Proceedings of CONF-FMCE 2026 Symposium: Smart City and Infrastructure Engineering
DOI: 10.54254/2755-2721/2026.CH35092

Defect Engineering and Interface Optimization of Wide Bandgap Semiconductors: A Multiscale Computational Study

JunJie Wang

School of Electronic and Information Engineering, Soochow University, Suzhou, China
1784865643@qq.com

Abstract. Wide bandgap semiconductors, particularly gallium nitride (GaN) and silicon carbide (SiC), are critical materials for next-generation electronic and high-frequency devices. However, material defects and interface states significantly degrade device performance and reliability. This study presents an integrated multiscale computational framework combining density functional theory, molecular dynamics simulations, and machine learning approaches to analyze defect formation and interface optimization in wide-bandgap semiconductor systems. Key findings include: (1) identification of dominant dislocation configurations in GaN heteroepitaxy and their impact on electron mobility; (2) development of a graph neural network model that predicts defect formation energies with an accuracy exceeding 85%; and (3) optimization of a SiC/SiO_2 interface passivation scheme that reduces interface state density by 60%. The proposed framework bridges atomic-scale defects to device-level performance, providing actionable insights for material synthesis and device fabrication.

Keywords: Wide bandgap semiconductor, Defect engineering, Interface optimization, Density functional theory, Machine learning

1. Introduction

The global transition toward carbon neutrality has accelerated the demand for high-efficiency power electronic systems [1]. Wide bandgap semiconductors, primarily gallium nitride (GaN) and silicon carbide (SiC), offer superior breakdown voltage, thermal conductivity, and switching speed compared to conventional silicon, making them key enablers for electric vehicles, renewable energy systems, and 5G infrastructure [1, 2]. Despite these advantages, material defects remain a fundamental bottleneck limiting the performance and commercial viability. In GaN heteroepitaxy, high dislocation densities (typically 10^8–10^{10} cm^{-2}) cause carrier scattering and leakage currents. For SiC power MOSFETs, bulk and interface defects significantly affect device performance [3]. in particular, defects at the SiO_2 interface trap charges, reducing channel mobility to less than 50% of theoretical values [4].

Traditional trial-and-error defect mitigation strategies are time-consuming and costly, creating an urgent need for predictive computational approaches [5]. Recent advances in machine learning (ML) have enabled high-throughput materials screening and device modeling [6], but their application to defect engineering in wide bandgap semiconductors remains limited [7].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.CH35092

This study establishes defect–property relationships by integrating first-principles calculations with machine learning algorithms. The research questions addressed are: (1) Which defect configurations dominate carrier transport degradation? (2) Can machine learning models accurately predict defect energetics without expensive DFT calculations? (3) Which interface passivation strategies optimally reduce trap states? By developing a multiscale computational framework spanning atomic to device scales, this work provides actionable insights for material synthesis and device manufacturing, ultimately supporting the widespread deployment of energy-efficient power electronics.

2. Method

2.1. Density functional theory (DFT) for defect calculations

All atomic-scale calculations were performed using DFT as implemented in the Vienna *Ab initio* Simulation Package [7]. The projector augmented wave method was employed with a plane-wave cutoff energy of 520 eV. Brillouin-zone integrations were performed using Monkhorst-Pack k-point grids with a density of $2\pi \times 0.03$ Å^{-1} [8]. Exchange–correlation effects were treated using the generalized gradient approximation in the Perdew–Burke–Ernzerhof formulation [9]. For defect formation energy calculations, the supercell approach was adopted, with correction schemes applied for charged defects and bandgap underestimation. The formation energy of a defect in charge state q is given by:

$$E^f[D^q] = E_{tot}[D^q] - E_{tot}[bulk] - \sum_i n_i\mu_i + q(E_{VBM} + E_F) + E_{corr} \quad (1)$$

Where $E^f[D^q]$ is the formation energy of defect D in charge state q. $E_{tot}[D^q]$ and $E_{tot}[bulk]$ are the total energies of the supercell containing the defect and the pristine bulk supercell, respectively. n_i and μ_i denote the number and chemical potential of atom species i removed from or added to the system. E_{VBM} is the valence band maximum. E_F is the Fermi level. and E_{corr} is the correction term for charged defects and finite-size effects.

2.2. Crystal graph convolutional neural network model

To accelerate defect property prediction, a crystal graph convolutional neural network (CGCNN) was developed [10]. In this model, crystal structures are represented as undirected graphs, where nodes correspond to atoms and edges represent interatomic bonds. Convolutional layers update node features by aggregating information from neighboring atoms:

$$v_i^{(t+1)} = v_i^{(t)} + \sum_{j\in N(i)} \sigma\left(z_{i,j}^{(t)} \odot g\left(z_{i,j}^{(t)}\right)\right) \quad (2)$$

Where $v_i^{(t)}$ is the feature vector of atom i at the t-th convolutional layer. N(i) denotes the set of neighboring atoms bonded to atom i. zi,j(t) is the concatenated feature vector of atoms i and j and their bond attributes; σ is the sigmoid activation function; g is a neural network layer; and ⊙ denotes element-wise multiplication.

The model was trained on 5,000 DFT-computed defect configurations, achieving a mean absolute error (MAE) of 0.12 eV and $R^2 > 0.85$ on the test set. For migration barrier prediction, the model achieved an MAE of 0.09 eV (Table 1). An active learning strategy combining uncertainty sampling

DOI: 10.54254/2755-2721/2026.CH35092

with diversity-based querying reduced DFT computational cost by 80%, achieving accuracy comparable to random sampling of 5,000 configurations using only 1,000 training samples.

Table 1. CGCNN model performance for defect property prediction

Property	MAE	R^2	Speedup vs. DFT
Formation energy	0.12 eV	0.87	~2000×
Migration barrier	0.09 eV	0.83	~2000×
Charge transition level	0.15 eV	0.81	~2000×

2.3. Multiscale modeling workflow

A multiscale framework bridging atomic to device scales was established. A machine learning potential using the moment tensor potential framework was trained on 10,000 DFT configurations, achieving a force MAE of 0.05 eV/Å. This enabled MD simulations of up to 10^6 atoms, allowing study of dislocation dynamics over nanosecond timescales. Defect parameters (density N_t, capture cross-sections σ_n/σ_p, energy levels E_t) were extracted and incorporated into TCAD models. The Shockley-Read-Hall recombination model gives carrier lifetime:

$$\tau = \frac{1}{\sigma V_{th} N_t} \tag{3}$$

where τ is the Shockley-Read-Hall recombination lifetime; σ is the capture cross-section of the defect; Vth is the thermal velocity of carriers; and Nt is the trap (defect) concentration.

TCAD simulations were performed using Synopsys Sentaurus with device geometries calibrated against experimental structures [11]. This hierarchical approach propagates atomic-scale defect information to macroscopic device performance metrics.

3. GaN heteroepitaxy dislocation screening

3.1. Dislocation formation mechanisms

In GaN heteroepitaxy, lattice mismatch induces strain leading to dislocation nucleation beyond a critical thickness. Threading dislocations (TDs) are the dominant defect type with edge, screw, and mixed character. The fundamental theory of dislocations in crystals was established by Cottrell and later extended to compound semiconductors [12]. DFT calculations revealed that edge-type TD core structures introduce deep levels within the GaN bandgap, acting as non-radiative recombination centers.

3.2. ML prediction of dislocation properties

Based on the CGCNN framework introduced in Section 2.2, a database of 3,000 dislocation configurations was constructed, including various dislocation types, core reconstructions, and impurity decorations (C, O, Si).

The CGCNN model was extended to predict dislocation properties using local atomic environments extracted from 5-Å-radius spheres around dislocation cores. Using this surrogate model, high-throughput screening of 50,000 hypothetical dislocation configurations was performed in 2 hours, a task requiring approximately 2,000 CPU years with DFT alone.

DOI: 10.54254/2755-2721/2026.CH35092

The screening identified that mixed-character dislocations with carbon decoration exhibit the highest thermal stability and lowest electrical activity, as carbon atoms preferentially segregate to tensile regions of dislocation cores, forming stable C-N bonds that passivate dangling bonds.

4. SiC MOS interface: ML-assisted process optimization

4.1. Interface structure and passivation screening

In SiC, dislocation dissociation behaviors have also been shown to influence carrier transport [13]. The SiC/SiO_2 interface contains a 1–2 nm transition region with carbon clusters. DFT calculations revealed that carbon dimers (C-C) and interstitials (C_i) introduce deep-level traps, with unpassivated D_it ≈ 10^{13} cm^{-2} eV^{-1} [4]. Nitridation (NO annealing) replaces carbon clusters with stable N-Si bonds, reducing D_it to 4×10^{12} cm^{-2} eV^{-1} (60% reduction). Al_2O_3 exhibits lower D_it (2×10^{12} cm^{-2} eV^{-1}) but reduced band offset (ΔE_c = 1.5eV vs. 2.7eV for SiO_2) (Table 2). A bilayer SiO_2/Al_2O_3 stack is proposed to balance interface quality and leakage performance.

Table 2. Comparison of SiC interface passivation schemes

Passivation method	D_it (cm^{-2} eV^{-1})	Reduction	Band offset (eV)
Unpassivated	1.0×10^{13}	—	2.7
NO annealing	4.0×10^{12}	60%	2.7
Al_2O_3 (high-κ)	2.0×10^{12}	80%	1.5
Bilayer SiO_2/Al_2O_3	3.0×10^{12}	70%	2.0

4.2. Bayesian optimization and predictive modeling

Based on the multiscale workflow described in Section 2.3, a Bayesian optimization framework identified optimal NO annealing conditions: 1150°C for 90 minutes, achieving predicted D_it = 3.5×10^{12} cm^{-2} eV^{-1}. A random forest regression model (cross-validated R^2 = 0.82) screened 1,000 candidate protocols. Feature importance analysis revealed annealing temperature as the most critical parameter. The top-ranked protocol (1175°C, 120 min NO annealing + 450°C forming gas anneal) achieved predicted D_it = 2.8×10^{12} cm^{-2} eV^{-1} (70% reduction).

5. Discussion: defect-performance correlation

5.1. Cross-scale parameter extraction

A key challenge in multiscale semiconductor modeling is translating atomic-scale defect information into effective device-level parameters suitable for TCAD simulations. In this work, an effective defect density was defined by combining atomic defect concentrations with defect clustering behavior obtained from MD simulations. This correction accounts for the fact that defect clusters often dominate carrier transport degradation rather than isolated point defects, consistent with dislocation dynamics simulations reported in the literature [14].

When applied to GaN HEMTs and SiC MOSFETs, this approach reproduced experimental I-V characteristics with deviations consistently below 15%. The predictive capability was further validated against measured threshold voltage drift in SiC MOSFETs under elevated temperature stress, where the simulation results aligned well with published experimental data [11]. This

DOI: 10.54254/2755-2721/2026.CH35092

agreement suggests that the dominant degradation mechanisms are indeed captured at the atomic scale.

5.2. Inverse design framework

Beyond explaining observed behavior, the proposed framework guides experimental efforts through inverse design—that is, specifying target device performance and deriving actionable defect control metrics. For a device targeting 1.2 kV breakdown voltage, the allowable dislocation density must remain below 10^7 cm^{-2}. To preserve channel mobility above 80% of its theoretical limit, carbon cluster concentrations should be suppressed to less than 10^{18} cm^{-3}. And to keep threshold voltage drift within 0.2 V, the interface state density needs to fall below 5×10^{11} cm^{-2} eV^{-1}. These thresholds provide concrete benchmarks for process development.

Using the machine learning surrogate models, a practical trade-off between interface quality and manufacturing cost was explored. The analysis indicates that industrial production may not require the laboratory-optimized extreme conditions. Instead, a more balanced process window—NO annealing at 1100 to 1150°C for 60 to 90 minutes—yields a D_it around 4×10^{12} cm^{-2} eV^{-1} while cutting energy consumption by 40%. This finding illustrates how computational screening can inform not just scientific understanding but also economically viable manufacturing strategies.

6. Conclusion

In this study, a multiscale computational framework integrating density functional theory, molecular dynamics simulations, and machine learning has been developed to address defect-related challenges in wide bandgap semiconductors, specifically GaN and SiC. The framework combines high-throughput defect screening using a crystal graph convolutional neural network with a cross-scale parameter extraction scheme that bridges atomic-scale defect properties to device-level TCAD simulations.

The core results demonstrate that machine learning can accelerate defect property prediction by approximately 2000-fold while maintaining near-DFT accuracy, with an active learning strategy reducing training costs by 80%. Application of this framework to GaN heteroepitaxy reveals that carbon-decorated mixed-character dislocations exhibit significantly lower electrical activity than other dislocation types, reframing defect engineering from dislocation elimination to controlled carbon incorporation. For SiC MOSFETs, the model identifies optimal NO annealing conditions that reduce interface state density by 60–70%, with temperature identified as the dominant process parameter. Critically, the cross-scale method reproduces experimental I-V characteristics with less than 15% error, confirming that dominant degradation mechanisms are captured at the atomic scale.

The broader significance of this work lies in demonstrating that predictive, computationally guided defect engineering is feasible for wide bandgap semiconductors, offering actionable benchmarks for process development and inverse design. However, several limitations must be acknowledged: the machine learning models degrade when applied to defect configurations outside the training domain, such as multi-species complexes, and the current framework assumes static defect structures, whereas real devices undergo dynamic degradation under operating stress. Future work should extend this methodology to emerging ultrawide bandgap materials including Ga_2O_3 and diamond, and integrate kinetic Monte Carlo simulations to enable long-term, time-dependent reliability forecasting. Ultimately, this multiscale approach provides a generalizable template for accelerating materials discovery and device deployment, supporting the transition toward more energy-efficient electronics.

DOI: 10.54254/2755-2721/2026.CH35092

References

[1] Millan, J., Godignon, P., Perpiñá, X., et al. (2014) A review of wide band-gap semiconductors for power electronics applications. IEEE Transactions on Power Electronics, 29(5): 2155–2163.
[2] Kapolnek, D., Keller, S., Vetury, R., et al. (1997) Anisotropic epitaxial lateral growth in GaN selective area epitaxy. Applied Physics Letters, 71(9)L: 1204–1206.
[3] Brotherton, S. D. (2001) Defect states in silicon carbide. Journal of Applied Physics, 89(12): 7934–7940.
[4] Pensl, G., Beljakowa, S., Frank, T., et al. (2005) Interface states at the SiO_2/SiC interface. Physica Status Solidi A, 202(4): 557–568.
[5] Butler, K. T., Davies, D. W., Cartwright, H., et al. (2018) Machine learning for molecular and materials science. Nature, 559(7715): 547–555.
[6] Marchetti, L. A., Ionescu, A. M., & Esseni, D. (2020) Machine learning for predictive modeling of 7 nm FinFET technology. IEEE Transactions on Electron Devices, 67(4): 1401–1409.
[7] Freysoldt, C., Grabowski, B., Hickel, T., et al. (2014) First-principles calculations for point defects in solids. Reviews of Modern Physics, 86(1): 253–305.
[8] Monkhorst, H. J., & Pack, J. D. (1976) Special points for Brillouin-zone integrations. Physical Review B, 13(12): 5188.
[9] Perdew, J. P., Ruzsinszky, A., Csonka, G. I., et al. (2008) Restoring the density-gradient expansion for exchange in solids and surfaces. Physical Review Letters, 100(13): 136406.
[10] Xie, T., & Grossman, J. C. (2018) Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. Physical Review Letters, 120(14): 145301.
[11] Rozen, J., Ahyi, A., Zhu, X., et al. (2019) 5.6 $m\Omega \cdot cm^2$, 1.2 kV rated, implanted 4H-SiC MOSFETs. IEEE Electron Device Letters, 40(1): 71–74.
[12] Cottrell, A. H. (1953) Dislocations and plastic flow in crystals. Oxford University Press.
[13] Kumagai, Y., Zair, A., Yokoi, T., et al. (2019) First-principles study of the dissociation of screw dislocations in 4H-SiC. Physical Review Materials, 3(9): 093603.
[14] Arsenlis, A. M., Cai, W., Tang, M., et al. (2007) Enabling strain hardening simulations with dislocation dynamics. Modelling and Simulation in Materials Science and Engineering, 15(6): 553.