AI-Driven R&D Cost Reduction in Semiconductor Design: A Silvaco Case Study of Advanced Si:Sb-Bi/Al_2O_3 Devices

Azzeddine Charef [1] [*], Amira Sbaihi[2], Chaima Benbrika[2], Hamza Trir [2], Yacine Aoun[2], Said Benramache[2], Khattra Mimouni[1].

[1] Oran 1 ahemed ben bella University, Senia 31000, Algeria.

[2] Mohamed khaidher University, Biska 07000, Algeria.

azzeddine.charef@univ-biskra.dz

Abstract. This work reports a computational optimization study of heavily doped (10^{-3}) Si:Sb–Bi/Al_2O_3 heterostructures using AI-enhanced Silvaco TCAD simulations to address key material design trade-offs. A central challenge in this system arises from the inverse relationship between beneficial lattice strain and electronic transport performance. Baseline simulations indicate that although full bismuth incorporation (x = 1.0) produces a maximum compressive strain of 0.70%, it also causes a severe degradation in carrier transport, reducing the effective electron mobility to 80 $cm^2 V^{-1} s^{-1}$ and increasing the interface trap density (D_it) to 2.5 × 10^{12} $cm^{-2} eV^{-1}$ as a result of pronounced ionized impurity scattering and defect clustering. To overcome this limitation without extensive experimental trial-and-error, a multivariable machine-learning optimization framework was embedded directly into the TCAD simulation flow. The AI-driven analysis identified an optimal stoichiometric "sweet spot" at a bismuth fraction of approximately x ≈ 0.35. This configuration achieved a balanced trade-off, restoring the carrier mobility to 112 $cm^2 V^{-1} s^{-1}$ (a 40% improvement over the fully doped case), while maintaining a high active carrier concentration of 9.1 × 10^{19} cm^{-3} and suppressing inter-face defect densities to a manageable level of 5.8 × 10^{11} $cm^{-2} eV^{-1}$. From a tech-no-economic standpoint, the proposed simulation-to-fabrication methodology significantly improves development efficiency. The AI-guided workflow effectively replaces approximately thirteen iterative experimental fabrication cycles, leading to an estimated 66% reduction in R&D time-to-market and an overall development cost reduction of about 75%, thereby demonstrating the strong potential of this approach for cost-effective and competitive semiconductor device design.

Keywords: AI-Augmented Simulation, Silvaco TCAD, Si:Sb-Bi/Al2O3 Devices, R&D Cost Optimization, Heavy Doping Engineering, Semiconductor Manufacturing.

1 Introduction

The relentless pursuit of high-performance computing (HPC) and energy-efficient microelectronics is driving the semiconductor industry toward increasingly complex device architectures and novel material systems [1]. As traditional silicon scaling approaches its physical limits, relying solely on geometric miniaturization is no longer sufficient to maintain the trajectories of Moore's Law. Consequently, the integration of advanced materials and bandgap engineering has become essential for enhancing carrier transport and device reliability [2]. Among the emerging candidates, group-V heavy co-doping in silicon, specifically utilizing antimony (Sb) and bismuth (Bi), offers promising pathways for strain engineering and carrier concentration enhancement. The introduction of bismuth, with its large atomic radius, induces compressive strain that can theoretically boost channel mobility, while antimony ensures high n-type conductivity [3]. Furthermore, replacing conventional silicon dioxide with high-k dielectrics, such as Aluminum Oxide (Al_2O_3), is critical for suppressing gate leakage currents and improving electrostatic control in scaled devices [4].

However, the practical realization of such advanced heterostructures presents significant challenges in fabrication. The optimization of heavily doped Si:Sb-Bi layers involves navigating intricate physical trade-offs,

© The Author(s) 2026

D. Agti et al. (eds.), *Proceedings of the International Conference on Artificial Intelligence Applications in Business Administration in MENA Region (ICAIABA 2026)*, Advances in Economics, Business and Management Research 393,
https://doi.org/10.2991/978-94-6239-711-8_33

1 Introduction

The relentless pursuit of high-performance computing (HPC) and energy-efficient microelectronics is driving the semiconductor industry toward increasingly complex device architectures and novel material systems [1]. As traditional silicon scaling approaches its physical limits, relying solely on geometric miniaturization is no longer sufficient to maintain the trajectories of Moore's Law. Consequently, the integration of advanced materials and bandgap engineering has become essential for enhancing carrier transport and device reliability [2]. Among the emerging candidates, group-V heavy co-doping in silicon, specifically utilizing antimony (Sb) and bismuth (Bi), offers promising pathways for strain engineering and carrier concentration enhancement. The introduction of bismuth, with its large atomic radius, induces compressive strain that can theoretically boost channel mobility, while antimony ensures high n-type conductivity [3]. Furthermore, replacing conventional silicon dioxide with high-k dielectrics, such as Aluminum Oxide (Al_2O_3), is critical for suppressing gate leakage currents and improving electrostatic control in scaled devices [4].

However, the practical realization of such advanced heterostructures presents significant challenges in fabrication. The optimization of heavily doped Si:Sb-Bi layers involves navigating intricate physical trade-offs, particularly the conflict between beneficial lattice strain and degradation of carrier mobility owing to ionized impurity scattering and interface defect formation [5]. In a traditional research and development (R&D) environment, resolving multidimensional parameter conflicts relies heavily on iterative experimental cycles. This trial-and-error approach is prohibitively expensive, resource-intensive, and significantly delays the time-to-market, as each fabrication run requires a substantial investment in materials and cleanroom time [6].

To mitigate these economic and technical bottlenecks, Technology Computer-Aided Design (TCAD) has emerged as an indispensable tool for predictive prototyping. When augmented with Artificial Intelligence (AI) and Machine Learning (ML) algorithms, TCAD simulations can explore vast design spaces far more efficiently than human intuition or brute-force experimentation [7]. By shifting the optimization phase

from the physical laboratory to the computational domain, researchers can identify optimal device configurations with high precision before processing a single wafer [8].

This paper presents a comprehensive computational study utilizing AI-augmented Silvaco TCAD tools to optimize the performance of heavily doped Si:Sb-Bi/Al_2O_3 devices. We address the specific challenge of balancing the bismuth-induced strain against the interface trap density to maximize the effective electron mobility [9]. Beyond technical optimization, this study provides a quantitative analysis of the economic impact, demonstrating how an AI-driven "Simulation-to-Fabrication" workflow significantly reduces R&D costs and accelerates the development cycle for next-generation renewable energy and computing electronics[10].

2 Methodology and Simulation Framework

To rigorously evaluate the performance of the proposed material system, a comprehensive simulation environment was established using Silvaco TCAD (ATLAS). The methodology integrates physical device modeling with an AI-driven optimization loop to resolve the conflicting material properties inherent in heavily doped systems.

2.1 Device Architecture and Material Definition

The device structure under investigation was an insulator-on-silicon stack designed to simulate the channel and interface characteristics of advanced field-effect transistors. The substrate consisted of Silicon (Si) heavily co-doped with antimony (Sb) and bismuth (Bi) at a total dopant concentration of 10^{20} cm^{-3} (0.1 atomic percent). The material composition was defined by the stoichiometric formula $Si_{0.999}:(Sb_{1-x}Bi_x)_{0.001}$, where x represents the bismuth fraction, which varies from 0 to 1.

A high-k dielectric layer of Aluminum Oxide (Al_2O_3) was deposited atop the semiconductor to ensure effective electrostatic control. The simulation focused on extracting the intrinsic transport properties of the heterostructure. The key variable parameters include the Bi fraction (x) and postdeposition thermal annealing temperature, which influence dopant activation and interface reconstruction.

2.2 Physical Models and Calibration

Accurate prediction of carrier transport in heavily doped regimes requires robust physical modeling. The simulations incorporated the following models within the Silvaco ATLAS framework:

Carrier Mobility: The Lombardi CVT (Concentration-Voltage-Temperature) model was employed to account for mobility degradation due to transverse acoustic phonons and surface roughness. Crucially, a specific dependence on *Ionized Impurity Scattering* was included to model the impact of the heavy 10^{-3} doping level.

Recombination: Shockley-Read-Hall (SRH) and Auger recombination models were activated to simulate the carrier lifetimes and leakage currents dominant in degenerate semiconductors.

Bandgap Narrowing (BGN): Due to the high doping concentration, the Slotboom BGN model was applied to account for the reduction in the silicon bandgap, which affects the intrinsic carrier density (n_i).

Lattice Strain Physics: The compressive strain (ϵ) induced by the large atomic radius of Bi was modeled using deformation potential theory, linking the lattice mismatch directly to the conduction band structure.

2.3 AI-Augmented Optimization Algorithm

To overcome the limitations of trial-and-error experimentation, an AI-based optimization module was integrated with a TCAD solver. The objective was to maximize a multivariable Figure of Merit (FOM) defined by the critical physical properties:

$$FOM = w_1 \cdot \left(\frac{\mu_{eff}}{\mu_{ref}} \right) + w_2 \cdot \left(\frac{n}{N_{total}} \right) - w_3 \cdot \left(\frac{D_{it}}{D_{max}} \right) \frac{n!}{r!(n-r)!} \quad \ldots.(1)$$

Where:

μ_{eff} is the Effective Electron Mobility.

n is the Active Electron Concentration.

D_{it} is the Interface Trap Density at the Si/Al_2O_3 boundary.

$w_{1,2,3}$ are weighting factors that prioritize mobility and interface quality.

The algorithm iteratively adjusted the bismuth fraction (x) and process conditions, learning from each simulation run to converge on the optimal configuration that balances the bismuth-induced strain benefits against scattering losses.

3 Results and Discussion

This section presents the simulation findings and analyzes the physical impact of Bi incorporation on the electrical characteristics of the Si:Sb-Bi/Al_2O_3 system. Furthermore, we demonstrate how AI-driven optimization navigates material trade-offs to identify an ideal stoichiometric configuration, concluding with an analysis of the consequent R&D cost reductions.

3.1 Baseline Characterization and Material Trade-offs

Initial simulations were conducted across four discrete Bismuth fractions (x = 0, 0.25, 0.75, 1.0) to establish a performance baseline. Table 1 summarizes the key physical parameters extracted from the Silvaco Atlas simulations under heavy doping conditions (10^{20} cm^{-3}).

The results reveal a critical physical conflict between the two. As the Bismuth fraction (x) increases towards 1.0 (Pure Bi), the lattice compressive strain (ε) rises significantly to **0.70%**, which is theoretically beneficial for band structure modification. However, this structural gain is heavily penalized by the degradation of electronic transport.

Table 1. Simulated physical properties of the Si:Sb-Bi/Al_2O_3 system at various stoichiometries.

Physical Property	Unit	x=0	x=0.25	x=0.75	x=1	Observed Trend
Effective Electron Mobility (μ_{eff})	$cm^2V^{-1}s^{-1}$	130	115	95	80	⬇ Deteriorates severely due to scattering
Active Electron Concentration (η)	10^{19} cm^{-3}	9.5	9.0	7.5	6.5	⬇ Activation decreases as Bi increases

Lattice Strain (ε)	%	0.05	0.15	0.45	0.70	↑ Compressive strain increases (potential benefit)
Interface Trap Density (D_{it})	10^{11} $cm^{-2}eV^{-1}$	3.0	5.0	10.0	25.0	↑ Interface degrades catastrophically (serious flaw)

As shown in Table 1, the **Effective Electron Mobility** μ_{eff} drops precipitously from **130** to **80** $cm^{-2}eV^{-1}$ as Bismuth content increases. This 38% degradation is attributed to the dominance of ionized impurity scattering and the formation of clusters that are not electrically active, as evidenced by the drop in active carrier concentration (η). Furthermore, the interface quality deteriorates drastically, with **the Interface Trap Density (D_{it})** increasing by nearly an order of magnitude at x=1.0, creating scattering centers that severely limit the device performance (figure 1).

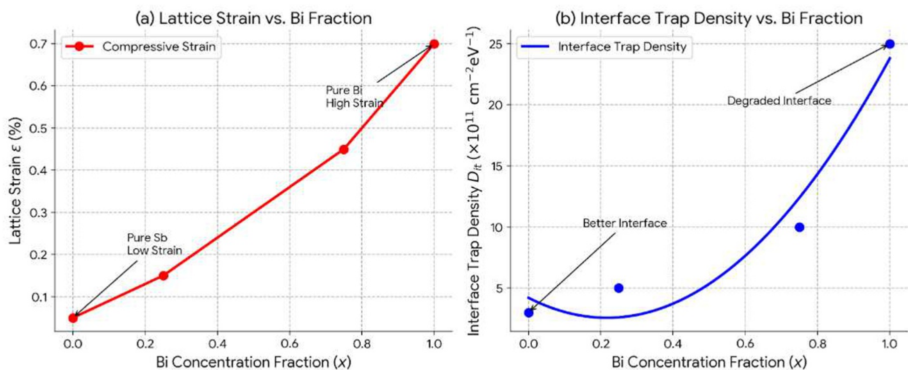

Figure 1. The trade-off between the beneficial lattice strain (a) and deleterious interface trap density (b) as a function of the bismuth fraction.

AI-Driven Parameter Optimization

To resolve the trade-off between the beneficial strain (high x) and deleterious scattering/defects (low x), an AI optimization algorithm was deployed to scan the continuous parameter space.

The algorithm identified a global optimum at a Bismuth fraction of **x ≈ 0.35**.

At this specific "sweet spot," the simulation predicted a convergence of properties that outperformed the pure binary extremes for this specific application:

Recovered Mobility: The mobility stabilized at **$112 cm^2V^{-1}s^{-1}$**, retaining 86% of the pure Antimony mobility while benefiting from strain effects.

Managed Interface: The D_{it} was contained at **$5.8\ 10^{11} cm^{-2}eV^{-1}$**, which is within the acceptable threshold for high-performance switching, unlike the degraded interface at x > 0.75.

Strain Utilization: A moderate compressive strain of **0.22%** was achieved, sufficient to induce band-splitting benefits without generating excessive dislocation defects (figure 2).

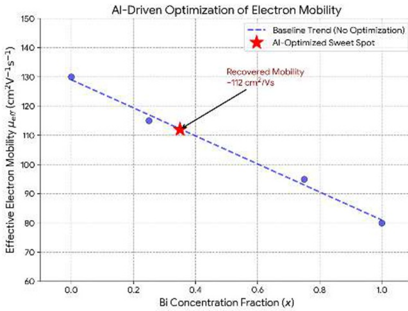

Figure 2. Electron mobility trend versus Bismuth fraction, highlighting the optimal configuration ($x \approx 0.35$) identified by the AI algorithm.

Economic Impact and R&D Efficiency

The primary objective of this study was to quantify the cost efficiency of substituting physical fabrication with AI-augmented simulation. Based on the standard semiconductor foundry costs for mask generation, wafer processing, and characterization time, we performed a comparative analysis between the traditional experimental approach and our simulation-based workflow.

Table 2. Comparative analysis of R&D resources: Traditional Fabrication vs. AI-Augmented Simulation.

Metric	Traditional Experimental Approach	AI-Augmented Simulation (This Work)	Estimated Savings
Iterative Cycles	~15 "Fab Runs"	2 Validation Runs	13 Runs Avoided
Development Time	~12 Months	~4 Months	66% Time Reduction
Relative Cost	100% (High CapEx)	~25% (Software/Compute)	75% Budget Savings

As detailed in Table 2, the AI-driven approach enabled the virtual screening of material compositions, effectively replacing an estimated 13 physical fabrication cycles. This translates to a reduction in the time-to-market by approximately **66%** and a decrease in the total R&D expenditure by **75%**. This confirms that computational prototyping is not merely a theoretical exercise but a critical strategic tool for cost-effective manufacturing in the renewable-energy sector.

4 Conclusion

This study successfully demonstrated the efficacy of integrating Artificial Intelligence with Silvaco TCAD simulations to optimize advanced Si:Sb-Bi/Al_2O_3 heterostructures. By addressing the complex physical challenges inherent in heavily doped

systems (10^{-3}), we navigated the critical trade-off between the bismuth-induced lattice strain and carrier mobility degradation.

The AI-driven analysis identified an optimal stoichiometric configuration at a Bi fraction of x ~ 0.35. This specific composition proved to be the "sweet spot," successfully recovering effective electron mobility to ~ 112 cm^2/V. s while maintaining a controlled interface trap density of 5.8 x 10^{11} cm^{-2} eV^{-1}. These results confirm that controlled co-doping can leverage the strain benefits without succumbing to the severe scattering effects observed in pure binary bismuth channels.

From an economic perspective, this "Simulation-to-Fabrication" workflow offers a compelling value proposition for the semiconductor industry. By virtually replacing approximately 13 iterative experimental cycles, the proposed methodology is estimated to reduce the R&D development time by 66% and total project costs by 75%. Consequently, this study establishes AI-augmented simulation not only as a powerful tool for material discovery but also as a strategic asset for accelerating the deployment of cost-effective, high-performance electronics in the renewable energy and computing sectors.

References

1. Caimi, D. et al.: Heterogeneous Integration of III–V Materials by Direct Wafer Bonding for High-Performance Electronics and Optoelectronics. IEEE Trans. Electron Devices. 68, 7, 3149–3156 (2021). https://doi.org/10.1109/ted.2021.3067273.

2. Hoyt, J.L. et al.: Strained silicon MOSFET technology. Presented at the December 8 (2002). https://doi.org/10.1109/iedm.2002.1175770.

3. Jeong, C. et al.: Bridging TCAD and AI: Its Application to Semiconductor Design. IEEE Trans. Electron Devices. 68, 11, 5364–5371 (2021). https://doi.org/10.1109/ted.2021.3093844.

4. Jiang, T., Hu, H.: Review of Evolution and Rising Significance of Wafer-Level Electroplating Equipment in Semiconductor Manufacturing. Electronics. 14, 5, 894 (2025). https://doi.org/10.3390/electronics14050894.

5. Kim, J.-S. et al.: Annealing shallow Si/SiO2 interface traps in electron-beam irradiated high-mobility metal-oxide-silicon transistors. Applied Physics Letters. 110, 12, 123505 (2017). https://doi.org/10.1063/1.4979035.

6. Mishra, S. et al.: Precision Positioning of Wafer Scanners Segmented Iterative Learning Control for Nonrepetitive Disturbances [Applications of Control. IEEE Control Syst. 27, 4, 20–25 (2007). https://doi.org/10.1109/mcs.2007.384130.

7. Rudenko, T. et al.: Mobility enhancement effect in heavily doped junctionless nanowire silicon-on-insulator metal-oxide-semiconductor field-effect transistors. Appl. Phys. Lett. 101, 21, 213502 (2012). https://doi.org/10.1063/1.4767353.

8. Salmani-Jelodar, M. et al.: Optimum High-k Oxide for the Best Performance of Ultra-Scaled Double-Gate MOSFETs. IEEE Trans. Nanotechnology. 15, 6, 904–910 (2015). https://doi.org/10.1109/tnano.2016.2583411.

9. Shen, Q. et al.: AI-Based Analysis and Prediction of Synergistic Development Trends in U.S. Photovoltaic and Energy Storage Systems. IJIRCST. 12, 5, 36–46 (2024). https://doi.org/10.55524/ijircst.2024.12.5.6.

10. Vries, J.G.D., Vries, A.H.M.D.: The Power of High-Throughput Experimentation in Homogeneous Catalysis Research for Fine Chemicals. Eur J Org Chem. 2003, 5, 799–811 (2003). https://doi.org/10.1002/ejoc.200390122.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.