Solid-State Electronics 197 (2022) 108438

Contents lists available at ScienceDirect

Solid State Electronics

journal homepage: www.elsevier.com/locate/sse

Tunneling leakage in ultrashort-channel MOSFETs—From atomistics to continuum modeling

Raphaël J. Prentki [a,b,*], Mohammed Harb [a], Chenyi Zhou [a], Pericles Philippopoulos [a], Félix Beaudoin [a], Vincent Michaud-Rioux [a], Hong Guo [b,a]

[a] Nanoacademic Technologies Inc., 666 Rue Sherbrooke Ouest, Suite 802, Montréal H3A 1E7, QC, Canada
[b] Department of Physics, McGill University, Ernest Rutherford Physics Building, 3600 Rue University, Montréal H3A 2T8, QC, Canada

ARTICLE INFO

Keywords:
Ultrashort-channel MOSFET
Direct source-to-drain tunneling
Bound-charge engineering
Nonequilibrium Green's function formalism
Atomistics
Finite-element method

ABSTRACT

The channel lengths of transistors are now nearing the nanometer, making these devices increasingly prone to direct source-to-drain tunneling (DSDT), a leakage mechanism commonly considered to set the end of Moore's law. In MOSFETs, the probability for a charge carrier to undergo DSDT decays exponentially with channel length, source depletion length, and drain depletion length. Bound-charge engineering (BCE) is a recently introduced scheme where the depletion lengths of FETs can be controlled through effective doping by surface bound charges residing on the interface between a semiconductor and an adjacent oxide. In this letter, BCE is applied to reduce DSDT leakage current down to acceptable levels in MOSFETs with channels as short as 2.3 nm; the higher the oxide permittivity, the lower the DSDT leakage. This idea is tested on ultrascaled Si nanowire MOSFETs via atomistic quantum transport simulations based on the nonequilibrium Green's function (NEGF) formalism and the tight-binding model, as well as on physically larger Si nanosheet MOSFETs via continuum NEGF–**k**·**p** simulations based on the finite-element method.

1. Introduction

The channel lengths of state-of-the-art Si MOSFETs are now being downscaled to $L \approx 10$ nm [1]. Quantum transport phenomena play a crucial role in the ultrashort-channel limit. For example, direct source-to-drain tunneling (DSDT) [2]—the quantum tunneling of charge carriers through a MOSFET's channel, which leads to significant leakage current—is generally considered to set the end of Moore's law [3]. DSDT can be inhibited by (1) engineering a high tunneling effective mass (EM), as may be done via selection of material, selection of crystal orientation, and strain [4], and (2) engineering a wide channel potential barrier, as may be done via localized channel doping [5]. Following the second alternative, we propose to inhibit DSDT through bound-charge engineering (BCE). BCE is a relatively simple and generally applicable scheme where low-dimensional semiconductors, such as Si nanowires (NWs) and nanosheets (NS), are effectively doped by surface bound charges residing on interfaces between Si and oxides; these charges are harnessed to reduce source–channel and channel–drain depletion lengths [6,7] and therefore tunneling lengths.

2. Device Structures and Simulation Methods

As vehicles for this study, we consider n-type Si NW and Si NS MOSFETs [Fig. 1(a)]. Their sources and drains are doped at $N_D = 1 \times 10^{20}$ cm^{-3} while their channels are intrinsic. The devices are gate-all-around with an oxide thickness of $t = 2$ nm. The Si NW and NS are grown in [110]; the NW has a diameter of $d = 2$ nm [Fig. 1(b)]; the NS has a width of $w = 15$ nm and a height of $h = 5$ nm [Fig. 1(c)]. Unless otherwise indicated, in all simulations, the drain voltage and device temperature are taken to be $V_{DS} = 50$ mV and $T = 300$ K, respectively.

Due to their ultrascaled dimensions, the NW devices are modeled via state-of-the-art self-consistent atomistic quantum transport simulations based on the nonequilibrium Green's function (NEGF) formalism [8] and the *spds** tight-binding (TB) model [9] as implemented in the first-principles quantum transport software NanoDCAL+ [6,10–12]. On the other hand, the silicon NS devices, which are physically larger, are simulated within continuum models. Specifically, we use the NEGF–**k**·**p** feature of the QTCAD software suite (Quantum-Technology Computer-Aided Design), a quantum hardware modeling tool based on the finite-element method (FEM) [12–14]. The parameters of the **k**·**p**

* Corresponding author.
E-mail address: rprentki@nanoacademic.com (R.J. Prentki).

https://doi.org/10.1016/j.sse.2022.108438
Received 14 July 2022; Accepted 18 August 2022
Available online 24 August 2022

model—notably the Si NS's EMs and bandgap energy—were extracted from TB bandstructure calculations.

3. Thermionic Emission and Tunneling

In relatively-long-channel MOSFETs, current is solely carried by electrons with energy $E > \phi$, where ϕ is the channel barrier energy [Fig. 2(a)]. In the off state, these electrons obey Boltzmann statistics; the drain current then satisfies [15]

$$I_{DS} \propto e^{-\frac{q\phi-\mu_S}{k_B T}}, \tag{1}$$

where $q > 0$ is the elementary charge, $\mu_{S,D}$ is the source/drain Fermi level, and k_B is Boltzmann's constant; this is known as thermionic emission. In ultrashort-channel MOSFETs [Fig. 2(b)], the Wentzel–Kramers–Brillouin approximation predicts that the tunneling probability of an electron of energy E through the source depletion region, channel, and drain depletion region (with lengths $L_{S,C,D}$, respectively) satisfies [15]

$$T_{\mathrm{DSDT}}(E) \propto e^{-\frac{1}{\hbar}\sqrt{m^*_{e,t}(\phi-E)}\sum_{i=S,C,D} a_i L_i}, \tag{2}$$

where $\hbar = h/(2\pi)$ is the reduced Planck constant, $m^*_{e,t}$ is the tunneling EM, and $a_{S,C,D}$ are dimensionless constants; for the channel's square barrier, $a_C = 2\sqrt{2}$.

In the BCE scheme, the depletion lengths $L_{S,D}$ can be controlled through the permittivity κ of the oxide adjacent to the source and drain depletion regions [6]. Specifically, it can be shown that $L_{S,D} \propto \sqrt{\kappa\phi}$ [7]. This calls for the MOSFET design shown in Fig. 1(a), where the source, channel, and drain are all adjacent to a high-κ oxide. A high-κ gate oxide enables high gate control [16]; a high-κ source and drain oxide enables low T_{DSDT} [Fig. 2(c)]. Parenthetically, in modern nanoelectronics, low-κ spacer oxides are often placed between the source and gate contacts to minimize load capacitance [17]. This is compatible with the DSDT inhibition scheme presented here, since only the source and drain depletion regions, which are typically only a few tens of nanometers in length, need to be surrounded by a high-κ oxide.

4. Simulations of Ultrashort-Channel MOSFETs

Fig. 3(a) shows the transfer characteristics of three Si NW MOSFETs with a channel length of $L = 2.3$ nm and various κ. The gate voltage V_{GS} is shown in reference to the off-state voltage V_{off}, defined so that the drain current I_{DS} of the three MOSFETs equals $I_{\mathrm{off}} \overset{\mathrm{def}}{=} 10^{-5}$ A·m^{-1} [1] at $V_{GS} - V_{\mathrm{off}} = 0$ V. A relevant metric for device performance is the subthreshold swing (STS):

$$S = \left[\underbrace{\left(\frac{\partial \log_{10} I_{DS}}{\partial \phi}\right)}_{\text{Transport factor}} \overbrace{\left(\frac{\partial \phi}{\partial V_{GS}}\right)}^{\text{Body factor}}\right]^{-1}. \tag{3}$$

The MOSFET with SiO_2 (red curve) has an STS roughly double that of the other devices [Fig. 4(b)], which significantly increases power dissipation [15]. The STSs of the devices with high-κ oxides approach those of long-channel devices [Fig. 3(b)] and the thermal limit of 60 mV·dec^{-1} [15], suggesting that charge transport in these devices is at least partly carried by high-energy electrons, as in thermionic emission, despite their ultrashort channels. In addition, the devices with high-κ oxides have power supply voltage $V_{DD} < 0.55$ V, which is well within the acceptable range for the next generations of transistors, as recommended by the IRDS$^{\mathrm{TM}}$[1]. By contrast, the Si NS devices have much poorer performance in the ultrashort-channel limit [Fig. 3(c)]. This is in great part due to the NS's large cross section, which results in poor gate control and small body factor [Eq. 3] when $L < 10$ nm [15].

Quantitatively, what degrades the STS of ultrashort-channel Si NW MOSFETs with low-κ oxides most strongly: poor gate control or high DSDT current? The DSDT current can be computed through the Landauer–Büttiker formula [8]:

$$I_{\mathrm{DSDT}} = \frac{q}{h}\int_{E_{C,S}}^{\phi} \overline{T}(E)[f_S(E) - f_D(E)]\,dE, \tag{4}$$

where $E_{C,S}$ is the value of the conduction band edge in the source, $\overline{T}(E)$ is the transmission function, and $f_{S,D}(E)$ are the source and drain Fermi–Dirac distributions. I_{DSDT}/I_{DS} is graphed in Fig. 4(a). Expectedly, the MOSFET with SiO_2 exhibits a higher proportion of DSDT. Furthermore, in the subthreshold regime, DSDT is by far the dominant charge transport mechanism, even in the devices with high-κ oxides. In terms of STS [Fig. 4(b)], both the transport [Fig. 4(c)] and body [Fig. 4(d)] factors are lower by around 20% in the MOSFET with SiO_2 compared to the MOSFETs with high-κ oxides.

What explains the significantly lower transport factor of the device with SiO_2? The transport factors of the MOSFETs with high-κ oxides are close to the thermionic emission value of $1/(k_B T \ln 10) \approx 16.8$ eV^{-1} [Eq. 1] even though most of their current is due to DSDT [Fig. 4(a)]. To understand this perplexing observation, consider that for DSDT, one has

$$\overline{T}(E) \propto e^{-\frac{\lambda}{\hbar}\sqrt{m^*_{e,t}(\phi-E)}}, \quad f_S(E) - f_D(E) \propto e^{-\frac{E-\mu_S}{k_B T}}, \tag{5}$$

where $\lambda = \sum_{i=S,C,D} a_i L_i$ [Eq. 2]. Note that $\overline{T}(E)$ $(f_S(E) - f_D(E))$ is an exponentially increasing (decreasing) function of E. In the low-κ (high-κ) case, λ is small (large) [Fig. 5(a)], so that most current is carried at low (high) energies close to $E = \mu_S$ $(E = \phi)$ [Fig. 5(b)]. Thus, the DSDT current in the high-κ case, which could be described as "thermal DSDT," obeys Eq. 1, thereby explaining the transport factor of 16.8 eV^{-1}. In the low-κ case, the channel energy barrier can be approximated to be a square barrier with width $L_S + L_D$ $(L \approx 0)$ and height ϕ. Thus, from Eq. 2 and an analytic formula for $L_{S,D}$ derived in a previous work [7], it can be approximated that

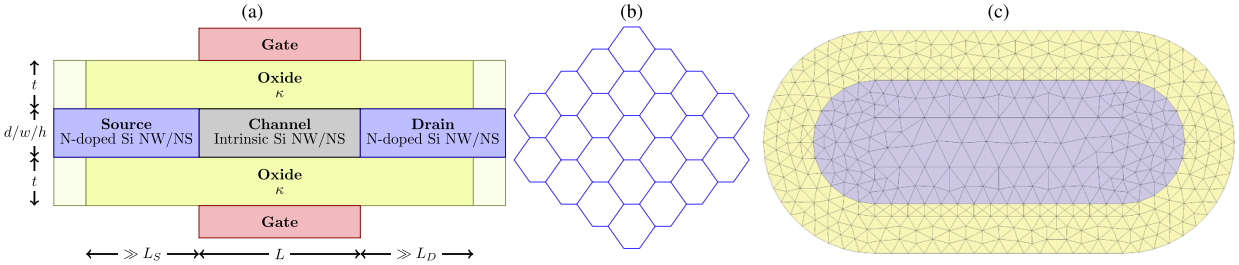

Fig. 1. (a) Schematic of the MOSFETs investigated in this letter. A cross section through the devices' axis of rotational symmetry is shown. An oxide with permittivity κ surrounds the channel as well as the source and drain depletion regions (with lengths $L_{S,D}$). The nature of the (spacer) oxide located far from these regions (light-yellow regions) does not affect DSDT. (b) Atomic structure of the unit cell of the Si NW. (c) FEM mesh of a cross section of the Si NS MOSFET.

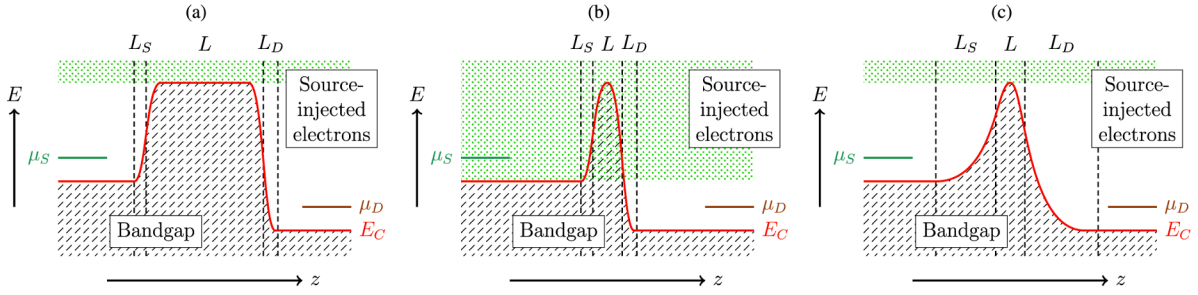

Fig. 2. Illustrations of off-state band diagrams (conduction band edge E_C as a function of position z along the transport direction) in the long-channel limit (a) and short-channel limit (b) of MOSFETs. Long-channel MOSFETs only exhibit thermionic emission, while short-channel MOSFETs exhibit both thermionic emission and DSDT (green-dotted region). If the source and drain depletion lengths $L_{S,D}$ are sufficiently long (c), DSDT can be significantly reduced.

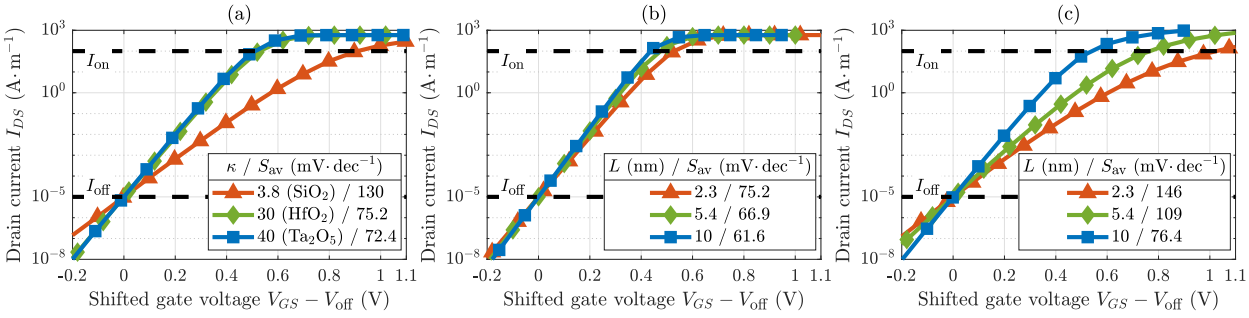

Fig. 3. (a) Transfer characteristics of Si NW MOSFETs with $L = 2.3$ nm and various κ as obtained from NEGF–TB simulations. (b) Transfer characteristics of Si NW MOSFETs with various L and $\kappa = 30$ (HfO_2) as obtained from NEGF–TB simulations. (c) Transfer characteristics of Si NS MOSFETs with various L and $\kappa = 30$ (HfO_2) as obtained from NEGF–**k**·**p** simulations. The average STS S_{av} is computed between I_{off} and $I_{on} \overset{def}{=} 10^2$ A·m^{-1} [1].

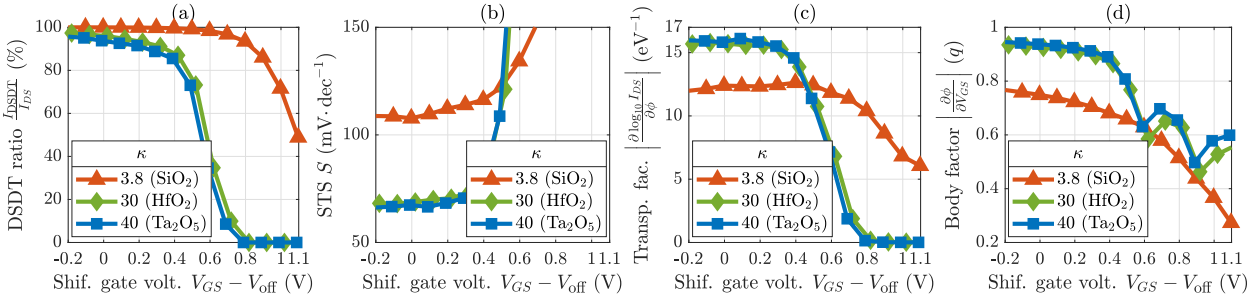

Fig. 4. Metrics pertaining to charge transport in Si NW MOSFETs with $L = 2.3$ nm and various κ as obtained from NEGF–TB simulations.

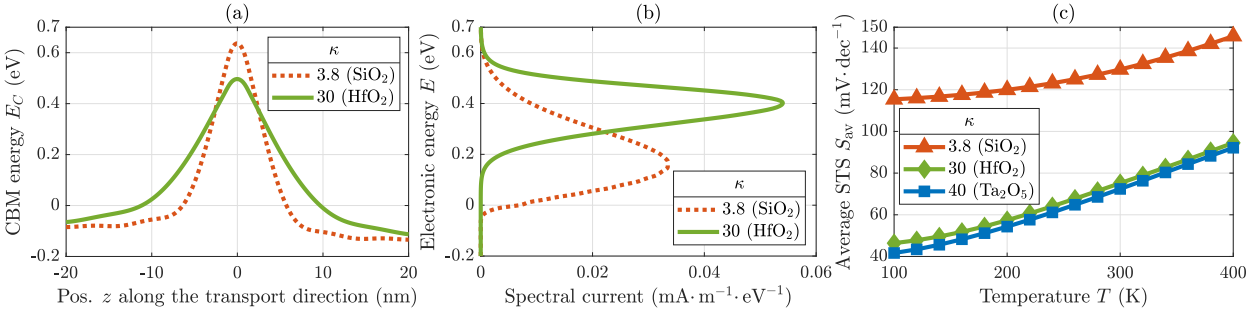

Fig. 5. (a) Band diagrams and (b) spectral current, i.e. $\frac{q}{h}T(E)\left[f_S(E) - f_D(E)\right]$, of Si NW MOSFETs with $L = 2.3$ nm, various κ, and $V_{GS} - V_{off} = 0$ V, as obtained from NEGF–TB simulations. It is assumed that $\mu_S = 0$. (c) Average STS of Si NW MOSFETs with $L = 2.3$ nm and various κ as a function of T as obtained by NEGF–TB simulations.

$$\left|\frac{\partial \log_{10} I_{DS}}{\partial \phi}\right| \approx \frac{1}{\ln 10}\frac{2a_C}{\hbar}\sqrt{\frac{2\kappa_{Si}\varepsilon_0 m^*_{e,t}}{q^2 N_D}}, \quad (6)$$

where $\kappa_{Si}\varepsilon_0$ is the Si permittivity. This results in a transport factor of 10.5 eV^{-1}, which is close to that observed in NEGF–TB simulations for the device with SiO_2 [Fig. 4(c)]. Furthermore, since Eq. 6 does not exhibit a *T*-dependence, charge transport in the low-κ case could be described as "athermal DSDT." To further corroborate this description of DSDT, the devices with HfO_2 and Ta_2O_5 exhibit a nearly linear dependence of average STS S_{av} on *T* [Fig. 5(c)]. In contrast, the device with SiO_2 exhibits a weaker dependence of S_{av} on *T*. In this case, charge transport fits somewhere between the two limiting cases of thermal and athermal DSDT. This is consistent with the observation that the mode of the spectral current for this device is at an energy slightly higher than μ_S, as can be seen in Fig. 5(b).

5. Conclusion

In this letter, we investigated ultrascaled MOSFETs with channels as small as 2.3 nm in length through atomistic (NEGF–TB) and FEM (NEGF–**k**·**p**) quantum transport simulations. When low-κ oxides are used, performance severely deteriorates as the channel is shortened. This is due to (1) poor gate control and (2) high DSDT current. High-κ oxides feed two birds with one seed in that they both increase gate control (due to greater gate capacitance) and reduce DSDT current (due to longer tunneling length, thanks to BCE). Overall, this BCE-inspired scheme to reduce DSDT is generally applicable to short-channel, ultrascaled MOSFETs, is compatible with low-κ spacer oxides, can be used in conjunction with EM engineering, and may pave a way toward improved performance in the ultrashort-channel limit. Finally, this work highlights the need for simulation platforms combining *ab initio* simulation tools (which can account for atomistic and quantum effects) with continuum models describing quantum transport at a mesoscopic scale, at which devices are too large to be described purely from first principles but small enough for tunneling to be significant.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The authors acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), the Fonds de Recherche du Québec - Nature et Technologies (FRQ-NT), the National Research Council Canada (NRC), the Digital Research Alliance of Canada, and Calcul Québec.

References

[1] International Roadmap for Devices and Systems, IEEE, 2020. https://irds.ieee.org/editions.
[2] Kawaura H, Sakamoto T, Baba T. Observation of source-to-drain direct tunneling current in 8 nm gate electrically variable shallow junction metal–oxide–semiconductor field-effect transistors. Appl Phys Lett 2000;76(25):3810–2. https://doi.org/10.1063/1.126789.
[3] Wang J, Lundstrom M, Does source-to-drain tunneling limit the ultimate scaling of MOSFETs?, in: International Electron Devices Meeting (IEDM) Technical Digest, IEEE, 2002, pp. 707–710. doi:10.1109/IEDM.2002.1175936.
[4] Mehrotra SR, Kim S, Kubis T, Povolotskyi M, Lundstrom MS, Klimeck G. Engineering nanowire n-MOSFETs at $L_g < 8$nm. IEEE Trans Electron Devices 2013;60(7):2171–7. https://doi.org/10.1109/TED.2013.2263806.
[5] Maassen J, Guo H. Suppressing leakage by localized doping in Si nanotransistor channels. Phys Rev Lett 2012;109(26):266803. https://doi.org/10.1103/PhysRevLett.109.266803.
[6] Prentki RJ, Harb M, Liu L, Guo H. Nanowire transistors with bound-charge engineering. Phys Rev Lett 2020;125(24):247704. https://doi.org/10.1103/PhysRevLett.125.247704.
[7] Prentki RJ. A surface potential model for field-effect transistors with bound-charge engineering. IEEE Trans Electron Devices 2021;68(9):4625–9. https://doi.org/10.1109/TED.2021.3096779.
[8] Datta S. Electronic Transport in Mesoscopic Systems. Cambridge University Press; 1997. https://doi.org/10.1017/CBO9780511805776.
[9] Niquet YM, Rideau D, Tavernier C, Jaouen H, Blase X. Onsite matrix elements of the tight-binding Hamiltonian of a strained crystal: Application to silicon, germanium, and their alloys. Phys Rev B 2009;79(24):245201. https://doi.org/10.1103/PhysRevB.79.245201.
[10] Taylor J, Guo H, Wang J. Ab initio modeling of quantum transport properties of molecular electronic devices. Phys Rev B 2001;63(24):245407. https://doi.org/10.1103/PhysRevB.63.245407.
[11] Harb M, Michaud-Rioux V, Zhu Y, Liu L, Zhang L, Guo H. Quantum transport modelling of silicon nanobeams using heterogeneous computing scheme. J Appl Phys 2016;119(12):124304. https://doi.org/10.1063/1.4944649.
[12] Nanoacademic Technologies Inc. [link]. https://nanoacademic.com/solutions/.
[13] Beaudoin F, Philippopoulos P, Zhou C, Kriekouki I, Pioro-Ladrière M, Guo H, Galy P. Robust technology computer-aided design of gated quantum dots at cryogenic temperature. Appl Phys Lett 2022;120(26):264001. https://doi.org/10.1063/5.0097202.
[14] Kriekouki I, Beaudoin F, Philippopoulos P, Zhou C, Lemyre JC, Rochette S, Mir S, Barragan MJ, Pioro-Ladrière M, Galy P. Interpretation of 28 nm FD-SOI quantum dot transport data taken at 1.4 k using 3D quantum TCAD simulations. Solid-State Electron 2022;194:108355. https://doi.org/10.1016/j.sse.2022.108355.
[15] R.J. Prentki, Theory and simulation of novel low-power nanotransistors, Ph.D. thesis, McGill University (Nov. 2021). https://escholarship.mcgill.ca/concern/theses/pg15bk850.
[16] Wilk GD, Wallace RM, Anthony JM. High-κgate dielectrics: Current status and materials properties considerations. J Appl Phys 2001;89(10):5243–75. https://doi.org/10.1063/1.1361065.
[17] Miller RD. In search of low-k dielectrics. Science 1999;286(5439):421–3. https://doi.org/10.1126/science.286.5439.421.