

Atomistic Quantum Transport for the Ångström Era

Philippe Blaise
Silvaco Inc.
Grenoble, 38000 France
philippe.blaise@silvaco.com

Abstract— This invited article addresses the transition from semi-classical to quantum TCAD, a shift necessary for predictive modeling in the Ångström era. Atomistic simulation approaches able to capture quantum effects will become increasingly important for highly scaled GAA-NS-FETs and their operation over a wide temperature range. We review atomistic quantum transport based on the NEGF formalism, its optimization and integration into a TCAD workflow, and its application to extracting predictive device metrics. The resulting framework provides accurate data, essential for the multi-scale modeling of future C-FET architectures.

Keywords—Ångström Era, GAA-NS-FET, Short Channel Effects, Quantum Confinement, Atomistic, Quantum Transport, NEGF, Low Temperature, FTCO.

I. INTRODUCTION

Recent technological developments in microelectronics encompass an unprecedented combination of manufacturing techniques, paving the way for the assembly of material layers at the nanoscale, with precise control over atomic defects and the achievement of near-perfect surfaces. This aspect alone could be considered the very essence of the Ångström era of the upcoming decade. At the same time, new architectures are emerging, such as the complementary FET (C-FET), which stacks n-channel transistors on top of p-channel transistors, all of which are based on gate-all-around nanosheet field-effect transistor (GAA-NS-FET) technology [1].

The advent of any new technology necessitates a comprehensive design framework, including a dedicated suite of process and device simulation tools. Process simulation is typically performed during the initial device design stage; it subsequently provides essential data regarding materials and device geometry for the device simulation stage. Such a multi-scale, multi-physics platform is offered by leading providers of Technology Computer-Aided Design (TCAD) tools.

In an industrial context, this technology development platform is coupled with a design exploration platform—based on compact modeling and associated capabilities (circuit layout, circuit simulation, parasitic extraction, packaging)—enabling full-circuit simulation. In recent years, a more systematic and synergistic approach has emerged to achieve an optimal Design-Technology Co-Optimization (DTCO) solution. Today, artificial intelligence (AI) adds a layer of abstraction, with digital twins (DT) enabling the implementation of a learning-based Fab-Technology Co-Optimization (FTCOTM) solution [2].

Current TCAD tools generally rely on continuum-based equations and laws. Consequently, to advance a viable simulation solution for GAA-NS-FET devices, priority must be given to adapting the TCAD platform for the Ångström era. Given the fabrication results achieved at the nanoscale, a naive view would be to incorporate as many atomistic and quantum effects as possible into TCAD. In practice, this represents a major transition that needs to be assessed for both process and device simulation.

In nanoscale process simulation, current continuum-based tools are considered highly satisfactory for optimizing

manufacturing conditions, including conformal deposition, etch uniformity, and yield control. However, the precise nature of atomic arrangements within a material, whether crystalline or disordered, decisively influences nanoscale device performance and reliability. Although an atomic-scale simulation solution would be extremely valuable, fully reproducing a manufacturing process using current quantum chemistry tools remains virtually impossible. This is due to the inherent complexity arising from the vast number of degrees of freedom, the extreme timescales, and the sequence of steps involving various physicochemical phenomena. For example, a standard etching step followed by atomic layer deposition (ALD) and then thermal annealing involves many time scales, phases, and reactants across several reactors. Hopefully, with the recent developments leveraging AI and neural networks to accelerate simulations, nanoscale process simulation can benefit, to some extent, from quantum chemistry software and associated methods, such as *ab initio* molecular dynamics (AIMD); this approach yields valuable insights into material defects and interfaces at the atomic scale, provided the process is tightly coupled with precise, specific physical characterization measurements that supply the simulator with relevant initial conditions [3].

In nanoscale device simulation, significant progress has been made and consolidated in the development of atomistic-scale charge-carrier transport solutions for GAA-NS-FET transistors. The typical cross-section of a nanosheet, of a few tens of square nanometers, induces quantum confinement resulting in a sub-band structure and a density of states that depend heavily on the nanosheet's exact dimensions, geometry, material, crystallographic orientation, and strain profile. Along the transport direction (orthogonal to the confinement directions), a self-consistent current solver combining the quantum and electrostatic confinements characteristic of the gate-all-around architecture generally relies on two very different approaches:

1. Semiclassical transport within a sub-band Boltzmann transport equation (SBTE) framework.
2. Atomistic quantum transport within the nonequilibrium Green's function (NEGF) framework.

Theoretically, SBTE equations can be derived from NEGF equations under the following assumptions: for a given sub-band along the transport direction, charge carriers are described by a classical average density of particles, and only weak interactions with a slowly varying potential, modeled as first-order perturbations (e.g., electron-phonon interactions or roughness scattering), are considered [4]. In this regard, SBTE equations are easier to solve than those based on the NEGF formalism, and the SBTE approach remains relatively efficient for predicting the on-state current of a long-channel transistor with moderate confinement. Furthermore, the SBTE method remains appealing due to its ease of implementation within a multi-scale framework, particularly when the aim is to remain closely aligned with conventional TCAD tools.

The transition from SBTE to NEGF in the Ångström era is analogous to the transition from geometric optics to full-

wave electromagnetics when optical structures become comparable to the wavelength. In both cases, the eigenmodes remain valid, but the propagation of waves—not merely their energies or occupations—must be described explicitly. Otherwise, several fundamental quantum transport effects would be completely neglected: coherent transport and quantum tunneling, interference of transmitted and reflected waves onto abrupt potential variations, energy-level broadening due to contacts, and resonant tunneling involving confined states.

Regarding transport along the channel, it is crucial to consider that, when the physical dimensions of transistor components, such as the channel and spacers, become comparable to the wavelength of the charge carriers (of the order of 10 nm), quantum effects can no longer be ignored. In the context of GAA-NS-FET device engineering, this atomistic quantum transport NEGF solution fully meets the optimization requirements for short-channel effects (SCE).

Furthermore, the NEGF formalism relies on spectral functions, which describe the available quantum states, and on correlation functions, which describe their non-equilibrium occupancy. In this regard, the NEGF formalism accounts for carrier interactions with the environment through specific self-energy terms. Standard self-energies are established for contacts, various types of phonons, and impurities, while other, more complex ones are defined for interactions such as electron-electron scattering [5]. Consequently, in the following sections, we will focus exclusively on this atomistic quantum transport approach based on the NEGF formalism.

Last but not least, the coupling of process and device simulations at the nanoscale remains a challenge. The direct and straightforward use of materials simulation and quantum chemistry software for device simulation is made possible by accurate band-structure calculations, provided that density functional theory (DFT) and many-body corrections beyond DFT are rigorously applied. This makes it possible, for example, to feed a device simulation with a precise sub-band structure corresponding to a given cross-section of the device, as well as with data regarding band offsets. Furthermore, electron-phonon coupling matrix elements can be determined using density functional perturbation theory (DFPT). Here too, device simulation exploits DFPT-derived data to calibrate electron-phonon self-energies [6]. Looking to the future,

thanks to rapid improvements in nanoscale process simulation, we can envision even more precise device simulation by accounting for all types of process-related atomic defects, from point to extended defects.

To summarize the atomistic quantum methodology, Fig. 1 illustrates the interplay between quantum transport based on the NEGF formalism (for device simulation), many-body perturbation theory (for physical characterization and spectroscopy), quantum chemistry (for process simulation), and conventional semi-classical TCAD simulation. Two metrics must be considered: the number of atoms and the coherence length; in this regard, NEGF quantum transport is compatible with the physical characteristics of GAA-NS-FET transistors.

II. EXAMPLES

This section presents significant results from NEGF atomistic simulations performed using the VICTORY ATOMISTIC quantum transport tool [7], which is derived from NEMO5 [8]. It should be noted that the DFT electronic band structures were obtained using the SIESTA quantum chemistry tool [9].

A. Nanosheet Band Structure

One of the most effective methods for describing the electronic band structure of a material is the Slater-Koster (SK) tight-binding (TB) model. An SK-TB parameter set is based on precise experimental measurements of the bulk material and then subsequently calibrated to DFT data, including many-body corrections. The transferability of an $sp^3d^5s^*$ -type TB model is evaluated for a silicon nanostructure oriented along the $\langle 100 \rangle$ axis, with a $3 \times 3 \text{ nm}^2$ square cross-section and strong confinement, as shown in Fig. 2. Complete passivation of the silicon dangling bonds with hydrogen is performed. The agreement between the DFT results and those obtained using the SK-TB method is very satisfactory for both the conduction and valence bands; this confirms the validity of the quantitative description of the sub-bands for subsequent use in atomistic quantum transport using the NEGF method.

Surface passivation is a critical issue in microelectronics technology, as it determines device performance and reliability. AIMD and position relaxation enable the modeling of a Si-SiO₂ interface, as illustrated in Fig. 3 for the previously

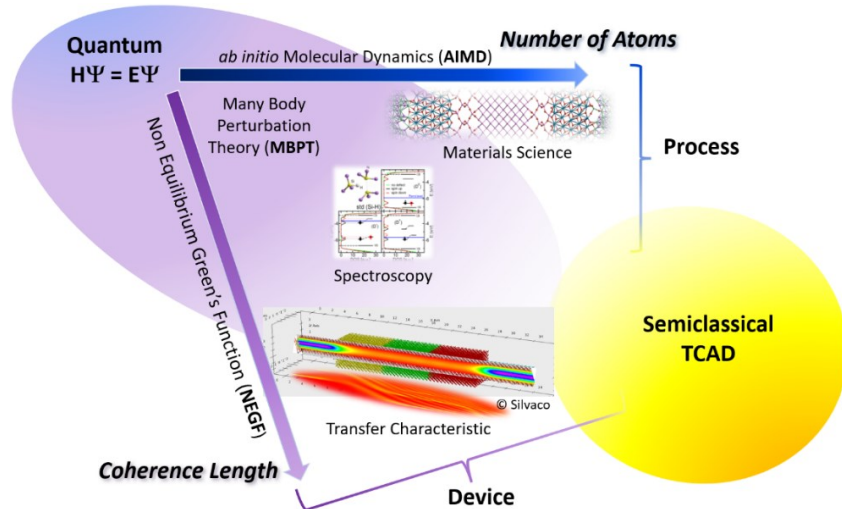


Fig. 1. Atomistic quantum simulation tools: ab initio molecular dynamics, many-body perturbation theory, NEGF in a multi-physics and multiscale environment connecting quantum mechanics with semiclassical TCAD.

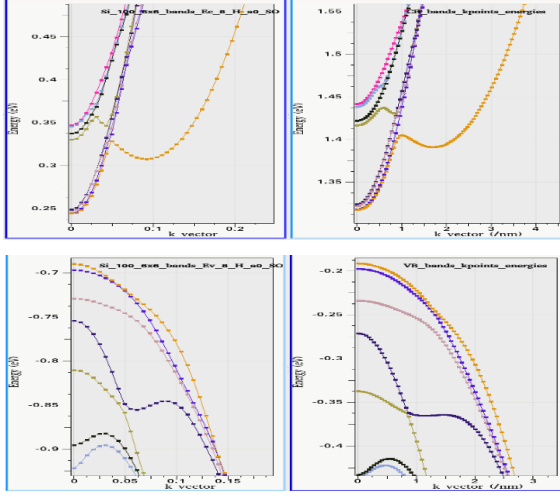


Fig. 2. Band structure for a $3 \times 3 \text{ nm}^2$ silicon-nanowire cross-section oriented $\langle 001 \rangle$ with hydrogen passivation. Conduction subbands are shown at the top, and valence subbands at the bottom. The left column presents results obtained using the DFT method. The right column shows those derived from SK-TB. Band alignment is performed at the band extrema. It should be noted that the 1.5 eV SK-TB band gap is direct at the Γ point, as confirmed by many-body calculations.

studied structure. The silicon surface is fully passivated by Si-O bonds, with the silicon surrounded by a cristobalite-like SiO_2 layer. Interface relaxation induces significant stress on the first atomic layer of silicon (at edges and corners), substantially altering the sub-band structure, see Fig. 3. This highlights the need for more accurate process simulations capable of generating realistic interface structures, followed by a DFT-TB parametrization procedure suited to quantum transport – an area of active research.

B. GAA-NS-FET at Low Temperature

In GAA-NS-FET technology, fringe-field effects at the gate edges are controlled through the introduction of low- k dielectric spacers. This directly leads to the potential emergence, at low temperatures and under low source-drain bias, of localized quantum states close to the band edges in the channel region.

Low-temperature simulations of GAA-NS-FET devices, based on the atomistic NEGF formalism, are performed at a few Kelvin or below, using an adaptive energy grid that provides an energy resolution far better than 1 meV. Furthermore, our NEGF solver calculates—with extreme precision—atomistic wavefunctions, energy- and space-resolved density of states, tunnel coupling between channel states and contacts, and, crucially, confined states.

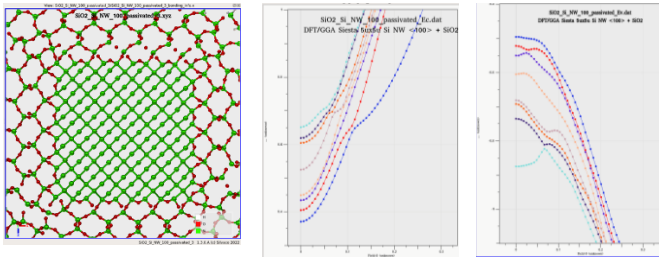


Fig. 3. $3 \times 3 \text{ nm}^2$ cross-section of a $\langle 001 \rangle$ -oriented silicon nanowire passivated with SiO_2 . The relaxed atomic structure is shown on the left. The DFT subbands are illustrated in the center for the conduction band and on the right for the valence band. They deviate significantly from the subband structures shown in Figure 2, primarily due to the deformation of the outermost silicon layer induced by oxidation of the nanowire.

Confined levels beneath the gate of a $2 \times 2 \text{ nm}^2$ NW-FET, featuring 7 nm long spacers and a 5 nm long gate, are illustrated in Fig. 4. Their exact position evolves as a function of the voltage applied to the gate. As parameters relevant to quantum technology, a lever arm of 27 meV/V and an effective gate-coupling capacitance of 2 aF are derived from the plots in Fig. 5.

Such a simulation capability, enabling the detection and characterization of localized states relevant to both cryo-CMOS and quantum-device applications, is unique. If the tunnel coupling to the source and drain electrodes is further reduced through device geometry or gate biasing, the system may enter a regime where the charging energy of a localized state exceeds both the contact-induced level broadening and the thermal energy $k_B T$. The localized state then no longer behaves as a resonant transport state but as a single-electron island operating in the Coulomb blockade regime. Capturing this transition requires extending the conventional single-particle NEGF formalism to include electron-electron interaction self-energies that account for the charging energy U . Transport is then governed by occupation-dependent addition energies, naturally reproducing the current suppression characteristic of Coulomb blockade.

C. GAA-NS-FET, NEGF Optimization, Output Parameters, Short-Channel Effects, Mobility, Injection Velocity

The computational cost of atomistic NEGF simulations is primarily determined by the matrix order N , which scales with the product of the number of atoms and the SK-TB basis size. For a full GAA-NS-FET, N can reach 10^5 degrees of freedom. The recursive Green's function (RGF) algorithm exploits the block-tridiagonal structure of the discretized Hamiltonian to compute only the Green's function blocks required for charge density and current, avoiding inversion of the full matrix and significantly reducing memory usage and computational cost. Its complexity is linear in the number of device slices (N_{slice} , typically a few tens) and cubic in the number of bands per slice (N_{bands} , typically a few thousand), with $N = N_{\text{slice}} \times N_{\text{bands}}$.

Because the dominant operations scale as $O(N_{\text{bands}}^3)$, detailed band structures are often simplified to limit the computational burden. Here, the Hamiltonian is projected onto a reduced basis via the generalized low-rank approximation (LRA), thereby eliminating degrees of freedom that have negligible impact on transport. This reduces the matrix rank by more than one order of magnitude while preserving the essential physics, accelerating NEGF simulations by several orders of magnitude and making systematic GAA-NS-FET studies and design-of-experiment (DoE) computationally feasible [10].

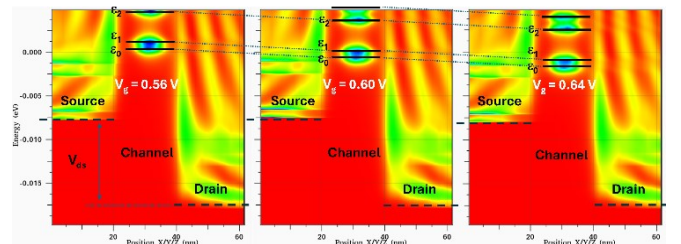


Fig. 4. Energy- and space-resolved density of states along the transport direction of a Si-GAA-NW-FET with a $2 \times 2 \text{ nm}^2$ cross-section, at low drain-source voltage ($V_{\text{ds}} = 10 \text{ mV}$) and low temperature ($T = 2 \text{ K}$). Several confined states $\epsilon_{1,2,3}$ are identified in the channel region and shift linearly from left to right as a function of the applied gate voltage $V_g = 0.56 \text{ V}, 0.60 \text{ V}, 0.64 \text{ V}$.

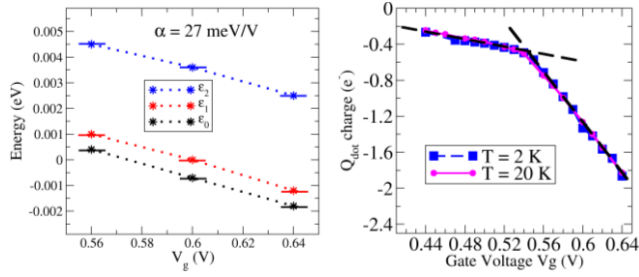


Fig. 5. (Left) Gate-voltage dependence of the first three localized energy levels identified in Fig. 4. (Right) Total electronic charge captured by the localized states of the lowest-energy function of V_g at $T = 2\text{K}$ and 20K .

Basic TCAD parameters are naturally extracted from Schrödinger or Poisson–Schrödinger simulations, including the transport and density-of-states effective masses for electrons and holes with parabolic or non-parabolic dispersion. Atomistic NEGF extends this capability by enabling the calibration of quantum-correction models, such as the density-gradient or Bohm quantum potential, and the accurate extraction of V_{th} , subthreshold swing (SS), and drain-induced barrier lowering (DIBL). Furthermore, simplified analytical models [11] for electron and hole mobilities $\mu_{e,h}$ and saturation and injection velocities $v_{sat,inj}$ can be replaced by quantum-mechanical values directly extracted from NEGF post-processing, as illustrated in Fig. 6 by the mobility curves and low- V_{ds} transfer characteristics. All these quantities inherently account for substrate orientation, mechanical strain, quantum confinement, detailed band structure, tunneling, interference, electron–phonon scattering, and self-consistent NEGF–Poisson coupling under open-boundary conditions.

Carrier–phonon and carrier–impurity interactions (charged impurities, alloy disorder, and remote scattering) are modeled using standard NEGF self-energy formulations, adapted to quantum confinement in GAA structures and compatible with the LRA optimization [10]. Further refinement of well-established bulk-material parameters requires systematic calibration against experimental transfer and output characteristics, complemented by density functional perturbation theory (DFPT) calculations of scattering matrix elements [6]. Although essential for predictive simulations, this calibration is still in its early stages because comprehensive GAA-NS-FET experimental data remain limited and are often unavailable due to industrial confidentiality.

III. MULTISCALE AND FAB-TECHNOLOGY CO-OPTIMIZATION

Rapid extraction of key device parameters—specifically V_{th} , SS, DIBL, mobility, and injection velocities—is achieved through the seamless integration of Victory Atomistic into Silvaco’s AI-driven FTCTM workflow via Victory DoE and Victory Analytics. Combined with digital twins that enable rapid review and design exploration, the FTCT solution provides a unified, high-speed quantum transport platform for optimizing n- and p-channel GAA-NS-FET components in an engineering environment [2].

IV. CONCLUSION AND PERSPECTIVE

Atomistic NEGF has become an essential predictive simulation tool for the Ångström era, providing a rigorous description of quantum confinement, tunneling, interference,

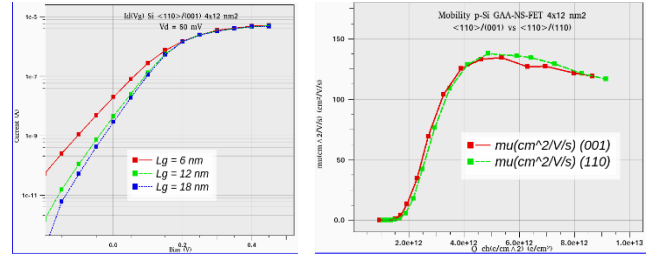


Fig. 6. (Left) $I_d(V_g)$ characteristic curves at low (50 mV) V_{ds} for an n-Si GAA-NS-FET of $4 \times 12 \text{ nm}^2$ cross-section oriented (001) and several gate lengths $L = 6, 12, 18 \text{ nm}$. (Right) Effective mobility of p-Si GAA-NS-FET of $4 \times 12 \text{ nm}^2$ cross-section, as a function of channel inversion charge for two substrate orientations (001) and (110).

and scattering in aggressively scaled GAA-NS-FETs over a wide temperature range, including cryogenic operation.

Combined with the Low-Rank Approximation and AI-assisted atomistic-TCAD workflows, it enables systematic exploration of device architectures while maintaining atomistic accuracy at practical computational cost. Beyond delivering key TCAD parameters, it provides a predictive framework for optimizing advanced CMOS technologies. Future developments will further strengthen its role as a cornerstone for the multiscale and multi-physics design of complementary FET (C-FET) architectures.

ACKNOWLEDGMENT

This work was performed within the framework of the research collaboration between Silvaco and the Purdue Research Foundation. The author sincerely thanks Prof. Tillmann Kubis, Katherine Ngai Pesic and Silvaco Associate Professor of Electrical and Computer Engineering at Purdue University for his collaboration and insightful discussions.

REFERENCES

- [1] International Roadmap for Devices and Systems (IRDS) 2024 Edition.
- [2] E. Guichard et al., "Driving Four Decades of TCAD Innovation: From Physical Simulation to AI-Powered Digital Twins," 2025 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Grenoble, France, 2025, pp. 1-6
- [3] Hyungmin An et al., "Etching-to-deposition transition in $\text{SiO}_2/\text{Si}_3\text{N}_4$ using CH_3Fy ion-based plasma etching: An atomistic study with neural network potentials", *Applied Surface Science*, Volume 716, 2026, p. 164601
- [4] L. P. Kadanoff and G. Baym, "Quantum Statistical Mechanics, Green's Function Methods in Equilibrium and Nonequilibrium Problems", New York: W. A. Benjamin, Inc., 1962
- [5] J. Cao et al., "Solving the Bethe-Salpeter Equation in the Nonequilibrium Green's Function Formalism," 2025 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), Grenoble, France, 2025, pp. 1-4
- [6] Samuel Poncé et al. "First-principles calculations of charge carrier mobility and conductivity in bulk semiconductors and two-dimensional materials", 2020, *Rep. Prog. Phys.*, 83, p. 036501
- [7] P. Blaise et al., "Nanoscale FET: How To Make Atomistic Simulation Versatile, Predictive, and Fast at 5nm Node and Below," *Proc. SISPAD*, Kobe, Japan, 2020, pp. 249-252
- [8] S. Steiger, et al., "Nemo5: A Parallel Multiscale Nanoelectronics Modeling Tool," *IEEE Trans Nano*, vol. 10, p. 1464, 2011
- [9] J. M. Soler et al., *The SIESTA method for ab initio order-N materials simulation*, *J. Phys. Condens. Matter* 14, 2745-2779 (2002)
- [10] Lemus, D.A., Charles, J., Kubis, T., "Mode-Space-Compatible Inelastic Scattering in Atomistic Nonequilibrium Green's Function Implementations," *J Comput Electron* 19, 1389–1398 (2020)
- [11] Lundstrom, Mark, "Fundamentals of Nanotransistors," *Lessons from Nanoscience: A Lecture Notes Series*, World Scientific, 2017