Device and Process Modeling:

20 years at Intel's other fab

Mark A. Stettler
Process Technology Modeling, Intel Corporation, Hillsboro, OR, USA
Email: mark.stettler@intel.com

***Abstract*—Embedding TCAD engineers in technology working groups has been an integral part of Intel's process development strategy since the company's inception. While this strategy remains the same, the challenges faced and the tools used by TCAD has undergone dramatic change over the last 20 years. This talk will discuss three recent trends in process and device TCAD: the rise in the use of "atomistic" scale simulations, the focus on modeling defects, and the continuing need to bridge new, more physically rigorous approaches to older, more computationally efficient methods. These trends will be illustrated with recent TCAD studies conducted at Intel.**

Keywords—TCAD; device modeling; process modeling; semiconductor simulation

I. INTRODUCTION

Device Engineering has undergone a profound change in the past 20 years and subsequently so has the infrastructure and expectations for the TCAD effort that supports it. In 1994, maintaining Dennard scaling for subsequent process generations was the primary strategy [1]. TCAD engineers helped enable this strategy by simulating the effects of implants, anneals, and geometry in controlling short channel effects as the gate length scaled, an effort that required a rather small toolset consisting of traditional process and device simulators. Twenty years later, the number of "knobs" introduced to maintain the scaling trend has increased substantially and so has the toolset required for TCAD support (Fig. 1). Strain engineering, the introduction of high-K and new gate materials, and the expansion into the 3rd physical dimension with the trigate architecture has greatly increased the process space explored by TCAD engineers. The challenge of scaling has naturally pushed TCAD into simulating smaller and more novel systems, where computationally expensive, quantum-based physical models are essential. This challenge has also propelled TCAD to simulate much larger systems—those above the device level—as the parasitic structures and circuits based upon novel devices can no longer be accurately simulated with existing compact models (Fig. 2). Solving problems of interest efficiently across the entire dimensional range has multiplied the number of simulation tools needed in the TCAD tool suite, which has grown well beyond those required 20 years ago. Fig. 3 shows the number and type of tools used on a daily basis at Intel, which has increased 5X since 1994, spanning from atomistic simulations to analytic compact models.

The remainder of this work will discuss three trends in process and device TCAD that are the result of recent scaling challenges: the rise in the use of "atomistic" scale simulations, the focus on modeling imperfection, and the continuing need to bridge new, more physically rigorous approaches to older, more computationally efficient methods. For all cases, examples based upon work done at Intel will be given.

II. ATOMISTIC SIMULATION

In the last 5 years, atomic scale device simulation—where each atom is explicitly modeled—has become an essential part of the TCAD toolbox for three reasons: 1) scaling has pushed the device physics needed to simulate strategic device options past Newtonian physics into the quantum regime, 2) fundamental material properties such as effective mass are modulated by seemingly minor dimensional changes and 3) the number of novel materials considered for new device options has grown considerably—beyond what is well-characterized in prior publications—with properties which now must be computed "on the fly."

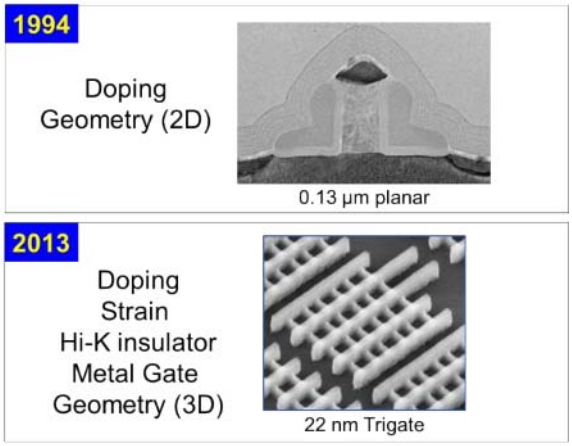

Fig. 1. Device engineering "knobs" in 1994 vs. 2014. *Images from Intel.*

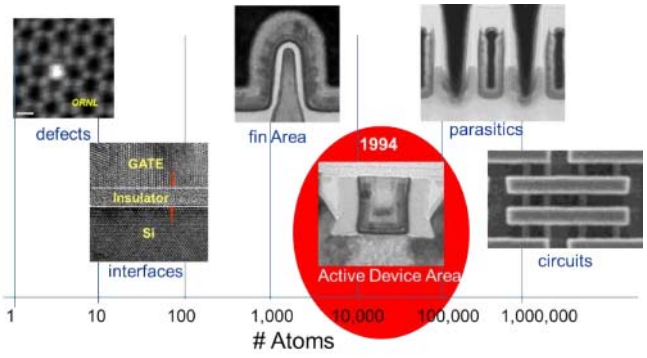

Fig. 2. Scales of systems expected to be addressed by TCAD simulation in 2014 compared to that expected in 1994 (the shaded red region). *Images from Intel and Oak Ridge National Laboratory.*

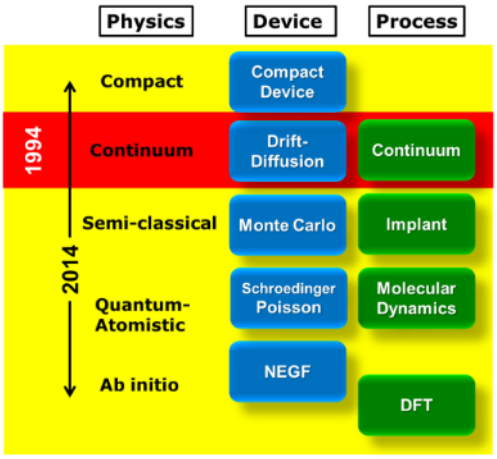

Fig. 3. The number of simulation tools used by TCAD to support technology development has increased considerably since 1994 (shaded in red).

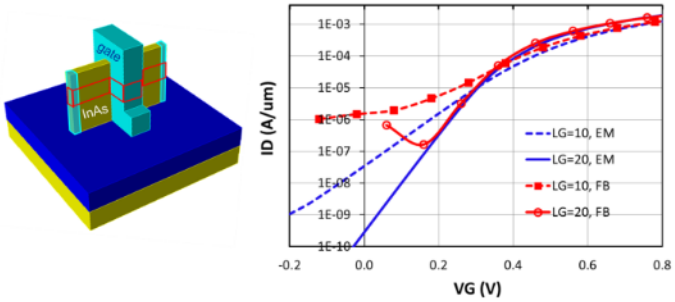

Fig. 4. Simulation of 10 and 20 nm InAs finfets using an Effective Mass (EM) and full-band Tight Binding (FB) NEGF approaches. The EM approach underestimates the leakage due to band-to-band tunneling [2].

An example illustrating the first reason is quantum mechanical tunneling. As device gate lengths approach 10 nm, tunneling will become the dominant leakage mechanism and any predictive simulation capability must include rigorous methods to compute these so-called "quantum effects." This is even more important when simulating III-V materials because the electron's lighter effective mass, which increases carrier velocity making III-V's an attractive alternative to silicon, also increases tunneling probability and thus leakage. Fig. 4 shows the necessity of using a very rigorous method when vetting strategic device options such as a InAs channel finfet. In this study, 10 and 20 nm gate length InAs channel finfets were simulated with the Non Equilibrium Greens Function (NEGF) based upon Tight-Binding and Effective Mass (EM) approaches [2]. Both approaches are solved on grid with atomic resolution.

In the subthreshold region of the finfet's I-V characteristic, there are two competing leakage mechanisms: direct tunneling through bandgap from source-to-drain which increases as the gate length decreases, and band-to-band tunneling from the conduction to the valence band, which increases with decreasing gate bias due to a greater overlap in the bands. The results show that the less accurate EM approach is not able to resolve leakage from this second mechanism, severely underestimating the off-current. In contrast, the full-band simulations predicts that band-to-band tunneling is substantial and is actually the dominant leakage mechanism at Vg=0 for the longer device, whose longer path reduces direct tunneling. This example demonstrates that less rigorous approaches, even those with atomic resolution, are often not sufficient to accurately evaluate a future technology, underscoring a need to routinely employ full-fledged solutions methods for validation. This is a recent development for industrial TCAD, where very rigorous tools were rarely used or even deemed necessary.

The next application demonstrates how atomistic tools can also be used to exploit quantum mechanical effects vs. mitigate them. In this example, a rigorous approach is used to "engineer" a better bandstructure for a tunnel FET (TFET), which is receiving attention as a future technology option because of its sharp turn-on characteristics at low gate biases due to a reliance on tunneling. One open issue with TFETs is the significant imbalance between the NMOS and PMOS characteristics making it very difficult to employ in CMOS technology. This difference is more pronounced in TFETs constructed with III-V materials. III-Vs are the material of choice because of their direct bandgap, which affords a high tunneling probability; however, PMOS TFET on-current suffers because of the low density-of-states in the III-V conduction band. A solution would be "engineer" a material which has a direct bandgap like III-Vs but has the conduction band density more similar to that in group IV material.

Fig. 5 shows the result of achieving that objective by applying mechanical stress to Ge, a column IV material. In this example, k.p pseuopotential theory is used to calculate the candidate bandstructures [3]. Mechanical stress modulates the atomic spacing, which in Ge differentially lowers the Γ and L valleys in the conduction band in relation to the valence band. At an applied biaxial tensile stress of 2500 MPa, the minimum energy of the Γ valley falls below the L valley, and the Ge transitions from being an indirect to a direct semiconductor. The same result was also found by adding a small amount of Sn to uncompressed Ge. Fig. 6 compares the I-V of a TFET constructed with this "designer" bandstructure with other candidates, all computed with NEGF. Only the GeSn NMOS and PMOS devices achieve comparable performance, and although the on-current of the GeSn is smaller at high Vg, it is greater at lower Vg making this device an excellent candidate for a low power technology.

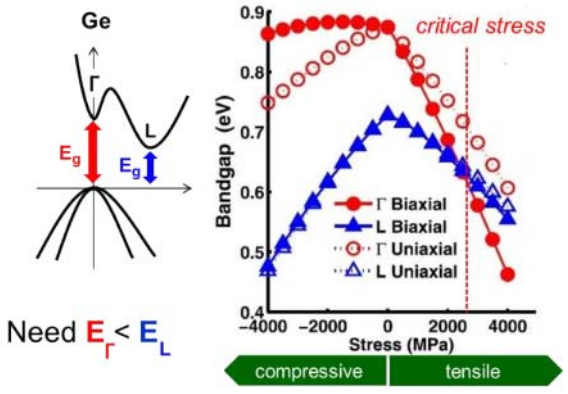

Fig. 5. The effect of stress on the bandgap of Ge calculated with k.p pseudopotential theory; stress differentially lowers the Γ and L conduction band valleys in relation to the valence band [3].

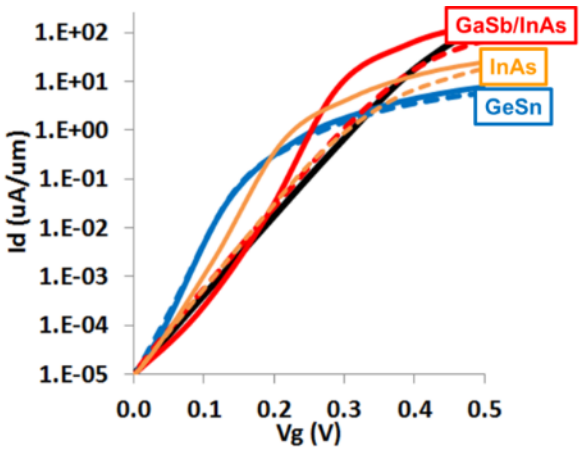

Fig. 6. Simulation of various NMOS and PMOS TFET devices using NEGF [4].

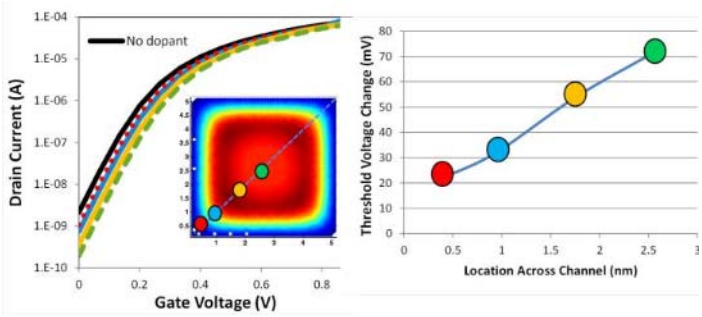

Fig. 7. Threshold voltage shift in a 10nm gate-length 5x5 nm silicon wire with a single dopant atom placed at various positions in the cross-section of the device [5].

III. Modeling Imperfection

Another recent trend originating from extreme scaling is that imperfections now play a much larger role in determining overall device behavior. In the nanoscale regime, any non-ideality—sometimes even a single atomic defect—can profoundly shift characteristics, resulting in the need for a new level of realism in modeling; TCAD must now provide methodologies and tools with sufficient resolution so that each new technology can be assessed for its vulnerability to imperfection and variation.

Fig. 7 shows an example how the placement of a single dopant can shift the threshold voltage (Vt) of a 10nm long, square silicon nanowire with a 5 X 5 nm cross-section [5], simulated with NEGF. One dopant placed in the middle of the channel can raise Vt 70 mV, completely compromising the on-current performance in a device targeted for a future low-voltage technology. The lone dopant still has a non-negligible effect even when moved to the corner.

The next example demonstrates how defects can also be exploited to gain performance. An edge dislocation is a missing plane of atoms which runs diagonally in the source and drain (S/D) regions (Fig. 8). This defect is created as the silicon lattice expitaxially regrows during annealing after heavy S/D ion implantation, which amorphizes the silicon. As shown in Fig. 8, this missing plane produces tensile strain lateral to the defect and compressive strain below. Tensile strain in the channel region improves NMOS performance by slightly increasing the lattice spacing of silicon atoms, modifying the bandstructure in such a way that electrons have higher mobility and thus move faster, increasing on-current. The more strain, the larger the performance, so maximizing these dislocations can be an important technology "knob."

Fig. 9 shows that both gate spacing and the order of the process flow can have a profound effect on the dislocation stress and resulting current gain [6]. To perform this study, first the electron mobility was rigorously calculated as a function of tensile stress and the results captured in a mobility model suitable for a drift-diffusion simulator. Second, the device structure and channel stress were computed from process simulations of the actual flow. The resulting structures and stresses were then used as inputs to a drift-diffusion simulator embedded with the custom mobility model, which calculated the on-current. As shown in the figure, reducing the gate pitch pushes the edge locations closer together increasing the tensile stress and thus the on-current. On-current also increases significantly when moving from a "gate first" (GF) process flow, where the gate is present throughout the regrowth anneal, to the "gate last" (GL) flow, where the gate has been removed, to be replaced after the anneal. The absence of a gate during the anneal introduces a free surface above the channel, allowing the atoms to be squeezed more tightly together, resulting in higher stress and more performance.

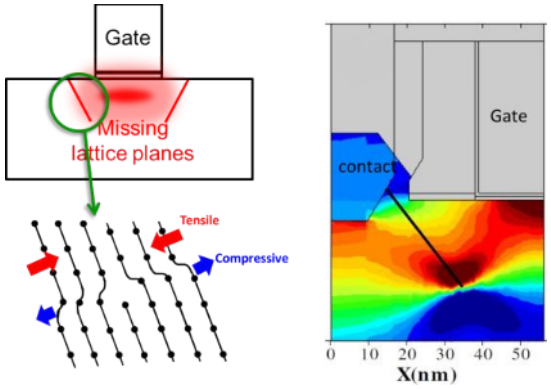

Fig. 8. Edge dislocations and resultant stress field (shown in half a MOSFET). Red denotes tensile stress and blue compressive stress.

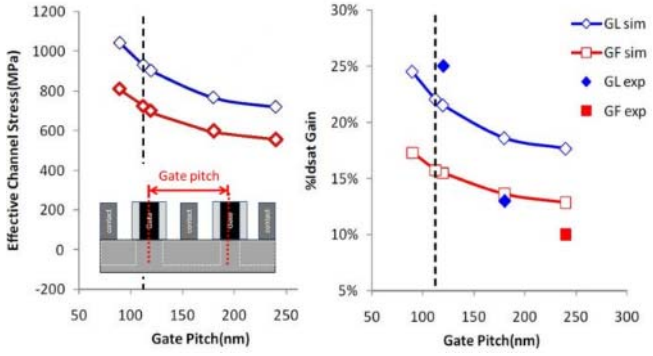

Fig. 9. The effect of gate pitch and gate first (GF) and gate last (GL) process flow on the stress and current gain from edge dislocation defects [6].

IV. Bridging The Gap

Although atomistic methods are necessary for scoping out the strategic device space, the incredible computation demand they require precludes them from day-to-day technology support. Fig. 10 shows how TCAD computation demand has increased roughly two orders of magnitude over the past decade. The first inflection point was the need to simulate in 3 (vs 2) dimensions for finfet simulation and also to calculate more fundamental material properties with the advent of strain engineering. The 2nd inflection point was the increased demand of atomistic simulation, which was discussed in the proceeding paragraphs. This has pushed TCAD engineers to develop efficient ways to capture more physically sophisticated effects in less computationally expensive codes such as Drift-diffusion (DD), which, after 30 years of seeming invalidity, is still the workhouse of technology development.

Fig. 11 shows how a Schrodinger-Poisson method can be combined with the DD approach to obtain the correct quantum mechanical carrier density throughout a 3D device. In this approach, an iterative method is used to link the two solutions, where 2D slices from the Schrodinger solution are employed to "shape" the carrier density in the DD solution, which solves for the transport. Fig. 12 shows an iterative scheme for including rigorous tunneling calculations in a DD device simulation framework [7]. In this approach, the bandstructure of the device at the dimensions of interest is first calculated using a rigorous 2D NEGF Tight Binding simulation. A two band representation of the bandstructure is then extracted and used as input to a 1D NEGF simulator which is coupled with the DD simulator. During device simulation, DD calculates the potential contours, which are used to extract the most probable tunneling paths. The 1D NEGF code then solves for the tunneling current along these paths, which is then introduced back into the DD as a generation term. Although this approach has a number of computation steps, it is much more efficient than calculating the entire I-V directly with the full NEGF approach. The approaches featured in Fig. 11 and 12 were shown to have sufficient accuracy for device studies.

Finally, it's worth noting that constructing hybrid solutions approaches featured above is greatly facilitated by having an open, modular simulation framework, where solution methods and output and input fields can easily be combined without having to devise an entirely new code. Intel has such as system, referred to as the Modular Device Simulator [8].

I. Conclusion

TCAD support at Intel and elsewhere has changed dramatically since the 1990's, with much of the change coming in just the last 5 years. Atomistic approaches are now a necessity to capture the essential physics and variation properties of future technologies, while more computationally efficient approaches that retain the essence of the underlying physics are still needed to keep throughput high in today's technology development.

References

[1] R. Dennard, et al., "Design of ion implanted MOSFETs with very small physical Dimensions," IEEE Journal of Solid State Circuits, vol. SC-9, no. 5, pp. 256-268, (1974).

[2] D. Basu, R. Kotlyar, C. Weber et al., "Ballistic Band-to-Band Tunneling in the OFF state in InGaAs MOSFETs," in press (2014).

[3] R. Kotlyar, U. Avci, S. Cea, R. Rios, T. Linton et al., "Bandgap engineering of group IV materials for complementary n and p tunneling field effect transistors," Applied Physics Letters 102, 113106 (2013).

[4] U Avci, D. Morris, S. Hasan et al., "Energy efficiency comparison of nanowire heterojunction TFET and Si MOSFET at Lg=13nm, including P-TFET and variation considerations," IEDM Tech. Digest (2013).

[5] M. Giles and R. Golizadeh, ITRS ERM Workshop on Deterministic, Conformal, and Monolayer Doping (2013).

[6] C. Weber, S. Cea et al., "Modeling of NMOS performance gains from edge dislocation stress," IEDM Tech. Digest (2011).

[7] S. Hasan, unpublished.

[8] T. Linton, K. Foley, F. Heinz, R. Kotlyar, P. Matagne *et al.*, "MDS—a new, highly extensible device simulator," SISPAD (2007).

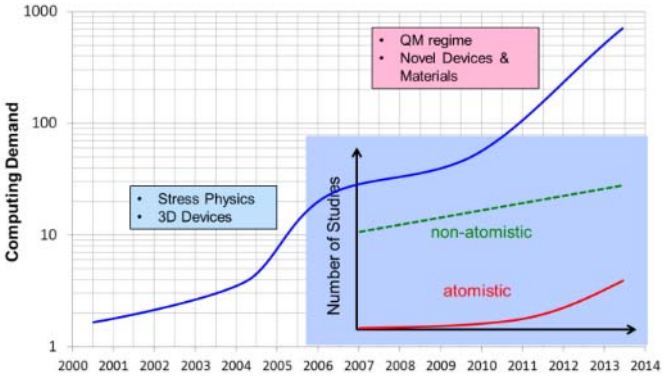

Fig. 10. Computational demand vs. year for TCAD technology support.

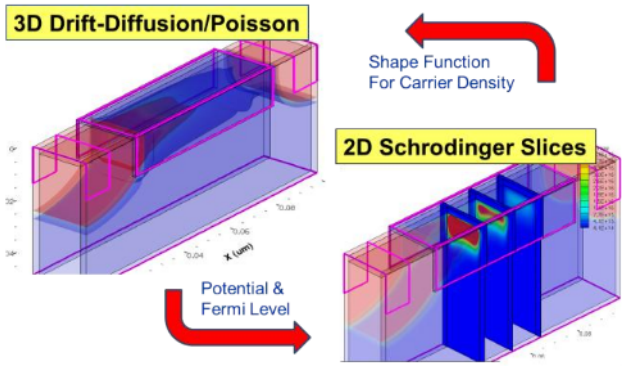

Fig 11. Iteration scheme in a combined Drift-Diffusion and Schrodinger Solver approach for efficiently obtaining accurate carrier densities.

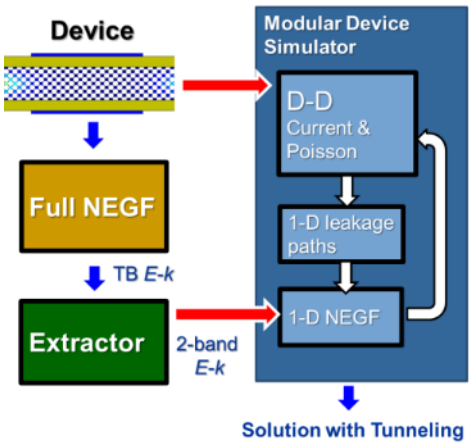

Fig 12. Iteration scheme in a combined Drift-Diffusion and NEGF Solver approach for efficiently obtaining accurate tunneling current.