SYNOPSYS®
Silicon to Software™

Newsletter for semiconductor process and device engineers

December 2019

TCAD news

Latest Edition

Welcome to the December 2019 edition of the TCAD News. This edition addresses two important topics in the scaling of memory and logic technology.

The first article presents a Sentaurus TCAD methodology for simulating the electrical characteristics of the polysilicon channel in 3D NAND, one of the key features of this technology whose detailed modeling is becoming critical to support the continued integration of more layers.

The second article describes the exploration of 2D-material FETs, emerging candidate transistors for future logic nodes, based on the atomic-scale modeling capabilities of QuantumATK. This article is a fitting example of the rising importance of QuantumATK in the down-selection of novel materials and device structures.

With the approaching holiday season, I would also like to take this opportunity to wish you happy holidays and a prosperous New Year.

With warm regards,

Terry Ma
VP Engineering, TCAD

Polysilicon Channel Modeling for 3D NAND Flash Memory Development

The advent of 3D NAND flash memory marks the most profound breakthrough in the memory industry over the last decade. It overcomes the lithography and few-electron limitations of 2D NAND flash architecture and enables continued density scaling by vertically stacking bit cells to tens or even hundreds of layers. The resulting arrays of tall cylindrical columns form NAND strings and scale up the bit density at decreasing cost with improved reliability and performance. The number of stacking layers has recently surpassed 100 and is expected to exceed 200 in the next few years.

One of the distinct features of 3D NAND flash memory is the cylindrical polysilicon channel that affects the device characteristics in several key aspects:

a. Its relatively poor electrical conductivity, as compared with single crystal silicon, poses a constraint to the sensing current and may eventually limit vertical scaling as the channel resistivity increases with increasing number of stacking layers;

b. Polysilicon grain size and charge traps at grain boundaries (GBs) are random in nature, causing intrinsic variation in drive current due to scattering at GBs or carrier scattering by interfacial traps;

c. Traps at GBs also cause Vth variation due to trap-induced carrier depletion in the channel;

d. Due to thermionic process, the channel drive current exhibits an inverse temperature effect, whereby the current increases with increasing temperature, as opposed to what is observed in single crystal silicon. This effect is induced by small potential barriers around traps at GBs, which carriers must overcome with higher thermal velocities;

e. The intermittent charge and discharge of a trap at a GB leads to random telegraph noise, causing cell Vth variation.

In this article we describe a Sentaurus TCAD flow for simulating the electrical behavior of the cylindrical polysilicon channel in the 3D NAND string. The simulation flow uses Sentaurus Structure Editor and Sentaurus Process for generating the Voronoi grains structure representative of the polysilicon grains, and Sentaurus Device for simulating electrical variability and the inverse temperature effect due to discrete traps at the GBs.

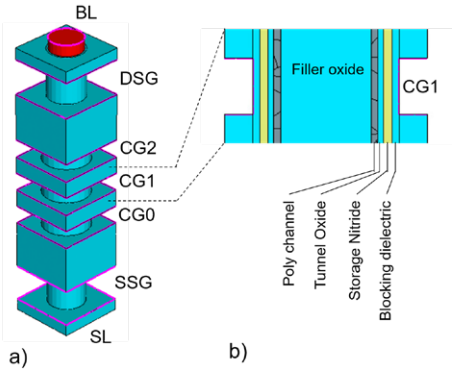

Figure 1: a) A cylindrical 3D NAND structure with 3 control gates (CG0, CG1, CG2), two select gates (SSG, DSG), and BL and SL at top and bottom; b) Cross-section showing SONOS dielectric stack for a flash cell.

Contact TCAD
For further information and inquiries:
tcad_team@synopsys.com

A simplified, yet representative, 3D NAND structure is used for the simulations (Figure 1a). It consists of 3 cylindrical control gates (CG0, CG1, CG3), bounded by source and drain select gates (SSG, DSG), with bitline (BL) and source line (SL) terminals at the two ends of a vertical stack. The wordline (WL) width (control gate length) and spacing are 40 and 38nm, respectively. The flash cell is of charge trap type (Figure 1b), with polysilicon channel, tunnel oxide, storage nitride, and blocking oxide (SONOS) thicknesses of 5, 4, 6, and 5nm, respectively. The filler oxide diameter is 65nm. Both the channel and source/drain regions are doped n-type at $1x10^{17}$ and $1x10^{20}$ cm^{-3}, respectively.

Polysilicon grains are generated with Voronoi tessellation in the polysilicon channel (Figures 2a and 2b). The average grain size is set to 17.6nm for the rest of the article, except otherwise stated.

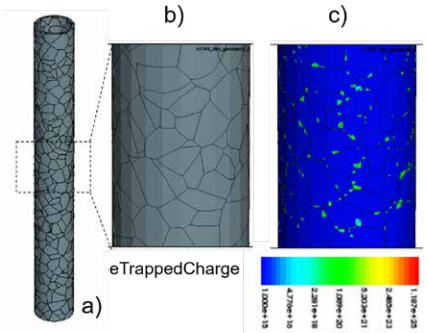

Figure 2: a) A cylindrical poly tube with Voronoi grains; b) Close-up view of the grain distribution; c) Random discrete traps placed at grain boundaries.

A polysilicon GB corresponds to the interface between two mutually misoriented crystal grains, often associated with atomic bond distortion or displacement, giving rise to interfacial defects. They can be modeled as discrete traps randomly placed at GBs in Sentaurus Device at a given effective interfacial density (Figure 2c).

For device simulations, mesh refinement is applied at the silicon-oxide interface and silicon-silicon interfaces at GBs. Default

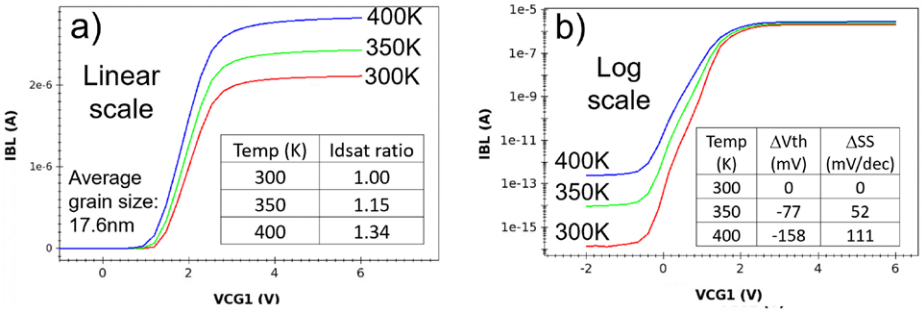

Figure 3: a) IBL vs. VCG1 in linear scale at different temperatures, along with inset for temperature dependency of drive current ratio; b) IBL vs. VCG1 in log scale, along with inset for temperature dependency of Vth and sub-threshold swing (SS) variations.

Sentaurus Device settings for bulk mobility, Lombardi interface degradation, and high field saturation models are used.

Read operation is performed to analyze the device I-V characteristics, with pass gates (CG0, CG2) held at 6.0V, select gates (DSG, SSG) at 3.0V, BL at 1.5V, SL at 0V, while sweeping CG1 from -2.0 to 6.0V.

Figure 3 shows temperature dependency of BL current vs. control gate voltage (IBL vs. VCG1) in linear and log scales, along with insets for current ratio (Figure 3a) and Vth and sub-threshold swing (SS) variation (Figure 3b). The inverse temperature effect is evident: as temperature increases from 300 to 400K, the drive current increases by 34%, while the Vth decreases at a rate of 1.58mV/K and SS increases at 1.11mV/dec/K.

The inverse temperature effect occurs when the interface trap density exceeds a threshold, between $1x10^{12}$ and $5x10^{12}$ cm^{-2}, and its magnitude depends on the interface trap density.

Because traps are discrete, carrier transport is governed by percolation, resulting in a significant reduction of drive current relative to the case when traps are absent. The trap energy levels are fixed in this study at mid-bandgap and 0.1eV offset from the conduction and valence band edges for acceptor-like or donor-like traps, respectively.

Figure 4 shows a conduction band diagram, cut along CG0 channel through discrete traps. The small bumps are energy barriers induced by trapped charges and are responsible for the afore-mentioned channel conductivity degradation and thermionic effect. The rise of the band edge at the right is due to lower voltage at the neighboring control gate (CG1).

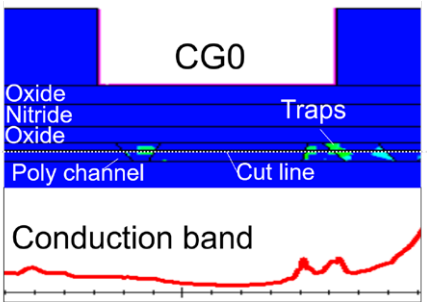

Figure 4: Conduction band diagram under CG0, showing small energy barriers induced by trapped charge. The rise of band edge at the right is due to lower voltage at the neighboring control gate (CG1).

In the absence of traps, there might be transport degradation at poly GBs due to, for instance, the difference in electron affinity from grain crystal orientation. To account for such degradation, a thermionic boundary condition at silicon-silicon interfaces is used in the device simulations. This boundary condition, however, should not to be confused with thermionic mechanism associated with trap-induced barriers.

Figure 5 compares IBL vs. VCG1 for two average grain sizes (17.6 vs. 13.0nm). As expected, the smaller the grain size, the more total number of GBs are in the channel, resulting in further reduction in drive current.

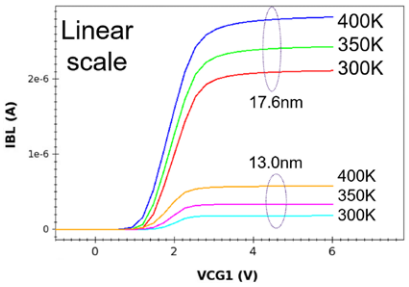

Figure 5: IBL vs. VCG1 at two average grain sizes (13.0 vs. 17.6nm).

Since poly grain formation and interface trap location are random in nature, device performance variations from random sampling are expected. Figure 6 shows the variation of drive current ratio for 10 samples of grain configuration and trap placement, obtained by varying random seeds in simulations. For a fixed grain configuration, the drive current variation due to random trap placement is significantly smaller in magnitude than that due to random poly grain configuration. For better statistics, a much larger sampling size is needed.

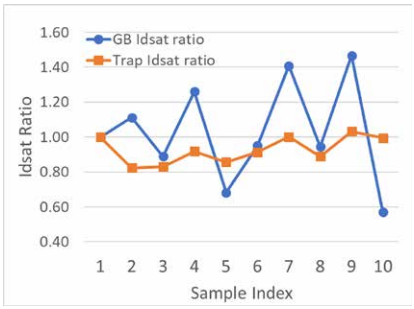

Figure 6: BL drive current variations induced by random sampling of poly grain configuration and trap placement at GBs.

In conclusion, enhancements in the Q-2019.12 release of Sentaurus TCAD enable robust generation of Voronoi grain boundaries in a polysilicon channel for 3D NAND flash memory. Device simulation of the impact of discrete traps at GBs capture the resulting degradation in drive current, the inverse temperature effect, and the variations in Vth and drive current, all of which are primary considerations for technologists working to develop and optimize the next-generation 3D NAND memory.

QuantumATK Enables the Early Exploration of 2D-Material FET Performance

Low dimensionality materials are important candidates for future logic devices owing to their excellent transport characteristics, as shown in Figure 7. These materials include two-dimensional (2D) materials such as transition metal dichalcogenides (TMDs), black phosphorene, graphene and 1D carbon nanotubes (CNTs). Early research indicates that planar FET structures with these materials integrated in the channel exhibit promising performance and electrostatic control.

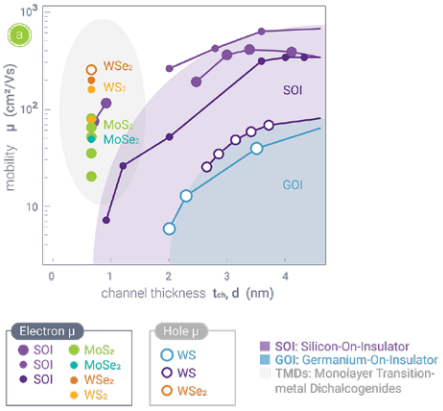

Figure 7: Mobility as a function of channel thickness for semiconducting channels based on TMDs, silicon-on-insulator (SOI) and germanium-on-insulator (GOI) devices [1].

Integration into silicon-based planar FET structures is facilitated by the excellent performance shown in films just a few monolayers in thickness, thereby enabling processing via atomic layer deposition or physical vapor deposition.

In this article we focus on the modeling of 2D-material FETs using the atomic-scale simulator QuantumATK, illustrated through four recently completed studies. Modeling of 2D-material FETs is motivated by the need to down-select among many candidate 2D materials and their heterostructures. QuantumATK simulates the characteristics of 2D-material FETs such as on-state current (I_{ON}), off-state current (I_{OFF}), transconductance, subthreshold swing (SS), contact resistance, Schottky barrier height, etc, thereby providing key insights on how to improve the performance of these devices.

QuantumATK can also be used to calculate bandstructures and phonon-limited electron mobilities for 2D materials. Parameters such as effective masses extracted from the bandstructures can be inserted in Sentaurus TCAD advanced transport simulations, enabling pathfinding Design-Technology Co-Optimization (DTCO) of 2D-material FETs.

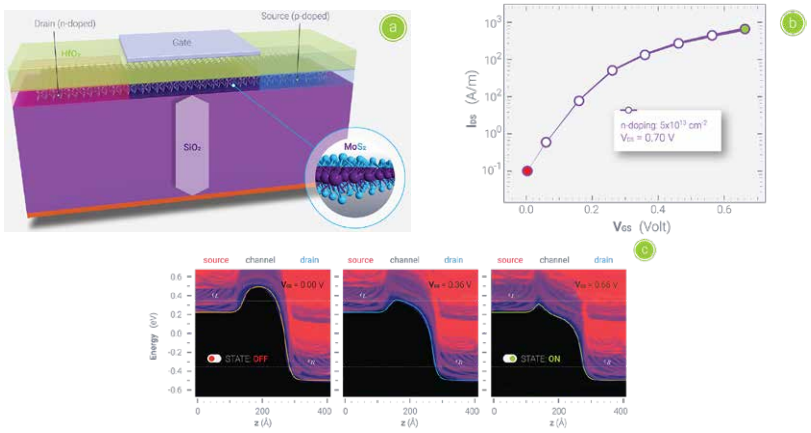

Figure 8: 2D-material FET comprised of a single monolayer of MoS_2. In this model device, the source and drain electrodes are made of heavily doped MoS_2 to form a n-i-n 2D FET. The device has a single metal gate (15 nm long) that is treated as a microscopic metallic region. The top gate is separated from the channel by a dielectric HfO_2 insulator. The MoS_2 monolayer is placed on a 20 nm thick dielectric region, which mimics the SiO_2 layer in a SOI wafer. I_{DS}-V_{GS} curve and PLDOS obtained with DFT-NEGF.

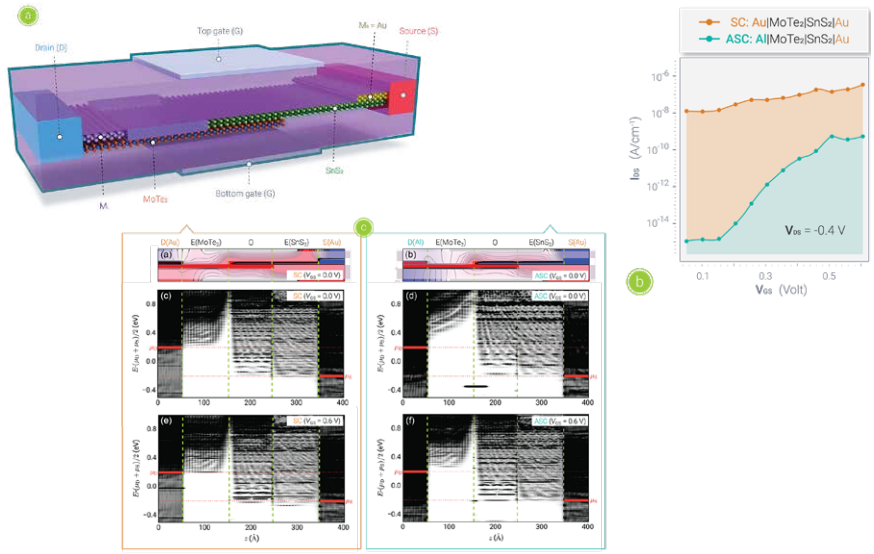

Figure 9: a) Double-gated 2D-material TFET based on the $MoTe_2/SnS_2$ heterojunction (with the 30 nm channel length) and two different metallic contacts. The device is encapsulated in a high-dielectric region, while a thin dielectric region is placed above the exposed $MoTe_2$ region. In the symmetrically contacted $M_L/MoTe_2/SnS_2/M_R$ device, Au is used for both the source (S) and drain (D) metallic contacts, whereas in the asymmetrically contacted device, drain and source are set to Al and Au, respectively, b) Transconductance: reverse-bias I_{DS}-V_{GS} curves at the drain source voltage -0.4V. c) Cut planes of the Hartree electrostatic potential and the PLDOS [2].

The first study investigates a single-gated 2D-material FET, comprised of a monolayer of MoS_2, as shown in Figure 8. The study aims to characterize the highest performance attainable from the structure given that contact resistance was not accounted for. The MoS_2 monolayer was placed on a 20 nm thick dielectric region, which mimics the SiO_2 layer in a SOI wafer. QuantumATK DFT-NEGF (non-equilibrium Green's function) simulation of I_{ds}-V_{gs} characteristics (Figure 8b) yields I_{on} approaching 10^3 A/m and SS of 88 mV/dec at the target V_{ds} = 0.7V. The projected local density of states (PLDOS) and electrostatic potential show the expected decrease in the source barrier and the attendant electron tunneling through the barrier as the device is switched on.

The second study considers a double-gated FET based on a heterojunction of 2D materials, which further expands the device design space and allows for a flexible tuning of the tunneling barrier height. The channel is formed by a $MoTe_2/SnS_2$ heterojunction and two different metallic contacts as shown in Figure 9 [2]. In the symmetrically contacted $M_L/MoTe_2/SnS_2/M_R$ device, Au is used for both the source and drain metallic contacts, whereas in the asymmetrically contacted device, drain and source are set to Al and Au, respectively, to have a rather large work function difference. The DFT-NEGF simulations show that source-drain contact materials can significantly change the output characteristics of FETs, as shown in Figure 9, and can be used to optimize device performance. Whereas work function differences in source-drain metals in the asymmetrically contacted device creates internal fields that can reduce subthreshold current, Al source contact reduces Fermi level pinning of $MoTe_2$ and increases space-charge modulation by the gate leading to higher transconductance.

It is also very important to investigate the performance of devices with aggressively scaled, sub-10 nm, 2D-material-based channel lengths. A QuantumATK study, performed by Imec and the Universities of Leuven, and Antwerp [3] investigated the performance of the MoS_2 based double gated FET in the sub-10 nm regime as shown in Figure 10. The study found that there is a tunneling current flowing between the source and the drain electrodes in 5 nm and 2nm channel length devices, which significantly increases the off-state (leakage) current, thus reducing the I_{ON}/I_{OFF} ratio and deteriorating subthreshold swing (SS) values as shown in Figure 10.

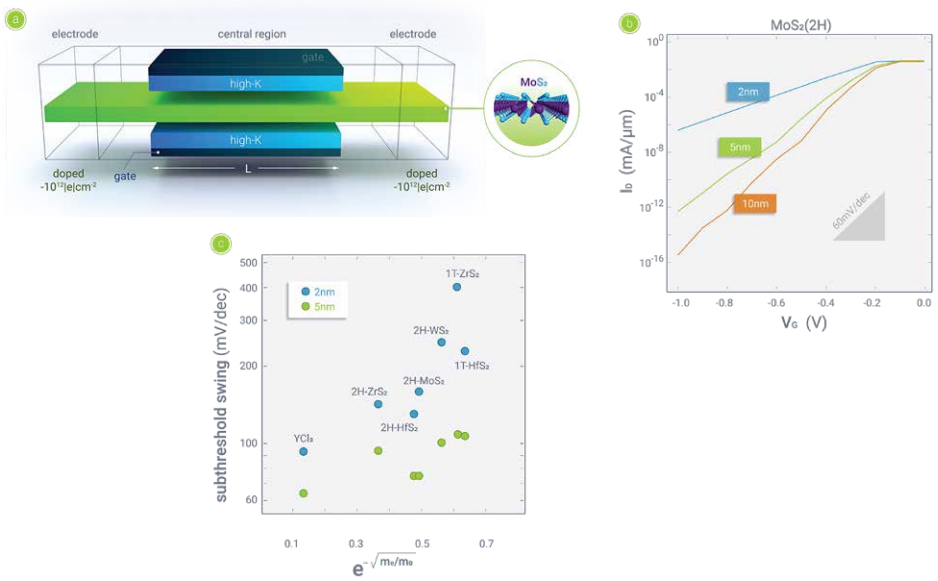

Figure 10: a) Double-gated MoS_2 based FET. b) I_{DS}-V_{GS} characteristics for the channel lengths of 2, 5 and 10 nm. C) SS as a function of the exponential of the $(m_b)^{1/2}$, electron effective mass on a semi log scale [3].

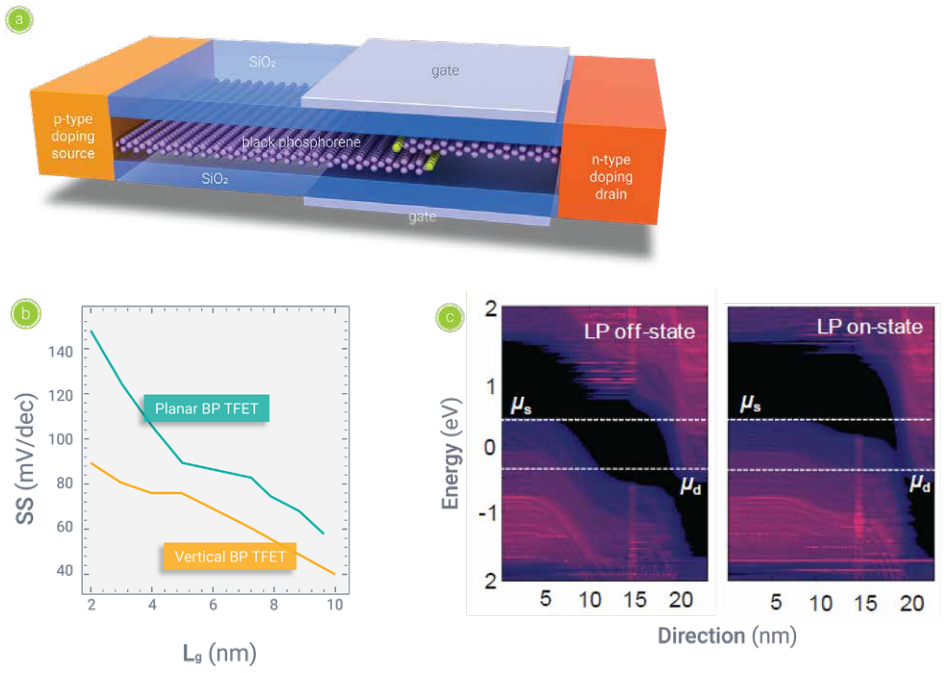

Figure 11: a) TFET based on a homojunction of layered black phosphorene. b) SS for different gate lengths. C) Calculated PLDOS reveal valence-to-conduction band tunneling [4].

The electron tunneling process depends on the effective mass of the 2D channel material, motivating the search for 2D materials with an effective mass larger than that of MoS_2 to reduce the source-to-drain tunneling, and improve device performance.

Yet another study performed by Peking University and North China University of Technology [4] uses QuantumATK to simulate the performance of a tunnel FET (TFET) based on a homojunction of layered black phosphorene (Figure 11) over a range of gate lengths, from 2 to 10 nm. The study concludes that a vertical TFET structure has lower leakage relative to a planar TFET structure in the same material system. This study is a good illustration of the need for the co-investigation of promising materials and specific device structures.

References

[1] D. Akinwande, C. Huyghebaert, C.-H. Wang, M. I. Serna, S. Goossens, L.-J. Li, H.-S. P. Wong and F. H. K. Koppens, "Graphene and two-dimensional materials for silicon technology", Nature 573, 507 (2019).

[2] S. Smidstrup, T. Markussen, P. Vancraeyveld, J. Wellendorf, J. Schneider, T. Gunst, B. Vershichel, D. Stradi, P. A. Khomyakov, U. G. Vej-Hansen, M.-E. Lee, S. T. Chill, F. Rasmussen, G. Penazzi, F. Corsetti, A. Ojanpera, K. Jensen, M. L. N. Palsgaard, U. Martinez, A. Blom, M. Brandbyge, and K. Stokbro, "QuantumATK: An integrated platform of electronic and atomic-scale modelling tools", J. Phys.: Condens. Matter 32, 015901 (2020). arXiv: 1905.02794v2.

[3] A. K. A. Lu, G. Pourtois, T. Agarwal, A. Afzalian, I. P. Radu, and M. Houssa, "Origin of the performances degradation of two-dimensional-based metal-oxide-semiconductor field effect transistors in the sub-10nm regime: A first-principles study", App. Phys. Lett. 108, 043504 (2016).

[4] H. Li and J. Lu, "Sub-10 nm vertical tunneling transistors based on layered black phosphorene homojunction", App. Sur. Sci. 465, 895 (2019).

SYNOPSYS®
Silicon to Software™

Synopsys, Inc. • 690 East Middlefield Road • Mountain View, CA 94043 • www.synopsys.com