Quasi-Ballistic Drift-Diffusion Simulation of SiGe Nanowire MOSFETs Using the Kinetic Velocity Model

KO-HSIN LEE[1], AXEL ERLEBACH[2], OLEG PENZIN[3], AND LEE SMITH[4] (Member, IEEE)

1 Department of TCAD, Synopsys Taiwan Company Limited, Hsinchu 302, Taiwan
2 Department of TCAD, Synopsys Switzerland LLC, 8050 Zurich, Switzerland
3 Department of TCAD, Synopsys Inc., Hillsboro, OR 97124, USA
4 Department of TCAD, Synopsys Inc., Mountain View, CA 94043, USA

CORRESPONDING AUTHOR: K.-H. LEE (e-mail: helen.lee@synopsys.com)

ABSTRACT This paper presents the calibration of the novel kinetic velocity model (KVM) in the drift-diffusion (DD) transport approach, which can account for the ballistic effect in short-channel devices. The KVM considers a thermionic emission limit and a free carrier acceleration limit for the mobility. We develop a methodology to extract the parameters for the KVM and for the high-field saturation velocity model for SiGe nanowires over the whole mole fraction range. The calibrated DD simulations with KVM show good agreement with Boltzmann transport equation results in terms of on-state current and carrier-weighted velocity distribution over a wide range of gate lengths for both linear and saturation regimes.

INDEX TERMS Ballistic mobility, drift diffusion, nanowire, silicon germanium, kinetic velocity, thermal velocity.

I. INTRODUCTION

Gate-all-around (GAA) transistors have the potential to further scale CMOS devices beyond present-day FinFET technology [1]. It has been reported that the short-channel effects can be lessened in GAA devices due to improved gate control [2]. Superior electrostatics and dynamic performance have been shown in the horizontally stacked nanosheet devices compared to FinFET [3]. Furthermore, the CMOS integration of n- and p-type devices has been successfully demonstrated on vertically GAA Si nanowire (NW) MOSFETs [4].

When scaling devices toward short channels, quasi-ballistic transport becomes increasingly important. However, the standard drift-diffusion (DD) approach is not able to fit the current-voltage characteristics over a wide gate length range using the same model parameters. Therefore, it is required to extend the DD approach by including a ballistic mobility model that accounts for the ballistic resistance and the finite injection velocity. The ballistic mobility term is often combined with the standard DD mobility using Matthiessen's rule [5]–[7]. Various expressions for the ballistic mobility have been proposed, such as simple compact models and equations that calculate the ballistic mobility from injection velocity, gradient of the quasi-Fermi potential, and carrier density locally [6], [8]–[11]. Most of them require an explicit specification of channel length and/or are generally not well designed for fitting to the local mesh-point based TCAD DD implementation. Furthermore, the standard Matthiessen's rule is not a good approach for combining ballistic with DD transport due to the underlying simple series resistance model. Therefore, a modified Matthiessen's rule is needed with weighting factors for ballistic and DD transport terms. An alternative formulation, based on the kinetic velocity model (KVM) [12], has been presented to account for the ballistic effects in the DD approach. In this model, the ballistic mobility, including both thermionic emission and free carrier acceleration terms, is expressed using mesh-point based DD variables and does not require an explicit dependence on the channel length. The ballistic and DD mobility are combined

using Frensley's rule [12]–[13], instead of Matthiessen's rule.

For short channel devices, not only the quasi-ballistic transport should be considered in current-voltage characteristics simulations, the tunneling effect is also becoming pronounced, especially for Ge-enriched transistors [14]. In this work, the focus is mostly to extend the DD simulation at the quasi-ballistic transport regime. We report the calibration and validation of KVM in the DD approach to account for the ballistic effect. The KVM and the high-field saturation (HFS) model parameters are calibrated for SiGe n- and p-type nanowire MOSFETs. The subband Boltzmann transport equation (subband-BTE) is employed as reference without implementation of tunneling models.

FIGURE 1. Geometry and doping profile of a SiGe n-NW. The height and width of NW are 5 nm and the silicon oxide thickness is 1 nm.

II. MODEL CALIBRATION FOR QUASI-BALLISTIC TRANSPORT

The extension of DD simulation to the quasi-ballistic transport regime requires calibrations of KVM and high-field saturation (HFS) model parameters. With the KVM approach, instead of Matthiessen's rule, the ballistic mobility (μ_{bal}) is combined with the total DD mobility (μ_{dd}) using the following Frensley's rule [12] to compute the final mobility (μ):

$$\frac{\mu}{\mu_{dd}} + \left(\frac{\mu}{\mu_{bal}}\right)^2 = 1 \tag{1}$$

It is found that Frensley's rule works better for the ballistic mobility in the saturation regime compared to the Matthiessen's rule [12]–[13].

In the KVM model, the total ballistic mobility (μ_{bal}) is composed of two terms: thermionic emission (μ_{bal}^T) and free carrier acceleration (μ_{bal}^B), and is expressed as follows [12]:

$$\frac{1}{\mu_{bal}} = \frac{\alpha_T}{\mu_{bal}^T} + \frac{\alpha_B}{\mu_{bal}^B} \tag{2}$$

where α_T and α_B are fitting parameters. Both μ_{bal}^T and μ_{bal}^B are inversely proportional to the square root of effective mass, $\sqrt{k_B T/m^*}$, and are given by:

$$\mu_{bal}^T = \frac{\sqrt{\frac{k_B T}{m^*}} y_T\left(\frac{q\psi}{k_B T}\right)}{|\nabla\psi|}, \quad y_T(x) = e^x\sqrt{\alpha_0 + 2\left(1 - e^{-x}\right)} \tag{3}$$

$$\mu_{bal}^B = \frac{\sqrt{\frac{k_B T}{m^*}} y_B\left(\frac{q\psi}{k_B T}\right)}{|\nabla\psi|}, \quad y_B(x) = \sqrt{\alpha_0 + 2x} \tag{4}$$

where ψ is the quasi-Fermi potential with reference to the source contact, m^* is the effective mass, k_B is the Boltzmann's constant, and T is the lattice temperature. The quasi-Fermi potential at the source contact is set to zero.

Most of the ballistic models in TCAD define a ballistic mobility, which is constant along the channel, in order to set the injection velocity at the source. For this correction to work with different channel length, the ballistic mobility itself depends on the channel length that must be entered in the model as a parameter. The KVM calculates the
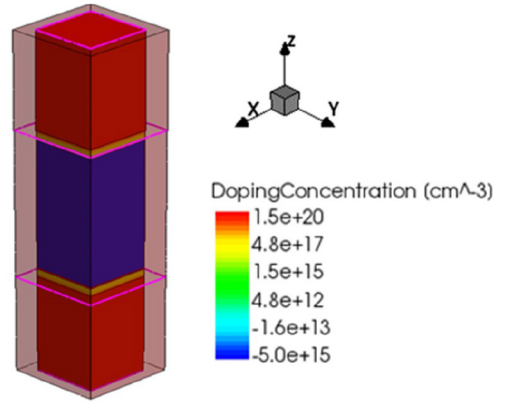
mobility locally for each mesh point in the device and provides with that an automatically increasing ballistic mobility with increasing distance from the source. This way the ballistic channel resistance does not increase linearly with the gate length but has a finite value and an explicit parameter for the gate length is not necessary. The KVM model avoids therefore the explicit dependency on channel length. It is implemented in the Sentaurus Device simulator [15], where m^* is introduced through the fitting parameter k and has the dependence with Ge mole fraction:

$$k = \frac{q v_T}{2 k_B T}, \quad v_T = \sqrt{2 k_B T / \pi m^*} \tag{5}$$

In the above equations, q is the electron charge and v_T is the thermal velocity. The k parameter is a lumped parameter that is used by many TCAD users since years. We therefore use this parameter here too.

In high driving field, the DD mobility is modeled based on Caughey-Thomas formula [16]:

$$\mu_{dd} = \frac{\mu_{low}}{\left(1 + \left(\frac{\mu_{low} F_{hfs}}{v_{sat}}\right)^\beta\right)^{\frac{1}{\beta}}} \tag{6}$$

where μ_{low} is the low-field mobility, F_{hfs} is the driving field, and v_{sat} is the saturation velocity, which has also dependence with Ge mole fraction.

III. DEVICE GEOMETRY AND SIMULATION SETTING

The parameter extraction and validation of the ballistic mobility model is performed on 3D square SiGe n- and p-type MOSFET nanowires with a cross-sectional area of 5x5 nm², as shown in Fig. 1. The channel direction is along the <110> crystal orientation with length varying from 7 nm to 200 nm and the channel doping is 1x10¹⁵ cm⁻³.

The KVM parameters are calibrated with the solution of subband-BTE [17] as reference, which considers the impact of strong quantum confinement as well as quasi-ballistic transport featured in small device configurations. In our simulations, phonon, surface roughness, Coulomb, and alloy scattering, are included. For nMOS, a 2kp-based ellipsoidal

band structure model is used for the X-valleys while an ellipsoidal, effective mass band structure model is used for L- and Gamma-valleys. For pMOS, the 6kp band structure model is used. The scattering parameters have been calibrated to give a consistent long-channel mobility compared to experimental results [18].

In our DD simulations, the low-field mobility is modeled using a combination of enhanced Lombardi model and Philips unified mobility model [15]. In high driving field, the DD mobility is modeled using the Canali model [16]. The KVM is implemented in all regions with both the thermionic emission and free carrier acceleration terms considered. Additionally, the bandgap widening effect due to volume quantization at small feature size is included and Fermi-Dirac statistics is considered for the strong inversion regime.

IV. CALIBRATION METHODOLOGY

The calibration and validation of the quasi-ballistic model is performed for gate lengths in the range of 7 nm to 200 nm. In a first step, the long-channel mobility in DD simulation is calibrated to the subband-BTE reference at on-state when drain voltage (V_d) and gate voltage (V_g) are 0.05V and 0.5V, respectively. Differences of the inversion charge integral within channel region between subband-BTE and DD simulations at on-state are compensated by adjusting the workfunction prior to the long-channel mobility calibration.

Second, the channel access resistance (R_0) is extracted from the gate length (L_{gate}) dependence of the total resistance (R) using the equation $R = R_0 + R'L_{gate}$. The access resistance difference between the subband-BTE and DD simulations can be added into the DD simulation either as an additional lumped resistance or by modifying the mobility model parameters in the source and drain regions. We then compensate this access resistance difference together with the introduction of the quasi-ballistic correction. Afterwards, the k parameter of the KVM is adjusted when keeping α_T and α_B as 1.52 [13] and 0.3, respectively, to match the on-state current at low-drain voltage. The α_T and α_B are proposed in [12], which have been demonstrated to achieve a good fit with respect to Monte Carlo and subband-BTE tools. The v_{sat} and β parameters of HFS model are optimized to match the on-state current at high-drain bias voltages ($V_d = 0.5V$; $V_g = 0.5V$) over the range of gate lengths. In general, several iterations through the calibration procedure previously mentioned to optimize the parameters are needed. It is also noted that the KVM and HFS model parameters can depend on the carrier polarity and the material composition.

Figure 2 presents the channel length dependent resistance for Si n-NW and p-NW at linear regime ($V_d = 0.05V$; $V_g = 0.5V$) between the subband-BTE reference and the DD simulations. The resistance fits well to subband-BTE result after going through the above calibration sequences for KVM and HFS model parameters. The results for DD simulation without quasi-ballistic correction and high-field saturation parameter adjustment (HFS model is with default
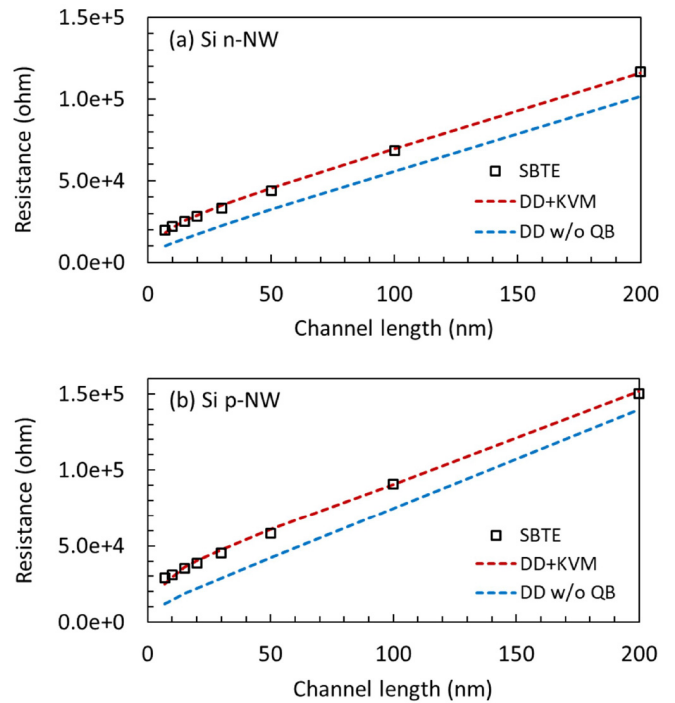

FIGURE 2. Channel-length dependent resistance for (a) Si n-NW and (b) Si p-NW at linear regime ($V_d = 0.05V$; $V_g = 0.5V$) simulated using subband-BTE, DD without quasi-ballistic correction, and DD+KVM.

TABLE 1. Extracted mobility (cm^2/Vs) for Si and Ge n- and p-NW with $5 \times 5\ nm^2$ cross-sectional area and 200 nm-long channel length using subband-BTE, DD without quasi-ballistic correction, and DD + KVM.

Type	Subband-BTE	DD without quasi-ballistic	DD+KVM
Si n-NW	207	218	212
Si p-NW	193	183	194
Ge n-NW	362	403	368
Ge p-NW	328	333	343

parameters [16], [19]) are plotted as well for comparison. The resistance difference between subband-BTE and DD without quasi-ballistic correction for NW with 200 nm-long channel length mostly comes from the un-calibrated mobility in source and drain regions.

Our calibration methodology is devoted to implement and calibrate KVM/HFS model parameters for quasi-ballistic correction at on-state. Therefore, when voltages other than the on-state voltage are considered, the inversion charge compensation has to be performed at this specific gate voltage prior to the drain current extraction.

V. RESULTS

Table 1 summarizes the extracted mobilities for Si and Ge n- and p- NW with a cross-sectional area of 5x5 nm² for 200 nm-long channel length. For Si, the mobilities for n- and p-NW are 207 cm^2/Vs and 193 cm^2/Vs, respectively, using subband-BTE. In DD simulation the mobilities for 200 nm-channel are 218 cm^2/Vs and 183 cm^2/Vs, respectively, for n- and p-NW prior to the quasi-ballistic correction. After the
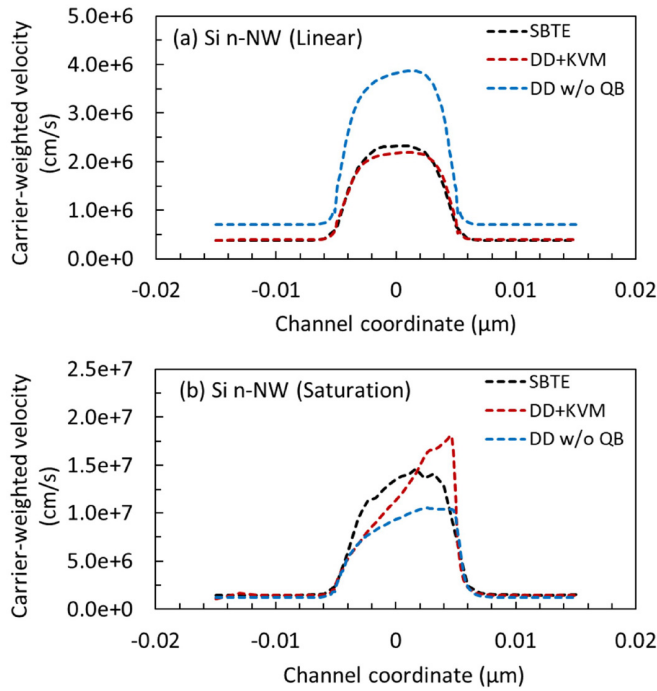
FIGURE 3. Carrier-weighted velocity along the channel for a Si n-NW with 10 nm-long channel length at (a) linear ($V_d = 0.05V$; $V_g = 0.5V$) and (b) saturation ($V_d = 0.5V$; $V_g = 0.5V$) regimes simulated using subband-BTE, DD without quasi-ballistic correction, and DD+KVM.
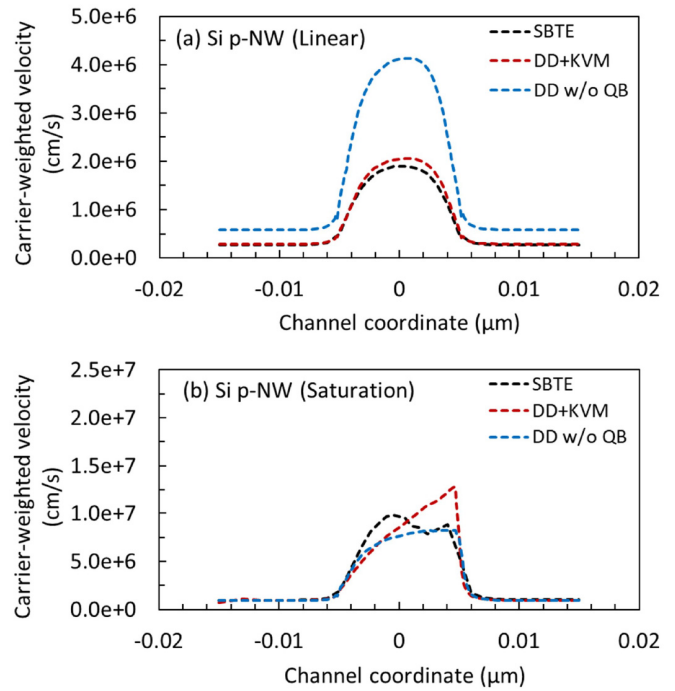

FIGURE 4. Carrier-weighted velocity along channel for a Si p-NW with 10 nm-long channel length at (a) linear ($V_d = 0.05V$; $V_g = 0.5V$) and (b) saturation ($V_d = 0.5V$; $V_g = 0.5V$) regimes simulated using subband-BTE, DD without quasi-ballistic correction, and DD + KVM.

implementation of KVM with calibrated model parameters, the mobilities are 212 cm^2/Vs and 194 cm^2/Vs, respectively, for n- and p-NW, which are nearly the same as those before the quasi-ballistic correction.

Figures 3 and 4 show the carrier-weighted velocity along the channel coordinate in linear and saturation regimes for 10 nm-long Si n- and p-NW's. For the DD simulations without quasi-ballistic correction, the velocity at low drain bias is overestimated at short channels. While the use of the KVM greatly improves the agreement, the KVM also reduces the high-field velocity. It is then necessary to increase accordingly the saturation velocity in the HFS model. In Fig. 3(b) and 4(b), the velocity profile with KVM is still slightly different to subband-BTE results for high drain bias voltage. This DD velocity profile can be better adjusted through the β parameter in the HFS model for a single device. Nevertheless, in our calibration methodology the priority is to fit the drain current over a wide range of channel length, instead of the velocity profile along channel coordinate for one device. This trade-off is a result of the local character of the high-field saturation model.

The channel length dependent on-state drain currents at low- and high-drain bias voltages with calibrated KVM and HFS parameters in DD simulations for Si n- and p-NW are shown in Fig. 5. A good agreement to subband-BTE results is obtained over a wide range of channel length with one parameter set for the KVM and HFS model combination. Without the quasi-ballistic correction,
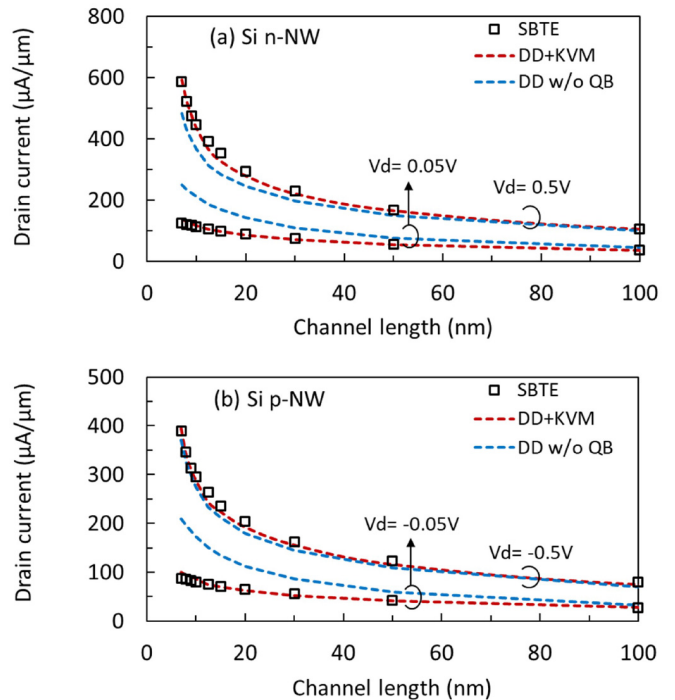

FIGURE 5. Channel length dependence of the drain current for (a) Si n-NW and (b) Si p-NW at linear and saturation regimes simulated using subband-BTE, DD without quasi-ballistic correction, and DD+KVM at $V_g = 0.5V$.

the DD velocity is too large at low drain bias and the drain current is correspondingly overestimated at short channel. The implementation of KVM reduces DD velocity and
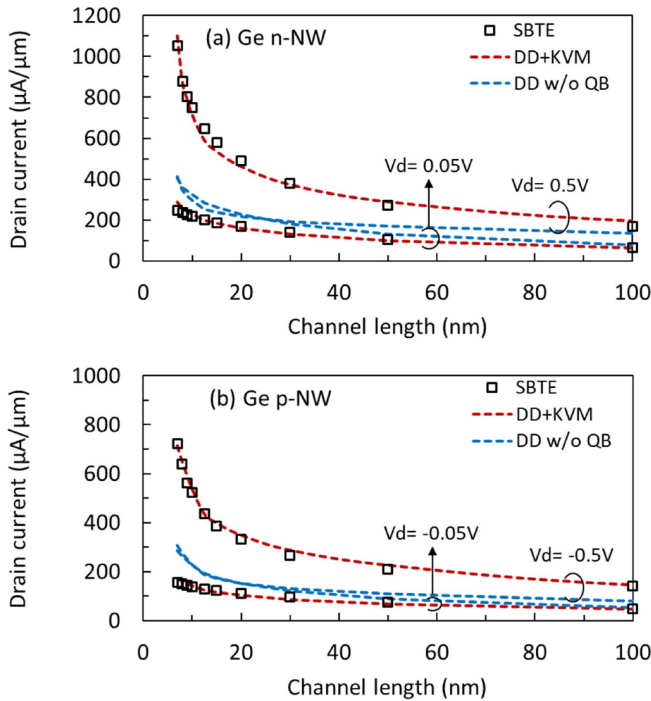
FIGURE 6. Channel length dependence of the drain current for (a) Ge n-NW and (b) Ge p-NW at linear and saturation regimes simulated using subband-BTE, DD without quasi-ballistic correction, and DD+KVM at $V_g = 0.5V$.

FIGURE 7. Calibrated k and v_{sat} parameters of the kinetic velocity model and high field saturation model for a 5x5 nm² SiGe NW.

leads to lower drain current at low drain bias. In order to consider the velocity overshoot due to missing scattering events in the channel, the saturation velocity should be increased from the equilibrium values (1.07e7 *cm/s* for electrons and 0.84e7 *cm/s* for holes in Silicon [16] and 0.65e7 *cm/s* for electrons and 0.60e7 *cm/s* for holes in Germanium [19]) to match the drain current at high drain bias. Without this increase from the equilibrium values, the increase of drain current with decreasing channel length is reduced which can be drastically seen for Ge. In Fig. 6, good agreement of on-state drain currents for DD simulations with calibrated KVM and HFS parameters is also seen with respect to subband-BTE for Ge n- and p-type nanowires.

In Fig. 7, the calibrated parameters of KVM and HFS model are shown for <110> SiGe NW. The k parameter for electrons does not change much for germanium mole fractions up to 0.7 and increases drastically above 0.85 which agrees well with the Ge-mole fraction dependence of the conduction band structure properties. For holes, the k parameter increases smoothly with the germanium mole fraction from 0 to 1, since k is proportional to the square of the inversion of the effective mass. The calibrated values for the saturation velocity used in combination with the KVM are in general higher than the ones obtained without using any ballistic model. From the results of the calibration of the saturation velocity for the long channel devices, one can see that the ballistic effect is an important part of the carrier transport even for medium channel lengths.
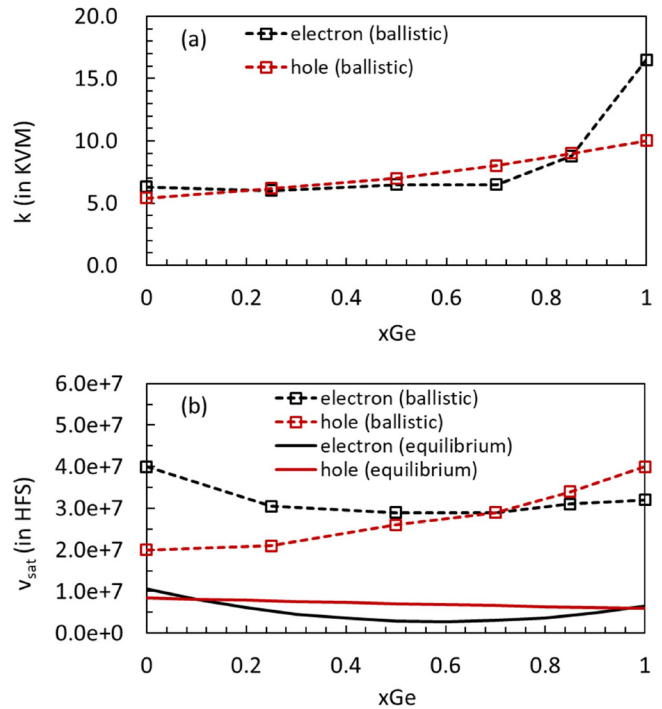
VI. CONCLUSION

We extended the DD simulation of SiGe NWs and related devices by the calibration of the KVM to subband-BTE results. A new methodology to calibrate KVM and HFS model parameters to the solutions of Boltzmann transport and Schrödinger equations has been established. The on-state drain current characteristics agree well with the subband-BTE results over a wide range of channel lengths with a single set of model parameters. In contrast to typical v_{sat} calibration without accounting for the ballistic effect in DD, it has been shown that such a good agreement of DD simulations with subband-BTE results can be achieved without an explicit gate length dependence of DD model parameters using a physically sound parameter set and KVM.

Mole fraction dependent parameters of both KVM and HFS models for <110> SiGe n- and p-type nanowires that can be used for DD simulations have been extracted and presented.

REFERENCES

[1] K. J. Kuhn, "Considerations for ultimate CMOS scaling," *IEEE Trans. Electron Devices*, vol. 59, no. 7, pp. 1813–1828, Jul. 2012, doi: 10.1109/TED.2012.2193129.

[2] H. Mertens *et al.*, "Vertically stacked gate-all-around Si nanowire CMOS transistors with dual work function metal gates," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2016, pp. 19.7.1–19.7.4, doi: 10.1109/IEDM.2016.7838456.

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

[4] H. Mertens *et al.*, "Vertically stacked gate-all-around Si nanowire transistors: Key process optimizations and ring oscillator demonstration," in *IEDM Tech. Dig.*, San Francisco, CA, USA, Dec. 2017, pp. 37.4.1–37.4.4, doi: 10.1109/IEDM.2017.8268511.

[5] M. Choi, V. Moroz, L. Smith, and J. Huang, "Extending drift-diffusion paradigm into the era of FinFETs and nanowires," in *Proc. SISPAD*, Washington, DC, USA, Sep. 2015, pp. 242–245, doi: 10.1109/SISPAD.2015.7292304.

[6] R. Kotlyar, R. Rios, C. E. Weber, T. D. Linton, M. Armstrong, and K. Kuhn, "Distributive quasi-ballistic drift diffusion model including effects of stress and high driving field," *IEEE Trans. Electron Devices*, vol. 62, no. 3, pp. 743–750, Mar. 2015, doi: 10.1109/TED.2015.2392717.

[7] A. Erlebach, K.-H. Lee, and F. M. Bufler, "Empirical ballistic mobility model for drift-diffusion simulation," *Proc. IEEE Conf. ESSDERC*, Lausanne, Switzerland, Sep. 2016, pp. 420–423, doi: 10.1109/ESSDERC.2016.7599675.

[8] M. S. Shur, "Low ballistic mobility in submicron HEMTs," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 511–513, Sep. 2002, doi: 10.1109/LED.2002.802679.

[9] M. Lundstrom and X. Sun, "Some useful relations for analyzing nanoscale MOSFETs operating in the linear regime," 2016. [Online]. Available: http://arxiv.org/abs/1603.03132.

[10] S. Carapezzi, E. Caruso, A. Gnudi, S. Reggiani, and E. Gnani, "TCAD low-field mobility model for InGaAs UTB MOSFETs including quasi-ballistic corrections," in *Proc. IEEE Conf. ESSDERC*, Lausanne, Switzerland, Sep. 2016, pp. 416–419, doi: 10.1109/ESSDERC.2016.7599674.

[11] A. Schenk and P. Aguirre, "TCAD models of the ballistic mobility in the source-to-drain tunneling regime," *Solid-State Electron.*, vol. 157, pp. 1–11, Jul. 2019, doi: 10.1016/j.sse.2019.03.065.

[12] O. Penzin, L. Smith, A. Erlebach, M. Choi, and K.-H. Lee, "Kinetic velocity Model to account for ballistic effects in the drift-diffusion transport approach," *IEEE Trans. Electron Devices*, vol. 64, no. 11, pp. 4599–4606, Nov. 2017, doi: 10.1109/TED.2017.2751968.

[13] W. R. Frensley, "Barrier-limited transport in semiconductor devices," *IEEE Trans. Electron Devices*, vol. 30, no. 12, pp. 1619–1623, Dec. 1983, doi: 10.1109/T-ED.1983.21421.

[14] K.-H. Kao, A. S. Verhulst, W. G. Vandenberghe, B. Soree, G. Groeseneken, and K. De Meyer, "Direct and indirect band-to-band tunneling in germanium-based TFETs," *IEEE Trans. Electron Devices*, vol. 59, no. 2, pp. 292–301, Feb. 2012, doi: 10.1109/TED.2011.2175228.

[15] *Sentaurus Device User Guide*, Synopsys, Inc., Mountain View, CA, USA, Jun. 2018.

[16] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field," *Proc. IEEE*, vol. 55, no. 12, pp. 2192–2193, Dec. 1967, doi: 10.1109/PROC.1967.6123.

[17] *Sentaurus Device QTX Guide*, Synopsys, Inc., Mountain View, CA, USA, Jun. 2018.

[18] M. Frey, J. Huang, F. Heinz, A. Erlebach, L. Smith, and V. Moroz, "Performance analysis of p-type silicon nanowire FETs with silicon-germanium cladding," in *Proc. IEEE Conf. SISPAD*, Oct. 2016, pp. 273–276, doi: 10.1109/SISPAD.2016.7605200.

[19] C. Jacoboni, F. Nava, C. Canali, and G. Ottaviani, "Electron drift velocity and diffusivity in germanium," *Phys. Rev. B, Condens. Matter*, vol. 24, pp. 1014–1026, Jul. 1981, doi: 10.1103/PhysRevB.24.1014.

KO-HSIN LEE, photograph and biography not available at the time of publication.

AXEL ERLEBACH received the M.S. degree in theoretical physics from the Dresden University of Technology, Dresden, Germany, in 1987, and the Ph.D. degree in physics from the University of Duisburg-Essen, Essen, Germany, in 1998.

He is currently involved in the extraction and validation of technology computer-aided design device simulation parameters.

OLEG PENZIN received the M.S. degree in electrical engineering from Novosibirsk Technical University, Novosibirsk, Russia, in 1984.

His current research interests include advanced semiconductor physics, Si and III-V devices, and numerical methods applicable for general TCAD device simulation.

LEE SMITH (Member, IEEE) received the B.S. degree in physics from the University of Florida and the Ph.D. degree in physics from Stanford University.

He is currently Research and Development Manager of the TCAD Device Simulation Group, Synopsys which is engaged in the development of advanced device modeling techniques for a variety of applications including CMOS, power, RF, and memory devices.