1)

SCIENCE CHINA
Information Sciences

• SUPPLEMENTARY MATERIAL •

Supplementary Material: Automatic TCAD Calibration Based on Bayesian Optimization for SiC MOSFET

Mengyao Zhao[1†], Yanji He[2†], Zhenya Wang[1], Yong Gu[1], Tianchun Nie[1], Yu Huang[3], Runhua Huang[3], Song Bai[3], Haining Shi[4], Wei Wang[2,5], Jiaxing Wei[1], Sheng Li[1], Yongjia Li[1], Jie Ma[1], Siyang Liu[1], Long Zhang[1*] & Weifeng Sun[1*]

[1]*Southeast University, Nanjing 210096, China;*
[2]*The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400, China;*
[3]*Nanjing Electronic Devices Institute, Nanjing 210042, China;*
[4]*Suzhou Nuclear Power Research Institute, Suzhou 215004, China;*
[5]*The Hong Kong University of Science and Technology, Hong Kong SAR, China*

Abstract

Keywords

Citation Supplementary Material. Sci China Inf Sci, for review

1 Appendix A: Fabrication Process Conditions

The SiC NMOS device is fabricated on an N-epi/N-sub wafer with a channel length of 2 μm, channel width of 60 μm, and gate oxide thickness of 45 nm. The gate operation voltage is designed to be 20 V to meet the driving requirements of SiC power MOSFETs. Fig. 1 shows the fabrication process flow. The P-well is formed by multiple Al ion implantations at energies ranging from 240 keV to 420 keV and doses ranging from 1×10^{12} cm^{-2} to 1.2×10^{13} cm^{-2}, followed by high-temperature annealing at 1650 °C for 30 minutes. The ohmic contacts are formed by Nickel alloy. For TCAD simulation, the ion implantation doping profiles are directly imported from Secondary Ion Mass Spectrometry (SIMS) measurements (Fig. 1(b)) to avoid additional calibration of the doping distribution.

2 Appendix B: TCAD Model Parameters

The gate material is set as N-type poly-Si with a default work function of approximately 4.08 eV, and the fixed positive charge density at the SiC/SiO_2 interface is initialized to 1×10^{12} cm^{-2}, which is a typical value for SiC simulation. Four mobility models are activated in the simulation: the Doping Dependence (DD) model describes carrier mobility dependence on doping concentration; the Enormal Lombardi (EL) and Enormal Interface Charge (EIC) models account for interface mobility degradation due to the vertical electric field and interface positive charges, respectively; and the High Field Saturation (HFS) model characterizes velocity saturation at strong electric fields.

Fourteen physical parameters are selected as calibration variables from these mobility and interface models. `Q_int` is the fixed positive charge density at the SiC/SiO_2 interface. Although its software default in Sentaurus TCAD is 0, the typical value adopted for SiC simulations is 1×10^{12} cm^{-2}. `WF` is the work function of the N-type poly-Si gate. `Ar_mud` is the bulk electron mobility pre-factor in the DD model. `A_Enormal`, `lambda`, and `C_Enormal` are EL model parameters characterizing interface

*† These authors contributed equally to this work.
* Corresponding author (email: longzh@seu.edu.cn, swffrog@seu.edu.cn)

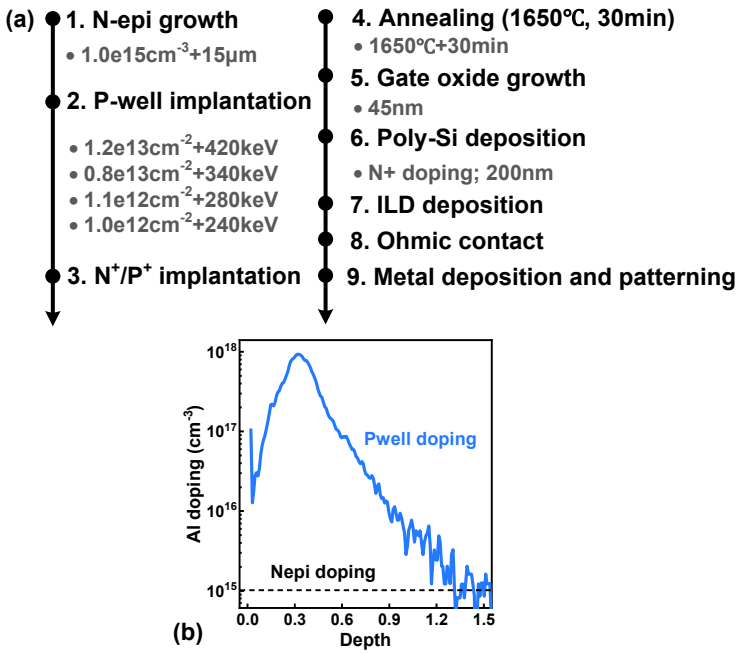


Figure 1 (a) Fabrication process flow of the SiC NMOS device. (b) Secondary Ion Mass Spectrometry (SIMS) doping profiles of P-well.

mobility degradation due to the vertical electric field. `mu_1`, `c_trans`, `Nc_exp1`, `Nc_exp2`, and `c_exp` characterize interface mobility degradation due to interface positive charges in the EIC model. `beta0`, `vsat0`, and `alpha` are velocity saturation parameters in the HFS model. Their physical interpretations, default values, calibration search bounds and final calibrated parameter value are listed in Table 1. The default parameter values provided by Sentaurus TCAD 2024 significantly overestimate the mobility near the SiC/SiO_2 interface, resulting in approximately 100% overestimation of the drain current relative to measurement. For instance, the larger calibrated interface charge density (`Q_int`) directly accounts for the more significant channel mobility degradation and the higher threshold voltage observed in the actual device.

Table 1 Key Interface and Mobility Parameters for I–V Calibration*

Parameter	Model	Meaning	default	value range	calibrated results
Q_int	/	SiO2/Si interface charge density	1e12**	[1e11, 2e12]	1.51e12
WF	/	gate poly workfunction	4.08	[3.9,4.3]	4.17
Ar_mud	DD***	bulk mobility	910	[300,1000]	572
A_Enormal	EL	interface E_F scattering coefficient	2	[1.5, 2.5]	1.86
lambda	EL	interface E_F scattering coefficient	0	[-0.05,0.05]	0.014
C_Enormal	EL	interface E_F scattering coefficient	1e4	[5e3,2e4]	5028
mu1	EIC	interface mobility	40	[10,50]	33
c_trans	EIC	interface charge scattering coefficient	1e18	[1e17,5e18]	1.12e18
Nc_exp1	EIC	interface charge scattering coefficient	1	[0.5,2]	1.87
Nc_exp2	EIC	interface charge scattering coefficient	0.5	[0.25,1]	0.685
c_exp	EIC	interface charge scattering coefficient	1.5	[0.5,3]	0.852
beta0	HFS	velocity saturation coefficient	1.2	[0.6, 1.8]	1.19
vsat0	HFS	carrier saturation velocity	2e7	[1e7, 3e7]	1.61e7
alpha	HFS	velocity saturation coefficient	0	[-0.9,2]	0.77

* provided by Synopsys Sentaurus TCAD 2024;
** 0 by default; but typically 1e12
*** DD, EL, EIC and HFS stand for Doping Dependence, Enormal Lombardi, Enormal Interface Charge and High Field Saturation

Fig. 2 compares the simulated electron mobility distributions before and after calibration. The peak mobility is reduced from over 600 $cm^2/V{\cdot}s$ to approximately 250 $cm^2/V{\cdot}s$, while the surface mobility drops from 70 $cm^2/V{\cdot}s$ to about 25 $cm^2/V{\cdot}s$, consistent with the stronger interface scattering captured by the calibrated model.

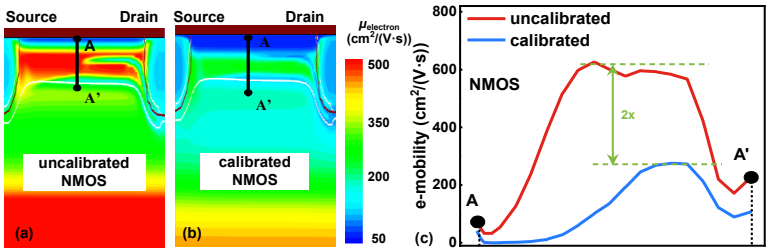


Figure 2 Simulated electron mobility distribution of (a) uncalibrated and (b) calibrated SiC NMOS. (c) Electron mobility extracted along the A–A′ line.

3 Appendix C: Collaborative Workflow of PCA, MOGP, and BO

The proposed calibration framework starts from an initial dataset composed of 100 TCAD simulations generated by Latin hypercube sampling. Each sample contains a set of TCAD model parameters (**TMP**) and the corresponding simulated I–V curves. The purpose of the PCA-MOGP-BO collaborative loop is to construct a sample-efficient surrogate model for the mapping from **TMP** to I–V curves, and then use this surrogate to identify the parameter set that minimizes the discrepancy between simulated and measured characteristics.

To reduce computational complexity and avoid over-fitting, principal component analysis (PCA) is first applied to the simulated I–V curves before surrogate-model construction. The original curves are mean-centered and projected onto a small number of principal components (**PC**), which preserve the dominant curve information and can be transformed back to the full I–V curves through PCA inverse reconstruction. These **PC** are compact statistical representations of the original curves rather than directly measured current points.

Table 2 PCA-MOGP surrogate model for I–V curves

Input:	M I–V curves $\{\mathbf{I}_i\}_{i=1}^{M}$, $\mathbf{I}_i \in \mathbb{R}^N$; TMPs $\{\mathbf{t}_i\}_{i=1}^{M}$, $\mathbf{t}_i \in \mathbb{R}^d$; k
PCA:	
1.	Mean-center data: $\mathbf{X} \in \mathbb{R}^{M\times N}$ with rows $\mathbf{I}_i^{\mathrm{T}}$; compute $\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}/(M-1)$
2.	Eigen-decomposition: select top k eigenvectors $\mathbf{V}_k \in \mathbb{R}^{N\times k}$ of $\mathbf{C}$
3.	Project to PC space: $\mathbf{p}_i = (\mathbf{I}_i - \bar{\mathbf{I}})^{\mathrm{T}}\mathbf{V}_k \in \mathbb{R}^k$; stack as $\mathbf{P} \in \mathbb{R}^{M\times k}$
MOGP:	
4.	Train: $\mathbf{p} = f(\mathbf{t}) + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$; predict $\mathbf{p}^*$, $\mathrm{Var}(\mathbf{p}^*)$ for $\mathbf{t}^*$
5.	Reconstruct: $\hat{\mathbf{I}}^* = (\mathbf{p}^*)^{\mathrm{T}}\mathbf{V}_k^{\mathrm{T}} + \bar{\mathbf{I}}$; propagate uncertainty
Output:	Predicted I–V curve $\hat{\mathbf{I}}^*$ with uncertainty

A multi-objective Gaussian process (MOGP) surrogate model is then trained to learn the mapping between **TMP** and the extracted **PC** based on the initial dataset. For a given **TMP** candidate, the MOGP model predicts the corresponding **PC** together with their uncertainty, and the predicted I–V curves are subsequently reconstructed through PCA inverse transformation. Because the initial dataset is still limited, the first surrogate model is not sufficiently accurate and must be progressively refined through Bayesian optimization (BO).

$$I\text{–}V \text{ curves} \underset{}{\overset{\text{PCA}}{\rightleftarrows}} \mathbf{PC} \overset{\text{MOGP}}{\rightleftarrows} \mathbf{TMP} \tag{1}$$

In the optimization loop, BO uses the current MOGP surrogate and its uncertainty to propose new sampling points that are expected to provide the largest improvement, balancing exploration of uncertain regions and exploitation near promising parameter combinations. After each batch of new TCAD simulations, the dataset is augmented and the MOGP surrogate is retrained, so that the mapping accuracy between **TMP** and **PC**, and thus between **TMP** and the reconstructed I–V curves, is gradually improved. The loop continues until the surrogate accuracy is sufficient and the optimal **TMP** minimizing the mismatch to the measured I–V curves is identified.

$$\mathrm{BO}: \mathrm{MOGP}_{i+1} \xleftarrow{\text{update}} \mathrm{MOGP}_i \tag{2}$$

4 Appendix D: Benchmark of Machine-Learning TCAD Calibration Methods

Table 3 summarizes the calibration accuracy and simulation cost of the proposed workflow in comparison with prior machine-learning-based TCAD calibration methods [1–5]. The proposed method achieves below-2% max NRMSE (the largest NRMSE among the three target I–V curves) for both transfer and output characteristics using only ∼250 simulations, outperforming all listed approaches. This efficiency advantage mainly originates from the integration of Bayesian optimization, which is specifically engineered for costly black-box optimization. The workflow is highly generalized and can be readily adapted to different devices and process platforms.

Table 3 Benchmark of Machine-Learning-Based TCAD Calibration Methods

Year-Ref	Calibrated Device	Main Algorithm	Simulations	Accuracy/Error	Parameters
This work	**SiC CMOS**	**MOGP & BO**	**250**	**NRMSE<2%; R^2>0.99*; MAE<50μA****	14
2023-[1]	Ga_2O_3 Schottky diode	Autoencoder	20,000	R^2=0.937	6
2022-[2]	SiC Vertical MOSFET	ANN	not reported	NRMSE=2.82%	not reported
2024-[3]	GaN HEMT	DNN	not reported	MAE=0.65Ω·mm	not reported
2024-[4]	Si CMOS	Transfer Learning	1000	95%***	20
2025-[5]	N-Type a-Si TFT	Autoencoder&FNN	1000	R^2=0.9993	8
2025-[5]	Poly-Silicon TFT	Autoencoder&FNN	1000	R^2=0.9852	18
2025-[5]	IGZO TFT	Autoencoder&FNN	1000	R^2=0.9859	16

* R^2 means the similarity between simulation and target curves here, instead of MOGP accuracy
** MAE: Mean Absolute Error;
*** calculation method not reported

References

1 M. Eng and H. Y. Wong, "Automatic TCAD model parameter calibration using autoencoder," in 2023 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Kobe, Japan: IEEE, Sept. 2023, pp. 277–280.
2 P. Yang, W. Ming, J. Liang, I. Ludtke, S. Berry, and K. Floros, "Hybrid data-driven modeling methodology for fast and accurate transient simulation of SiC MOSFETs," IEEE Trans. Power Electron., vol. 37, no. 1, pp. 440–451, Jan. 2022.
3 Z. Wang, L. Li, R. C. C. Leon, J. Yang, J. Shi, T. van der Laan, and M. Usman, "Improving semiconductor device modeling for electronic design automation by machine learning techniques," IEEE Trans. Electron Devices, vol. 71, no. 1, pp. 263–271, Jan. 2024.
4 Y. Choi, S. Myung, K. Kim, G. Kang, B. Jung, Y. Jeon, K. Yeom, S. Han, J. Jeong, and D. S. Kim, "Real time TCAD calibration via transfer learning," in 2024 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), San Jose, CA, USA: IEEE, Sept. 2024, pp. 1–4.
5 M. H. Ng et al., "AI-based general solution for automated TCAD parameter calibration: case studies on TFT devices," IEEE Trans. Electron Devices, vol. 72, no. 12, pp. 6757–6764, Dec. 2025.