electronics

Article

Symbolic Regression Based on Kolmogorov–Arnold Networks for Gray-Box Simulation Program with Integrated Circuit Emphasis Model of Generic Transistors

Yiming Huang [1,2], Bin Li [1], Zhaohui Wu [1,*] and Wenchao Liu [2]

[1] School of Microelectronics, South China University of Technology, Guangzhou 510641, China
[2] Primarius Technologies Co., Ltd., Building C, No. 888, Huanhu Xi Er Road, Lingang New Area, Pilot Free Trade Zone of China, Shanghai 430048, China
* Correspondence: phzhwu@scut.edu.cn

Abstract: In this paper, a novel approach to symbolic regression using Kolmogorov–Arnold Networks (KAN) for developing gray-box Simulation Program with Integrated Circuit Emphasis models of generic transistors is proposed. Unlike traditional black-box models, such as artificial neural networks, (ANN), the developed KAN-based model enhances interpretability by generating explicit mathematical expressions while maintaining high accuracy in device modeling. By combining the computational efficiency of neural network approaches with the transparency of formula-based modeling, the SPICE model generation is significantly accelerated, thereby improving the efficiency of the design technology co-optimization (DTCO) process. The experimental results demonstrate that the expressions derived from the KAN model accurately represent the current–voltage (I–V) characteristics of the BSIM–CMG compact model and provide nearly symmetric results. To further validate the effectiveness and versatility of the approach, we embedded the trained I–V KAN model into a 12 nm FinFET SPICE model and performed 11-stage ring oscillator (RO) simulations. The results indicate that the KAN-based SPICE model achieves accuracy comparable to the original 12 nm FinFET SPICE model, demonstrating its potential to streamline device modeling for advanced technology nodes.

Keywords: symbolic regression; KAN; SPICE model; generic transistors; design technology co-optimization; I–V modeling

Academic Editor: Frédérique Ducroquet

Received: 10 January 2025
Revised: 5 March 2025
Accepted: 13 March 2025
Published: 16 March 2025

Citation: Huang, Y.; Li, B.; Wu, Z.; Liu, W. Symbolic Regression Based on Kolmogorov–Arnold Networks for Gray-Box Simulation Program with Integrated Circuit Emphasis Model of Generic Transistors. *Electronics* **2025**, *14*, 1161. https://doi.org/10.3390/electronics14061161

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In the post-Moore era of semiconductor industry, Design Technology Co-Optimization (DTCO) has emerged as a critical framework for developing advanced semiconductor devices [1–4]. By integrating design and technology considerations, DTCO aims to optimize performance and efficiency as technology scaling approaches fundamental physical limits. In this process, the Simulation Program with Integrated Circuit Emphasis (SPICE) model is the core technology, which serves as an essential intermediary between the fabrication process and circuit design, enabling accurate simulations of transistor behavior and providing insights into device performance in practical applications [5–7]. However, generating accurate and reliable SPICE models is a time-consuming and labor-intensive process, often requiring several days to weeks to complete. The iterative nature of design and optimization further exacerbates these delays, significantly slowing down product development cycles and hindering the rapid exploration of new technologies [8–10].

Typically, industrial applications using SPICE models aim to keep the relative error within approximately 5% for critical transistor I–V characteristics (e.g., $I_d - V_{gs}$, $I_d - V_{ds}$ curves). This level of accuracy is generally considered sufficient for practical design decisions, ensuring confidence in device performance predictions during circuit simulation. Consequently, SPICE model generation approaches must not only be rapid but must also maintain a precision that meets or surpasses this 5% threshold.

To meet the growing demand for faster SPICE model generation, various artificial neural network (ANN)-based methods have been introduced as potential alternatives [11–15]. These black-box models have demonstrated considerable success in achieving high fitting accuracy, making them suitable for device simulation. Nonetheless, ANN-based models lack explicit mathematical expressions and rely on complex neuron mappings, which limits their interpretability. Consequently, these models typically require implementation through extensible lookup tables and Verilog-A code for device substitution in simulations [16–18]. Moreover, although ANN models effectively capture the nonlinear behavior of transistors, their lack of transparency and absence of closed-form expressions present challenges when integrating them into existing SPICE simulation frameworks.

Given these limitations, there is a clear need for a method that enables rapid SPICE model creation while providing interpretable, formula-based expressions. To address this challenge, a SPICE modeling approach based on Kolmogorov–Arnold Networks (KAN) is proposed. Leveraging the Kolmogorov–Arnold representation theorem, KAN has the ability to decompose complex multivariate functions into combinations of univariate functions, gaining increasing recognition across various domains [19–22]. This capability allows the KAN model to generate explicit mathematical expressions that accurately represent the current-voltage (I–V) characteristics of transistors.

The KAN-based model combines the computational efficiency of neural network approaches with the transparency of formula-based modeling. For emerging process nodes, developing accurate SPICE models traditionally requires extensive device characterization and complex parameter extraction, often taking several months of iterative refinement. Our approach accelerates this process by automatically generating compact mathematical expressions from limited measurement data. The expressions derived from the KAN model can be directly embedded into existing SPICE frameworks, reducing SPICE model development time for new technology nodes. This acceleration is particularly valuable in advanced nodes, where complex physical effects make traditional analytical modeling increasingly challenging. By streamlining the transition from device measurements to simulation-ready models, our approach enhances the efficiency of DTCO workflows and shortens PDK delivery timelines. Additionally, the KAN model provides nearly symmetric results in I–V characteristics, improving the accuracy and reliability of device simulations across both nominal and corner cases.

The rest of this paper is organized as follows: Section 2 introduces the proposed KAN-based SPICE modeling methodology, detailing the data preprocessing steps, the construction of the KAN architecture inspired by the Kolmogorov–Arnold representation theorem, the training process with hyperparameter tuning, and the symbolic regression approach for deriving interpretable mathematical expressions. Section 3 presents the experimental results and analysis. In addition to evaluating the proposed model on the BSIM-CMG compact model dataset, this section demonstrates the KAN-based model's performance on a 12 nm FinFET SPICE model. It includes an 11-stage ring oscillator (RO) simulation to further validate its practical applicability. Section 4 outlines potential future developments needed to extend KAN-based modeling to high-frequency circuit design. Finally, Section 5 concludes the paper by summarizing the key findings and emphasizing

the potential of the proposed KAN approach for accelerating SPICE model development in the semiconductor industry.

2. Modeling

2.1. Data Preparation and Preprocessing

The foundation of an accurate transistor model is built upon a comprehensive dataset covering diverse operating conditions. The data preparation and preprocessing steps are as follows.

Two complementary sets of I–V characteristic data, specifically drain current (I_d) as a function of gate-source voltage (V_{gs}) and drain-source voltage (V_{ds}), are collected for this study. The first dataset is sourced from simulations using a BSIM-CMG Compact model to establish baseline performance. The second dataset consists of de-identified measurements from 12 nm FinFET devices, obtained through standard parameter extraction procedures. This dual-dataset approach enables comprehensive validation of the KAN methodology on both idealized simulation data and real-world manufacturing data, ensuring practical applicability in commercial semiconductor design workflows.

Input voltages V_{gs} and V_{ds} are normalized to a consistent range [0, 1] using min–max scaling. This normalization improves numerical stability during the training process and prevents features with larger magnitudes from dominating the learning process. The normalization is applied as shown in Equation (1):

$$V_{norm} = \frac{V - V_{min}}{V_{max} - V_{min}}, \tag{1}$$

where V represents each voltage variable independently, V_{min} is the minimum value and V_{max} is the maximum value for that variable in the dataset. For both datasets, V_{min} = 0 V, while V_{max} = 0.8 V for the BSIM-CMG dataset and V_{max} = 2.0 V for the experimental 12 nm FinFET dataset.

Due to the wide dynamic range of I_d, particularly in the subthreshold region, a logarithmic transformation is applied to the current values. A small constant (ε, e.g., 10^{-13} A) is added to avoid taking the logarithm of zero. This transformation ensures that small current variations in the subthreshold region are adequately represented during training. The transformation is performed using Equation (2):

$$log(I_d) = log_{10}(I_d + \varepsilon). \tag{2}$$

The preprocessed datasets are then split into training and testing sets with a ratio of 9:1. Stratified sampling is applied to ensure that both the training set and the unseen test set maintain a balanced representation of all operating regions, supporting robust model evaluation and generalization performance, as illustrated in Figure 1.

For the BSIM-CMG simulation dataset shown in Figure 1a, the V_{gs} and V_{ds} values range from 0 to 0.8 V, with V_{gs} sampled at intervals of 0.1 V, and V_{ds} sampled at intervals of 0.01 V. For the experimental 12 nm FinFET dataset shown in Figure 1b, an extended voltage range of 0 to 2 V is used, with V_{gs} sampled at intervals of 0.2 V and V_{ds} sampled at intervals of 0.05 V. This extended range ensures model robustness outside normal operating regions. All data points where V_{gs} = 0.4 V, along with additional randomly selected points to maintain the 9:1 ratio, are used as the test dataset, while the remaining data are used for training.

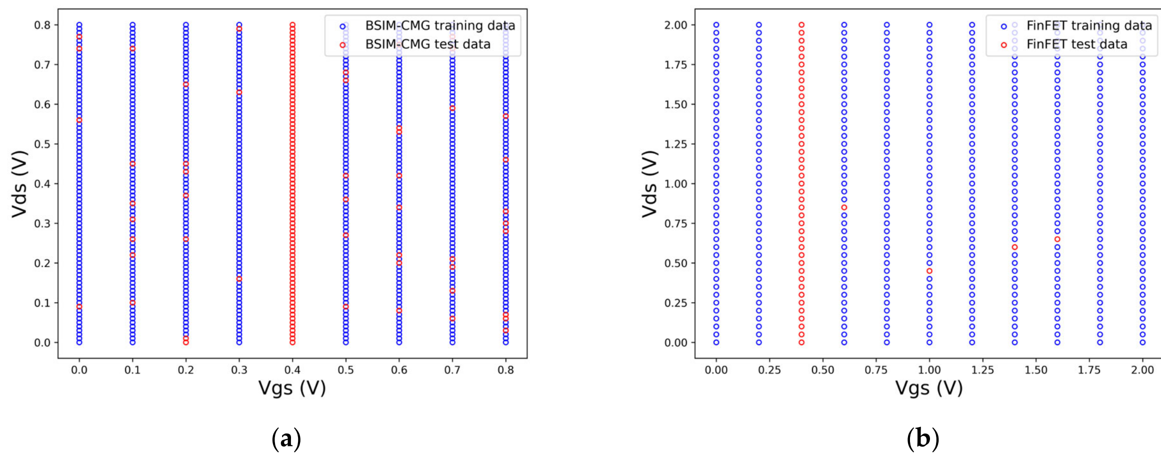

Figure 1. Datasets used for the KAN-based model: (**a**) BSIM-CMG compact model discrete measured points; (**b**) 12 nm FinFET discrete measured points.

2.2. KAN Model Architecture

The core of the proposed approach is the KAN model, which is inspired by the Kolmogorov–Arnold representation theorem. In this architecture, complex multivariate functions are decomposed into combinations of simpler univariate functions.

As depicted in Figure 2, the KAN model employed in this study is designed with a multi-layer structure adapted specifically for transistor modeling. The specific shape of the KAN is defined as [input, [4, 1], [4, 1], [4, 1], output].

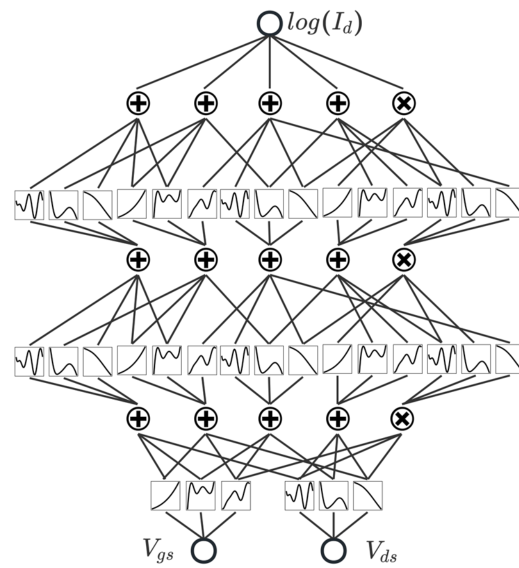

Figure 2. KAN model architecture.

Input refers to the two input features, V_{gs} and V_{ds}, [4, 1], [4, 1], [4, 1] representing three hidden layers. These layers consist of four neurons for sum connections, which represent additive combinations of the inputs or intermediate outputs from the previous layer, and one neuron for product connections, which represent multiplicative interactions between intermediate outputs. The output node corresponds to the log-transformed drain current ($log(I_d)$).

The specific layer structure described above allows the KAN model to capture both additive and multiplicative interactions between the input voltages, providing a robust framework for modeling the complex behavior of transistors.

Each edge in the KAN is associated with a learnable activation function implemented as a cubic B-spline. Unlike traditional neural networks with fixed activation functions (such as ReLU or sigmoid), these B-splines consist of piecewise cubic polynomials joined at knot points, offering continuous second derivatives. Each B-spline activation is parameterized by a set of learnable control points that are optimized during training, typically 3–4 per edge in our implementation. This approach provides local control over the function's shape and allows the model to adapt precisely to the nonlinear relationships between voltages and currents across different operating regions.

By leveraging both sum and product connections across multiple hidden layers with flexible B-spline activations, the KAN model can effectively capture intricate relationships between input voltages and their impacts on the output drain currents. This combination enables efficient approximation of the underlying device physics without being constrained to predetermined functional forms, making it a tool for accurate and interpretable transistor modeling.

To further enhance the model's smoothness and maintain Gummel symmetry, we replace the original spline network used for I_d modeling with a set of continuous analytical functions. This transition ensures smooth behavior across the entire V_{ds} range, particularly around V_{ds} = 0, where maintaining symmetry is crucial. Unlike piecewise spline-based approaches, which may introduce discontinuities in higher-order derivatives, the adoption of continuous analytical functions enables a seamless and physically consistent representation of transistor behavior across all operating regions.

2.3. Model Training and Hyperparameter Tuning

Training the KAN model involves optimizing its parameters (spline coefficients) to minimize the difference between predicted and actual drain current values. The mean squared error (MSE) between the predicted and actual log-transformed drain currents serves as the loss function, guiding the optimization process. The Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm is employed for optimization due to its efficiency in handling a large number of parameters [19]. L1 and L2 regularization are applied to the spline coefficients to prevent overfitting and encourage sparsity in the model. The regularization strengths are tuned through experimentation. Key hyperparameters include the learning rate for the L-BFGS optimizer, the regularization coefficients, the number of grid points for the spline representation (G), and the number of neurons in each hidden layer. These hyperparameters are carefully tuned using a validation set to achieve optimal performance.

In KAN model training, regularization plays a critical role in preventing overfitting while preserving the model's expressiveness. We employ a composite regularization strategy incorporating both L1 and L2 terms in the loss Equation (3):

$$L(\theta) = MSE(y, \hat{y}) + \lambda_1 \sum |\theta_i| + \lambda_2 \sum \theta_i^2, \quad (3)$$

where MSE(y, ŷ) represents the mean squared error between the predicted and actual log-transformed drain currents, θ_i are the spline coefficients, and λ_1 and λ_2 control the strength of L1 and L2 regularization, respectively.

The L1 term induces sparsity by reducing small coefficients to zero, effectively performing feature selection within the B-spline representation. This is particularly important for KAN models, where the high flexibility of cubic B-splines could otherwise lead to unnecessarily complex expressions. The L2 term prevents coefficient explosion by penalizing large weights, thereby stabilizing the optimization process and improving the model's numerical behavior, especially in regions between data points.

Through cross-validation on a separate validation set comprising 15% of the training data, we systematically evaluated regularization strengths across five orders of magnitude. The optimal values were determined to be $\lambda_1 = 10^{-4}$ and $\lambda_2 = 5 \times 10^{-5}$, which minimize validation error while maintaining model expressiveness. This regularization configuration significantly improved the model's generalization ability, reducing test error by approximately 28% compared to unregularized training.

In the spline activation functions, we use G = 3 B-spline grid points for each neuron. Our experiments indicate that G = 3 strikes a suitable compromise between expressive power and computational efficiency, enabling sufficient function approximation without incurring excessive training overhead. The key hyperparameters in our approach include the learning rate for the L-BFGS optimizer, the regularization coefficients, the B-spline grid size (G), and the number of neurons within each hidden layer. These hyperparameters are finely tuned on a validation set to achieve an optimal balance between accuracy and model complexity.

2.4. Symbolic Regression and Model Interpretation

After training, the learned spline activation functions are converted into symbolic expressions. This transformation involves mapping the spline coefficients and basis functions to a combination of elementary functions, such as polynomials, exponentials, and trigonometric functions, chosen from a scalable predefined library, as shown in Table 1.

Table 1. Function categories and types used in symbolic regression.

Category	Functions
Polynomial functions	x, x^2, x^3, x^4, x^5
Reciprocal functions	$x^{-1}, x^{-2}, x^{-3}, x^{-4}$
Root functions	$x^{0.5}, x^{1.5}, x^{-1.5}$
Exponential and Logarithmic	$e^x, \ln(x), \lg(x)$
Trigonometric and Hyperbolic	$\sin(x), \cos(x), \tan(x), \tanh(x)$
Inverse trigonometric	$\arcsin(x), \arccos(x), \arctan(x)$
Other functions	$\lvert x \rvert, sigmoid(x)$

The resulting symbolic expression is represented as an explicit function of V_{gs} and V_{ds}, providing a transparent and interpretable model of the transistor's behavior.

3. Results

3.1. Generated Expressions by KAN

The explicit mathematical expression generated by the KAN model is presented. The expression was derived through symbolic regression, based on the input variables, V_{gs} and V_{ds}. Due to the length and complexity of the expression, a portion of the expression is summarized in Table 2 for clarity.

Table 2. Partial representation of the I–V curve explicit expression generated by KAN.

$$log(I_d) = -2.3212 \times ((-0.0227 \times (-(0.026 \times ((0.0036 \times (0.3631 - V_{gs})^2 - 0.0004 \times (0.7952 - V_{ds})^2 + 0.2137) \times (0.0145 \times (0.3803 - V_{gs})^2 - 0.0264 \times (1 - 0.266 \times V_{ds})^2 + 0.0401) \times (-0.0017 \times (0.4688 - V_{ds})^2 + 0.0013 \times (-V_{gs} - 0.3064)^2 + \ldots + 0.0003 \times (0.6716 - V_{ds})^2 + 1)^2 - 0.0023 \times ((0.6001 - V_{ds})^2 - 0.1208 \times (1 - 0.9566 \times V_{gs})^2 + 0.7343)^2 + 0.0002 \times (-0.1413 \times (0.6124 - V_{ds})^2 + (1 - 0.7689 \times V_{gs})^2 - 0.2985)^2 + 0.0006 \times (0.0089 \times (0.7548 - V_{ds})^2 - 0.0008 \times (-V_{gs} - 0.1455)^2 + 1)^2 - 0.1514)^2 - 1)^2 - 3.0158$$

Although the generated expression is redundant and lengthy, its clear and explicit form makes it suitable for direct embedding into SPICE models. This facilitates more transparent and efficient simulations compared to black-box models, which lack such interpretability. The explicit nature of the expression ensures seamless integration into existing SPICE frameworks without the need for complex conversions or approximations. Additionally, the efficiency of the KAN architecture allows for rapid generation of symbolic models, making the modeling process faster and more streamlined.

As summarized in Table 2, this gray-box structure is advantageous for integrating I–V equations directly within traditional SPICE frameworks, reducing the reliance on extra interfaces or interpolation modules. Consequently, this approach not only improves simulation efficiency but also enhances the maintainability of the overall modeling system.

Compared to traditional SPICE model generation methods, which typically require several days to weeks due to the iterative nature of parameter tuning and optimization, the KAN-based approach offers a significant reduction in model development time. The KAN-based approach can produce accurate symbolic models within 6–12 h for small datasets or 1–2 days for larger, more complex datasets. This represents a 3–10× acceleration, achieved without compromising accuracy, providing a practical advantage over conventional methods in terms of speed.

3.2. Comparison of Fitting Performance: KAN vs. ANN

To evaluate the performance of the KAN model, a comparison was conducted against a traditional ANN model for both the BSIM-CMG model and the 12 nm FinFET SPICE model. The ANN model was designed with a simple architecture consisting of five layers: an input layer, three hidden layers, and an output layer. Each hidden layer consists of 50 neurons, and the Sigmoid activation function was used between layers. Xavier normal initialization was applied to the weights to ensure proper convergence and stability during training.

The training of the ANN model was performed using the Adam optimizer with a learning rate of 0.01. The MSE loss function was employed as the task involves regression. The model was trained for 20,000 steps, and the training process involved minimizing the MSE between the model's predictions and the true values on the training dataset. The training was conducted on a GPU when available, ensuring optimal performance.

After training, the model was evaluated on the test dataset. Both the MSE and mean absolute error (MAE) metrics were used to assess the fitting accuracy of the ANN and KAN models. The results are summarized in Table 3.

Table 3. Performance comparison between KAN and ANN models.

Model	MSE	MAE
ANN for BSIM-CMG model	0.0001192	0.0085191
KAN for BSIM-CMG model	0.0001056	0.0082065
ANN for 12 nm FinFET SPICE model	0.0000892	0.0088232
KAN for 12 nm FinFET SPICE model	0.0000093	0.0027809

As shown in Table 3, the KAN model achieves lower MSE and MAE values compared to the ANN model for both the BSIM-CMG model and the 12 nm FinFET SPICE model, indicating superior fitting accuracy. Notably, the performance advantage of KAN is particularly significant for the 12 nm FinFET SPICE model, where it reduces the MSE by approximately an order of magnitude compared to the ANN approach. Additionally, the KAN model provides interpretable expressions, whereas the ANN model remains a black-box approach, limiting transparency.

As illustrated in Figure 3, both the KAN and ANN models demonstrate high accuracy in predicting the I–V characteristics of the SPICE model. Across the entire voltage range, both models exhibit strong alignment with the SPICE model's curve, effectively capturing the nonlinear behavior of the transistor. Figure 3a presents the comparison for the BSIM-CMG model, while Figure 3b shows the results for the 12 nm FinFET SPICE model. This indicates that both models are well-suited for generalizing to unseen test data, providing reliable predictions for transistor behavior.

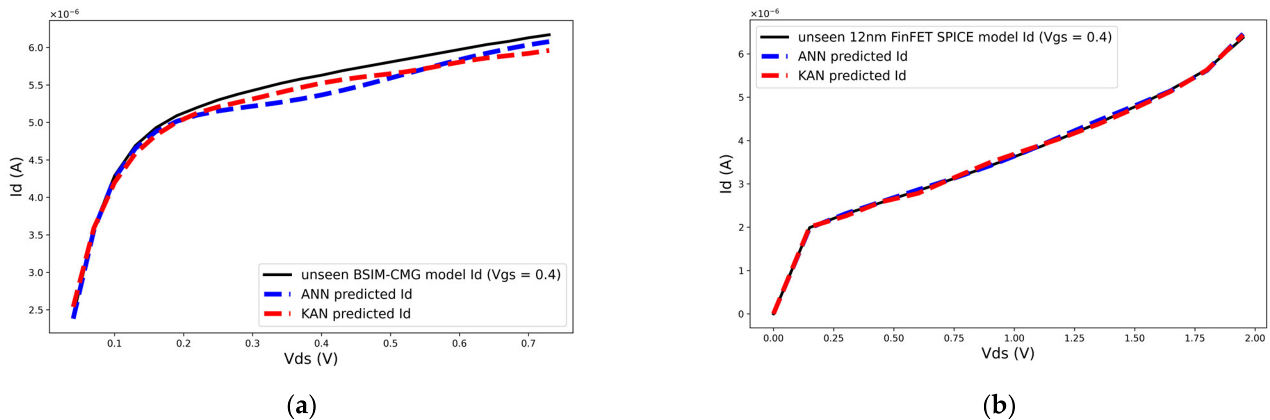

Figure 3. Fitting performance on unseen test data: (**a**) comparison between KAN and ANN models for the BSIM-CMG model, (**b**) comparison between KAN and ANN models for 12 nm the FinFET SPICE model.

However, the KAN model exhibits a slight advantage over the ANN model in regions with pronounced nonlinearity, such as when V_{gs} is near the threshold or V_{ds} becomes relatively large. In these regions, the KAN model achieves better alignment with the SPICE model's I–V curve, as highlighted by the red arrows in Figure 3. Quantitatively, the KAN model maintains a relative error within 5%, which is consistent with industry standards for SPICE modeling. By contrast, the ANN model shows slightly larger deviations in these regions, particularly when capturing rapid changes in the I–V relationship.

The advantage of the KAN model in nonlinear regions can be attributed to the local refinement capability of B-spline basis functions. Unlike ANN models, which rely on a global distribution of weights, the KAN approach leverages localized approximations to better adapt to the rapid variations in the I–V characteristics. This localized refinement allows the KAN model to effectively handle complex transistor behavior while maintaining high accuracy, even in areas with significant nonlinearities.

Overall, the results in Figure 3 confirm that the KAN model not only generalizes well to unseen data but also demonstrates superior fitting accuracy in capturing the nonlinear characteristics of the transistor across both theoretical models (BSIM-CMG) and industrial-grade models (12 nm FinFET). Beyond its modeling accuracy, the KAN model's interpretable mathematical structure provides a practical advantage over traditional black-box models, making it highly suitable for industrial applications such as DTCO. This capability highlights the KAN model's potential to accelerate SPICE model development while maintaining compliance with industry standards, bridging the gap between accuracy and interpretability in transistor modeling.

3.3. Symmetry Testing and RO Simulation

Gummel testing is commonly used in the SPICE model to assess the symmetry and accuracy of the I–V characteristics, particularly for transistors. Symmetry in the I–V curve is a critical property for ensuring that the model accurately represents the physical behavior

of the device across different operating regions. Therefore, the Gummel test was conducted to evaluate the KAN model's ability to replicate this essential feature.

As plotted in Figure 4, the test results are divided into three parts: (a) shows the comparison of the smoothed I–V curve between the KAN model and the SPICE model, (b) illustrates the first derivative of the I–V curve, and (c) presents the second derivative. In all figures, the mathematical expressions generated by the KAN model closely align with the results from the SPICE model, demonstrating near-perfect symmetry in the I–V characteristics.

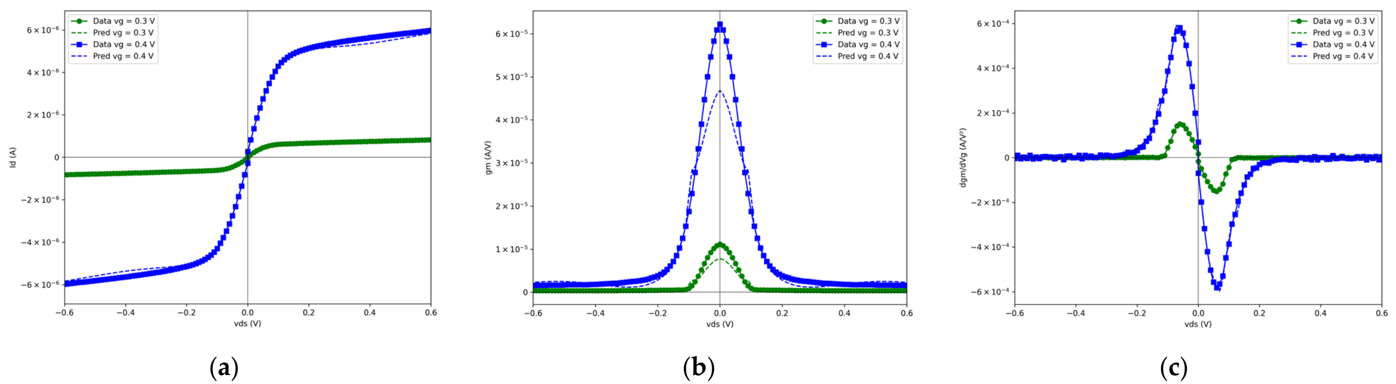

Figure 4. Symmetry performance in the Gummel test. (**a**) shows the comparison of the smoothed I–V curves between the KAN model and the SPICE model. (**b**) illustrates the first derivative of the I–V curve. (**c**) presents the second derivative of the I–V curve.

Across these comparisons, the KAN model captures not only the basic I–V characteristics but also preserves critical derivative properties throughout the operating range. Particularly noteworthy is the model's ability to accurately represent the transconductance ($g_m = dI_d/dV_{gs}$) and its derivative (dg_m/dV_{gs}), which are fundamental parameters in transistor circuit design. The consistency between the KAN model and SPICE model extends to these higher-order derivatives, underlining that the KAN approach preserves essential physical properties such as symmetry during operation. From a gray-box perspective, this fidelity in representing both the direct behavior and its derivatives is indispensable for ensuring reliable circuit simulations in both time and frequency domains, thereby validating the efficiency of the KAN approach in SPICE modeling.

Furthermore, the Gummel test demonstrates that the KAN model is suitable for incorporating into SPICE simulations, as it provides an equivalent level of accuracy and reliability as the traditional SPICE model.

To further validate the practical applicability of the KAN model in circuit design applications, we embedded the KAN-based I_d model for 12 nm n-type FinFET into the standard 12 nm FinFET SPICE model framework. An 11-stage ring oscillator (RO) circuit was then constructed and simulated using both the original commercial SPICE model and the KAN-embedded model. The comparison of the characteristic waveforms from both models is shown in Figure 5.

The simulation results demonstrate that the output voltage waveforms of the 11-stage RO simulated using the KAN-embedded model are highly consistent with those from the commercial SPICE model. Both models exhibit stable oscillation behavior, with the KAN model accurately capturing the oscillation frequency and amplitude characteristics. This close agreement confirms that the KAN model not only preserves the device-level symmetry properties but also maintains high fidelity when integrated into complex circuit simulations, making it suitable for practical circuit design and verification workflows.

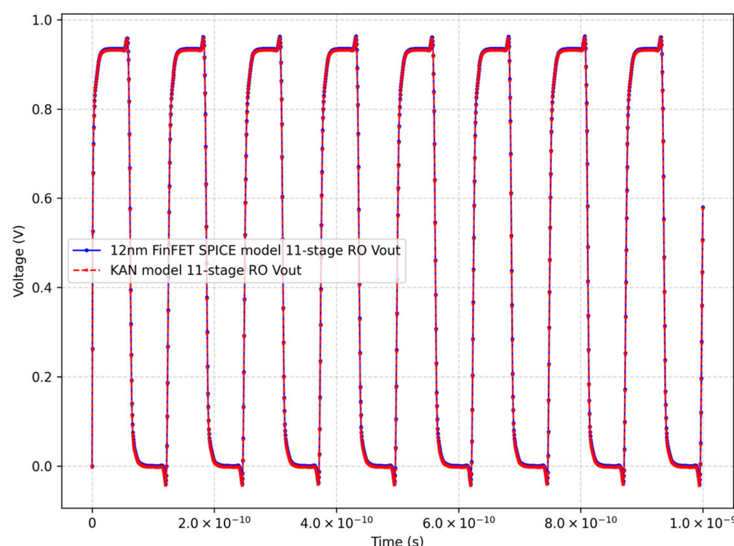

Figure 5. Comparison of 11-stage ring oscillator simulation waveforms between KAN-embedded model and commercial SPICE model for 12 nm FinFET technology.

4. Discussion

While our KAN-based model demonstrates excellent performance for DC and transient simulations, further investigation is needed for RF applications. Specifically, RF circuit design would require accurate preservation of higher-order derivatives beyond the second order for reliable harmonic balance analysis. Though our model ensures basic symmetry conditions and differentiability, comprehensive validation in high-frequency domains remains unexplored. The model's behavior under large signal swings typical in RF applications would need additional verification. Future work will focus on extending the KAN framework to explicitly address RF performance metrics, incorporating noise modeling, and validating through complex RF circuit benchmark simulations to establish the approach's viability across the full spectrum of analog, digital, and RF applications in advanced technology nodes.

5. Conclusions

In this paper, a novel symbolic regression approach using KAN is introduced for developing gray-box SPICE models of generic transistors. The KAN model has effectively combined the computational efficiency of neural networks with the interpretability of formula-based modeling by generating explicit mathematical expressions that integrate seamlessly into SPICE frameworks. The results indicate that the KAN model can accurately capture the transistor's I–V characteristics and significantly accelerate SPICE model generation, thereby enhancing the efficiency of the DTCO process.

The experimental evaluations demonstrate that the expressions derived from the KAN model provide nearly symmetric I–V characteristics, improving the accuracy and reliability of device simulations. Furthermore, the 11-stage ring oscillator simulations validate the model's high performance in circuit-level applications, with oscillation frequency and power consumption predictions that closely match reference data. This confirms the KAN model's capability to maintain accuracy across both device and circuit levels without compromising simulation efficiency. Compared to traditional black-box models, such as ANN, the KAN model can not only achieve superior fitting accuracy but also offer enhanced interpretability, enabling direct embedding into existing SPICE models without the need for complex conversions or approximations. By addressing the limitations of existing methods, the developed KAN-based approach can offer an effective and efficient

solution for SPICE modeling of generic transistors, holding substantial practical application value in the semiconductor industry.

Author Contributions: Conceptualization, Y.H.; methodology, Y.H.; software, Y.H.; validation, Y.H.; formal analysis, Y.H.; investigation, Y.H.; resources, B.L., Z.W. and W.L.; data curation, Y.H.; writing—original draft preparation, Y.H.; writing—review and editing, B.L. and Z.W.; visualization, Y.H.; supervision, B.L. and Z.W.; project administration, B.L. and W.L.; funding acquisition, B.L., Z.W. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Guangdong S&T Programme, China [Grant number 2022B0101180001].

Data Availability Statement: The datasets presented in this article are not readily available because PDK files are confidential and require declassification before use. Requests to access the datasets should be directed to the corresponding author.

Conflicts of Interest: Author Wenchao Liu was employed by the company Primarius Technologies Co. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Han, G.; Hao, Y. Design technology co-optimization towards sub-3 nm technology nodes. *J. Semicond.* **2021**, *42*, 020301. [CrossRef]
2. Tatum, L.P.; Sikder, U.; Liu, T.J.K. Design technology co-optimization for back-end-of-line nonvolatile NEM switch arrays. *IEEE Trans. Electron Devices* **2021**, *68*, 1471–1477. [CrossRef]
3. Sun, Y.; Zhang, Z.; Liu, T.J.K. Improved MEOL and BEOL parasitic-aware design technology co-optimization for 3 nm gate-all-around nanosheet transistor. *IEEE Trans. Electron Devices* **2021**, *69*, 462–469. [CrossRef]
4. Zhang, S.; Zhang, L.; Ding, C.; Wang, L.; Zhang, H.; Ding, M.; Zhang, S.; Shi, W.; Wei, Y. DTCO optimizes critical path nets to improve chip performance with timing-aware OPC in deep ultraviolet lithography. *Appl. Opt.* **2023**, *62*, 7216–7225. [CrossRef] [PubMed]
5. Ghasemi, M.; Sam, M.; Moaiyeri, M.H.; Khosravi, F.; Navi, K. A new SPICE model for organic molecular transistors and a novel hybrid architecture. IEICE Electron. *Express* **2012**, *9*, 926–931.
6. Takahashi, Y.; Sekine, T.; Yokoyama, M. SPICE model of memristive device using Tukey window function. *IEICE Electron. Express* **2015**, *12*, 20150149. [CrossRef]
7. Zhang, C.Y.; Fu, H.P.; Zhu, Y.Y.; Ma, J.G.; Zhang, Q.J. A hybrid model of III–V FETs with accurate high-order derivatives. *IEICE Electron. Express* **2017**, *14*, 20170448. [CrossRef]
8. Liu, X.; Ytterdal, T.; Kachorovskii, V.Y.; Shur, M.S. Compact terahertz SPICE/ADS model. *IEEE Trans. Electron Devices* **2019**, *66*, 2496–2501. [CrossRef]
9. Zhang, Z.; Wang, R.; Chen, C.; Huang, Q.; Wang, Y.; Hu, C.; Wu, D.; Wang, J.; Huang, R. New-Generation Design-Technology Co-Optimization (DTCO): Machine-Learning Assisted Modeling Framework. In *2019 Silicon Nanoelectronics Workshop (SNW)*; IEEE: New York, NY, USA, 2019; pp. 1–2.
10. Dai, X.; Jha, N.K. Improving convergence and simulation time of quantum hydrodynamic simulation: Application to extraction of best 10-nm FinFET parameter values. *IEEE Trans. Very Large Scale Integr. Syst.* **2016**, *25*, 319–329. [CrossRef]
11. Wang, J.; Kim, Y.H.; Ryu, J.; Jeong, C.; Choi, W.; Kim, D. Artificial neural network-based compact modeling methodology for advanced transistors. *IEEE Trans. Electron Devices* **2021**, *68*, 1318–1325. [CrossRef]
12. Jeong, H.; Woo, S.; Choi, J.; Cho, H.; Kim, Y.; Kong, J.T.; Kim, S. Fast and expandable ANN-based compact model and parameter extraction for emerging transistors. *IEEE J. Electron Devices Soc.* **2023**, *11*, 153–160. [CrossRef]
13. Chavez, F.; Tung, C.T.; Kao, M.Y.; Hu, C.; Chen, J.H.; Khandelwal, S. Deep learning-based IV global parameter extraction for BSIM-CMG. *Solid-State Electron.* **2023**, *209*, 108766. [CrossRef]
14. Guo, G.; You, H.; Li, C.; Tang, Z.; Li, O. A physics-informed automatic neural network generation framework for emerging device modeling. *Micromachines* **2023**, *14*, 1150. [CrossRef] [PubMed]
15. Huang, Y.; Li, B.; Wu, Z.; Liu, W. ResNet modeling for 12 nm FinFET devices to enhance DTCO efficiency. *Electronics* **2024**, *13*, 4040. [CrossRef]
16. Chalkiadaki, M.A.; Valla, C.; Poullet, F.; Bucher, M. Why-and how-to integrate Verilog-A compact models in SPICE simulators. *Int. J. Circuit Theory Appl.* **2013**, *41*, 1203–1211. [CrossRef]
17. Aishwarya, K.; Lakshmi, B. Radiation study of TFET and JLFET-based devices and circuits: A comprehensive review on the device structure and sensitivity. *Radiat. Eff. Defects Solids* **2023**, *178*, 229–257. [CrossRef]

18. Brinson, M.E.; Kuznetsov, V. A new approach to compact semiconductor device modelling with Qucs Verilog-A analogue module synthesis. *Int. J. Numer. Model. Electron. Netw. Devices Fields* **2016**, *29*, 1070–1088. [CrossRef]
19. Liu, Z.; Wang, Y.; Vaidya, S.; Ruehle, F.; Halverson, J.; Soljačić, M.; Hou, T.Y.; Tegmark, M. KAN: Kolmogorov-Arnold Networks. *arXiv* **2024**, arXiv:2404.19756.
20. Sulaiman, M.H.; Mustaffa, Z.; Saealal, M.S.; Saari, M.M.; Ahmad, A.Z. Utilizing the Kolmogorov-Arnold Networks for chiller energy consumption prediction in commercial buildings. *J. Build. Eng.* **2024**, *96*, 110475. [CrossRef]
21. Koenig, B.C.; Kim, S.; Deng, S. KAN-ODEs: Kolmogorov-Arnold Network Ordinary Differential Equations for learning dynamical systems and hidden physics. *arXiv* **2024**, arXiv:2407.04192. [CrossRef]
22. Yu, T.; Qiu, J.; Yang, J.; Oseledets, I. Sinc Kolmogorov-Arnold Network and its applications on Physics-Informed Neural Networks. *arXiv* **2024**, arXiv:2410.04096.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.