

Automatic Prediction of Metal–Oxide–Semiconductor Field-Effect Transistor Threshold Voltage Using Machine Learning Algorithm

Seoyeon Choi, Dong Geun Park, Min Jung Kim, Seain Bang, Jungchun Kim, Seunghee Jin, Ki Seok Huh, Donghyun Kim, Jerome Mitard, Cheol E. Han,* and Jae Woo Lee*

A fast and precise threshold voltage (V_{th}) extraction method is required for the process design of electronic systems using metal–oxide–semiconductor field-effect transistors (MOSFETs) and its immediate on-site analysis during fabrication. The selection of a suitable V_{th} extraction method is a complicated task because it involves a trade-off between accuracy and simplicity according to the device scheme. Herein, an automatic-prediction method of the MOSFET V_{th} using machine learning (ML) is proposed. The ML model is trained with V_{th} , extracted using different methods (2nd derivative, constant current, and Y-function) and from various kinds of FETs (finFET, 2D FET, and metal–oxide thin-film transistors). The concept of threshold ratio (R_{th}) for universal V_{th} prediction, which considers the normalized V_{th} within certain V_G ranges, is suggested. The precision and accuracy of ML models are statistically verified by calculating the root mean square error (RMSE), mean absolute error, and mean coefficients of determination (R^2) values. The universal ML model (k -nearest neighbor (kNN)) achieves 1.35% of RMSE and 0.98 of R^2 for the best score. The ML model eliminates the ambiguity in V_{th} extraction and provides objective V_{th} prediction for most FET schemes used in the semiconductor industry and research field.

1. Introduction

Metal–oxide–semiconductor field-effect transistors (MOSFETs) are crucial components of electronic systems such as logic, memory, sensor, and neuromorphic devices.^[1–3] Therefore, to improve the design convenience and reliability of the application, quick and accurate estimation of MOSFET device parameters is required.

To obtain the primary MOSFET device parameters such as threshold voltage (V_{th}), subthreshold swing, and mobility, the drain current versus gate voltage (I_D – V_G) characteristics are utilized. For MOSFET modeling, V_{th} is a key factor, determining the design margin of the system as an on-off standard, and it is used as a reference factor while calculating other electrical parameters.^[4–6]

To extract the V_{th} from I_D – V_G characteristics, the “constant current” (CC) method was generally applied in industry^[1,2,7] because it simply designates the corresponding V_G with a prespecified current

level, as shown in **Figure 1a**. However, the accuracy of V_{th} extraction can be affected by noise, humps, or unwanted current degradation, caused from various short-channel effects in the subthreshold region (off-current, I_{OFF}).^[7–9]

Numerous alternatives to the CC method were proposed to extract a more accurate V_{th} . For example, the “extrapolation in linear region” method^[1] is often used instead of the CC method. The method is not affected by I_{OFF} because it uses I_D turned on enough (called on-current). However, it is significantly affected by parasitic-series resistance and the mobility degradation of the carrier at the short channel. To avoid the parasitic-series resistance, the “transconductance linear extrapolation”^[10,11] and “2nd derivative” methods^[12] were suggested. Additionally, the “Y-function” method^[13–16] was investigated to solve mobility degradation and series resistance issues. Also, integration-based method has been used to extract V_{th} without influence of series resistance and noise.^[17]

Compared to the CC method, the alternatives use additional numerical analysis such as differentiation and linear

S. Choi, D. G. Park, M. J. Kim, S. Bang, J. Kim, S. Jin, K. S. Huh, D. Kim, C. E. Han, J. W. Lee


Department of Electronics and Information Engineering
Korea University

2511 Sejong-ro, 30019 Sejong, Republic of Korea
E-mail: cheolhan@korea.ac.kr; orion627@korea.ac.kr

J. Mitard

imec

Kapeldreef 75, 3001 Leuven, Belgium

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202200302>.

© 2022 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202200302

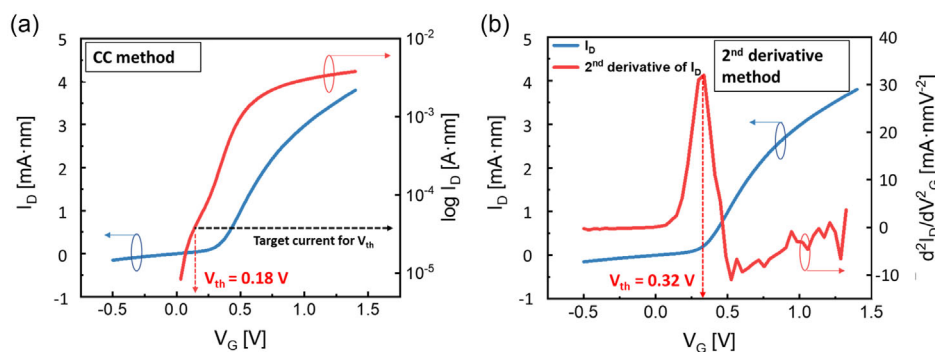


Figure 1. Comparison of V_{th} extraction between a) CC and b) 2nd derivative method with the same I_D - V_G curve. Depending on the extraction method, the result of V_{th} is different.

extrapolation. The complexity of the extraction process is a disadvantage when prompt and intuitive interpretations of device analysis are required. To solve the trade-off between simplicity and accuracy of V_{th} analysis, different extraction algorithms were selectively applied and merged with the CC method based on the faced case.^[7] Due to the significance of the resultant V_{th} in determining the device performance and reliability,^[18] many researchers have attempted to obtain an optimal V_{th} extraction algorithm, and it has not been solved thus far.

In semiconductor engineering, machine learning (ML) has gradually replaced the role of humans.^[19] For example, ML was used to detect and repair defects in the 7 nm mask design.^[20] Currently, automatic circuit design with ML is being developed. Additionally, ML algorithms were used to operate devices efficiently in FET-based sensor applications and optimize the fabrication processes.^[21–26] However, V_{th} estimation by ML has not been investigated thus far.

In this study, we propose a novel method by supervised ML to determine the reliable V_{th} . This is a pilot attempt at adopting the ML for extracting the device parameters of MOSFETs.^[26–29] Regardless of selecting extraction methods and device schemes, the automatic V_{th} prediction method was designed to recognize

the shape of I_D - V_G characteristics. **Figure 2** shows the process of extracting V_{th} using ML.

To train the ML model for the prediction of V_{th} , three different device schemes, fin field-effect transistors (finFETs), metal-oxide thin-film transistors (TFTs), and two-dimensional (2D) FETs, and three different V_{th} extraction methods (2nd derivative, Y-function, and CC method) were used. A threshold ratio (R_{th}) was suggested to make a universal ML model. To achieve precise prediction results, a preprocessing process of input data was investigated. Finally, the R_{th} of various device schemes was precisely predicted by the ML model.

2. Result and Discussion

2.1. Optimized Preprocessing and ML Algorithm using FinFET

Four different preprocessing cases were tested with finFET data to obtain the optimized ML algorithm: 1) decision tree regression with only normalized V_G , 2) k -nearest neighbor (kNN) regression with only normalized V_G , 3) decision tree regression with the normalized dataset for both V_G and I_D , and 4) kNN regression

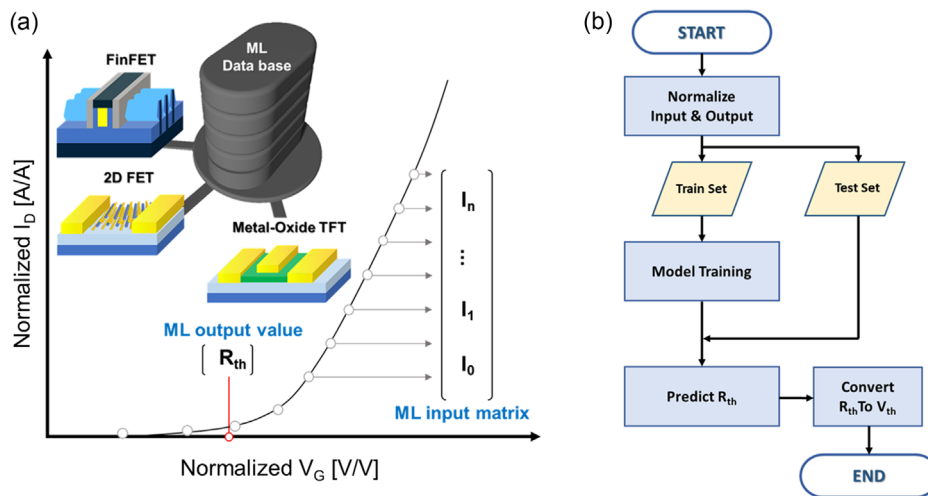


Figure 2. a) Schematic of preprocessing for the ML model. Normalized I_D - V_G characteristics of finFET, 2D FET, metal-oxide TFT were used. R_{th} is the output of sample and normalized drain current is input feature matrix. b) Flowchart of the predicting V_{th} process using ML.

with the normalized dataset for both V_G and I_D . To examine the reliability of prediction, tests were repeatedly conducted 50 times with randomly selected 30 finFET devices that are not included in training set. The best case of the four different preprocessing methods was observed by comparing the mean value and standard deviation (STD) for each evaluation metric.

Figure 3 shows evaluation metrics for all cases of preprocessing methods and ML algorithms. Both root mean squared error (RMSE) and mean absolute error (MAE) indicated an error between the output (extracted by 2nd derivative method) and predicted R_{th} values, indicating overall prediction accuracy: the smaller, the better. In the RMSE point of view, kNN algorithm with both V_G and I_D normalization showed the lowest mean value of error (Figure 3j, 1.14×10^{-2} , ≈ 17 mV for finFET). The largest mean value of RMSE was shown in decision tree algorithm with only V_G normalization, and it was only 1.47×10^{-2} (Figure 3a).

range, RMSE was $< 1.47\%$, which was ≈ 22 mV for the prediction of V_{th} in finFET and still small. Mean coefficients of determination (R^2) values indicate the accuracy of a model in predicting data. The value was large enough (from 0.95 to 0.97, Figure 3c,f,i,l) for all cases of preprocessing methods and ML algorithms.

The STD of the metrics captured the precision of the prediction: the smaller STD value has the more stable prediction. The kNN algorithm with V_G and I_D normalization showed the most reliable RMSE and MAE (Figure 3j,k). Any algorithms with normalized V_G and I_D (Figure 3g–l) showed smaller STD than those with only V_G normalization (Figure 3a–f). Although mean values of errors and R^2 were not significantly different, the kNN algorithm with both V_G and I_D normalization was the most stable and reasonably accurate model for the R_{th} prediction of finFET.

Figure 4 shows the best results among 50 test sets using the kNN algorithm with V_G and I_D normalization. In Figure 4a,

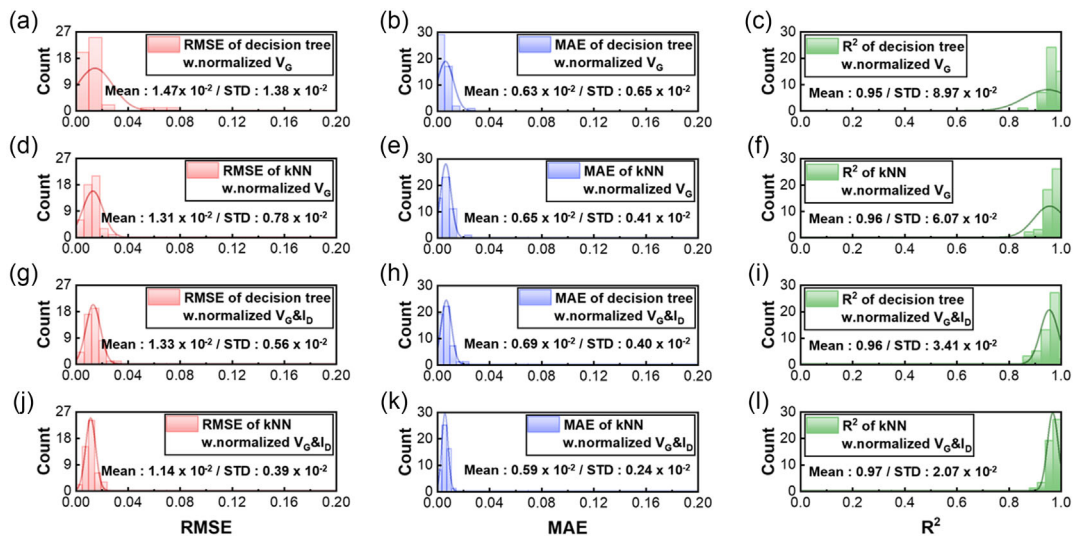


Figure 3. Histograms of R_{th} prediction about finFET with four different preprocessing cases. Each vertical column shows a,d,g,j) RMSE, b,e,h,k) MAE, and c,f,i,l) R^2 of preprocessing cases, respectively. a–c) is decision tree with normalized V_G . d–f) is kNN with normalized V_G . g–i) is decision tree with normalized V_G and I_D . j–l) is kNN with normalized V_G and I_D .

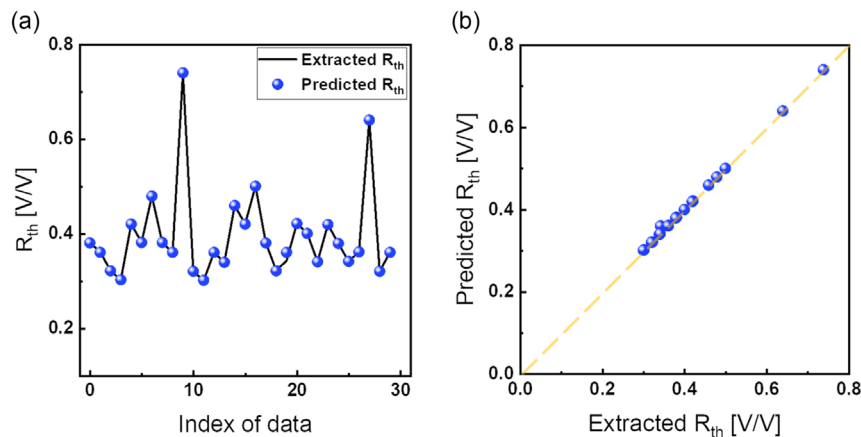


Figure 4. a) Comparison between extracted R_{th} (by 2nd derivative method, black line) and predicted R_{th} using kNN with V_G and I_D normalized finFET dataset (blue circle). b) Correlation between extracted R_{th} and predicted R_{th} . It shows 0.99 of R^2 .

predicted R_{th} by ML (circle) and extracted R_{th} by 2nd derivative method (line) are compared for each test. In this test set, R_{th} of finFET varied from 0.30 to 0.74 (from 0.07 to 0.91 V for V_{th} variability). Besides the R_{th} variability in finFET, the ML model using the kNN algorithm with V_G and I_D normalization accurately predicted R_{th} , and RMSE and MAE were 0.33×10^{-2} and 0.08×10^{-2} . In Figure 4b, the correlation plot is shown with the same data as Figure 4a. In this case, R^2 was 0.99.

2.2. Universal ML Model for R_{th} Prediction

The selected ML model using the kNN algorithm with V_G and I_D normalization was extended to the other devices. The universal model was trained with randomly selected 1290 I_D - V_G curves of finFETs, metal-oxide TFTs, and 2D FETs. One test set included R_{th} of 30 devices (ten for each type of device, not used for training). The prediction results of 30 test sets are shown in Figure 5a. Mean values of RMSE, MAE, and R^2 were 3.46×10^{-2} (in mV, ≈ 69 mV for finFET case), 2.01×10^{-2} , and 0.93, respectively, and they were larger than the prediction result in the previous section that was applied to only finFET (Figure 3j,k), however, still comparable.

In Figure 5b,c, the superior result of predicted R_{th} was compared with R_{th} estimated by the conventional V_{th} extraction method. In this case, RMSE and MAE were 1.35×10^{-2} and 0.84×10^{-2} , respectively. As shown in Table 1, the RMSE and MAE of 2D FETs and metal-oxide TFTs are larger than those of finFETs, and the largest error is shown in the 2D FET case. Similar to RMSE and MAE, R^2 of the 2D FET showed the worst value (0.97). In the universal ML model, the numbers of devices used for the training set were different (finFETs: metal-oxide TFTs: 2D FETs = 1217:47:26), and three different V_{th} extraction methods were used, thereby making the ML model biased to the finFET and 2nd derivative method. With enough number of metal-oxide TFT and 2D FET, the prediction result of the universal ML model could be improved. Also, the more device type could make the better performance of ML model. More details are discussed in Supporting Information (Figure S3, Supporting Information).

Table 1. RMSE, MAE, and R^2 of the superior universal ML model for 30 test sets. The evaluation metrics for each device are investigated.

	Index in Figure 5b	RMSE	MAE	R^2
Overall for best set	0–29	1.35×10^{-2}	0.84×10^{-2}	0.98
FinFET	0–9	0.10×10^{-2}	0.04×10^{-2}	0.99
Metal-oxide TFT	10–19	1.38×10^{-2}	1.09×10^{-2}	0.98
2D FET	20–29	1.88×10^{-2}	1.40×10^{-2}	0.97

3. Conclusion

In this study, the very first ML approach for V_{th} extraction was proposed. The I_D - V_G characteristics and extracted V_{th} with CC, 2nd derivate, and the Y-function method were used as input and output of ML model to predict V_{th} . To select the proper ML algorithm, four types of preprocessing methods were tested with the kNN and decision tree regression about finFETs. The kNN regression with normalized for both V_G and I_D was selected for the best scores of prediction performance and the lowest STD. For a mixed dataset with finFET, metal-oxide TFT, and 2D FET, the universal ML model achieved 1.35% of RMSE, 0.84% of MAE, and 0.98 of R^2 for the best score. The approach was unconstrained by the structure of the FET device, channel material/polarity, gate-bias range, etc., exhibited reliable results even in repetitive learning and prediction using each different test set. Therefore, extracting V_{th} with the ML approach can reduce time and effort. In addition, the objective V_{th} can be obtained beside researcher's personal consideration of data characteristics. Although experiments indicate high performance for extracting V_{th} , much higher performance is expected when different kinds of devices' data are used for the training dataset.

4. Experimental Section

Preprocessing of Data (Data Description): In this study, I_D of a transistor and its V_{th} were used as an input and output of the ML model, respectively. The V_{th} was extracted from each I_D - V_G characteristic dataset. For the universal ML model, various V_{th} extraction methods and different FET

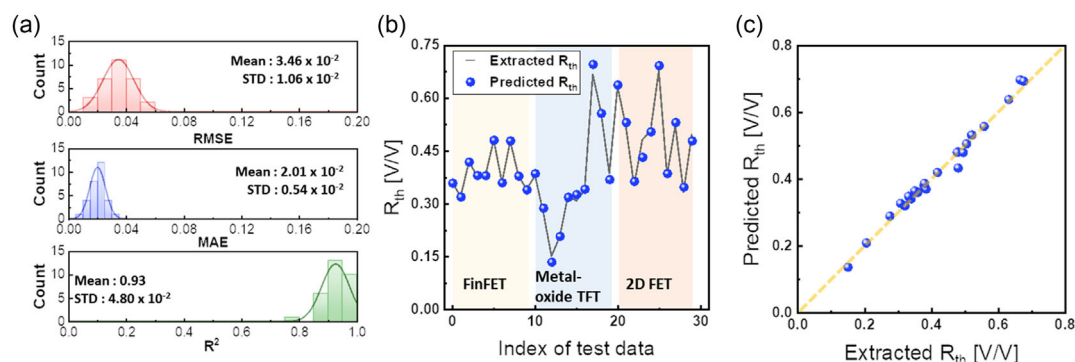


Figure 5. a) Histograms of R_{th} prediction using selected preprocessing (V_G and I_D normalization) and the kNN ML model. Ten finFETs, ten metal-oxide TFTs, and ten 2D FETs were randomly tested 30 times. b) The best result of a) and c) shows the correlation plot of b). For each device scheme, different V_{th} extraction methods were used; however, ML predicted accurate R_{th} regardless of selecting extraction methods and device schemes.

samples were used. For different FET samples, channel materials, device structures, and operating-bias range were the variables.

For finFET, 1227 I_D - V_G data were provided by imec (Belgium), finFET which had a three-dimensional fin-shaped Si channel, operated at $0 \approx \pm 2$ V gate-bias range since it was used in logic devices for low-power applications. The 2nd derivative method was used to extract V_{th} of finFET. The metal-oxide TFTs and 2D FET are operated in a larger V_G bias range than finFET. Metal-oxide TFTs with an indium-gallium-zinc-oxide (IGZO) channel were used for display application and operated in the -40 to 60 V gate-bias range.^[30–34] 57 IGZO TFT devices were used, and the Y-function method was applied for extracting V_{th} . For device operation, 36 2D (WS_2) FET devices, with -20 to 60 V gate-bias range were used. V_{th} of 2D FET was extracted by the CC method.^[35–37]

The ML input data, composed of only an I_D dataset, do not include the information of the corresponding V_G range. Therefore, real V_{th} as an output of the ML model causes the overfitting when ML model learns the input data. Instead of a real V_{th} value, the proportion of V_{th} (thresholding ratio, R_{th}) in the normalized V_G range (0–1) was used for the output of the ML model, so that the ML model is universally implemented without considering V_G range of device (Figure S1, Supporting Information).

The scale of I_D is different based on the device type. For example, a few 2D FETs have several nA levels of I_D , whereas I_D of some metal-oxide TFTs has tens of μ A level. Thus, the dataset was preprocessed in two ways for examining the impact of I_D normalization on extraction performance: 1) normalizing only V_G and 2) normalizing both V_G and I_D (0–1). The final V_{th} is obtained by converting the predicted R_{th} into the real value of V_{th} with the actual operating bias range of each device. The converting formula is

$$V_{th} = R_{th} \times (V_{Gmax} - V_{Gmin}) + V_{Gmin} \quad (1)$$

where V_{Gmax} and V_{Gmin} are the maximum and minimum value of V_G range, respectively.

Machine Learning Algorithm: The kNN and decision tree ML regression algorithms were used to predict V_{th} in this study. There are two types of tasks in supervised learning: classification and regression. Classification is applicable when the output value is categorical, consisting of a fixed number of classes. In contrast, the regression algorithm is used when the output value is continuous. The output of ML is a continuous quantity of R_{th} . Thus, the regression is appropriate; predicting a continuous value directly would be more accurate than classifying it into a few discrete values (classes). More details are provided in Figure S2, Supporting Information.

The value of “ k ” in kNN and “max depth” in the decision tree were designated as hyper parameters, which improved the performance of the model by appropriate values. “ k ” is the number of training samples to calculate the output value of the target test sample. “max depth” is the maximum distance between the top and terminal nodes. The best number of “ k ” (nearest neighbor samples of kNN) and “max depth” (the longest tree path’s nodes of the decision tree), inducing the highest regression accuracy, were investigated by comparing in the range 1–30. Three evaluation metrics, RMSE, MAE, and R^2 , were used for this experiment.

Python was used to conduct the experiment. Training and evaluation processes in ML algorithms were implemented by “scikit-learn” package, which provided various functions for ML.^[38]

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

S.C. and D.G.P. contributed equally to this work. This research was supported by the National Research Foundation (NRF-2022R1A2C1010447, 2021R1F1A1063342), European Nanoelectronics Access (ASCENT+),

WISET (Korea Foundation for Women In Science, Engineering and Technology), and Korea University Grant.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Research data are not shared.

Keywords

decision tree, k -nearest neighbors, machine learning, MOSFET, threshold-voltage extraction

Received: September 7, 2022

Revised: November 20, 2022

Published online: December 20, 2022

- [1] D. Schroeder, in *Semiconductor Material and Device Characterization*, John Wiley, New York, NY.
- [2] A. Ortiz-Conde, F. G. Sánchez, J. J. Liou, *Acta Cient. Venez.* **2000**, *51*, 176.
- [3] Y. Hwang, J. H. Jeon, J. Lee, J. Yoon, F. S. Kim, H. Kim, *J. Nanosci. Nanotechnol.* **2021**, *21*, 4303.
- [4] A. Ortiz-Conde, J. Rodriguez, F. G. Sánchez, J. Liou, *Solid-State Electron.* **1998**, *42*, 1743.
- [5] J. Benson, N. V. D’Halleweyn, W. Redman-White, C. A. Easson, M. J. Uren, O. Faynot, J.-L. Pelloie, *IEEE Trans. Electron Devices* **2001**, *48*, 1019.
- [6] X. Zhou, K. Lim, W. Qian, *Solid-State Electron.* **2001**, *45*, 507.
- [7] A. Ortiz-Conde, F. J. G. Sanchez, J. J. Liou, A. Cerdeira, M. Estrada, Y. Yue, *Microelectron. Reliab.* **2002**, *42*, 583.
- [8] Y. Taur, D. Zicherman, D. Lombardi, P. J. Restle, C. Hsu, H. Nanafi, M. R. Wordeman, B. Davari, G. G. Shahidi, *IEEE Electron Device Lett.* **1992**, *13*, 267.
- [9] B. Cretu, T. Boutchacha, G. Ghibaudo, F. Balestra, *Electron. Lett.* **2001**, *37*, 717.
- [10] M. Tsuno, M. Suga, M. Tanaka, K. Shibahara, M. Miura-Mattausch, M. Hirose, in *Proc. of 1998 Asia and South Pacific Design Automation Conf.*, IEEE, Piscataway, NJ **1998**, p. 111–116.
- [11] M. Tsuno, M. Suga, M. Tanaka, K. Shibahara, M. Miura-Mattausch, M. Hirose, *IEEE Trans. Electron Devices* **1999**, *46*, 1429.
- [12] H.-S. Wong, M. H. White, T. J. Krutsick, R. V. Booth, *Solid-State Electron.* **1987**, *30*, 953.
- [13] C. Diouf, A. Cros, S. Monfray, J. Mitard, J. Rosa, D. Gloria, G. Ghibaudo, *Solid-State Electron.* **2013**, *85*, 12.
- [14] G. Ghibaudo, *Electron. Lett.* **1988**, *24*, 543.
- [15] D. Fleury, A. Cros, H. Brut, G. Ghibaudo, in *IEEE Int. Conf. on Microelectronic Test Structures*, IEEE, Piscataway, NJ **2008**, p. 160–165.
- [16] N. Subramanian, G. Ghibaudo, M. Mouis, in *Proc. of the European Solid State Device Research Conference*, IEEE, Piscataway, NJ **2010**, p. 309–312.
- [17] F. G. Sánchez, A. Ortiz-Conde, G. De Mercato, J. Salcedo, J. Liou, Y. Yue, *Solid-State Electron.* **2000**, *44*, 673.
- [18] G. Espinera, D. Nagy, A. Garcia-Loureiro, N. Seoane, G. Indalecio, *Solid-State Electron.* **2019**, *159*, 165.

- [19] I. H. Sarker, *SN Comput. Sci.* **2021**, 2, 1.
- [20] Y. Ma, F. Wang, Q. Xie, L. Hong, J. Mellmann, Y. Sun, S. W. Gao, S. Singh, P. Venkatachalam, J. Word, in *Design-Process-Technology Co-Optimization for Manufacturability XIII*, SPIE, Bellingham, WA **2019**, p. 38.
- [21] V. Shirmohammadli, B. Bahreyni, *IEEE Sens. J.* **2022**, 22, 10183.
- [22] J.-H. Ahn, *J. Sens. Sci. Technol.* **2021**, 30, 1.
- [23] Z. Ballard, C. Brown, A. M. Madni, A. Ozcan, *Nat. Mach. Intell.* **2021**, 3, 556.
- [24] S. Namuduri, B. N. Narayanan, V. S. P. Davuluru, L. Burton, S. Bhansali, *J. Electrochem. Soc.* **2020**, 167, 037552.
- [25] J. Yang, S. Liu, Y. Meng, W. Xu, S. Liu, L. Jia, G. Chen, Y. Qin, M. Han, X. Li, *ACS Appl. Mater. Interfaces* **2022**, 14, 25629.
- [26] T.-L. Wu, S. B. Kutub, *IEEE Trans. Electron Devices* **2020**, 67, 5448.
- [27] H. Carrillo-Núñez, N. Dimitrova, A. Asenov, V. Georgiev, *IEEE Electron Device Lett.* **2019**, 40, 1366.
- [28] K. Ko, J. K. Lee, M. Kang, J. Jeon, H. Shin, *IEEE Trans. Electron Devices* **2019**, 66, 4474.
- [29] D. Glowienka, Y. Galagan, *Adv. Mater.* **2022**, 34, 2105920.
- [30] Y. G. Park, D. Y. Cho, R. Kim, K. H. Kim, J. W. Lee, D. H. Lee, S. I. Jeong, N. R. Ahn, W. G. Lee, J. B. Choi, M. J. Kim, D. Kim, S. Jin, D. G. Park, J. Kim, S. Choi, S. Bang, J. W. Lee, *Adv. Electron. Mater.* **2022**, 2101273.
- [31] J. M. Chung, X. K. Zhang, F. Shang, J. H. Kim, X. L. Wang, S. Liu, B. G. Yang, Y. Xiang, *Nanoscale Res. Lett.* **2018**, 13, 164.
- [32] J. H. Kim, E. K. Park, M. S. Kim, H. J. Cho, D. H. Lee, J. H. Kim, Y. Khang, K. Park, Y. S. Kim, *Thin Solid Films* **2018**, 645, 154.
- [33] S. Hong, S. Lee, M. Mativenga, J. Jang, *IEEE Electron Device Lett.* **2014**, 35, 93.
- [34] D. H. Cho, S. H. Yang, J. H. Shin, C. W. Byun, M. K. Ryu, J. I. Lee, C. S. Hwang, H. Y. Chu, *J. Korean Phys. Soc.* **2009**, 54, 531.
- [35] H. S. Wang, Q. Y. Liu, X. M. Feng, Z. Zhang, K. Wang, Z. J. Liu, J. F. Dai, *Mater. Res. Express* **2020**, 7, 076302.
- [36] A. R. Wang, Y. H. Wang, J. F. Li, N. Xu, S. L. Li, X. R. Wang, Y. Shi, F. Q. Wang, *Appl. Phys. Lett.* **2021**, 118, 121104.
- [37] L. M. Yang, K. Majumdar, H. Liu, Y. C. Du, H. Wu, M. Hatzistergos, P. Y. Hung, R. Tieckelmann, W. Tsai, C. Hobbs, P. D. Ye, *Nano Lett.* **2014**, 14, 6275.
- [38] J. G. Hao, T. K. Ho, *J. Educ. Behav. Stat.* **2019**, 44, 348.