Novel neural-network-based method for advanced termination design in power semiconductor devices

Nicolò Ripamonti, *Hitachi Energy*, Switzerland
Jhelum Chakravorty, *Hitachi Energy*, Canada
Alexey Sapozhnik, *Hitachi Energy*, Switzerland
Gaurav Gupta, *Hitachi Energy*, Switzerland
Edyta Kuk, *Hitachi Energy*, Poland
Luca De Michielis, *Hitachi Energy*, Switzerland

***Abstract*—This paper addresses the limitations in conventional analytical/TCAD simulation based termination design approaches by presenting an innovative two-phase framework that uses machine learning (ML) to create surrogate models for design optimization. By integrating advanced neural network techniques—including convolutional layers, depth attention mechanisms, and neural operators—the approach enables rapid generation of optimized termination designs across multiple objectives simultaneously. The framework demonstrates high-level of prediction accuracy, $> 97\%$, for breakdown voltage (V_{BD}), leakage current (I_{ces}), and electric field distribution (E) and shows superior performance when compared to previously reported analytical approaches. While validated using ring-termination design for 1.7 kV Si IGBT as a case study, the framework architecture is designed for broad applicability across different termination concepts, voltage classes, materials and target various design optimization objectives. This versatility, combined with significant acceleration of development cycles, represents a powerful new approach for power semiconductor device optimization.**

***Index Terms*—power semiconductors, IGBT, diode, Silicon, edge termination, floating field-limiting rings, machine learning.**

I. Introduction

In high-power semiconductor devices, the edge termination is crucial for preventing premature breakdown by mitigating electric field crowding at curved junctions along device periphery [1]. Among various termination techniques, floating field-limiting rings (FFLR) are widely adopted in silicon power devices due to their integration simplicity, as they can be implemented using the same process steps and masks as active area implants. However, optimizing FFLR designs involves balancing multiple parameters, including ring count, geometry, spacing, and doping profiles—factors that significantly impact electric field distribution (E), breakdown voltage (V_{BD}), and termination length (L).

Traditional FFLR optimization methods face significant trade-offs. Analytical approaches offer computational efficiency but often fail to fully capture the complex device behavior [2]. While TCAD simulations provide high accuracy, they require extensive computation time, making comprehensive design space exploration impractical [3], [4]. These limitations become particularly challenging when considering multiple design objectives simultaneously, such as maximizing V_{BD} while minimizing termination length and optimizing electric field distribution.

Machine learning presents a promising alternative, capable of recognizing physical relationships from simulation data while dramatically accelerating optimization processes [5]. While previous work has primarily focused on V_{BD} optimization, a comprehensive framework enabling simultaneous optimization of multiple performance metrics, including V_{BD}, E, termination length, and I_{ces}, remains an open challenge.

This paper introduces a novel two-phase framework for optimizing termination designs. Through a systematic approach combining offline model training and online optimization, the framework enables rapid exploration of design configurations while maintaining physical accuracy. While demonstrated on a 1.7 kV Silicon IGBT FFLR termination, the framework's architecture is applicable to a wide range of termination concepts, including variation of lateral doping (VLD) for higher voltage classes, junction termination extension (JTE) in SiC devices, and hybrid approaches typically beyond the reach of conventional design methods.

II. Design Methodology

A. Problem Formulation

The optimization of the FFLR termination with boron implantation can be formulated as a constrained multi-objective optimization problem. Let $x \in \mathbb{R}^m$ represent the complete vector of design parameters $x = [s_1, \ldots, s_n, w_1, \ldots, w_n, N, d, N_d]$ (see Figure 1), where w_i and s_i are the ring widths and spacings, respectively, N is the peak doping concentration at the Si interface, d is the doping profile depth, and N_d is the doping concentration at d. In the following analysis, we assume that the doping profile is the same for all rings, and also the number of rings can be optimized. The general optimization problem can be formulated as:

$$\begin{aligned} \min_x \quad & f(x) = \sum\nolimits_i \alpha_i f_i(x) \\ \text{s.t.} \quad & g_j(x) \leq 0, \quad j = 1, \ldots, p \end{aligned} \tag{1}$$

where the potential objective function components $f_i(x)$ and constraint functions $g_j(x)$ include:

- Breakdown voltage $V_{BD} = V(I = I_{th})$ defined at a current threshold (I_{th}).
- Termination length $L = \sum_{i=1}^{n}(s_i + w_i) + FFZ$, where FFZ is the field-free zone.
- Leakage current I_{ces}, defined at the rated voltage.

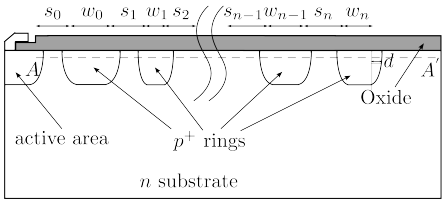

Fig. 1. Cross-sectional view illustrating key parameters of the p-type FFLR doping profile in the termination. The doping concentration decreases both vertically and laterally from each implanted region. Electric field is studied at the cutline AA' at 0.5µm from the top interface.

- Electric field flatness $F = \frac{\overline{|E_i - \overline{E_i}|}}{\overline{E_i}}$, which quantifies the uniformity of field peaks close to the Si interface, where E_i are the local maxima of the electric field.
- E_n refers to the last electric field peak magnitude, closest to the chip edge.

The weights in Eq. (1) can be adjusted to emphasize different design priorities, such as prioritizing V_{BD} maximization or termination length minimization. This flexibility is particularly important as termination designs with similar V_{BD} can exhibit significantly different E, as illustrated in Fig. 2. By incorporating field profile characteristics through objectives like F and E_n, our framework allows for controlling the electric field profile — a critical factor for preventing localized breakdown effects and improving the device reliability.

B. TCAD Simulation Framework

Device simulations were performed using Synopsys Sentaurus TCAD [6]. The V_{BD} was calculated at a temperature of $125°C$ using the external resistor method. The doping dependence and high-field saturation mobility models were included in the physics section among others to properly account for the required accurate analysis of the junction breakdown.

To generate a comprehensive training dataset for the surrogate model, 512 samples were generated by sampling the design space using Sobol quasi-random sequences.

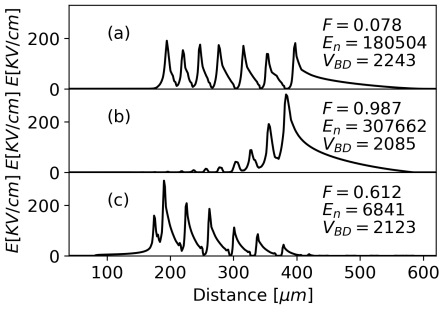

Fig. 2. Example electric field distributions along termination width illustrating: (a) uniform peak distribution (low F value), (b) high last peak (large E_n value), (c) low last peak (small E_n value).

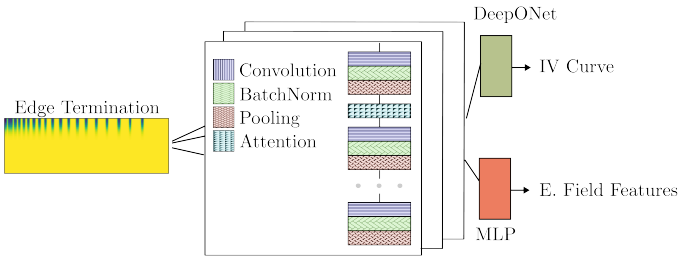

Fig. 3. Overview of the proposed ML architecture for the approximation of termination electrical properties.

C. ML-based surrogate model

The ML-based surrogate model takes a fundamentally different approach from previous methods by processing the 2D doping concentration data directly on a uniform mesh grid as input, focusing specifically on the $30\mu m$-deep region next to the top interface where termination behavior is most critical. This mesh-based approach makes the model inherently agnostic to the type of termination structure (FFLR, VLD, JTE, etc.), unlike methods requiring specific parameters to characterize each termination concept, enabling a truly general solution for termination design optimization. To effectively process this spatial doping data, the neural network architecture employs parallel processing paths: convolutional layers capture local spatial features while an attention mechanism identifies long-range dependencies across the termination region. These processed features then flow into two specialized prediction paths, a multi-layer perceptron for electric field characteristics and a DeepONet [7] architecture for modeling I-V characteristics as continuous functions (Fig. 3). The model's exceptional performance is demonstrated through complementary visualizations in Fig. 4 and 5. While Fig. 4 shows excellent agreement between predicted and simulated I-V characteristics in both sub-breakdown and breakdown regions for the termination designs from the test dataset, Fig. 5 extends this validation across key electrical parameters including V_{BD}, F, and electric field metrics. The strong correlation along the diagonal

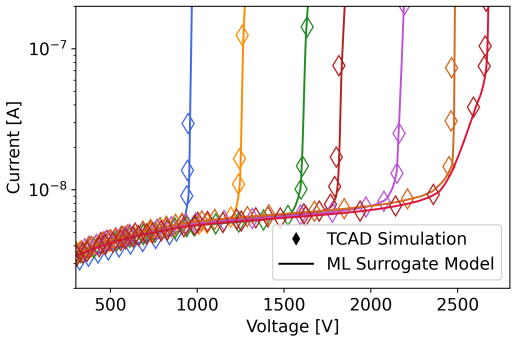

Fig. 4. Comparison between ML surrogate model predictions and TCAD simulation results for various FFLR TE designs at a temperature of $125°C$ for 1.7kV voltage class in the test set, demonstrating the model's ability to accurately capture both leakage current and breakdown behavior.

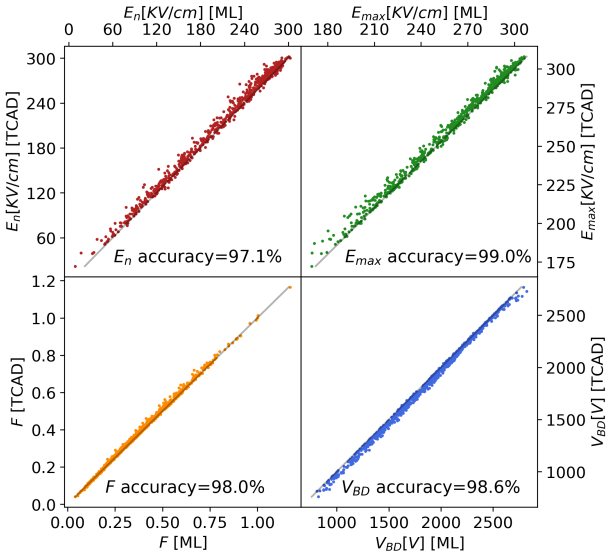

Fig. 5. Comparison between predicted and TCAD-simulated values for V_{BD} and electric field characteristics, demonstrating strong correlation across all metrics.

with only minor, randomly distributed deviations confirms the absence of systematic bias. This specialized mesh-based architecture achieves remarkable consistency, with accuracy reaching 98.6% for V_{BD}, 98% for flatness indicator, 99% for maximum peak electric field, and 97.3% for last peak electric field measurements.

D. Optimization framework

The optimization framework leverages the neural network surrogate model to obtain the I-V curve and the V_{BD} within milliseconds, compared to traditional $10-20$ minute TCAD simulations. Global exploration begins with a genetic algorithm generating thousands of candidates across defined parameter ranges. Each candidate is rapidly evaluated using the surrogate model and ranked according to the fitness function, after which evolutionary operations perform recombination and perturbation of parameters within bounds of high-performing designs. The most promising candidates then undergo local refinement using the Constrained Optimization BY Linear Approximation (COBYLA) method. An adaptive sampling strategy ensures the continuous improvement in critical design regions and enhances the surrogate model accuracy by incorporating the best designs into the training dataset and fine-tuning the model (Fig. 6). This hybrid approach balances exploration and refinement, iterating until either target performance metrics are achieved or the specified maximum iteration number is reached.

III. Results and Discussion

Figure 7 compares the proposed ML optimization approach with previously reported TCAD techniques [3], [4], [8]. While conventional techniques can only optimize geometry (spacings and ring widths) to increase V_{BD}, they do not simultaneously optimize doping profiles or consider additional parameters

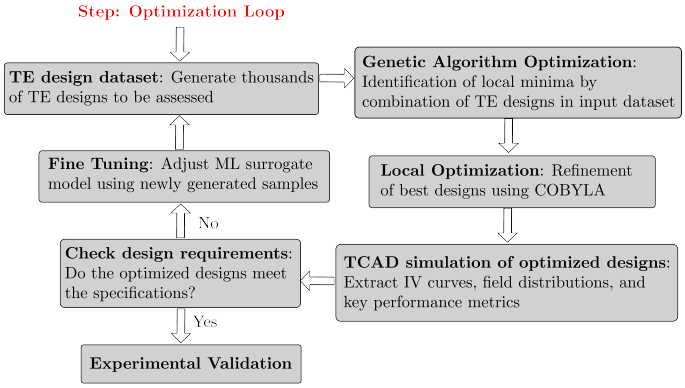

Fig. 6. Overview of the hybrid optimization framework showing interaction between genetic algorithm-based global exploration, COBYLA local refinement, and adaptive sampling strategy for surrogate model improvement. Arrows indicate workflow and data flow between components.

like I_{ces} and E. In contrast, the ML framework demonstrates advantages in two key scenarios:

- Geometry-only optimization (Fig. 7(b)): When focusing solely on maximizing V_{BD}, the ML method outperformed traditional approaches. Even with additional constraints like minimizing termination length or shaping electric field profiles, it provides designs outperforming the previously reported methods.
- Combined geometry and doping optimization (Figure 7(a)): The ML framework achieved its most significant improvements when simultaneously optimizing both ring geometry and doping profiles, surpassing the previously reported optimization methods even under additional design constraints. This highlights the ML algorithm ability to explore a broader design space.

Approach	TE $[\mu m]$	V_{BD} $[V]$	Comp. time $[min]$
ML (Geometry and doping)			≈ 5600
Max. V_{BD}	537.4	2992	≈ 4
ML (Geometry)			≈ 5600
Max. V_{BD}	531.3	2803	≈ 4
Cheng (generating all designs took $\approx$ 250 min)			
Max. V_{BD} (n rings)	516	2163	≈ 250
Max. V_{BD} ($n+2$ rings)	574	2463	≈ 250
Max. V_{BD} ($n+4$ rings)	626	2665	≈ 250
Bauer			
Max. V_{BD} (n rings)	476	2286	≈ 1800
Max. V_{BD} ($n+2$ rings)	531	2504	≈ 2500
Max. V_{BD} ($n+4$ rings)	583	2663	≈ 3300
Baradai			
Max. V_{BD} (n rings)	482	2542	≈ 2800
Max. V_{BD} ($n+2$ rings)	533	2685	≈ 3600
Max. V_{BD} ($n+4$ rings)	582	2805	≈ 4400

TABLE I
Comparison of termination performance metrics between traditional optimization methods and proposed ML-based approach.

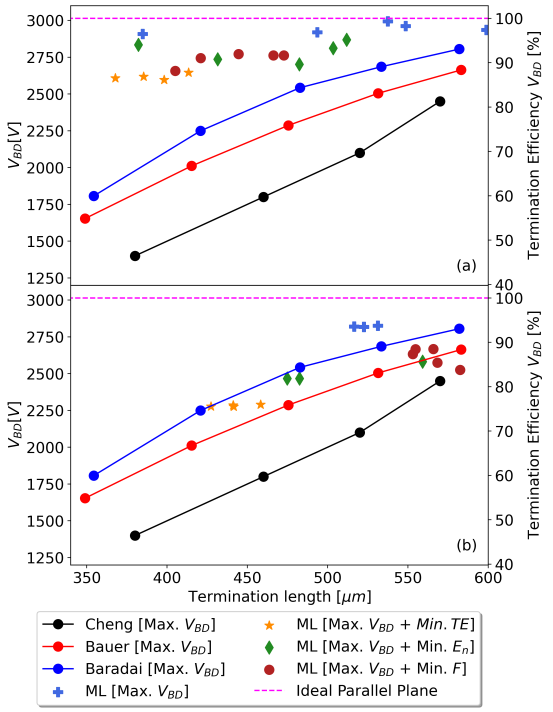

Fig. 7. Optimization performance across different objective weightings: Comparing traditional and ML methods for geometry+doping (a) and geometry-only (b) termination design, with each point representing a unique combination of optimization weights in Eq. 1.

Table I provides the performance comparison between our ML-based framework and other termination design methods. The ML approach with combined geometry and doping optimization achieves the highest efficiency of almost 100% while maintaining a compact termination length of $537.4\mu m$. Even when restricted to geometry-only optimization for direct comparison, it reaches an efficiency of 93% with comparable termination lengths. While other methods can achieve higher breakdown voltages by adding rings, this comes at the cost of larger termination areas. A key advantage of our framework is its computational efficiency after initial training. Though generating the training dataset requires approximately 5600 computer minutes, subsequent optimization runs take only about 4 minutes. This contrasts sharply with traditional methods requiring hundreds to thousands of minutes per optimization attempt. This efficiency makes our framework particularly valuable for design space exploration and optimization under varying constraints, as new scenarios can be rapidly evaluated using the trained model.

Figure 8 shows simulation results examining termination sensitivity to interface charges at the Si/SiO_2 interface. Designs from the training set exhibiting high V_{BD} are compared to the designs optimized specifically for V_{BD} with various electric field constraints. The results demonstrate that termination performance is inherently sensitive to the charge density, as it significantly alters the electric field profile compared to the ideal interface with no charge. However, the degree

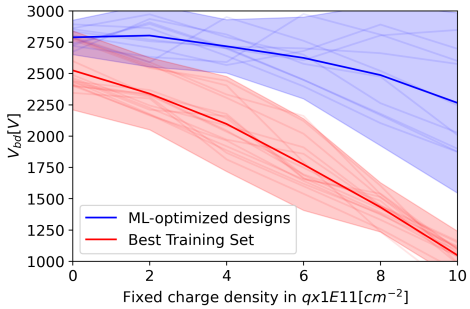

Fig. 8. Breakdown voltage sensitivity analysis for different termination designs as a function of surface fixed charge density.

of sensitivity varies notably between designs. Configurations optimized with specific electric field constraints (E_n and F) show improved robustness against interface charge. This highlights how optimization constraints can help balance maximum blocking performance with the reliability requirements.

IV. Conclusion

This paper introduces a ML framework that advances the optimization of power semiconductor device termination designs by integrating neural network techniques with physics-informed modeling. Our two-phase approach enables multi-objective exploration of termination configurations across semiconductor technologies, demonstrated on a 1.7 kV Silicon IGBT with generalizability to variable lateral doping, junction termination extension, and hybrid termination concepts. Achieving prediction accuracies over 97% and reducing optimization time by multiple orders of magnitude, the framework provides a tool capable of generating efficient termination designs with various constraints. As semiconductor technologies evolve towards more complex architectures, this approach provides a versatile computational tool for systematically expanding device performance optimization.

References

[1] B. Jayant Baliga. *Fundamentals of Power Semiconductor Devices*. 2019.

[2] Jin He, Mansun Chan, Xing Zhang, and Yangyuan Wang. A new analytic method to design multiple floating field limiting rings of power devices. *Solid-State Electronics*, 50:1375–1381, 2006.

[3] Nabil El Baradai, Carmelo Sanfilippo, Rossano Carta, Federica Cappelluti, and Fabrizio Bonani. An improved methodology for the CAD optimization of multiple floating field-limiting ring terminations. *IEEE Transactions on Electron Devices*, 58:266–270, 2011.

[4] Friedhelm Bauer and Umamaheswara Vemulapati. Area efficient floating field ring termination, EP 3 353 814 B1, Sep. 2016.

[5] Jingyu Li, Haoyue Zhao, Hao Yuan, Fengyu Du, Zixi Wang, Xiaoyan Tang, Qingwen Song, and Yuming Zhang. An innovative and efficient approach for field-limiting ring design of power devices based on deep neural networks. *IEEE Electron Device Letters*, 45(3):488–491, 2024.

[6] *Sentaurus Device User Guide*, volume Version V-2024.03 (March 2024). 2024.

[7] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature machine intelligence*, 3(3):218–229, 2021.

[8] Xu Cheng, J.K.O. Sin, J. Shen, Yong jin Huai, Rui zhen Li, Yu Wu, and Bao wei Kang. A general design methodology for the optimal multiple-field-limiting-ring structure using device simulator. *IEEE Transactions on Electron Devices*, 50:2273–2279, 2003.