.24638216/v1

[12] T. Mohamed, V. M. van Santen, L. Alrahis, O. Sinanoglu, and H. Amrouch, "Graph attention networks to identify the impact of transistor degradation on circuit reliability," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 71, no. 7, pp. 3269–3281, 2024.

[13] D. Kim and A. Oh, "How to find your friendly neighborhood: Graph attention design with self-supervision," 2022.

[14] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," 2019.

[15] NVIDIA, "Cuda, release: 10.2.89," 2020. [Online]. Available: https://developer.nvidia.com/cuda-toolkit

[16] Q. Liu, M. Vinet, J. Gimbert, N. Loubet, R. Wacquez, L. Grenouillet, Y. Le Tiec, A. Khakifirooz, T. Nagumo, K. Cheng, H. Kothari, D. Chanemougame, F. Chafik, S. Guillaumet, J. Kuss, F. Allibert, G. Tsutsui, J. Li, P. Morin, S. Mehta, R. Johnson, L. F. Edge, S. Ponoth, T. Levin, S. Kanakasabapathy, B. Haran, H. Bu, J.-L. Bataillon, O. Weber, O. Faynot, E. Josse, M. Haond, W. Kleemeier, M. Khare, T. Skotnicki, S. Luning, B. Doris, M. Celik, and R. Sampson, "High performance utbb fdsoi devices featuring 20nm gate length for 14nm node and beyond," in *2013 IEEE International Electron Devices Meeting*, 2013, pp. 9.2.1–9.2.4.

[17] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," vol. 14, 03 2001.

[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: http://tensorflow.org/

[19] M. Casse, B. C. Paz, F. Bergamaschi, G. Ghibaudo, F. Balestra, and M. Vinet, "(invited) cryogenic electronics for quantum computing ics: What can bring fdsoi," *ECS Transactions*, vol. 111, no. 1, p. 149, may 2023. [Online]. Available: https://dx.doi.org/10.1149/11101.0149ecst

[20] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[21] V. M. van Santen, H. Amrouch, and J. Henkel, "Modeling and mitigating time-dependent variability from the physical level to the circuit level," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 7, pp. 2671–2684, 2019.

[22] ——, "New worst-case timing for standard cells under aging effects," *IEEE Transactions on Device and Materials Reliability*, vol. 19, no. 1, pp. 149–158, 2019.

[23] L. T. Clark, V. Vashishtha, L. Shifren, A. Gujja, S. Sinha, B. Cline, C. Ramamurthy, and G. Yeric, "ASAP7: A 7-nm finFET predictive process design kit," *Microelectronics Journal*, 2016.

[24] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, p. 1929–1958, jan 2014.

[25] V. M. van Santen, S. Thomann, C. Pasupuleti, P. R. Genssler, N. Gangwar, U. Sharma, J. Henkel, S. Mahapatra, and H. Amrouch, "Bti and hcd degradation in a complete 32 × 64 bit sram array – including sense amplifiers and write drivers – under processor activity," in *2020 IEEE International Reliability Physics Symposium (IRPS)*, 2020.

[26] S. Mishra, H. Amrouch, J. Joe, C. K. Dabhi, K. Thakor, Y. S. Chauhan, J. Henkel, and S. Mahapatra, "A simulation study of nbti impact on 14-nm node finfet technology for logic applications: Device degradation to circuit-level interaction," *IEEE Transactions on Electron Devices*, vol. 66, no. 1, pp. 271–278, 2019.

[27] N. Parihar, N. Goel, S. Mukhopadhyay, and S. Mahapatra, "Bti analysis tool—modeling of nbti dc, ac stress and recovery time kinetics, nitrogen impact, and eol estimation," *IEEE Transactions on Electron Devices*, vol. 65, no. 2, pp. 392–403, 2018.

[28] U. Sharma and S. Mahapatra, "A spice compatible compact model for hot-carrier degradation in mosfets under different experimental conditions," *IEEE Transactions on Electron Devices*, vol. 66, no. 2, pp. 839–846, 2019.