

RESEARCH ARTICLE | FEBRUARY 29 2024

Machine learning investigation of high-k metal gate processes for dynamic random access memory peripheral transistor

Namyong Kwon; JoonHo Bang; Won Ju Sung; Jung Hoon Han; Dongin Lee; Ilwoo Jung; Se Guen Park; Hyodong Ban; Sangjoon Hwang; Won Yong Shin; Jinhye Bae ; Dongwoo Lee  



APL Mater. 12, 021131 (2024)
<https://doi.org/10.1063/5.0191100>



Articles You May Be Interested In

Impact of both metal composition and oxygen/nitrogen profiles on p-channel metal-oxide semiconductor transistor threshold voltage for gate last high-k metal gate

J. Vac. Sci. Technol. B (January 2013)

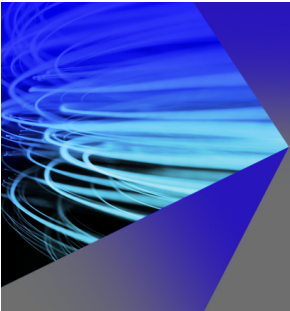
Intermixing reduction in ultra-thin titanium nitride/hafnium oxide film stacks grown on oxygen-inserted silicon and associated reduction of the interface charge dipole

J. Appl. Phys. (November 2021)

Modeling the effects of lanthanum, nitrogen, and fluorine treatments of Si-SiON-HfO₂-TiN gate stacks in 28 nm high-k-metal gate technology

J. Appl. Phys. (June 2017)

29 January 2026 05:43:44



AIP Advances

Why Publish With Us?



21DAYS
average time
to 1st decision



OVER 4 MILLION
views in the last year



INCLUSIVE
scope

[Learn More](#)

 AIP
Publishing

Machine learning investigation of high-k metal gate processes for dynamic random access memory peripheral transistor

Cite as: APL Mater. 12, 021131 (2024); doi: 10.1063/5.0191100

Submitted: 11 December 2023 • Accepted: 19 January 2024 •

Published Online: 29 February 2024



Namyong Kwon,¹ JoonHo Bang,² Won Ju Sung,¹ Jung Hoon Han,¹ Dongin Lee,¹ Ilwoo Jung,¹ Se Guen Park,¹ Hyodong Ban,¹ Sangjoon Hwang,¹ Won Yong Shin,³ Jinhye Bae,⁴  and Dongwoo Lee^{2,a)} 

AFFILIATIONS

¹ Department of Memory Business, Samsung Electronics, Samsungjeonja-ro, Hwasung-si 18448, Republic of Korea

² School of Mechanical Engineering, Sungkyunkwan University, Seobu-ro, Suwon 16419, Republic of Korea

³ School of Mathematics and Computing, Yonsei University, Seoul 03722, Republic of Korea

⁴ Department of NanoEngineering, University of California San Diego, La Jolla, California 92093, USA

^{a)} Author to whom correspondence should be addressed: dongwoolee@skku.edu

ABSTRACT

Dynamic random access memory (DRAM) plays a crucial role as a memory device in modern computing, and the high-k/metal gate (HKMG) process is essential for enhancing DRAM's power efficiency and performance. However, integration of the HKMG process into the existing DRAM technology presents complex and time-consuming challenges. This research uses machine learning analysis to investigate the relationships among the process parameters and electrical properties of HKMG in DRAM. The expectation-maximization imputation was utilized to fill in the missing data, and the Shapley additive explanations analysis was employed for the regression models to predict the electrical properties of HKMG. The impact of the process parameters on the electrical properties is quantified, and the important features that affect the performance of the HKMG transistor are characterized by using the explainable AI algorithm.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0191100>

I. INTRODUCTION

The demand for high-speed data communication has experienced an explosive increase in recent years, necessitating a corresponding increase in the bandwidth of the dynamic random access memory (DRAM). To operate at a higher speed and with a minimized leakage current, it is essential for peripheral transistors in DRAM to reduce the off-current (I_{off}). However, conventional polysilicon oxynitride (pSiON) transistors have encountered limitations in controlling the drain-induced barrier lowering (DIBL) as the chip sizes shrink because of the short-channel effect. This issue leads to increased power consumption, posing a significant challenge in the DRAM technology. In response to these challenges, researchers have explored the use of high-k/metal gate (HKMG) transistors for DRAM peripheral transistors.^{1–3} In comparison with the pSiON technology, HKMG offers several advantages, including improved speed and better control of the leakage

current. Furthermore, HKMG presents potential solutions to the increased power consumption problem caused by DIBL in pSiON transistors, thus enhancing the power efficiency and performance of DRAM. We compare the structures [Fig. 1(a)] and the performances [Figs. 1(b)–1(d)] of the 10 nm third generation DRAM products of Samsung Electronics that utilize the HKMG transistors and the conventional pSiON transistors. The HKMG transistors were found to exhibit a 38% reduction in I_{off} [Fig. 1(b)], a 60% increase in the on-current [I_{on} , Fig. 1(c)], and a 26% decrease in the propagation delay time (tPD) as compared to pSiON transistors. These results suggest that HKMG transistors maintain a lower I_{off} while achieving a higher I_{on} , indicating their potential to reduce power consumption in high-speed DRAM applications.

Implementation of HKMG technology in DRAM comes with its own set of challenges. Reliability concerns, such as the negative bias thermal instability (NBTI), interface trap problems related to hafnium oxide (HfO), and V_{th} variations due to Fermi level pin-

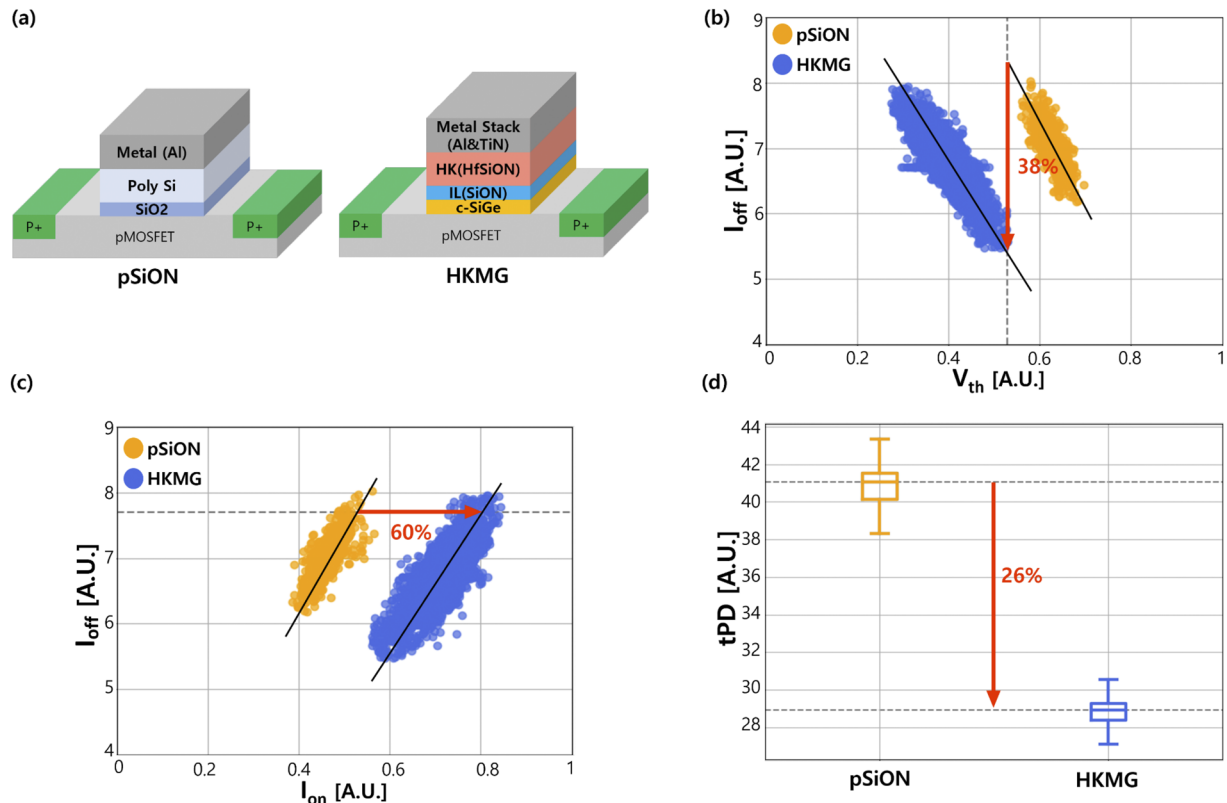


FIG. 1. (a) Schematic diagrams of the pSiON (left) and HKMG (right) p-type transistors. Comparison of the (b) I_{off} - V_{th} , (c) I_{off} - I_{on} characteristics, and (d) tPD values for the HKMG and pSiON transistors.

ning, significantly impact the performances of the DRAM devices, hindering the achievement of optimal power efficiency.⁴ Although previous studies reported the effects of nitrogen (N) concentration and channel silicon germanium (c-SiGe) layer on the performances of the HKMG transistors,⁵⁻⁷ a deeper understanding of the relationships among the process parameters and the electrical properties is required to further optimize the performances. We propose an implementation of machine learning (ML) analysis to investigate the multi-dimensional and complex correlations among the process parameters and the electrical properties.⁸⁻¹¹ By constructing ML regression models and utilizing the explainable artificial intelligence (XAI) algorithms, the degree of impact of the process parameters on the electrical properties is quantified and the main parameters affecting the performance of the HKMG transistor are identified.¹² We also discuss the use of the imputation ML algorithms to fill in the missing data, a very common issue in the semiconductor industry to conserve time and resources.^{13,14}

II. EXPERIMENTAL

A. Device Preparation

The HKMG transistors in this study were prepared using a gate-first process to avoid the Fermi-level pinning caused by the

high heat of DRAM fabrication procedures [Fig. 1(a)].^{3,4,15} First, a c-SiGe layer was epitaxially grown on the silicon substrate to enhance the hole mobility.² This process reduces both threshold voltage (V_{th}) and off-current (I_{off}), which improves the performance of p-type metal oxide semiconductor (PMOS) devices and makes them suitable for low-power DRAM applications. The formation of the interlayer (IL), which includes silicon oxynitride (SiON), was achieved by using oxidation and radio frequency nitridation (RFN). IL prevents reaction between the Hf-based high-k (HK) layer and the Si substrate. By adjusting the N concentration within IL, Si-N bonds are controlled, reducing the interface traps that easily capture the hole carriers.⁵⁻⁷ This indicates that RFN instigated radical-induced re-oxidation, leading to an increase in capacitance-equivalent thickness (CET).^{2,3,15,16} Subsequently, a layer of hafnium silicate (HfSiON) was deposited on top of IL as a HK gate dielectric. HfSiON is known to be stable compared to HfO_2 in a high thermal budget.^{2,17} Next, a de-coupled nitridation (DPN) process was used to form the HK dielectric layer. Since N atoms play a role in filling the oxygen vacancies of HK dielectric layers as well as in controlling the dielectric constant, fine-tuning the DPN process is essential to obtain suitable gate dielectrics.^{18,19} Aluminum (Al)-titanium nitride (TiN) stacks were then subsequently deposited on top of the HfSiON layer to control the work function. The application of a metal gate with high-k becomes crucial to avoid the V_{th} control

TABLE I. Process parameters used as input and electrical properties as output in the machine learning analysis and their missing rates.

Process parameters	Missing data rate (%)
Interlayer nitrogen concentration (IL N%)	57
Interlayer thickness (IL Thk)	63
Nitrogen concentration of the HfSiON layer (HK N%)	69
Channel length (L_{ch})	41
Gate spacer offset width (L_{offset})	59
Electrical properties	
Propagation delay time (tPD)	0
Threshold voltage (V_{th})	3
On-current (I_{on})	2
Off-current (I_{off})	8
Overlap capacitance (Cov)	2
Gate oxide thickness (Tox)	1
Threshold voltage shift by negative-bias temperature instability (NBTI)	8

challenges and interface disparities present when using conventional polysilicon.²⁰ Upon completing the fabrication of the HKMG stacks, the devices underwent the standard DRAM fabrication processes.

B. Processing parameters and electrical properties of HKMG

For machine learning analysis in this study, we consider the five critical process parameters as input features, which are the nitrogen concentration of the interlayer (*IL N%*), thickness of the interlayer (*IL Thk*), nitrogen concentration of the HfSiON layer (*HK N%*), channel length (L_{ch}), and the gate spacer offset width (L_{offset}). The seven electrical properties of the HKMG module considered for machine learning analysis are the propagation delay time (*tPD*), threshold voltage (V_{th}), on-current (I_{on}), off-current (I_{off}), overlap capacitance (*Cov*), gate oxide thickness (*Tox*), and threshold voltage shift by negative-bias temperature instability (*NBTI*). The processing parameters of ion implantation conditions (channel, source, and drain doping), annealing, and device profile were kept consistent.

Typically, not all the process parameters and electrical properties are monitored and measured for every wafer to conserve time and resources in the semiconductor industry. Consequently, missing data in the measurement dataset is inevitable. Of the available 17 241 wafers, the measurements of both process parameters and electrical properties were carried out only for 353 wafers. Additionally, the data of the seven electrical properties showed an average missing rate of 4%, while the five process parameters exhibited a high average missing rate of 59%(Table I), each of which is missing completely at random (MCAR). The MCAR characteristic of the missing values in DRAM can be attributed to random variations in the manufacturing processes and measurement conditions. The imputation methods using machine learning have been utilized to fill in the missing data.

C. Machine learning

Machine learning analysis has been carried out as schematically shown in Fig. 2. First, the five process parameters and the seven electrical properties of the wafers containing HKMG were collected and tabulated. The missing data rate for each parameter was identified (Table I). Second, the imputation machine learning algorithms were employed to fill in the missing data. Since the process parameters are expected to have local correlations with each other, we applied imputation algorithms by incorporating all the features of electrical properties and process parameters. Prior to the imputation process, we randomly removed 1%, 5%, 10%, and 30% of the known values for each feature item from the original data. Then, nine different imputation algorithms of the expectation–maximization (EM),²¹ mean imputation (Mean), median imputation (Median), iterative imputation with automatic model selection (Hyperimpute),²² nonparametric random forest (Missforest),²³ imputation by chained equations (Ice),²⁴ multiple imputation by chained equations (Mice),²⁵ low-rank matrix approximation (Softimpute),²⁶ and the generative adversarial imputation nets (GAIN)²⁷ were constructed to fill in the missing data with different rates. Following the imputation (data-filling) process, the data for each feature were normalized (min-max scaler) to handle the varying distribution ranges among individual features. The RMSE values were then calculated between the original dataset and the imputed dataset for each removal rate and each imputation algorithm. The imputation algorithm with the lowest RMSE value was then selected and used to fill in the missing data. In the third step (Fig. 2), machine learning based regression models were constructed to predict the electrical properties with the input features of the process parameters. All the models were created using the data that had been imputed using the selected imputation algorithm. The regression models of random forest (RF),²⁸ extremely randomized trees (Extratrees),²⁹ extreme gradient boosting (XGBoost),³⁰ gradient boosting (GBoost),³¹ categorical boosting (CatBoost),³² and the support vector regression (SVR)³³ were built. The hyper-parameters of each model were selected based on the grid search³⁴ (Table S1). For the dataset from the 353 wafers,

29 January 2026 05:43:44

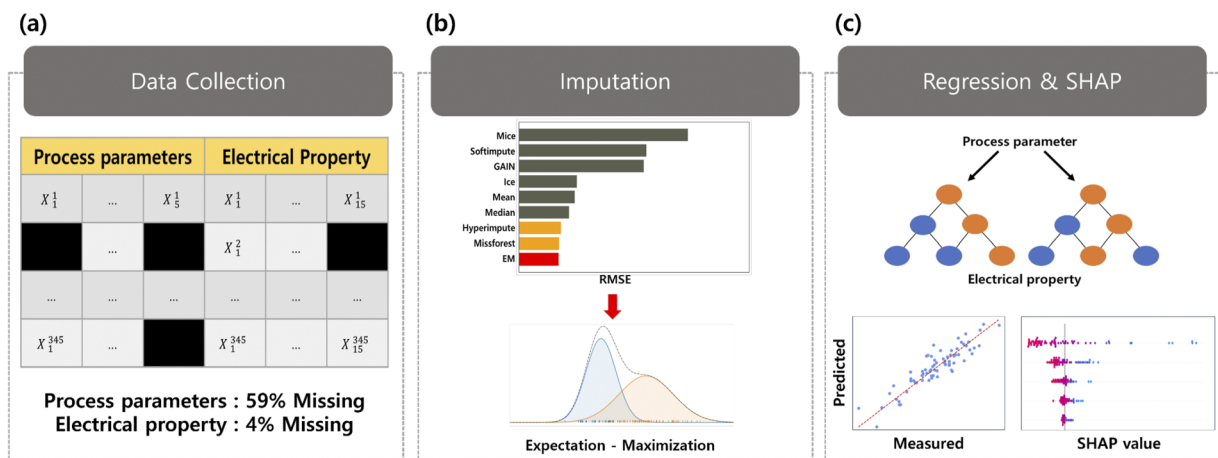


FIG. 2. Flowchart of the machine learning based analysis for the HKMG DRAM parameters. (a) In the data collection step, the parameters to be analyzed were tabulated, and missing rates for the data are identified. (b) Next, the imputation machine learning modeling was carried out to fill in the missing data. (c) Finally, regression models were built, and explainable AI algorithms were employed to investigate the correlation between the input features (process parameters) and the output parameters (electrical properties).

an 80% (training)–20% (test) split was used with a fivefold cross-validation.³⁵ A regression model with the smallest RMSE value was selected. SHAP (Shapley additive explanations) analysis was then conducted for the selected regression model³⁶ to quantify the influence of individual variables on the model’s predictions.

III. RESULTS AND DISCUSSION

A. Missing data imputation

Figure 3 and Table S2 present the imputation errors in RMSE determined from the nine imputation ML models for the datasets

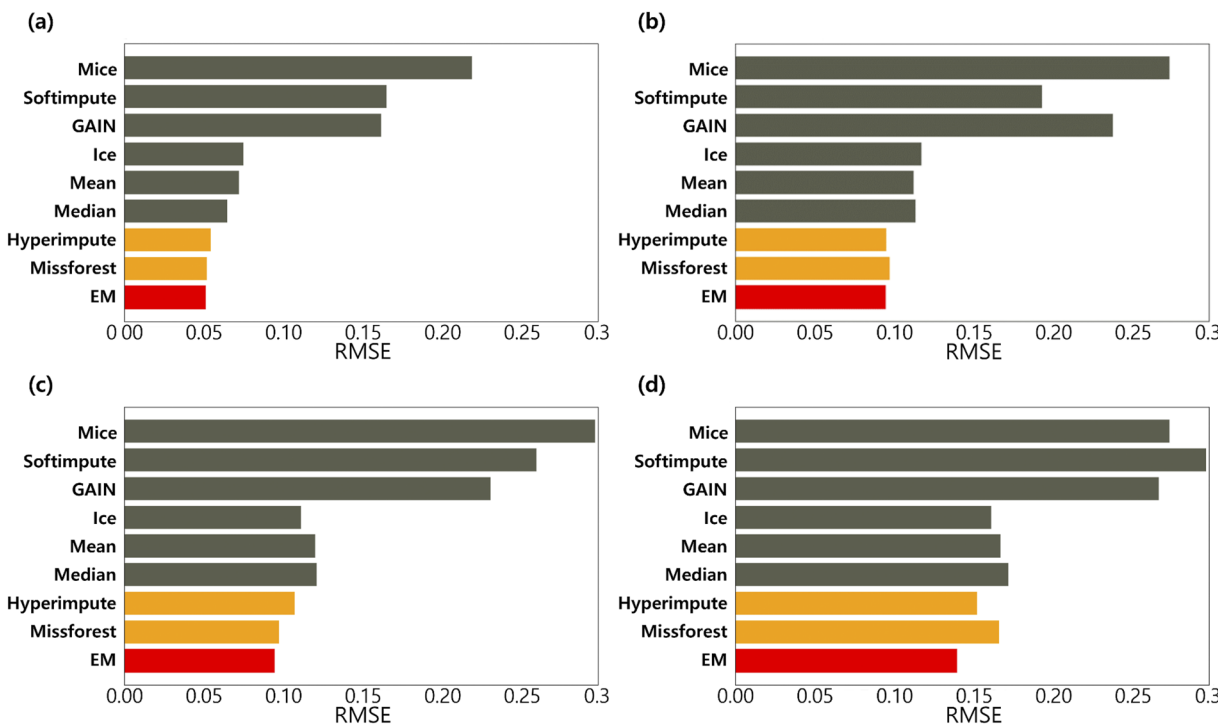


FIG. 3. RMSE values determined from various imputation algorithms for the dataset with (a) 1%, (b) 5%, (c) 10%, and (d) 30% removal rates.

with data removal rates of 1%, 5%, 10%, and 30%. Notably, algorithms such as Mice, Ice, GAIN, and Softimpute consistently showed relatively high errors, with RMSE values exceeding 0.15. The iterative imputation algorithms, such as Ice and Mice, are more appropriate for the MAR dataset and thus less suited for our MCAR dataset.³⁷ In contrast, Missforest, Hyperimpute, and EM exhibited robust performances, maintaining RMSE values below 0.16 even when dealing with the dataset with a 30% removal rate. The MCAR characteristic of the data in this work would allow EM to effectively offer statistically robust and practical imputation and result

in a low RMSE value (Fig. 3).^{38–40} In addition, tree-based Missforest and Hyperimpute algorithms seem to address the complex relation between the missing data and other parameters well, leading to low RMSE values.

Simultaneously, tree-based Missforest and Hyperimpute algorithms show compatibility with the complex interaction patterns observed in our missing data. However, when using the Missforest or Hyperimpute techniques for imputation, the filled-in data can significantly influence the model training process, especially for tree-based regression models, potentially leading to overfitting.⁴¹

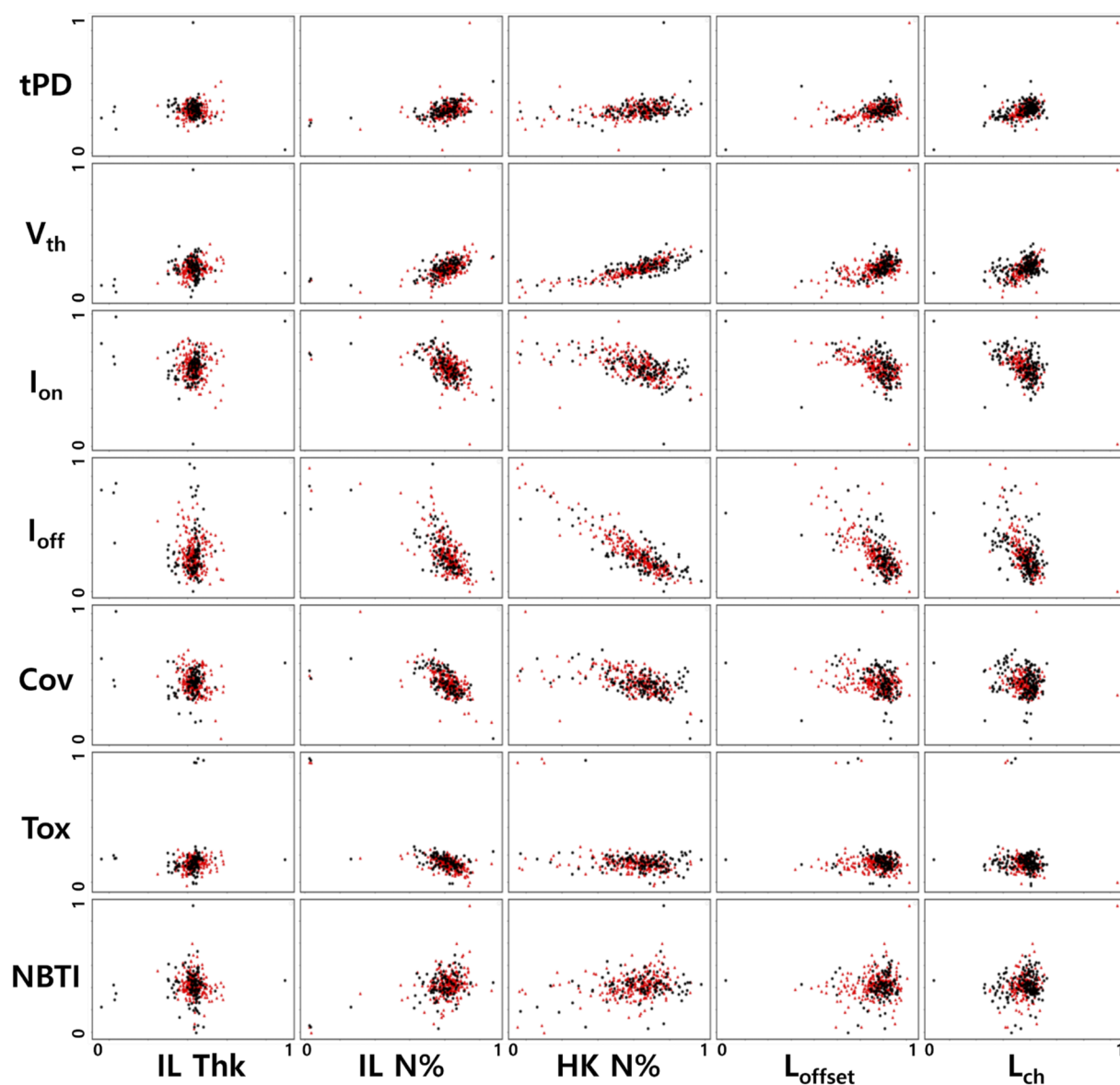


FIG. 4. Scatter plots for the normalized original (black) and imputed data (red) for seven electrical properties (y axis) and five process parameters (x axis) by the expectation–maximization (EM) imputation algorithm.

In contrast, the EM algorithm takes an iterative approach with two key steps: the expectation step (E-step) and the maximization step (M-step).^{21,42} These steps work within the parametric models to iteratively estimate model parameters, enabling the algorithm to maximize the likelihood estimates for imputation. Therefore, as a statistical and generative model, EM operates independently of various regression algorithms. Therefore, we selected EM as the imputation algorithm and all the missing data were filled using EM, as shown in Fig. 4. The interpolated datasets are found to closely resemble the distribution of the original data.

B. Regression modeling

Within the hyper-parameter ranges presented in Table S1, various regression models, including RF, Extratrees, XGBoost, GBoost, CatBoost, and SVR, were optimized using the grid search method. The optimized models were then validated through the fivefold cross-validation, and the average R^2 scores and RMSE were determined as presented in Table S3. It is evident from the table that the overall R^2 performance is not exceptionally high. Considering the inherent complexity of semiconductor manufacturing processes involving numerous steps, the construction of highly reliable regression models may be a considerable challenge. Despite this challenge, we observe promising results, with the highest R^2 scores of 0.765 for V_{th} [Extratrees, Fig. 5(a)], 0.923 for I_{off} [CatBoost, Fig. 5(b)], and 0.790 for Tox [CatBoost, Fig. 5(c)]. The R^2 scores for Cov , I_{on} , $NBTI$, and tPD are lower than 0.6 (Table II and Fig. S1), indicating that these variables inadequately account for variance in the data,⁴³ making them unsuitable for SHAP analysis. The lower R^2 scores of these properties, including Cov , I_{on} , and tPD , can be attributed to their weaker correlation with the HKMG process parameters. These properties may exhibit a substantial sensitivity to other descriptors that are not used in this paper. We, therefore, focus our SHAP analysis exclusively on V_{th} , I_{off} , and Tox .

C. Process factor analysis using SHAP

1. Threshold voltage (V_{th})

Figure 6(a) shows the SHAP mean absolute values for all the features of the prediction model for V_{th} , and Fig. 6(b) shows the SHAP individual value summary plot. It can be seen from the figures that $HK\ N\%$ exhibits the most dominant influence on V_{th} . In addition, both $HK\ N\%$ and $IL\ N\%$ exhibit a positive correlation with V_{th} . Injection of nitrogen into both the HK layer and the IL is essential for reducing pre-existing oxygen vacancies and regulating the dielectric constant. However, an excessive nitrogen content can lead to the generation of the interface states and positive fixed charges, increasing the number of interface traps (Nits), which subsequently leads to an increase in V_{th} .²⁰ The HK layer has a higher permittivity compared to the IL [permittivity: $HfSiOx$ (HK): 15, $SiON$ (IL): 4–8], hence, $HK\ N\%$ exerts a more significant influence on the formation of Nits compared to $IL\ N\%$. Increasing L_{ch} reduces the V_{th} roll-off phenomenon, thus contributing to an increase in V_{th} . The positive relation between $IL\ thk$ and V_{th} is a well-known behavior, as can be confirmed by the result in Fig. 6(b). L_{offset} is observed to have a positive impact on the effective channel length, which can be attributed to the occurrence of the lightly doped drain (LDD) overlap, thus,

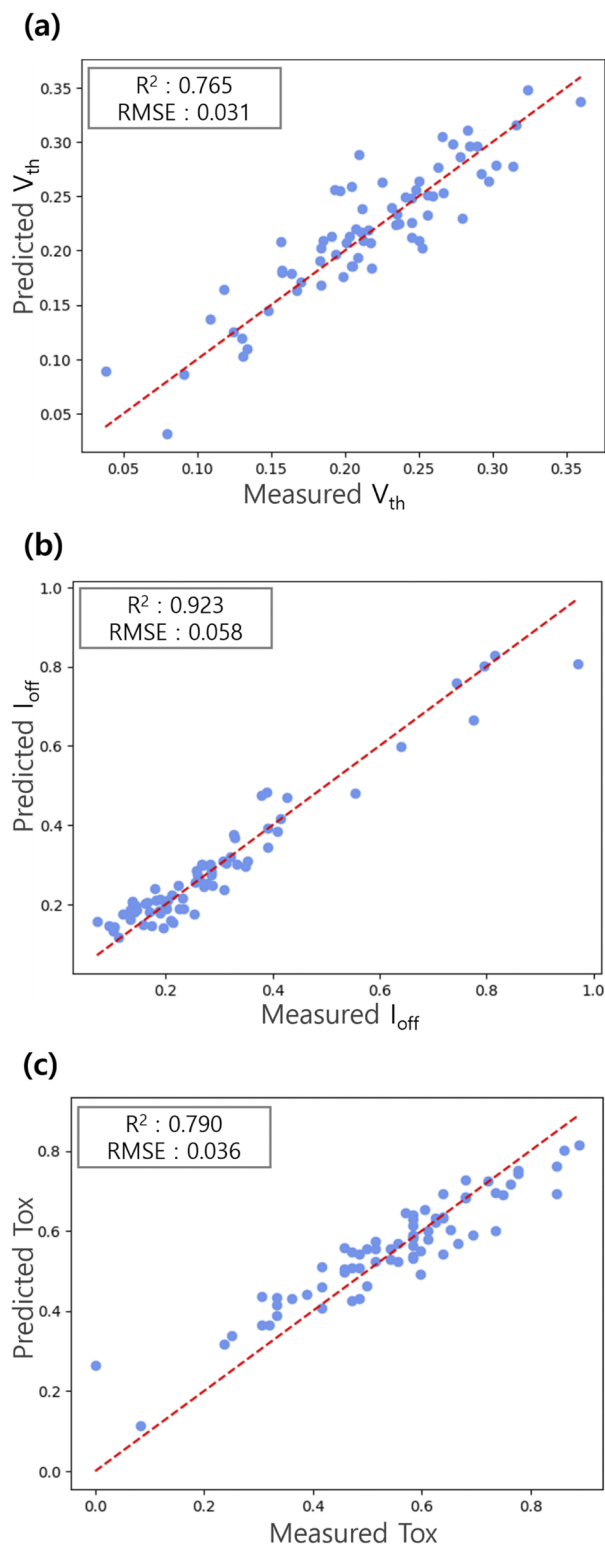


FIG. 5. Regression ML model performances for (a) threshold voltage (V_{th}), (b) off-current (I_{off}), and (c) gate oxide thickness (Tox).

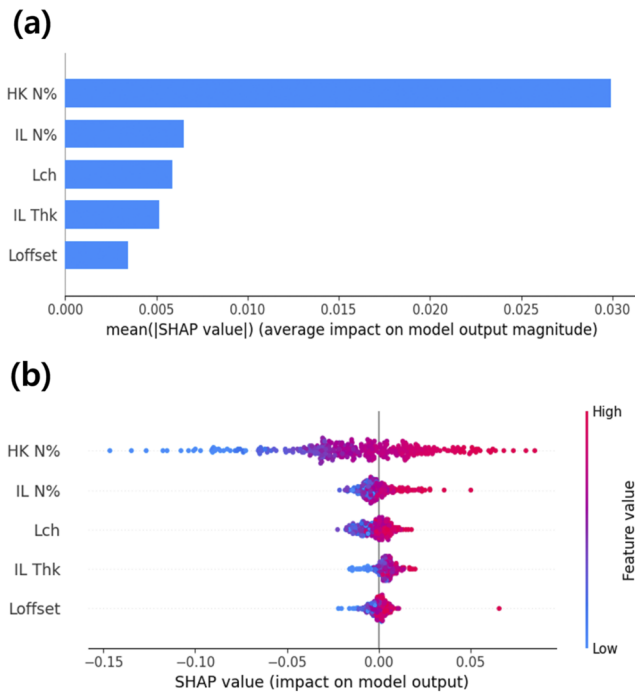


FIG. 6. Shapley additive explanation summary plots for (a) the mean absolute SHAP value and (b) individual SHAP values of the Extratrees model for threshold voltage (V_{th}).

confirming the positive influence on V_{th} . Nevertheless, it is evident from the results shown in Fig. 6 that L_{ch} , $IL\ Thk$, and L_{offset} have a relatively minor impact on V_{th} .

Figure 7 shows the SHAP double-feature dependence plots for V_{th} . As shown in Fig. 7(a), for the normalized $IL\ N\%$ range of 0–0.7, the feature does not significantly contribute to the increase in V_{th} . However, beyond 0.7, an increase in $IL\ N\%$ has a noticeable contribution to the increase in V_{th} by closely interacting with $HK\ N\%$. Similarly, it is shown in Fig. 7(b) that the SHAP values of L_{ch} on V_{th} vary depending on the L_{ch} range, and they appear to have a combined effect with L_{offset} to have a positive impact on V_{th} . Double-feature dependencies are less pronounced in other feature pairs, as shown in Fig. S2.

2. Off-current (I_{off})

From the SHAP analysis on the regression model for I_{off} , we notice that the influence of $HK\ N\%$ is dominant, followed by that of L_{offset} , $IL\ N\%$, $IL\ Thk$, and L_{ch} [Fig. 8(a)]. $HK\ N\%$, L_{offset} , and $IL\ N\%$ are found to exhibit negative correlations with I_{off} [Fig. 8(b)]. As previously discussed, the increase in $IL\ N\%$ and $HK\ N\%$ leads to an increase in Nits, subsequently leading to an increase in V_{th} and a decrease in I_{off} . In addition, the influence of $HK\ N\%$ on V_{th} and I_{off} is more pronounced than that of $IL\ N\%$ because of higher permittivity of the HK layer in comparison with the IL. Figure 8 shows that I_{off} exhibits a substantially negative correlation with L_{offset} , and a minor correlation with L_{ch} . An increase in $IL\ Thk$ would

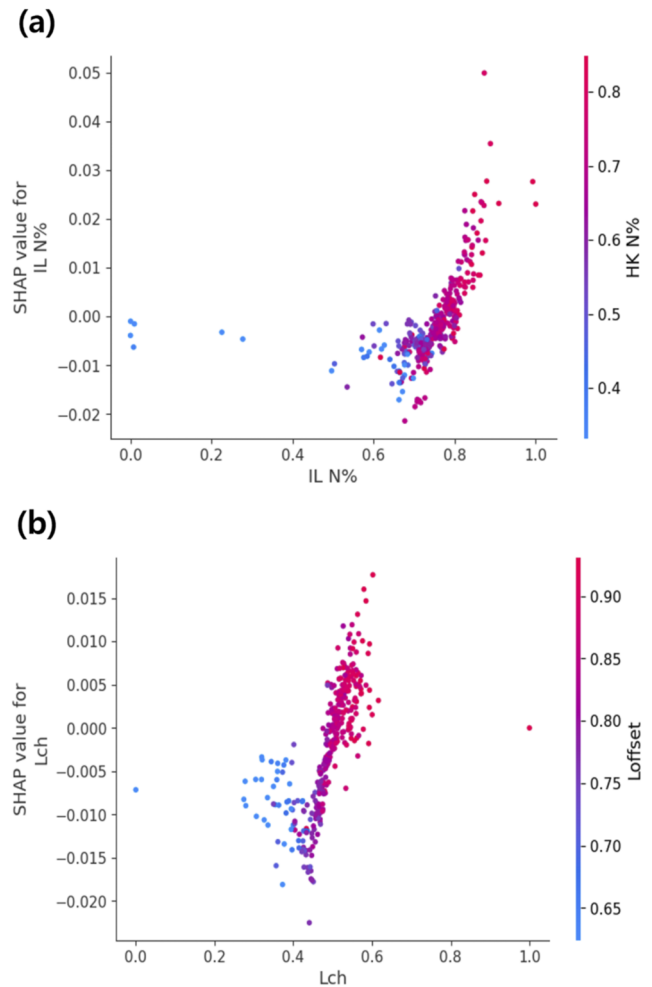


FIG. 7. Shapley additive explanation (SHAP) double-feature dependence plots for the Extratrees model for threshold voltage (V_{th}), showing the feature relationships between (a) interlayer nitrogen concentration ($IL\ N\%$) and HfSiON layer nitrogen concentration ($HK\ N\%$), (b) channel length (L_{ch}) and gate spacer offset width (L_{offset}). All the features shown are normalized values.

theoretically lead to an increase in Tox , which, in turn, might contribute to a reduction in I_{off} .⁴⁴ Due to the presence of numerous factors influencing I_{off} beyond Tox , however, the contributions of $IL\ Thk$ on I_{off} were found to be minor. Double-feature dependency is not pronounced for the regression model of I_{off} (Fig. S3).

3. Gate oxide thickness (Tox)

The result of the SHAP analysis for Tox shown in Fig. 9(a) identifies the two most significant contributors of $IL\ N\%$ and $IL\ Thk$. As shown in Fig. 9(b), $IL\ N\%$ exhibits a negative correlation with Tox , while $IL\ Thk$ appears to have a positive correlation with Tox . The directional influence of these factors is consistent with our prior observation: increase in $IL\ N\%$ results in an increase in Nits and

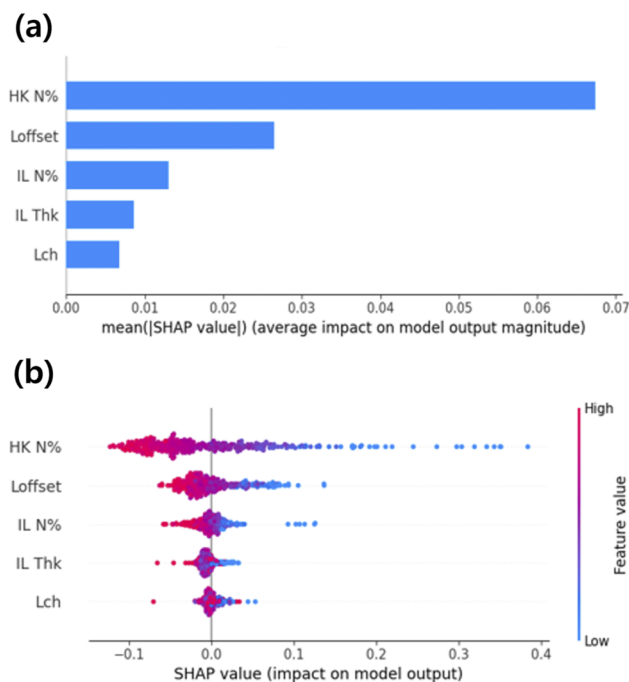


FIG. 8. Shapley additive explanation summary plots for (a) the mean absolute SHAP value and (b) individual SHAP values of the CatBoost model for off-current (I_{off}).

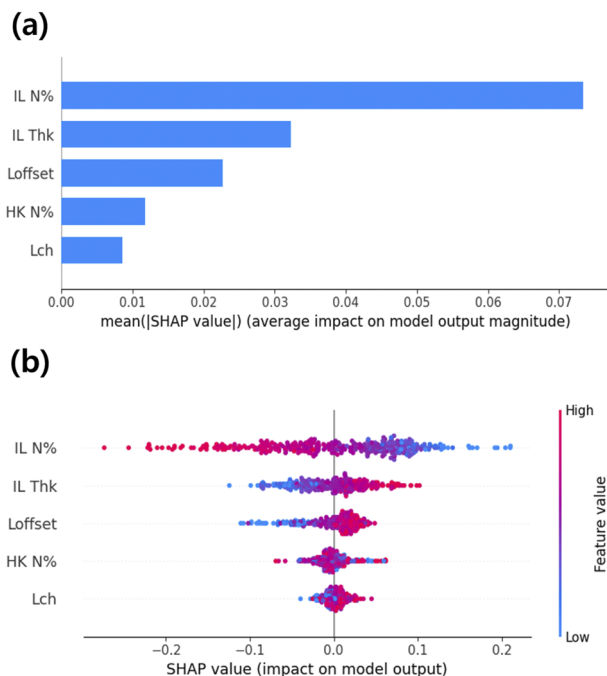


FIG. 9. Shapley additive explanation summary plots for (a) the mean absolute SHAP value and (b) individual SHAP values of the CatBoost model for Gate oxide thickness (Tox).

a decrease in $IL\ Thk$, ultimately resulting in a reduction in Tox . It is expected that $HK\ N\%$ has a more significant impact on Tox compared to $IL\ N\%$, considering that the HK layer is designed to be thicker than the IL. However, the analysis revealed that $IL\ N\%$ has a greater mean absolute SHAP value than $HK\ N\%$. The double-feature dependency is not pronounced for the regression model of Tox (Fig. S4).

The results shown in Figs. 6–9 highlight the ability of SHAP analysis to provide individual factor influences, which have been challenging to characterize through traditional linear graph-based analyses or by black-box machine learning models.⁴⁵ The consistency between the SHAP results and domain knowledge underscores the effectiveness of EM imputation to fill in the missing data and the capability of XAI to analyze the complex relation between the process parameters and the electrical properties.⁴⁶ The achievement of desired electrical properties through process parameter control can be effectively fulfilled when considering the correlations among factors with multiple interdependencies, rather than relying solely on the limited number of primary causes.

IV. CONCLUSION

In this paper, we utilized machine learning to explore the intricate relationships among the process parameters and electrical properties in the HKMG DRAM technology. Various imputation ML algorithms were tested, and the EM imputation adeptly addressed the challenge associated with missing data. The application of machine learning regression models unveiled non-linear and intricate relationships. The prediction models for V_{th} , I_{off} , and Tox consistently achieved high R^2 scores, allowing reliable SHAP analysis. The results of the XAI model found that the most significant process parameters for V_{th} , I_{off} , and Tox are all related to the concentration of nitrogen. The directional relations and double-feature dependencies have been analyzed for the features and the outputs in conjunction with the domain knowledge of the semiconductor industry. We believe that a similar strategy that combines the imputation and XAI algorithms can be used to unveil other complex relations among the process parameters and product characteristics, which may provide an opportunity for the rapid development of new DRAM products.

SUPPLEMENTARY MATERIAL

The supplementary material provides the hyper-parameters and additional results of the machine learning analysis used in this work.

ACKNOWLEDGMENTS

This research was supported by Samsung Electronics Co., Ltd. (Grant No. IO201211-08077-01), the Korea Institute for Advancement of Technology (KIAT) (Grant No. P0008458), and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (Grant No. 2021R1A4A1029780).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Namyong Kwon and JoonHo Bang contributed equally to this work.

Namyong Kwon: Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **JoonHo Bang:** Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Won Ju Sung:** Data curation (equal). **Jung Hoon Han:** Data curation (equal). **Dongin Lee:** Writing – review & editing (equal). **Ilwoo Jung:** Writing – review & editing (equal). **Se Guen Park:** Writing – review & editing (equal). **Hyodong Ban:** Writing – review & editing (equal). **Sangjoon Hwang:** Writing – review & editing (equal). **Won Yong Shin:** Supervision (equal); Validation (equal). **Jinhye Bae:** Writing – review & editing (equal). **Dongwoo Lee:** Project administration (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

REFERENCES

- ¹K. Mistry, R. Chau, C.-H. Choi, G. Ding, K. Fischer, T. Ghani, R. Grover, W. Han, D. Hanken, M. Hattendorf, J. He, C. Allen, J. Hicks, R. Huessner, D. Ingerly, P. Jain, R. James, L. Jong, S. Joshi, C. Kenyon, K. Kuhn, K. Lee, C. Auth, H. Liu, J. Maiz, B. McIntyre, P. Moon, J. Neiryck, S. Pae, C. Parker, D. Parsons, C. Prasad, L. Pipes, B. Beattie, M. Prince, P. Ranade, T. Reynolds, J. Sandford, L. Shifren, J. Sebastian, J. Seiple, D. Simon, S. Sivakumar, P. Smith, D. Bergstrom, C. Thomas, T. Troeger, P. Vandervoorn, S. Williams, K. Zawadzki, M. Bost, M. Brazier, M. Buehler, and A. Cappellani, “A 45 nm logic technology with high-k + metal gate transistors, strained silicon, 9 Cu interconnect layers, 193 nm dry patterning, and 100% Pb-free packaging,” in *2007 IEEE International Electron Devices Meeting* (IEEE, 2007), pp. 247–250.
- ²M. Sung, S.-A. Jang, H. Lee, Y.-H. Ji, J.-I. Kang, T.-O. Jung, T.-H. Ahn, Y.-I. Son, H.-C. Kim, S.-W. Lee, S.-M. Lee, J.-H. Lee, S.-B. Baek, E.-H. Doh, H.-J. Cho, T.-Y. Jang, I.-S. Jang, J.-H. Han, K.-B. Ko, Y.-J. Lee, S.-B. Shin, J.-S. Yu, S.-H. Cho, J.-H. Han, D.-K. Kang, J. Kim, J.-S. Lee, K.-D. Ban, S.-J. Yeom, H.-W. Nam, D.-K. Lee, M.-M. Jeong, B. Kwak, J. Park, K. Choi, S.-K. Park, N.-J. Kwak, and S.-J. Hong, “Gate-first high-k/metal gate DRAM technology for low power and high performance products,” in *2015 IEEE International Electron Devices Meeting* (IEEE, 2015), pp. 26.6.1–26.6.4.
- ³S. H. Jang, S. Yamada, K. Lee, J. Lim, J. Han, J. Jang, J. Yeo, C. Lee, S. Baek, J. Lee, and J.-H. Lee, “A fully integrated low voltage DRAM with thermally stable gate-first high-k metal gate process,” in *2019 IEEE International Electron Devices Meeting* (IEEE, 2019), pp. 28.4.1–28.4.3.
- ⁴R. Ritzenthaler, T. Schram, E. Bury, J. Mitard, L.-A. Ragnarsson, G. Groeseneken, N. Horiguchi, A. Thean, A. Spessot, C. Caillat, V. Srividya, and P. Fazan, “Low-power DRAM-compatible replacement gate high-k/metal gate stacks,” in *2012 Proceedings of the European Solid-State Device Research Conference* (IEEE, 2012), pp. 242–245.
- ⁵C. Ortolland, D. Jaeger, T. J. Mcardle, C. Dewan, R. R. Robison, K. Zhao, J. Cai, P. Chang, Y. Liu, V. Varadarajan, G. Wang, A. I. Chou, D. P. Ioannou, P. Oldiges,

- P. Agnello, S. Narasimha, V. Narayanan, and G. Freeman, “2nd generation dual-channel optimization with cSiGe for 22 nm HP technology and beyond,” in *2013 IEEE International Electron Devices Meeting* (IEEE, 2013), pp. 9.4.1–9.4.4.
- ⁶A. Khakifirooz, K. Cheng, N. Loubet, T. Nagumo, A. Reznicek, Q. Liu, T. M. Levin, L. F. Edge, H. He, J. Kuss, F. Allibert, B.-Y. Nguyen, B. Doris, and G. Shahidi, “Hole transport in strained and relaxed SiGe channel extremely thin SOI MOSFETs,” *IEEE Electron Device Lett.* **34**, 1358–1360 (2013).
- ⁷K.-H. Chen, C.-Y. Lin, M.-C. Chen, Y.-S. Lin, Y.-C. Chang, Y.-H. Lin, F.-Y. Jin, F.-M. Ciou, K.-C. Chang, W.-C. Hung, T.-T. Kuo, C.-H. Yeh, P.-H. Chen, and T.-C. Chang, “Advanced low-temperature-high-pressure hydrogen treatment for interface defect passivation in Si- and SiGe-channel MOSCAPs,” *IEEE Trans. Electron Devices* **67**, 5403–5407 (2020).
- ⁸D. Lee, D. You, D. Lee, X. Li, and S. Kim, “Machine-learning-guided prediction models of critical temperature of cuprates,” *J. Phys. Chem. Lett.* **12**, 6211–6217 (2021).
- ⁹D. You, S. Ganorkar, S. Kim, K. Kang, W.-Y. Shin, and D. Lee, “Machine learning-based prediction models for formation energies of interstitial atoms in HCP crystals,” *Scr. Mater.* **183**, 1–5 (2020).
- ¹⁰J. Song, H. Jo, T. Kim, and D. Lee, “Experimental data management platform for data-driven investigation of combinatorial alloy thin films,” *APL Mater.* **11**, 091117 (2023).
- ¹¹Z. Pan, Y. Liu, Z. Li, Q. Xun, and Y. Wu, “Roughness detection method based on image multi-features,” *Proc. Inst. Mech. Eng., Part E* (published online) (2023).
- ¹²C. K. Shin and S. C. Park, “A machine learning approach to yield management in semiconductor manufacturing,” *Int. J. Prod. Res.* **38**, 4261–4271 (2000).
- ¹³S.-Y. Lee, T. P. Connerton, Y.-W. Lee, D. Kim, D. Kim, and J.-H. Kim, “Semi-GAN: An improved GAN-based missing data imputation method for the semiconductor industry,” *IEEE Access* **10**, 72328–72338 (2022).
- ¹⁴D.-S. Kwak and K.-J. Kim, “A data mining approach considering missing values for the optimization of semiconductor-manufacturing processes,” *Expert Syst. Appl.* **39**, 2590–2596 (2012).
- ¹⁵R. Ritzenthaler, T. Schram, A. Spessot, C. Caillat, M. Aoulaiche, M. J. Cho, K. B. Noh, Y. Son, H. J. Na, T. Kauerauf, B. Douhard, A. Nazir, S. A. Chew, A. P. Milenin, E. Altamirano-Sanchez, G. Schoofs, J. Albert, F. Sebai, E. Vecchio, V. Paraschiv, W. Vandervorst, S.-G. Lee, N. Collaert, P. Fazan, N. Horiguchi, and A. V.-Y. Thean, “A low-power HKMG CMOS platform compatible with dram node 2x and beyond,” *IEEE Trans. Electron Devices* **61**, 2935–2943 (2014).
- ¹⁶S. Hlali, N. Hizem, L. Militaru, A. Kalbousi, and A. Soufi, “Effect of interface traps for ultra-thin high-k gate dielectric based MIS devices on the capacitance-voltage characteristics,” *Microelectr. Reliab.* **75**, 154–161 (2017).
- ¹⁷M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, A. Shanware, and L. Colombo, “Application of HfSiON as a gate dielectric material,” *Appl. Phys. Lett.* **80**, 3183–3185 (2002).
- ¹⁸S. S. Tan, T. P. Chen, J. M. Soon, K. P. Loh, C. H. Ang, and L. Chan, “Nitrogen-enhanced negative bias temperature instability: An insight by experiment and first-principle calculations,” *Appl. Phys. Lett.* **82**, 1881–1883 (2003).
- ¹⁹C.-H. Chen, Y.-K. Fang, S.-F. Ting, W.-T. Hsieh, C.-W. Yang, T.-H. Hsu, M.-C. Yu, T.-L. Lee, S.-C. Chen, C.-H. Yu, and M.-S. Liang, “Downscaling limit of equivalent oxide thickness in formation of ultrathin gate dielectric by thermal-enhanced remote plasma nitridation,” *IEEE Trans. Electron Devices* **49**, 840–846 (2002).
- ²⁰W. J. Sung, H. S. Kim, J. H. Han, S. G. Park, J.-H. Oh, H. Ban, and J. Lee, “Investigation on NBTI control techniques of HKMG transistors for low-power DRAM applications,” in *2023 IEEE International Reliability Physics Symposium* (IEEE, 2023), pp. 1–5.
- ²¹A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *J. R. Stat. Soc., Ser. B: Methodol.* **39**, 1–22 (1977).
- ²²D. Jarrett, B. Cebere, T. Liu, A. Curth, and M. Van Der Schaar, “HyperImpute: Generalized iterative imputation with automatic model selection,” in *Proceedings of the 39th International Conference on Machine Learning* (PLMR, 2022), Vol. 162, pp. 9916–9937.
- ²³D. J. Stekhoven and P. Bühlmann, “MissForest—Non-parametric missing value imputation for mixed-type data,” *Bioinformatics* **28**, 112–118 (2012).
- ²⁴P. Royston and C. Group, “Multiple imputation of missing values: Update of ice,” *Stata J.: Promot. Commun. Stat. Stata* **5**, 527–536 (2005).

- ²⁵P. Royston and I. White, "Multiple imputation by chained equations (MICE): Implementation in *Stata*," *J. Stat. Software* **45**, 1 (2011).
- ²⁶E. J. Candes and Y. Plan, "Matrix completion with noise," *Proc. IEEE* **98**, 925–936 (2010).
- ²⁷J. Yoon, J. Jordon, and M. Van Der Schaar, "GAIN: Missing data imputation using generative adversarial nets," *Proc. Mach. Learn. Res.* **80**, 5689–5698 (2018).
- ²⁸L. Breiman, "Random forests," *Mach. Learn.* **45**, 5 (2001).
- ²⁹P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.* **63**, 3–42 (2006).
- ³⁰T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery 2016), pp. 785–794.
- ³¹J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Stat.* **29**, 1189–1232 (2001).
- ³²L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," *Neural Inf. Process. Syst.* **18**, 6639–6649 (2018).
- ³³D. Ruppert, "The elements of statistical learning: Data mining, inference, and prediction," *J. Am. Stat. Assoc.* **99**, 567 (2004).
- ³⁴P. M. Lerman, "Fitting segmented regression models by grid search," *Appl. Stat.* **29**(1), 77 (1980).
- ³⁵M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. R. Stat. Soc., Ser. B: Methodol.* **36**, 111–133 (1974).
- ³⁶S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.* **30**, 4768–4777 (2017).
- ³⁷M. J. Azur, E. A. Stuart, C. Frangakis, and P. J. Leaf, "Multiple imputation by chained equations: What is it and how does it work?," *Int. J. Methods Psychiatr. Res.* **20**(1), 40–49 (2011).
- ³⁸R. Little and D. Rubin, *Statistical Analysis with Missing Data*, 3rd ed. (Wiley, 2019).
- ³⁹J. L. Schafer and J. W. Graham, "Missing data: Our view of the state of the art.," *Psychol. Methods* **7**(2), 147–177 (2002).
- ⁴⁰P. D. Allison, *The Sage Handbook of Quantitative Methods in Psychology* (Sage Publications, Ltd., Thousand Oaks, CA, 2001).
- ⁴¹D.-H. Lee, J.-K. Yang, C.-H. Lee, and K.-J. Kim, "A data-driven approach to selection of critical process steps in the semiconductor manufacturing process considering missing and imbalanced data," *J. Manuf. Syst.* **52**, 146–156 (2019).
- ⁴²T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Magn.* **13**, 47–60 (1996).
- ⁴³D. S. Moore, W. I. Notz, and M. A. Flinger, *The Basic Practice of Statistics* (W. H. Freeman and Company, New York, 2013).
- ⁴⁴N. M. Radzi and R. Sanudin, "Effect of oxide thickness variation in sub-micron NMOS transistor," *IOP Conf. Ser.: Mater. Sci. Eng.* **226**, 012145 (2017).
- ⁴⁵D. V. Fryer, I. Strumke, and H. Nguyen, "Model independent feature attributions: Shapley values that uncover non-linear dependencies," *PeerJ Comput. Sci.* **7**, e582 (2021).
- ⁴⁶J. A. Lee, J. Park, Y. T. Choi, R. E. Kim, J. Jung, S. Lee, M. H. Seo, and H. S. Kim, "Influence of tensile properties on hole expansion ratio investigated using a generative adversarial imputation network with explainable artificial intelligence," *J. Mater. Sci.* **58**, 4780–4794 (2023).