Received 27 November 2025; revised 29 January 2026; accepted 14 May 2026. Date of publication 18 May 2026; date of current version 29 May 2026.
The review of this article was arranged by Editor P. Pavan.

Digital Object Identifier 10.1109/JEDS.2026.3694652

The Effect of Gate Scaling on Drain Current in Ultra-Scaled Nanosheet and Nanowire FETs: A 3-D Monte Carlo Simulation Study

MURAD G. K. ALABDULLAH[1,2], N. SEOANE[3], A. J. GARCÍA-LOUREIRO[3] (Member, IEEE), AND K. KALNA[1] (Senior Member, IEEE)

[1]NanoDeCo Group, Department of Electronic and Electrical Engineering, Faculty of Science and Engineering, Swansea University, Swansea, SA1 8EN Wales, U.K.
[2]Department of Electronic Techniques, Kirkuk Technical Institute, Northern Technical University, Kirkuk 36001, Iraq
[3]CITIUS, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain

CORRESPONDING AUTHOR: M. G. K. ALABDULLAH (m.g.k.alabdullah@swansea.ac.uk)

ABSTRACT The gate scaling to an ultimate length of 8 nm in Si nanosheet (NS) and nanowire (NW) field-effect transistors (FETs) results in a notable reduction in the drain drive current (I_{DD}), despite conventional scaling theory predicting an increase. Using advanced 3D finite element ensemble Monte Carlo (MC) simulations with Schrödinger equation quantum corrections, we investigate the impact of gate scaling from 22 nm to 8 nm in gate-all-around (GAA) NS and NW FETs. Our results indicate that while I_{DD} initially increases for a gate length scaled from 22 nm to 16 nm as expected, further scaling to sub-16 nm lengths leads to a decline of up to 18% in NS FETs and 20% in NW FETs at 8 nm. This I_{DD} reduction is mainly due to enhanced long-range Coulomb interactions between the source and the drain, inducing fringing electric fields at the source gate-edge in addition to already present fringing electric fields at the drain gate-edge, leading to increased channel back-scattering.

INDEX TERMS Nanosheet FET, nanowire FET, quantum confinement, fringing electric fields, back-scattering effects.

I. INTRODUCTION

Gate-all-around field-effect transistors (GAA FETs) have emerged as a dominant next-generation architecture [1], [2], [3] replacing the existing FinFET architecture for integrated circuits beyond the 3 nm complementary metal-oxide-semiconductor (CMOS) technology node. GAA FETs will be the second major 3D innovation in transistor architecture, following FinFETs, where the industry adopts a fundamentally different architecture. Among the proposed designs, nanosheet (NS) and nanowire (NW) architectures are the leading candidates for fabrication by foundries in the coming years [4]. Compared to FinFETs, NS FET and NW FET architectures offer enhanced gate control and greater suppression of short-channel effects. This advantage stems from the complete gate enclosure around the carrier transport channel, enhancing electrostatic integrity [5], [6], [7]. The NS and NW architectures allow for fabrication into a vertical structure by stacking the channels between the source and the drain in a vertical design [8], leading to a larger on-current while maintaining a lower leakage current, thanks to all-around gate control compared to the existing FinFET architectures [9]. Previous studies on nanosheet (NS) and nanowire (NW) FETs have focused on scaling benefits as improved electrostatic control, low off-currents, and increased drive currents at smaller dimensions [7], [8], [10], [11].

This paper investigates the reasons behind the unexpected reduction in drain drive current (I_{DD}) occurring between gate lengths (L_{G}) of 14 nm and 8 nm. Our approach employs state-of-the-art 3D finite element (FE), quantum corrected ensemble Monte Carlo (MC) simulations incorporating 2D Schrödinger equation (SchE) quantum correction (QC) [12], accurately simulating highly non-equilibrium electron transport in operating Si-based n-type NS and NW FETs. Our study identifies and explains the physical processes responsible for the I_{DD} degradation, when the device gate length is scaled down to the source-to-drain tunneling limit of 8 nm [13].

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

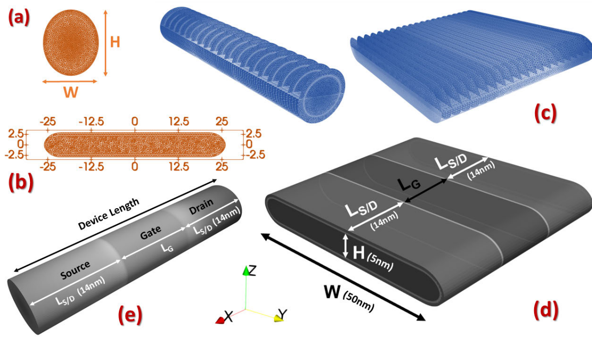

FIGURE 1. Finite-Element (FE) mesh of the (a) NW FET, (b) NS FET. (c) 3D view of the 2D slices across the channel for the NW (left), and the NS (right), used in solving Schrödinger equation. (d) 3D schematic of the NS with dimensions used for the simulations: width (W), height (H), gate length (L_G), and source/drain length ($L_{S/D}$). (e) A 3D view of the NW and the dimensions used for the MC simulations.

II. DEVICE SIMULATION TECHNIQUE

A. 3D MONTE CARLO SIMULATION FRAMEWORK

The nanoscale dimensions of GAA FETs require that the simulation domain, including the device active region, be described using a 3D FE mesh to correctly capture quantum-mechanical confinement in order to assess their performance. Figs. 1(a) and (b) illustrate 2D cuts of the finite-element (FE) tetrahedral meshes, while Figs. 1(d) and (e) show the actual 3D meshed shapes of nanosheet (NS) and nanowire (NW) FETs. This mesh allows for the solution of the Poisson equation and particle movement and scattering in **k**-space and in real-space [12], [14]. The use of FE mesh in particle simulations must address the issue of artificial self-forces arising from the assignment of charges from the particles to the mesh nodes [15]. To accurately capture quantum-mechanical confinement in the transistor channel, the simulation MC toolbox integrates the solutions of the 2D Schrödinger equation [16]. These quantum corrections are solved along 2D slices as illustrated in Fig. 1(c) throughout the device channel.

The ensemble MC engine employs a non-parabolic and anisotropic band structure for silicon, accounting for the Γ, L, and X valleys, and includes all relevant electron scattering mechanisms [12]. These comprise interactions with acoustic and non-polar optical phonons (intra- and inter-valley), ionized impurity scattering via third-body exclusion using the Ridley model [17], and interface roughness (IR) scattering based on Ando's formulation [18]. The IR scattering requires an effective electric field, which is computed dynamically at each electron position. The IR is characterized by a root mean square (RMS) height $\Delta_{\mathrm{RMS}} = 1.5$ nm and a correlation length $\lambda_c = 1.7$ nm [11]. The comparatively high Δ_{RMS} may reflect fabrication-process-induced degradation of interface quality in stacked-GAA integration, as well as limitations of the quantum-mechanical Ando IR scattering model, which relies on 'second-order perturbation' theory and does not account for the proximity of other interfaces in NS and NW device geometries [19]. Reported RMS heights range from $\Delta_{\mathrm{RMS}} \approx 0.3$ nm for well-optimized processes [20] to $\Delta_{\mathrm{RMS}} \approx 0.8$ nm for more aggressive or less selective etches [21]. In practice, Δ_{RMS} can reach approximately 5.0 nm if etch chemistry, processing conditions, or tool control are not well optimized [22]. Since the simulation domain encompasses the full source and drain (S/D) extensions and access regions, series resistance is inherently captured without additional post-processing. Electron transport in heavily doped S/D regions is resolved through electron–ionized impurity scattering. The scattering uses static screening with self-consistent calculations of the Fermi level and electron temperature, derived from local carrier density and energy at each time step, thus ensuring realistic injection into the channel. This method has shown excellent agreement with experimental I_D–V_G data in prior MC studies on SOI FinFETs, nanosheet, and nanowire FETs [11], [12], and the same approach is used here to ensure accurate modelling of access and contact resistances [23].

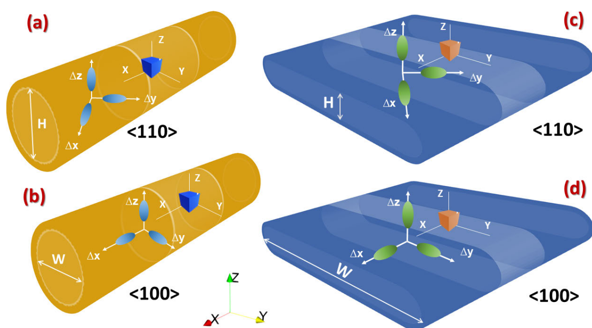

FIGURE 2. A 3D view of NW and NS FETs, highlighting the device geometries and the constant-energy ellipsoids of the conduction band Δ valleys in silicon. Each set of ellipsoids is doubly degenerate. Valley positions are shown for: (a) NW FET with a ⟨110⟩ channel, (b) NW FET with a ⟨100⟩ channel, (c) NS FET with a ⟨110⟩ channel, and (d) NS FET with a ⟨100⟩ channel.

B. IMPACT OF CRYSTAL ORIENTATION ON TRANSPORT PROPERTIES

Figure 2 shows the device geometry and conduction band Δ valley alignment in NW and NS FETs with various channel orientations. For ⟨110⟩ channels, valleys involve both longitudinal (m_l) and transverse effective masses (m_t) in transport and confinement directions. In ⟨100⟩ channels, a m_t from two valleys and m_l from one valley aligns with the transport direction, while the confinement cross-section has m_l and m_t from two valleys and two m_t from one valley. NS FETs, with larger cross-sections, are simulated using the effective mass approximation (EMA) [24]. For NW FETs, the stronger quantum confinement of their cross-section necessitates a full-band tight-binding (TB) model to account for band structure changes and valley splitting [14]. Further insights into the orientation-dependent behavior in NW FETs are provided in Section IV, where the valley contributions to the drain current are analyzed for ⟨100⟩ and ⟨110⟩ orientations (Fig. 8), illustrating the role of anisotropic masses and valley configuration in current variations.

III. NANOSHEET AND NANOWIRE DEVICE STRUCTURES

A. COMPARISON OF THE MC SIMULATIONS WITH EXPERIMENT

We model two experimental devices: a 12 nm gate-length NS FET [8] and a 22 nm gate-length NW FET [10], both

TABLE 1. Dimensions and parameters for the simulated NS and NW FETs in subsection III-B and Section IV. Here, S/D stands for source/drain, and m_0 denotes the free-electron mass.

Dimensions	NS	NW
Channel height (H) [nm]	5	7.17
Channel width (W) [nm]	50	5.7
Perimeter (P) [nm]	110	20.28
S/D region length ($L_{S/D}$) [nm]	14	14
Gate length scaling (L_G) [nm]	22 → 8	
Channel orientation	⟨110⟩, ⟨100⟩	
Gate dielectric: high-κ thickness (EOT) [nm]	5.0 (1.0)	
Max n-type S/D doping [cm^{-3}]	5×10^{19}	
Uniform p-type channel doping [cm^{-3}]	1×10^{15}	
Bulk → TB energy bandgaps (E_G) for Si Δ-valleys in NW FETs [14]		
E_G: Longitudinal Δ_{long} [eV]	1.12 → 1.282	
E_G: Transverse Δ_{trans} [eV]	1.12 → 1.247	
Bulk → TB electron effective masses for Si Δ-valleys in NW FETs [14]		
Longitudinal ($m_l \times m_0$)	0.916 → 0.92	
Transverse ($m_t \times m_0$)	0.19 → 0.233	

featuring a ⟨110⟩ channel orientation. The NS FET has a 12 nm gate length, a 5 nm channel height, a 50 nm channel width, and a total perimeter of 110 nm. The NW FET has an elliptical cross-section due to the fabrication process, with a major diameter of 14.2 nm, a minor diameter of 11.3 nm, and a perimeter of 40.2 nm. The dimensions of both devices are summarized in Table 1. The doping profiles are extracted through reverse engineering of the subthreshold region in the experimental I_D-V_G data, following the methodology in [25]. The I_D-V_G characteristics from our 3D MC simulations are compared to experimental data at both high and low drain biases in Fig. 3. All drain currents are normalized to the device perimeter to ensure a fair comparison. Excellent agreement is observed, validating the accuracy of the simulation framework.

B. CROSS-SECTION SCALING OF NW FET

While simulations of NS FETs employ bulk Si effective masses due to their larger dimensions (a width of 50 nm and a height of 5 nm) [24], making the effective mass approximation (EMA) method [26] an appropriate choice, the scaled NW FET simulations use energy bandgaps and electron effective masses derived from tight-binding (TB) calculations [14], [27], [28]. These energy bandgaps and effective masses are then employed in 3D MC transport simulations and in solutions of the 2D SchE. Such quantum-confinement-affected parameters are necessary because the strong quantum confinement occurs in scaled NW FETs with a cross-sectional area of 7.17 × 5.7 nm^2, influencing the Si band structure [14], [29]. The band structure parameters obtained from TB calculations are summarized in Table 1. The gate dielectric is composed of a high-κ, HfO_2-based dielectric [30] with a physical thickness of 5 nm [31] and an equivalent oxide thickness (EOT) of 1.0 nm [11], consistent with the experimental device [8].

In this study, only the gate length is scaled, while the source/drain lengths ($L_{S/D}$), channel width, and channel height remain unchanged across all gate lengths for both NS and NW FETs. This approach isolates the impact of gate length scaling on the drive drain current (I_{DD}), while ruling out any additional changes in other device dimensions or the doping profile. Although source/drain doping and oxide thickness can be optimized to partially mitigate the reduction in I_{DD} [7], this study intentionally keeps these parameters constant to isolate the intrinsic effect of channel length scaling on device performance.

IV. GATE LENGTH SCALING

A. DRIVE CURRENT

The gate length scaling of GAA FETs at the end of the roadmap [4] is limited by the source-to-drain tunnelling in nanoscale transistors, which restricts the distance between the source and the drain to about 8 nm [13] and the maximum active n-type doping of the source/drain regions achievable by ion implantation to 1×10^{21} cm^{-3} [32]. Consequently, this raises the question of the maximum drive current that GAA FETs can deliver within these constraints.

Figs. 4 and 5 show I_{DD}, defined at $V_G - V_T = 0.7$ V and $V_D = 0.7$ V [4], as a function of the gate length in the NS and NW FETs, respectively. L_G scaling starts with a 22 nm

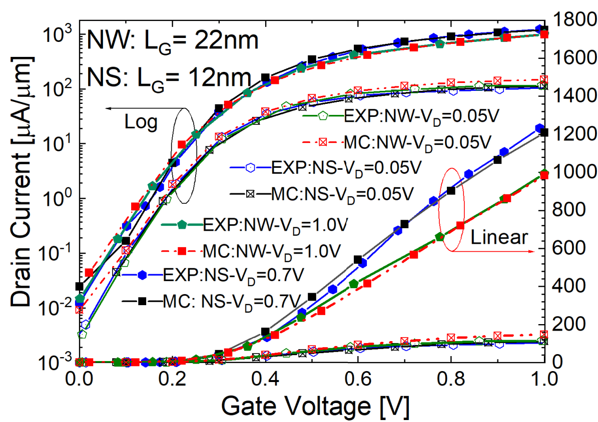

FIGURE 3. I_D-V_G **characteristics of the 12 nm L_G NS FET and the 22 nm L_G NW FET, obtained from 3D MC simulations, at both high and low drain biases, for the ⟨110⟩ channel orientation. The MC results are validated against experimental data (EXP) extracted from Ref. [8] for NS FET and Ref. [10] for NW FET, both fabricated with the ⟨110⟩ orientation. The gate workfunction is assumed to be 4.33 eV for NS FET and 4.51 eV for NW FET.**

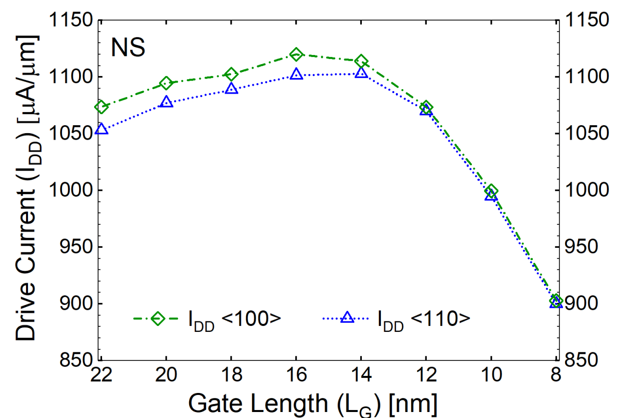

FIGURE 4. The effect of scaling the gate length to the ultimate dimension on the drain-drive current I_{DD} at V_G-V_T = 0.7 V and V_D = 0.7 V in n-type NS FETs for two channel orientations: ⟨100⟩ and ⟨110⟩.

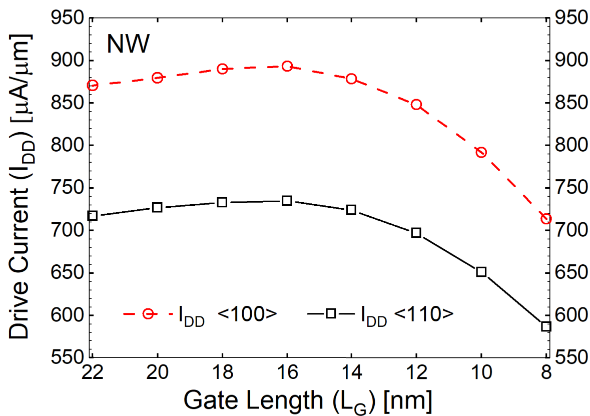

FIGURE 5. The effect of scaling the gate length to the ultimate dimension on the drain-drive current I_{DD} at V_G-V_T = 0.7 V and V_D = 0.7 V in *n*-type NW FETs for two channel orientations: ⟨100⟩ and ⟨110⟩.

and continues till its ultimate length of 8 nm, considering the device dimensions and parameters of the NS FET and NW FET highlighted in Table 1. Note that the 22 nm L_G NW FET shown in Fig. 5 has been scaled down [23] to a smaller cross-section than the experimental 22 nm L_G NW FET described in Subsection III-A. The purpose of using the 22 nm NW FET is to validate the accuracy of the simulation toolbox against the experimental IBM data [10], while the smaller cross-section NW FET is used to align with the IRDS [4] roadmap for future technology. Finally, both crystallographic channel orientations, ⟨110⟩ and ⟨100⟩, are simulated for NS and NW FETs.

As the gate length scales from 22 nm to 16 nm in both NS and NW FETs, I_{DD} increases, consistent with the predictions of scaling theory [33]. However, as the scaling continues to 14 nm L_G and further down to 8 nm L_G, I_{DD} begins to significantly decrease, contrary to the scaling theory as illustrated in Figs. 4 and 5. For the NS FET, the 16 nm L_G with the ⟨110⟩ channel orientation exhibits the highest drive current, reaching 1101 μA/μm. Further scaling of the gate length results in a reduction in I_{DD}. At a gate length of 12 nm, I_{DD} is 1070 μA/μm, which is approximately 3% lower than that of the 16 nm L_G NS FET (see Fig. 4). The 10 nm device exhibits a further reduction to 995 μA/μm, a decline of about 10%, while the 8 nm device shows an 18% drop to 900 μA/μm. The transistors with a ⟨100⟩ channel orientation deliver higher on-currents compared to those with a ⟨110⟩ orientation, thanks to the smaller electron effective mass along the transport direction, as shown in Fig. 2 [7].

On the other hand, the NW FETs with the ⟨110⟩ orientation exhibit an I_{DD} of 696 μA/μm for L_G = 12 nm, 650 μA/μm for L_G = 10 nm, and 587 μA/μm for L_G = 8 nm, as shown in Fig. 5. These values represent reductions of 5%, 11.5%, and 20%, respectively, compared to the 16 nm L_G NW FET, which achieves an I_{DD} of 735 μA/μm. The ⟨100⟩ orientation of the NW FET channel delivers a significantly higher I_{DD}, with similar reduction trends as the gate length scales down. For example, the 12 nm L_G device achieves an I_{DD} of 848 μA/μm, corresponding to a 5% reduction compared to the maximum I_{DD} of 893 μA/μm for the 16 nm L_G device. Further scaling results in I_{DD} values of 791 μA/μm (an 11.4% reduction) for 10 nm L_G and 713 μA/μm (a 20% reduction) for 8 nm L_G. Furthermore, when Δ_{RMS} is reduced from 1.5 nm to 0.3 nm, the behavior of drain-current degradation with gate length scaling remains the same for both NS FETs and NW FETs. The reduced roughness increases the drain drive current, resulting in an increase of about 29% in I_{DD} for NS FETs and 32% for NW FETs at V_D = 0.7 V and $V_G - V_T$ = 0.7 V. This increase is consistent with the reduction in momentum relaxation expected for a smoother interface.

B. IMPACT OF FRINGING FIELDS ON DEVICE CURRENT

The increase in I_{DD} with gate length scaling up to 16 nm is well understood [34], [35]. This enhancement is attributed to the increase in average electron velocity, which results from the stronger fringing electric fields beneath the drain edge of the gate. However, at very short channel lengths starting from 14 nm, Coulomb interaction between the source and drain leads also to an increase in the fringing electric field at the source side of the gate because the source-to-drain distance is so small that the source side starts to be affected by the drain side potential [35].

Figs. 6 and 7 show the electric field distribution in the channel of NS and NW FETs with a ⟨110⟩ channel orientation. As the gate length decreases, the fringing fields at the source and drain gate edges become more prominent for L_G < 16 nm due to the long-range Coulomb interaction between the source and the drain. Therefore, the increase in the fringing electric field at the source side increases electron *back-scattering* [35], resulting in a decrease in I_{DD} because a larger fraction of carriers is scattered back toward the source rather than being transmitted to the drain. Note that NW FETs with the ⟨100⟩ channel orientation exhibit a higher drain current than those with a ⟨110⟩ channel orientation (see Fig. 5). This increase is attributed to the higher average electron velocity and enhanced mobility in ⟨100⟩*n*-type Si FETs across all gate lengths [36].

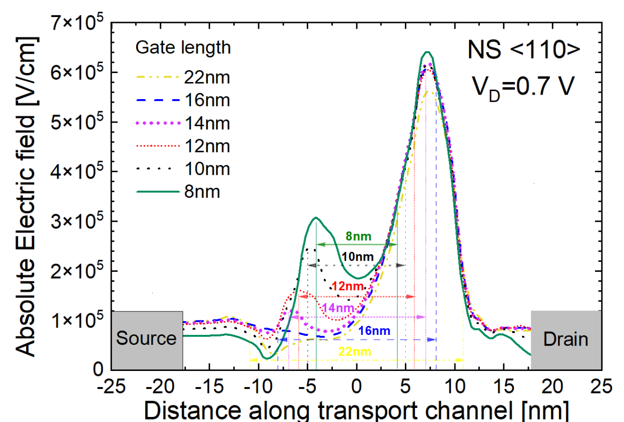

FIGURE 6. Distribution of the absolute electric field, $|\mathcal{E}(x)|$, along the transport direction x in the channel of NS-FETs with varying gate lengths (x < 0 toward the source; x = 0 at mid-channel; x > 0 toward the drain). The fringing electric field increases near the source gate edge as the gate length is scaled for L_G < 16 nm. This increased fringing field is due to drain-induced, long-range Coulomb interaction, contributing to enhanced electron back-scattering.

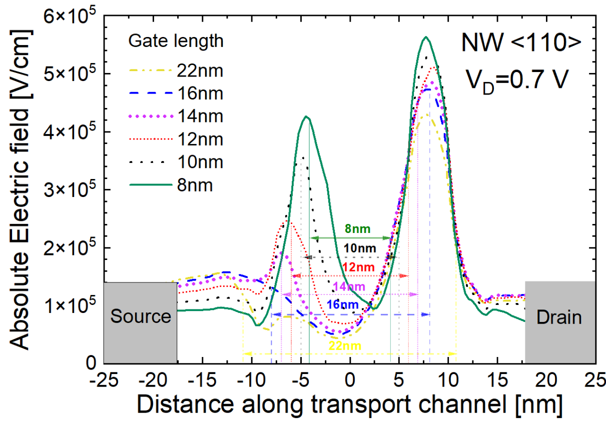

FIGURE 7. Distribution of the absolute electric field, $|\mathcal{E}(x)|$, along the transport direction x in the channel of NW-FETs with varying gate lengths ($x < 0$ toward the source; $x = 0$ at mid-channel; $x > 0$ toward the drain). The fringing electric field increases near the source gate edge as the gate length is scaled for $L_G < 16$ nm. This increased fringing field is due to drain-induced, long-range Coulomb interaction, contributing to enhanced electron back-scattering.

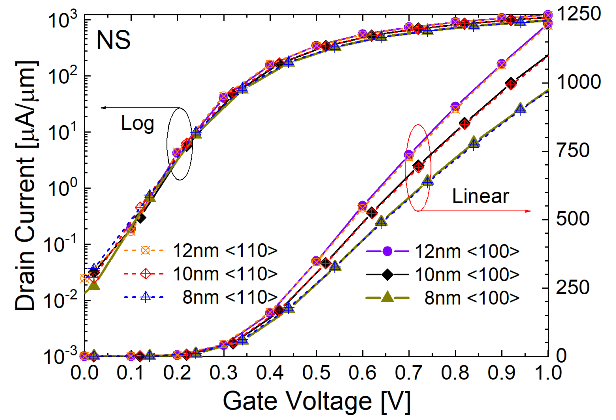

FIGURE 9. Simulated I_D-V_G characteristics of NS FETs with gate lengths of 12 nm, 10 nm, and 8 nm, obtained from 3D MC simulations at a high drain bias of 0.7 V for the ⟨110⟩ and ⟨100⟩ channel orientations.

The drift-velocity contribution from each Δ valley is governed mainly by its transport effective mass m_{trans}. Figure 8 shows that the Δ_z valley provides the largest contribution to I_{DD} for both channel orientations, whereas the Δ_x valley contributes the least. The figure further resolves the contributions of the Δ valleys (Δ_x, Δ_y, and Δ_z) to the total I_{DD} in the 12 nm L_{G} NW FET for the ⟨110⟩ and ⟨100⟩ channel orientations. For the ⟨100⟩ orientation, Δ_y and Δ_z contribute more strongly because their transport mass along the channel is the transverse mass m_t (where m_t is the transverse effective mass), while Δ_x has a heavier transport mass along the channel that is the longitudinal mass m_l (where m_l is the longitudinal effective mass). For the ⟨110⟩ orientation, the Δ_x and Δ_y contributions decrease because the transport direction mixes the longitudinal and transverse mass components of these valleys, which increases the effective m_{trans} and reduces the drift velocity. In contrast, the Δ_z valley retains a comparatively lighter m_{trans} along the channel and therefore remains dominant. Accordingly, the valley-resolved current contributions are $\Delta_z > \Delta_y > \Delta_x$ for the ⟨110⟩ orientation. For the ⟨100⟩ orientation, the Δ_z and Δ_y valleys dominate, and Δ_x contributes the least.

Figs. 9 and 10 show the complete I_{D}–V_{G} characteristics for NS and NW FETs under a high drain bias of 0.7 V, with device geometries summarized in Table 1, for gate lengths of 12 nm, 10 nm, and 8 nm. All extracted currents are normalized to the device perimeter to enable a fair comparison. The drain currents of the NW FETs are smaller than the equivalent drain currents of the NS FETs because of a stronger electron interaction with IR [7], [11]. The effect of the channel orientation is also larger in NW devices because of the stronger quantum confinement of their channel. In contrast, the NS FETs, which exhibit weaker quantum confinement in the channel, show a smaller dependence on the channel orientation. Note that the channel orientation (e.g., ⟨100⟩) is related to a surface orientation arising from different crystal planes at the top/bottom and sidewall interfaces in fabricated structures.

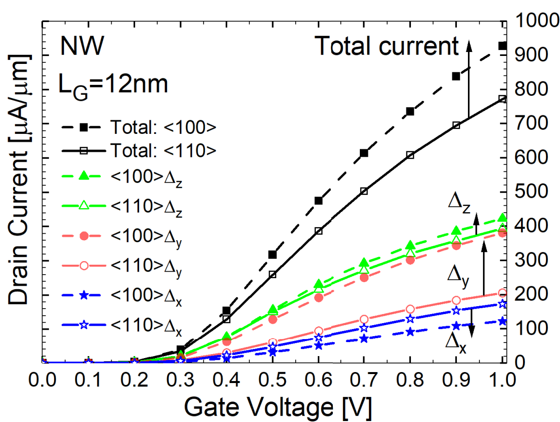

FIGURE 8. The contribution of the Δ-valleys to the drain current in the 12 nm gate length NW FET at a high drain bias of 0.7 V for two channel crystallographic orientations: ⟨110⟩ and ⟨100⟩.

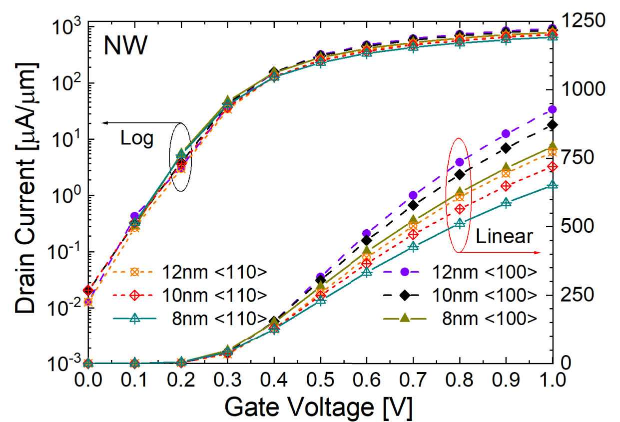

FIGURE 10. Simulated I_D-V_G characteristics of NW FETs with gate lengths of 12 nm, 10 nm, and 8 nm, obtained from 3D MC simulations at a high drain bias of 0.7 V for the ⟨110⟩ and ⟨100⟩ channel orientations.

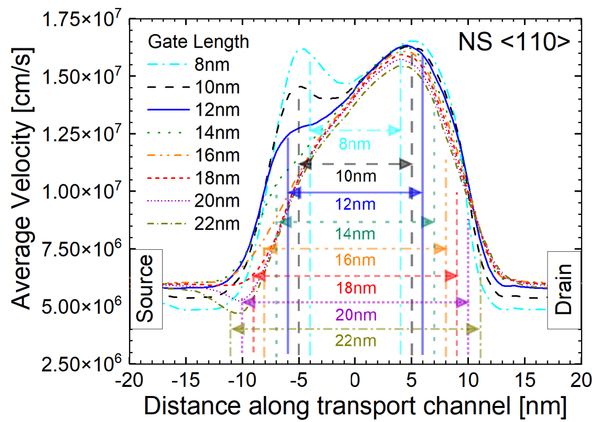

FIGURE 11. The average electron velocity at V_G-V_T = 0.7 V and V_D = 0.7 V for the ⟨110⟩ channel orientation in NS FETs across varying gate lengths (from 22 nm to 8 nm). The location of the gate is indicated by arrows.

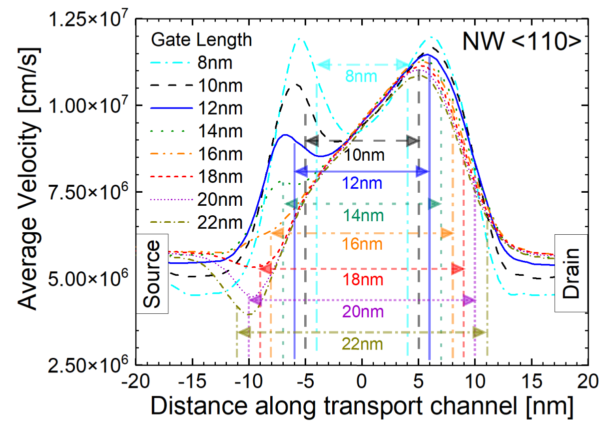

FIGURE 12. The average electron velocity at V_G-V_T = 0.7 V and V_D = 0.7 V for the ⟨110⟩ channel orientation in NW FETs for varying gate lengths (from 22 nm to 8 nm). The location of the gate is indicated by arrows.

C. ELECTRON VELOCITY IN THE DEVICE CHANNEL

The average electron velocity along the channel with the ⟨110⟩ channel orientation in NS and NW FETs is depicted in Figs. 11 and 12, at a high drain bias of 0.7 V and a gate overdrive (V_G-V_T) of 0.7 V. At larger gate lengths of 22 nm and 20 nm, electrons lose velocity near the source gate-edge, because the fringing electric field in that region is small. However, as the gate length scales down, this initial loss of velocity disappears, and the average electron velocity increases. This increase results from the acceleration of electrons by the intensifying fringing electric field near the source gate-edge. In the gate-controlled region, devices with larger gate lengths (22 nm to 14 nm) exhibit a single velocity peak at the drain gate-edge.

As gate lengths decrease to 12 nm, 10 nm, and 8 nm, a second peak emerges at the source gate-edge due to back-scattering of electrons injected from the source. This back-scattering is driven by a significant rise in the fringing electric field at the source gate-edge, as the potential from the drain penetrates toward the source due to long-range Coulomb interactions when the source-to-drain distance reduces [35].

The presence of back-scattering is evident from a large increase in the electron velocity at the source spacer, while the electron kinetic energy (KE) decreases at the same location, as seen in Figs. 13 and 14. The decrease in KE, alongside a large increase in velocity, can only be explained by the electrons traveling back to the source simultaneously when they are injected into the channel. These large average peak electron velocities are a signature of so-called velocity overshoot [37], [38], [39]. The larger velocity in NS FETs is attributed to IR scattering, the strength of which exponentially depends on the distance from the interface. This scattering is, on average, less operative in NS FETs because electrons scatter less with the interface in the nanosheet cross-section compared to the cross-section of a nanowire.

Figs. 13 and 14 present the average electron KE profiles along the channel for scaled NS and NW FETs with a ⟨110⟩ channel orientation at a drain voltage of V_D = 0.7 V and a gate overdrive (V_G-V_T) of 0.7 V. At L_G < 14 nm, drain-induced long-range Coulomb interactions increase the fringing electric field in the source/spacer region. The higher source-side electric field accelerates injected electrons, increasing their KE and thereby increasing the emission and absorption of acoustic and non-polar optical phonons, which in turn increases the probability of back-scattering toward the source [35]. This increased back-scattering in the source/spacer region also reduces the average electron KE at the drain/spacer region [40].

Electrons achieve an average peak velocity of 1.2 × 10^7 cm/s in the 8 nm NW FET, while in the 8 nm NS FET, the peak velocity is 1.62 × 10^7 cm/s. When electrons enter the gate-controlled region, they are accelerated by the rising fringing electric field. Their KE remains very similar, increasing only by a small amount until the 16 nm gate length. When the L_G becomes smaller than 16 nm, the KE increases by larger amounts as the fringing electric field at the source/spacer region increases due to drain-induced long-range Coulomb interactions. The peak KE occurs near the drain/spacer region, where the fringing electric field is largest, similar to the average electron velocities in Figs. 11 and 12. The average KE increases with gate scaling

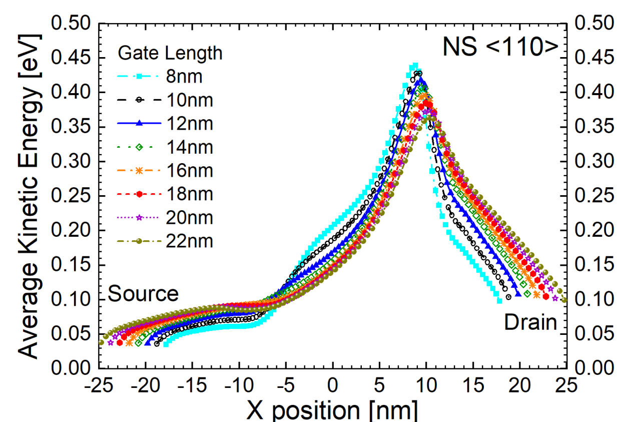

FIGURE 13. Average electron kinetic energy in NS FETs along the channel at V_G-V_T = 0.7 V and V_D = 0.7 V, with the ⟨110⟩ crystallographic orientation, for scaled gate lengths ranging from 22 nm to 8 nm.

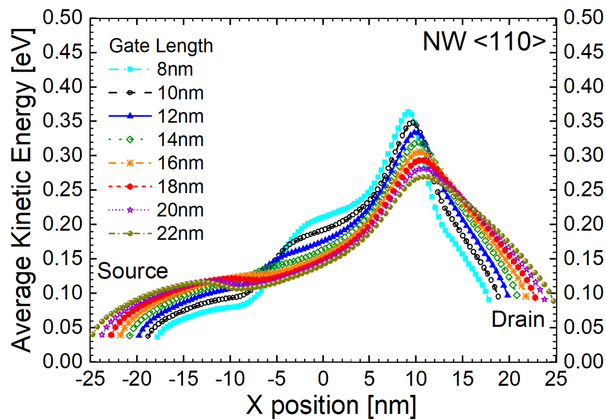

FIGURE 14. Average electron kinetic energy in NW FETs along the channel at V_G-V_T = 0.7 V and V_D = 0.7 V, with the ⟨110⟩ crystallographic orientation, for scaled gate lengths ranging from 22 nm to 8 nm.

to 8 nm. This is due to an increase in quasi-ballistic transport, where electrons retain more of the energy gained from the increasing fringing electric field despite scattering with phonons. Once electrons enter the heavily doped drain region, their average KE decreases due to strong momentum randomization from ionized-impurity scattering, together with phonon emission. These dissipation processes accelerate energy relaxation toward the drain and reduce the average electron KE in the drain region [40]. Contacts in source/drain are treated as a thermal reservoir at $T = 300$ K. Therefore, carriers injected from the contacts follow a thermal energy distribution. These injected electrons contribute to a lower local average kinetic energy near the source/drain.

Finally, although the scaling of gate length to extremely small dimensions enhances device density and integration, it also brings substantial fabrication challenges. Precise control over crucial parameters such as channel dimensions and doping uniformity becomes increasingly difficult at these scales. Furthermore, such extreme miniaturization exacerbates short-channel effects, particularly source-to-drain tunnelling [13], and increases variability due to metal grain variability and line edge roughness [12]. Additionally, maintaining the required drive current in nanoscale NS and NW FETs demands novel techniques, such as advanced dielectric engineering, strain engineering, and optimized doping strategies [7], [41].

These challenges may be mitigated through design modifications and the exploration of new materials or fabrication schemes that are better aligned with future technology requirements [42], [43]. In particular, GAA nanosheet performance can be further enhanced through process innovations such as bottom dielectric-isolation, wider nanosheets, and advanced multi-V_T modules [44], while recent experimental work has shown that CMOS performance can be improved even further by three-dimensional stacking (FET-over-FET) of complementary nanosheet tiers [45]. Resolving these fabrication and integration issues is therefore critical to enable the transition of ultra-scaled transistor technologies from research to commercial manufacturing.

V. CONCLUSION

Our 3D quantum corrected ensemble MC device simulations of *n*-type Si NS and NW FETs reveal that when their gate is scaled from 22 nm to 16 nm, the drive drain current increases as expected according to the scaling theory. However, when the gate length is scaled further to an ultimate length of 8 nm, the drain current starts to decrease by approximately 18% for NS FETs and 20% for NW FETs with the ⟨110⟩ channel orientation at 8 nm gate length. This decrease is driven by increased electron back-scattering in the source spacer region, resulting from the increased fringing electric field at the source gate-edge induced by the long-range Coulomb interaction between electrons in the source and drain regions due to their close proximity [35]. This effect is more pronounced in NW FETs because of their stronger quantum confinement [46]. Further exploration of innovative doping techniques, alternative channel materials, and advanced dielectric engineering could provide pathways to mitigate the identified performance limitations [7].

REFERENCES

[1] J. Jeong et al., "World's first GAA 3nm foundry platform technology (SF^3) with novel multi-bridge-channel-FET (MBCFET) process," in *Proc. IEEE Symp. VLSI Technol. Circuits (VLSI Technol. Circuits)*, Jun. 2023, pp. 1–2, doi: 10.23919/VLSITechnologyandCir57934.2023.10185353.

[2] J. Wang et al., "Challenges and opportunities for stacked transistor: DTCO and device," in *Proc. Symp. VLSI Technol.*, Jun. 2021, pp. 1–2.

[3] P. B. Vyas, C. Zhao, S. Dag, A. Pal, E. M. Bazizi, and B. Ayyagari-Sangamalli, "Next generation gate-all-around device design for continued scaling beyond 2 nm logic," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Sep. 2023, pp. 57–60, doi: 10.23919/SISPAD57422.2023.10319624.

[4] IEEE. (2023). *International Roadmap for Devices and Systems (IRDS): 2023 Edition*. [Online]. Available: https://irds.ieee.org/editions/2023.

[5] N. A. Kumari and P. Prithvi, "Performance evaluation of GAA nanosheet FET with varied geometrical and process parameters," *Silicon*, vol. 14, no. 15, pp. 9821–9831, Oct. 2022, doi: 10.1007/s12633-022-01695-7.

[6] M. Ehteshamuddin, S. A. Loan, and M. Rafat, "Stacked nanosheet based reconfigurable FET," in *Proc. 32nd Int. Conf. Microelectron. (ICM)*, Dec. 2020, pp. 1–4, doi: 10.1109/ICM50269.2020.9331813.

[7] M. G. K. Alabdullah, M. A. Elmessary, D. Nagy, N. Seoane, A. J. García-Loureiro, and K. Kalna, "Scaling challenges of nanosheet field-effect transistors into sub-2 nm nodes," *IEEE J. Electron Devices Soc.*, vol. 12, pp. 479–485, 2024, doi: 10.1109/JEDS.2024.3416200.

[8] N. Loubet et al., "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," in *Proc. Symp. VLSI Technol.*, Jun. 2017, pp. T230–T231, doi: 10.23919/VLSIT.2017.7998183.

[9] S. Barraud et al., "7-levels-stacked nanosheet GAA transistors for high performance computing," in *Proc. IEEE Symp. VLSI Technol.*, Jun. 2020, pp. 1–2, doi: 10.1109/VLSITechnology18217.2020.9265025.

[10] S. Bangsaruntip et al., "Density scaling with gate-all-around silicon nanowire MOSFETs for the 10 nm node and beyond," in *IEDM Tech. Dig.*, Dec. 2013, pp. 20.2.1–20.2.4, doi: 10.1109/IEDM.2013.6724667.

[11] D. Nagy, G. Espineira, G. Indalecio, A. J. Garcia-Loureiro, K. Kalna, and N. Seoane, "Benchmarking of FinFET, nanosheet, and nanowire FET architectures for future technology nodes," *IEEE Access*, vol. 8, pp. 53196–53202, 2020, doi: 10.1109/ACCESS.2020.2980925.

[12] N. Seoane, D. Nagy, G. Indalecio, G. Espiñeira, K. Kalna, and A. García-Loureiro, "A multi-method simulation toolbox to study performance and variability of nanowire FETs," *Materials*, vol. 12, no. 15, pp. 2391–2406, Jul. 2019, doi: 10.3390/ma12152391.

[13] N. Seoane and A. Martinez, "A detailed coupled-mode-space non-equilibrium Green's function simulation study of source-to-drain tunnelling in gate-all-around Si nanowire metal oxide semiconductor field effect transistors," *J. Appl. Phys.*, vol. 114, no. 10, Sep. 2013, Art. no. 104307, doi: 10.1063/1.4820390.

[14] N. Seoane, K. Kalna, X. Cartoixà, and A. García-Loureiro, "Multilevel 3-D device simulation approach applied to deeply scaled nanowire field effect transistors," *IEEE Trans. Electron Devices*, vol. 69, no. 9, pp. 5276–5282, Sep. 2022, doi: 10.1109/TED.2022.3188945.

[15] M. Aldegunde and K. Kalna, "Energy conserving, self-force free Monte Carlo simulations of semiconductor devices on unstructured meshes," *Comput. Phys. Commun.*, vol. 189, pp. 31–36, Apr. 2015, doi: 10.1016/j.cpc.2014.11.020.

[16] J. Lindberg et al., "Quantum corrections based on the 2-D Schrödinger equation for 3-D finite element Monte Carlo simulations of nanoscaled FinFETs," *IEEE Trans. Electron Devices*, vol. 61, no. 2, pp. 423–429, Feb. 2014, doi: 10.1109/TED.2013.2296209.

[17] B. K. Ridley, "Reconciliation of the conwell-weisskopf and brooks-herring formulae for charged-impurity scattering in semiconductors: Third-body interference," *J. Phys. C, Solid State Phys.*, vol. 10, no. 10, pp. 1589–1593, May 1977, doi: 10.1088/0022-3719/10/10/003.

[18] D. K. Ferry, *Semiconductor Transport*. London, U.K.: CRC Press, 2000, doi: 10.4324/9781315267548.

[19] D. Nagy, M. A. Elmessary, M. Aldegunde, J. Lindberg, A. J. Garcia-Loureiro, and K. Kalna, "Multi-subband interface roughness scattering using 3D finite element Monte Carlo with 2D Schödinger equation for simulations of sub-16nm FinFETs," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Sep. 2015, pp. 377–380, doi: 10.1109/SISPAD.2015.7292338.

[20] N. Loubet et al., "A novel dry selective etch of SiGe for the enablement of high performance logic stacked gate-all-around nanosheet devices," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2019, pp. 11.4.1–11.4.4, doi: 10.1109/IEDM19573.2019.8993615.

[21] H. W. Lee, Y. Choi, D. Shin, D.-S. Byeon, and D.-H. Ko, "Epitaxial growth of a silicon capping layer to mitigate roughness after the selective chemical etching of Si1-xGex," *Thin Solid Films*, vol. 707, Aug. 2020, Art. no. 138048, doi: 10.1016/j.tsf.2020.138048.

[22] A. I. Hochbaum et al., "Enhanced thermoelectric performance of rough silicon nanowires," *Nature*, vol. 451, no. 7175, pp. 163–167, Jan. 2008, doi: 10.1038/nature06381.

[23] M. A. Elmessary et al., "Scaling/LER study of Si GAA nanowire FET using 3D finite element Monte Carlo simulations," *Solid-State Electron.*, vol. 128, pp. 17–24, Feb. 2017, doi: 10.1016/j.sse.2016.10.018.

[24] A. Martinez, K. Kalna, P. V. Sushko, A. L. Shluger, J. R. Barker, and A. Asenov, "Impact of body-thickness-dependent band structure on scaling of double-gate MOSFETs: A DFT/NEGF study," *IEEE Trans. Nanotechnol.*, vol. 8, no. 2, pp. 159–166, Mar. 2009, doi: 10.1109/TNANO.2008.917776.

[25] D. Nagy, M. Aldegunde, M. A. Elmessary, A. J. García-Loureiro, N. Seoane, and K. Kalna, "Modelling of nanoscale multi-gate transistors affected by atomistic interface roughness," *J. Phys., Condens. Matter*, vol. 30, no. 14, Apr. 2018, Art. no. 144006, doi: 10.1088/1361-648x/aab10f.

[26] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*. Les Ulis Cedex, France: Les Editions de Physique, 1988.

[27] J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, "Empirical spds* tight-binding calculation for cubic semiconductors: General method and material parameters," *Phys. Rev. B, Condens. Matter*, vol. 57, no. 11, pp. 6493–6507, Mar. 1998, doi: 10.1103/physrevb.57.6493.

[28] V. Sverdlov, G. Karlowatz, S. Dhar, H. Kosina, and S. Selberherr, "Two-band k·p model for the conduction band in silicon: Impact of strain and confinement on band structure and mobility," *Solid-State Electron.*, vol. 52, no. 10, pp. 1563–1568, Oct. 2008, doi: 10.1016/j.sse.2008.06.019.

[29] S. Horiguchi, "Validity of effective mass theory for energy levels in Si quantum wires," *Phys. B, Condens. Matter*, vol. 227, nos. 1–4, pp. 336–338, Sep. 1996, doi: 10.1016/0921-4526(96)00435-8.

[30] M.-J. Zhao et al., "Dielectric properties of hafnium oxide film prepared by HiPIMS at different O_2/Ar ratios and their influences on TFT performance," *J. Science: Adv. Mater. Devices*, vol. 9, no. 2, Jun. 2024, Art. no. 100722, doi: 10.1016/j.jsamd.2024.100722.

[31] N. Breil et al., "Highly-selective superconformal CVD Ti silicide process enabling area-enhanced contacts for next-generation CMOS architectures," in *Proc. Symp. VLSI Technol.*, Jun. 2017, pp. 216–217, doi: 10.23919/VLSIT.2017.7998177.

[32] M. Lee, H.-Y. Ryu, E. Ko, and D.-H. Ko, "Effects of phosphorus doping and postgrowth laser annealing on the structural, electrical, and chemical properties of phosphorus-doped silicon films," *ACS Appl. Electron. Mater.*, vol. 1, no. 3, pp. 288–301, Mar. 2019, doi: 10.1021/acsaelm.8b00057.

[33] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. Rideout, E. Bassous, and A. R. LeBlanc, "Design of ion-implanted MOSFET's with very small physical dimensions," *IEEE J. Solid-State Circuits*, vol. SSC-9, no. 5, pp. 256–268, Oct. 1974, doi: 10.1109/JSSC.1974.1050511.

[34] T. Skotnicki, J. A. Hutchby, T. King, H.-S.-P. Wong, and F. Boeuf, "The end of CMOS scaling," *IEEE Circuits Devices Mag.*, vol. 21, no. 1, pp. 16–26, Jan. 2005, doi: 10.1109/MCD.2005.1388765.

[35] M. V. Fischetti, "Long-range Coulomb interactions in small Si devices. Part II. Effective electron mobility in thin-oxide structures," *J. Appl. Phys.*, vol. 89, no. 2, pp. 1232–1250, Jan. 2001, doi: 10.1063/1.1332424.

[36] K. Akarvardar, C. D. Young, M. O. Baykan, and C. C. Hobbs, "Understanding the FinFET mobility by systematic experiments," in *Toward Quantum FinFET*, W. Han and Z. M. Wang, Eds., Cham, Switzerland: Springer, 2013, pp. 55–79, doi: 10.1007/978-3-319-02021-1_3.

[37] J. Hattori, K. Fukuda, T. Ikegami, and Y. Hayashi, "Effects of electron quantum confinement on velocity overshoot in Si nanosheet gate-all-around FETs," *IEEE Trans. Electron Devices*, vol. 71, no. 6, pp. 3498–3503, Jun. 2024, doi: 10.1109/TED.2024.3389633.

[38] M. Lundstrom, *Fundamentals of Carrier Transport*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2000, doi: 10.1017/CBO9780511618611.

[39] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects," *Phys. Rev. B, Condens. Matter*, vol. 38, no. 14, pp. 9721–9745, Nov. 1988, doi: 10.1103/physrevb.38.9721.

[40] M. V. Fischetti, P. D. Yoder, M. M. Khatami, G. Gaddemane, and M. L. Van de Put, "'Hot electrons in Si lose energy mostly to optical phonons': Truth or myth?" *Appl. Phys. Lett.*, vol. 114, no. 22, Jun. 2019, Art. no. 222104, doi: 10.1063/1.5099914.

[41] F. M. Bufler, R. Ritzenthaler, H. Mertens, G. Eneman, A. Mocuta, and N. Horiguchi, "Performance comparison of *n*-type Si nanowires, nanosheets, and FinFETs by MC device simulation," *IEEE Electron Device Lett.*, vol. 39, no. 11, pp. 1628–1631, Nov. 2018, doi: 10.1109/LED.2018.2868379.

[42] W. Liao et al., "High-performance gate-all-around field effect transistors based on orderly arrays of catalytic Si nanowire channels," *Nano-Micro Lett.*, vol. 17, no. 1, p. 154, Feb. 2025, doi: 10.1007/s40820-025-01674-8.

[43] L. Li et al., "Stacked gate-all-around nanosheet transistors with full-air-spacers for reducing parasitic capacitance to improve device and circuit performance," *Microelectron. J.*, vol. 156, Feb. 2025, Art. no. 106535, doi: 10.1016/j.mejo.2024.106535.

[44] N. Loubet et al., "Enablement of next generation nanosheet technology," in *Proc. Extended Abstract Int. Conf. Solid State Devices Mater.*, Sep. 2021, pp. 63–64, doi: 10.7567/ssdm.2021.a-7-01.

[45] S. Reboh et al., "NanoStack transistor architecture for CMOS 7A node and beyond," in *Proc. Symp. VLSI Technol. Circuits (VLSI Technol. Circuits)*, Jun. 2025, pp. 1–3, doi: 10.23919/VLSITechnologyandCir65189.2025.11074866.

[46] C. S. Soares, A. C. J. Rossetto, M. A. Pavanello, D. Vasileska, and G. I. Wirth, "Impact of a charged oxide trap on *n*-type SOI nanowire FETs: From a quantum-corrected Monte Carlo device simulation perspective," *IEEE Trans. Electron Devices*, vol. 73, no. 2, pp. 739–745, Feb. 2026, doi: 10.1109/TED.2025.3641823.