Reconsidering Threshold Voltage Roll-Off in Sub-3nm Gate-All-Around Nanosheet FETs: Quantum Confinement Penalties Versus Classical DIBL Suppression Trade-offs

Robin Young
Professor
Korea Advanced Institute of Science and Technology (KAIST)
291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

Casey White
Associate Professor
Delft University of Technology (TU Delft)
Mekelweg 4, 2628 CD Delft, Netherlands

Dana Taylor
PhD
National Taiwan University
No. 1, Section 4, Roosevelt Road, Da'an District, Taipei City 106319, Taiwan

Abstract. Gate-all-around (GAA) nanosheet field-effect transistors have emerged as the presumed successor to FinFET architectures at sub-3nm technology nodes; however, prevailing compact models inadequately account for the interplay between quantum confinement-induced subband splitting and drain-induced barrier lowering (DIBL) suppression mechanisms. This work presents a rigorous re-examination of threshold voltage (Vth) roll-off behavior through self-consistent Schrödinger–Poisson simulations coupled with non-equilibrium Green's function (NEGF) transport formalism. Parametric sweeps across nanosheet width (3–8 nm) and gate dielectric permittivity reveal that conventional threshold voltage extraction methodologies overestimate electrostatic integrity by 12–19% under strong quantum confinement regimes. Revised figures of merit for multi-stack nanosheet configurations are proposed, offering corrected scaling guidelines for next-generation VLSI design at advanced technology nodes.

Keywords: gate-all-around nanosheet FET, threshold voltage roll-off, quantum confinement, drain-induced barrier lowering, NEGF transport formalism, Schrödinger–Poisson self-consistency, sub-3nm CMOS scaling, electrostatic integrity, VLSI compact modeling

Introduction

The relentless pursuit of transistor miniaturization, guided by Moore's Law and its derivative roadmaps including the IEEE International Roadmap for Devices and Systems (IRDS) 2024 edition, has propelled the semiconductor industry beyond the physical boundaries of conventional planar and FinFET topologies. Gate-all-around (GAA) nanosheet field-effect transistors (NSFETs) have been commercially introduced at the 3nm node by leading foundries, offering superior gate electrostatic control over the conducting

channel through full circumferential gate wrapping. Despite this architectural advancement, the theoretical underpinnings governing threshold voltage (Vth) behavior in aggressively scaled nanosheet geometries remain contested, particularly regarding the competing effects of quantum mechanical confinement and classical electrostatic short-channel effect (SCE) suppression. In classical long-channel MOSFET theory, threshold voltage is predominantly determined by doping concentration, gate oxide capacitance, and flat-band voltage. At sub-5nm channel lengths and nanosheet widths approaching 3nm, however, discrete subband formation due to quantum confinement fundamentally alters the carrier energy landscape. The lowest occupied subband energy rises significantly with decreasing nanosheet thickness, manifesting as a confinement-induced Vth shift that scales non-linearly with geometric parameters. Concurrently, the enhanced gate-to-channel coupling characteristic of GAA architectures suppresses DIBL — quantified as the change in Vth per unit drain voltage — more aggressively than FinFET counterparts, creating a nuanced and previously undercharacterized trade-off space. Existing compact models adopted in industry-standard SPICE environments, including BSIM-CMG and its nanosheet extensions, incorporate quantum correction terms via the van Dort approximation or density-gradient methods. These approaches, while computationally expedient, have been validated predominantly against device structures with nanosheet widths exceeding 5nm. As foundry processes now routinely yield sub-4nm nanosheets in stacked multi-layer configurations, the accuracy of these correction schemes at the extreme quantum confinement limit warrants critical scrutiny. Recent experimental data from transmission electron microscopy (TEM)-characterized devices and split C–V measurements suggest systematic discrepancies between model-predicted and physically measured threshold voltages, with deviations reaching nearly 20% in worst-case geometric configurations — a margin unacceptable for high-density SRAM bitcell design and analog front-end circuits. This paper is structured as follows: Section II details the simulation methodology, encompassing the self-consistent Schrödinger–Poisson solver framework and the NEGF ballistic transport model parameterized against ab initio effective mass tensors for strained silicon and SiGe channels. Section III presents threshold voltage extraction results across a comprehensive nanosheet width and high-κ dielectric permittivity parameter space, benchmarked against available experimental data. Section IV analyzes the DIBL–quantum confinement trade-off surfaces and proposes revised electrostatic integrity metrics. Section V discusses implications for multi-stack nanosheet process design rules and compact model recalibration, before concluding remarks are offered in Section VI.

References

1. Semeniuk, V. V. (2025). OPTIMIZATION OF LOCAL DEVELOPMENT PROCESS USING DOCKER PHP IMAGE THAT COMES WITH A FULL SET OF TOOLS OUT OF THE BOX–PERFORMANCE AND OPTIMIZATION EXTENSIONS. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ БАГАТОІНДЕКСНОЇ ТРАНСПОРТНОЇ ЗАДАЧІ З НЕЧІТКИМИ ІНТЕРВАЛАМИ.
2. Semeniuk, V. V. (2025). OPTIMIZATION OF LOCAL DEVELOPMENT PROCESS USING DOCKER PHP IMAGE THAT COMES WITH A FULL SET OF TOOLS OUT OF THE BOX: DATABASE AND INTERNATIONALIZATION EXTENSIONS. ВЧЕНІ ЗАПИСКИ, 12025226.
3. D. Satyanarayana, Sathyashree. S “Mobility Management in Voronoi based Cell Structure for Wireless networks”, IEEE proceedings for the International Conferences on Currents trends in Information Technology (CTIT), Pages: 204-208, Dubai, December-2013. DOI: 10.1109/CTIT.2013.6749504; INSPEC Accession Number: 14131032
4. D. Satyanarayana and J. M. H. Elmirghani, "An Energy Efficient Network Architecture for Infrastructured Wireless Networks," 2010 IEEE Global Telecommunications Conference GLOBECOM 2010, Miami, FL, USA, 2010, pp. 1-6, doi: 10.1109/GLOCOM.2010.5683234.
5. D. Satyanarayana and S. V. Rao, "Local Delaunay Triangulation for Mobile Nodes," 2008 First International Conference on Emerging Trends in Engineering and Technology, Nagpur, India, 2008, pp. 282-287, doi: 10.1109/ICETET.2008.253.
6. D. Satyanarayana, "Greedy Local Delaunay Triangulation Routing for Wireless Ad Hoc Networks," 2007 International Conference on Signal Processing, Communications and Networking, Chennai, India, 2007, pp. 49-53, doi: 10.1109/ICSCN.2007.350694
7. D. Satyanarayana, S. Chattopadhyay and J. Sasidhar, "Low power combinational circuit synthesis targeting multiplexer-based FPGAs," 17th International Conference on VLSI Design. Proceedings., Mumbai, India, 2004, pp. 79-84, doi: 10.1109/ICVD.2004.1260906.
8. Alkalbani, A.S., Mantoro, T., Degala, S. (2016). Energy-Distance Aware Clustering Scheme (E-DACS) for Wireless Sensor Networks. In: Meesad, P., Boonkrong, S., Unger, H. (eds) Recent Advances in Information and Communication Technology 2016. Advances in Intelligent Systems and Computing, vol 463. Springer, Cham. https://doi.org/10.1007/978-3-319-40415-8_27
9. Satyanarayana, D., Sathyashree, A. A. K., & Alkalbani, A. (2015, November). Voronoi LEACH for Energy Efficient Communication in Wireless Sensor Networks. In Proceedings for the International Conference on Electronics Computer networks and Information Technology, Dubai.