

Monte Carlo Simulation of the Silicon-Channel Thickness Dependence of Nanosheet Performance

F. M. Bufler, S. Baudot, S. Yang, P. Matagne, N. Horiguchi, and G. Hellings

IMEC

Leuven, Belgium

Fabian.Bufler@imec.be

Abstract—The silicon-channel thickness (t_{Si}) dependence of n -type and p -type long-channel mobility and short-channel nanosheet performance is investigated by Monte Carlo (MC) simulation. MC is found to be able to reproduce the experimental nanosheet mobility within 5.5 % down to $t_{Si}=2.5$ nm for NMOS and within 4.8 % down to $t_{Si}=3.5$ nm for PMOS without t_{Si} -dependent MC parameters. The mobility degradation for decreasing t_{Si} is due to an increase of the surface-to-volume ratio which leads to a stronger surface impact for the same roughness. Below a gate length (L_G) of 14 nm the propagation delay time (τ_{delay}) for nanosheets with $t_{Si}=3$ nm becomes smaller than τ_{delay} for $t_{Si}=5$ nm. The drain-induced barrier lowering (DIBL) at $t_{Si}=3$ nm is half the DIBL at $t_{Si}=5$ nm below $L_G=14$ nm.

Index Terms—Silicon thickness t_{Si} , contacted poly-pitch (CPP) scaling, gate-all-around (GAA) nanosheet, Monte Carlo, surface roughness.

I. INTRODUCTION

The nanosheet has replaced the FinFET in order to continue scaling by exploiting its GAA nature to reduce the short-channel effect, i.e. the increase of the off-current (IOFF) for decreasing L_G or, equivalently, the decrease of the on-current (ION) at constant IOFF. One possibility for further scaling and keeping the short-channel effect under control is to shrink t_{Si} of the nanosheet. This approach, however, involves a mobility degradation. In order to quantify this penalty nanosheet devices with t_{Si} down to 1.5 nm have recently been fabricated and measured [1].

Short-channel performance is affected by more uncertainties (e.g. on doping) and device simulation offers a way to systematically quantify on-state performance upon scaling, but only drift-diffusion simulation with t_{Si} -dependent calibration has been done [2]. Semiclassical approaches involve band structures which deviate below $t_{Si}=5$ nm from atomistic band models as used in NEGF simulation [3]. At the same time there is a need for device simulations which describe quasi-ballistic transport on a physical basis, are relatively fast and can treat realistic device structures [4]. We therefore investigate to what extent MC simulation is able to reproduce the measured thin-channel nanosheet mobilities and, based on this validation, assess the short-channel performance of thin-channel nanosheet devices.

II. MONTE CARLO APPROACH

The MC simulations in this work [5] use nonlocal empirical pseudopotential band structures including spin-orbit interaction. Acoustic and optical phonon scattering with constant

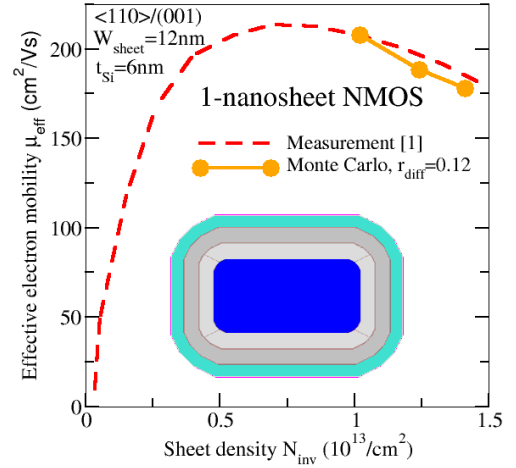


Fig. 1. Long-channel ($L_G=0.5 \mu\text{m}$ in MC simulation [4]) effective electron mobility as a function of inversion sheet density. MC simulations of the nanosheet are compared to measurements [1]. The ratio for diffusive surface scattering r_{diff} is calibrated at an inversion sheet density of $N_{inv}=1\text{e}13 \text{ cm}^{-2}$ to 12 % (corresponding to 88 % specular surface scattering).

deformation potentials, ionized impurity scattering according to the Brooks–Herring model and a combination of specular and diffusive surface scattering [4] is included. Size quantization is considered in terms of modified oxide permittivity and work functions extracted from density–gradient simulation of the device cross-section which accurately reproduces the quantum reference [6]. Source-to-drain tunneling is not taken into account.

III. EFFECTIVE MOBILITY

Our approach consists of calibrating process-related parameters at one point (here: surface roughness in terms of the diffusive ratio r_{diff} upon surface scattering at $t_{Si}=6$ nm and a sheet density of $N_{inv}=1\text{e}13 \text{ cm}^{-2}$) and then to simulate without further MC parameter changes dependencies (here: on N_{inv} or on t_{Si}) for the purpose of validation. The calibration leads in Fig. 1 to $r_{diff}=0.12$ and is close to $r_{diff}=0.15$ reproducing measured mobilities in FinFETs with a gate stack composed of SiO_2 and polysilicon [4] (these values for r_{diff} correspond to root-mean-square (rms) heights of $\Delta_{rms}=0.28$ nm and $\Delta_{rms}=0.32$ nm, respectively, as detailed in the Appendix). This

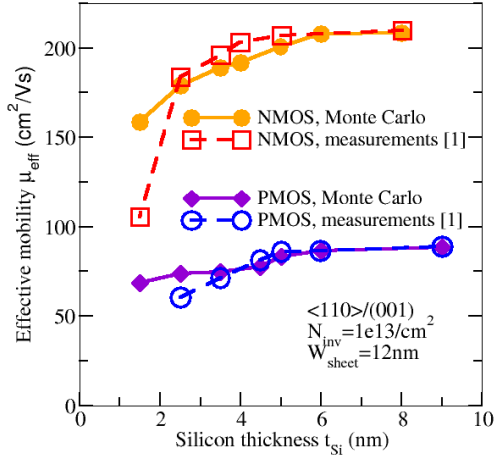


Fig. 2. Effective electron and hole mobilities as function of silicon thickness t_{Si} according to measurements [1] and MC simulations with a diffusive ratio of $r_{diff}=0.12$. The relative mobilities reported in measurements [1] are multiplied with the simulated absolute mobilities at $t_{Si}=6\text{nm}$ where the electron mobility has been calibrated to the measured absolute mobility as shown in Fig. 1.

suggests that surface roughness is also for nanosheets with a $\text{SiO}_2/\text{HfO}_2$ gate stack the dominant mobility degradation mechanism instead of e.g. remote Coulomb scattering from charges at the $\text{SiO}_2/\text{HfO}_2$ interface.

Fig. 2 shows good agreement between measurements [1] and MC simulations for the mobility dependence on t_{Si} (the same surface roughness as for NMOS is assumed for the PMOS MC mobilities) except for an overestimation by MC at the smallest t_{Si} . Note, however, that this may also be related to a process-related roughness increase at very small t_{Si} . This might indeed be indicated by the irregular shape of the thinnest nanosheets seen in the TEM images in Fig. 5 of [1].

The mobility degradation depends in our MC model apart from corner rounding only on the sidewall orientation and the surface-to-volume ratio given by

$$r_{s/v} = (2t_{Si} + 2W_{sheet}) / (t_{Si}W_{sheet}). \quad (1)$$

The electron mobility decreases despite the smaller unfavorable (110) sidewall contribution [4]. This shows that the increase of $r_{s/v}$ (e.g. 0.83 at $t_{Si}=3\text{ nm}$ versus 0.5 at $t_{Si}=6\text{ nm}$) is responsible for the electron mobility degradation which increases the surface impact for the same surface roughness.

IV. SHORT-CHANNEL PERFORMANCE

Fig. 3 shows how the increase of the subthreshold swing (SS) upon L_G reduction decreases ION for constant IOFF (note that our definition of L_G includes 1.3 nm sidewall HfO_2 on both spacer sides). This reverses the initial ION increase upon reducing L_G seen in Fig. 4 in the long-channel regime with negligible SS degradation, which results from a higher driving field at smaller L_G and hence a larger channel velocity. Note, however, that Fig. 4 refers for smaller L_G to the maximum achievable ION, since source-to-drain tunneling is

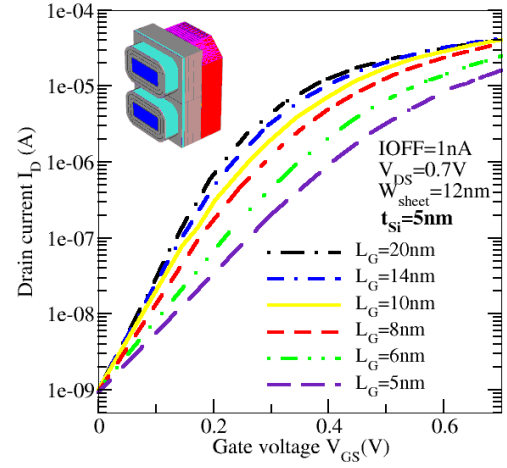


Fig. 3. MC NMOS transfer characteristics at $V_{DS}=0.7\text{ V}$ of the device in the inset for $t_{Si}=5\text{ nm}$ with gate lengths from $L_G=20\text{ nm}$ down to $L_G=5\text{ nm}$. The subthreshold swing (SS) increases upon L_G scaling and then degrades ION at constant IOFF. Contact resistivity and doping profiles are as calibrated to state-of-the-art FinFET measurements [7].

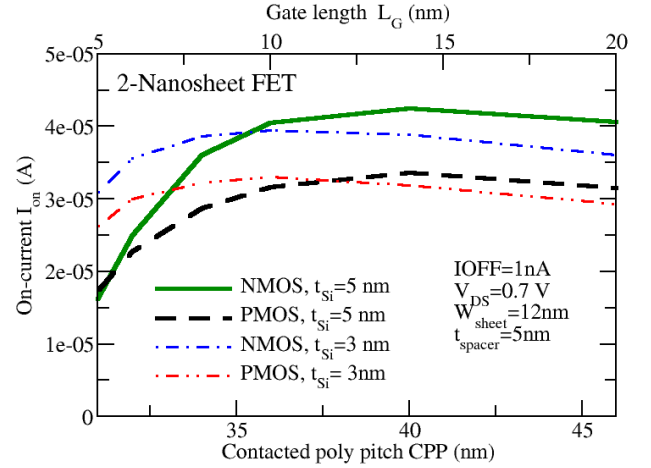


Fig. 4. ION as a function of gate length L_G (top) and contacted poly pitch CPP (bottom) for the device in the inset of Fig. 3 at $t_{Si}=5\text{ nm}$ ($W_{eff}=68\text{ nm}$) and $t_{Si}=3\text{ nm}$ ($W_{eff}=60\text{ nm}$). The maximum ION is shifted from $L_G=14\text{ nm}$ at $t_{Si}=5\text{ nm}$ to $L_G=10\text{ nm}$ at $t_{Si}=3\text{ nm}$. The only difference in doping compared to [7] is that here also the PMOS channel is undoped.

neglected which degrades SS and would therefore reduce ION at constant IOFF.

As performance metric in device simulation we then study the propagation delay time [8]

$$\tau = \frac{CV}{I}, \quad (2)$$

but in contrast to [8] not only the inversion capacitance, but also the parasitic gate-drain overlap capacitance [9] is considered in terms of

$$C = C_{inv} + 2C_{gd,over}, \quad (3)$$

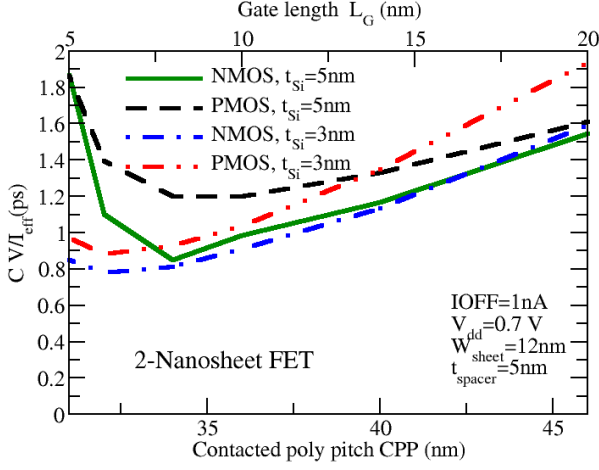


Fig. 5. Propagation delay time defined as gate capacitance C at $V_{GS}=V_{dd}$ and $V_{DS}=0$ V times supply voltage V divided by the effective current I_{eff} as function of L_G (top) and CPP (bottom). The propagation delay for $t_{Si}=3$ nm becomes smaller below $L_G=14$ nm.

which are calculated from the gate capacitance C_g according to $C_{inv}=C_g(V_{dd}) - C_g(0)$ and $C_{gd,over}=C_g(0)/2$. We also use in (3) for V the supply voltage V_{dd} and for the current I instead of I_{ON} [8] the effective current I_{eff} [10]. The resulting propagation delay time in Fig. 5 reveals superior performance for $t_{Si}=3$ nm below $L_G=14$ nm.

Finally, Fig. 6 shows that the DIBL is reduced by more than 50 % for $t_{Si}=3$ nm below $L_G=14$ nm.

V. CONCLUSIONS

In this work, quantum-corrected semiclassical MC device simulations have been performed for nanosheets with t_{Si} down to 1.5 nm. Good agreement with measured electron and hole mobilities for t_{Si} down to around 3 nm was found without t_{Si} -dependent MC parameters. This shows that MC is still accurate in this regime and can be used for performance estimations. The study of propagation delay and DIBL suggests that performance for $t_{Si}=3$ nm becomes below $L_G=14$ nm superior to that at $t_{Si}=5$ nm.

APPENDIX

Based on the original work of Fuchs for electrons in metals [11] surface roughness scattering is modeled in this work as a combination of specular scattering, where energy and parallel wave-vector are conserved, and diffusive scattering where only the energy is conserved and the direction of the wave-vector after hitting the surface is chosen at random. The surface roughness is here specified by the diffusive ratio r_{diff} or equivalently by the specular ratio

$$r_{spec} = 1 - r_{diff}. \quad (A.1)$$

Another approach uses an explicit scattering rate in terms of an autocorrelation function [12] where surface roughness is governed by the root-mean-square (rms) height Δ_{rms} .

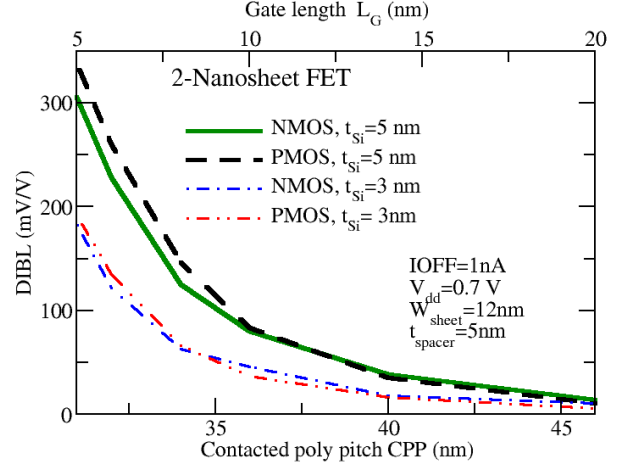


Fig. 6. DIBL as function of L_G (top) and CPP (bottom) for $t_{Si}=5$ nm and $t_{Si}=3$ nm. Shrinking t_{Si} from 5 nm to 3 nm roughly reduces the DIBL by a factor of two at and below $L_G=14$ nm.

In order to facilitate the comparison between these two approaches we specify below the general relation of Ziman between r_{spec} and Δ_{rms} which is given by [13]

$$r_{spec} = \exp\left(-\frac{16\pi^2\Delta_{rms}^2}{\lambda^2}\right). \quad (A.2)$$

Here, λ is the wavelength of the particle under consideration. In our case it is the De Broglie wavelength of electrons in the conduction band of silicon with a standard value of 10 nm [14]. For example, $r_{diff}=0.12$ corresponds to $\Delta_{rms}=0.28$ nm and $r_{diff}=0.15$ to $\Delta_{rms}=0.32$ nm. Note that (A.2) also opens the possibility to use measured values of Δ_{rms} [15] in MC simulations which employ a combination of specular and diffusive surface scattering.

ACKNOWLEDGMENT

The authors would like to thank B. Vermeersch and A. Pondini for pointing them to the references [13] and [15], respectively.

REFERENCES

- [1] A. Agrawal, W. Chakraborty, W. Li, H. Ryu, B. Markman, S. H. Hoon, R. K. Paul, C. Y. Huang, S. M. Choi, K. Rho, A. Shu, R. Iglesias, P. Wallace, S. Ghosh, K. L. Cheong, J. L. Hockel, R. Thorman, L. Baumgartel, L. Shoer, V. Mishra, S. Berrada, A. Ashita, C. Weber, B. Obradovic, A. A. Oni, Z. Brooks, N. Franco, J. Kavalieros, and G. Dewey, "Silicon RibbonFET CMOS at 6nm gate length," in *IEDM Tech. Dig.*, San Francisco (USA), Dec. 2024, pp. 2–2.
- [2] K. K. Bhuwalka, L.-C. Hung, G. Rzepa, F. Schanovsky, Z. Stanojevic, L. Hou, W. Li, M. Karner, and Z. Yu, "Contacted-poly-pitch (CPP) scaling limitations of Si gate-all-around nanosheet (GAA-NS) devices," *IEEE Trans. Electron Devices*, vol. 72, no. 12, pp. 6437–6446, Dec. 2025.
- [3] M. Luisier, N. Vetsch, V. Maillou, A. Winka, L. Deuschle, C. H. Xia, M. Kaniselman, M. Mladenović, J. Cao, and A. N. Ziogas, "Acceleration of atomistic NEGF: Algorithms, parallelization, and machine learning," in *Proc. Int. Conf. Simulation Semiconductor Process. Devices (SISPAD)*, Grenoble (France), Sep. 2025, pp. 288–291.
- [4] F. M. Bufler, F. O. Heinz, and L. Smith, "Efficient 3D Monte Carlo simulation of orientation and stress effects in FinFETs," in *Proc. Int. Conf. Simulation Semiconductor Process. Devices (SISPAD)*, Glasgow (UK), Sep. 2013, pp. 172–175.

- [5] Synopsys Inc., “Sentaurus Device Monte Carlo User Guide,” Release X-2025.09, Mountain View (USA), 2025.
- [6] F. M. Bufler and L. Smith, “3D Monte Carlo simulation of FinFET and FDSOI devices with accurate quantum correction,” *J. Comput. Electron.*, vol. 12, no. 4, pp. 651–657, Dec. 2013.
- [7] F. M. Bufler, D. Jang, G. Hellings, G. Eneman, P. Matagne, A. Spesot, and M. H. Na, “Monte Carlo comparison of n-type and p-type nanosheets with FinFETs: Effect of the number of sheets,” *IEEE Trans. Electron Devices*, vol. 67, no. 11, pp. 4701–4704, Nov. 2020.
- [8] M. Bohr, “MOS transistors: Scaling and performance trends,” *Semiconductor International*, vol. 18, pp. 75–79, Jun. 1995.
- [9] K. Nose, S.-I. Chae, and T. Sakurai, “Voltage dependent gate capacitance and its impact in estimating power and delay of CMOS digital circuits with low supply voltage,” in *Proc. Int. Symp. Low Power Electronics Design (ISLPED)*, Rapallo (Italy), Jul. 2000, pp. 228–230.
- [10] M. H. Na, E. J. Nowak, W. Haensch, and J. Cai, “The effective drive current in CMOS inverters,” in *IEDM Tech. Dig.*, San Francisco (USA), Dec. 2002, pp. 121–124.
- [11] K. Fuchs, “The conductivity of thin metallic films according to the electron theory of metals,” *Math. Proc. Camb. Philos. Soc.*, vol. 34, pp. 100–108, 1938.
- [12] T. Ando, A. B. Fowler, and F. Stern, “Electronic properties of two-dimensional systems,” *Rev. Mod. Phys.*, vol. 54, no. 2, pp. 437–672, Apr. 1982.
- [13] H. J. Lee, J. S. Jin, and B. J. Lee, “Assessment of phonon boundary scattering from light scattering standpoint,” *J. Appl. Phys.*, vol. 112, no. 6, 063513, 2012.
- [14] H. Watanabe, “A theoretical study on characteristic wavelength of silicon conduction electrons,” *AIP Advances*, vol. 15, no. 5, 055239, 2025.
- [15] S. Karapetyan, S. E. Zeltmann, G. Wilk, T.-K. Chen, V. D.-H. Hou, and D. A. Muller, “3D atomic-scale metrology of strain relaxation and roughness in gate-all-around transistors via electron ptychography,” *Nat. Commun.*, vol. 17, no. 1, 3561, Feb. 2026.