MODELING AND SIMULATION OF ADVANCED FLOATING BODY Z-RAM MEMORY CELLS

Viktor Sverdlov and Siegfried Selberherr

Institute for Microelectronics
Technische Universität Wien
Gusshausstrasse 27–29
1040 Vienna, Austria
Email: {sverdlov|selberherr}@iue.tuwien.ac.at

KEYWORDS
Floating body memory, TCAD simulations, scalability

ABSTRACT

A modeling approach to study advanced floating body Z-RAM memory cells is developed. In particular, the scalability of the cells is investigated. First, a Z-RAM cell based on a 50 nm gate length double-gate structure corresponding to state of the art technology is studied. A bi-stable behavior essential for Z-RAM operation is observed even in fully depleted structures. It is demonstrated that by adjusting the supply source-drain and gate voltages the programming window can be adjusted. The programming window is appropriately large in voltage as well as in current.
We further extend our study to a Z-RAM cell based on an ultra-scaled double-gate MOSFET with 12.5 nm gate length. We demonstrate that the cell preserves its functionality by providing a wide voltage operating window with large current differences. An appropriate operating window is still observed at approximately 25-30% reduced supply voltage, which is an additional benefit of scaling. The relation of the obtained supply voltage to the one anticipated in an ultimate MOSFET with quasi-ballistic transport is discussed.

INTRODUCTION

Standard DRAM cell scaling is hampered by the presence of a capacitor which is difficult to reduce in size. Recently, a revolutionary concept of a DRAM memory cell based on a transistor alone was introduced [1–3]. The ultimate advantage of this new concept is that it does not require a capacitor, and, in contrast to traditional 1T/1C DRAM cells, it thus represents a 1T/0C cell named Z (for zero)-RAM. While keeping all advantages of the first Z-RAM generation, the most recent generation of Z-RAM cells [4] is characterized by a significantly enlarged programming window and much longer retention times.
With CMOS downscaling continuing the question obviously arises whether a Z-RAM cell is also scalable. The goal of our study is to demonstrate that a Z-RAM cell preserves its functionality and remains operational for scaled MOSFETs.
In order to reach this goal, several issues must be addressed. Multi-gate FETs and finFETs are the most promising candidates for the upcoming CMOS MOSFETs beyond the 22nm technology node. The functionality of both generations of Z-RAM cells on partially depleted SOI structures was recently demonstrated [4]. With channel length reduced, maintaining control over the channel becomes increasingly challenging, and several measures must be taken to preserve it. Apart from improving electrostatic control by downscaling oxide thickness, the channel width can be reduced. This is achieved by artificially confining carriers within an ultra-thin silicon film. Due to the small dimensions of the silicon body there will be only few impurities inside. This results in unacceptably large threshold voltage fluctuations [5]. Fully depleted double-gate MOSFETs with undoped intrinsic silicon body are perfectly functionable [6, 7]. They preserve a good channel control, reasonable DIBL, large I_{on}/I_{off} ratio, and gain down to a channel length as short as 5 nm [8]. It is not clear, however, whether a Z-RAM cell based on a fully depleted scaled MOSFET would be operational.
Z-RAM cell functionality is based on charging the channel body with the majority carriers generated due to impact ionization. Therefore, for Z-RAM operation, namely for writing, the source-drain voltages are higher than for CMOS logic. An important question is whether this voltage can be reduced while scaling the device down. We demonstrate that, as for CMOS devices, this is generally true. Calculated voltages for scaled Z-RAM cells are around 1.4 V. This value is also in agreement with the writing voltage in a quasi-ballistic MOSFET with an ultra-short channel, which is considered as a good candidate for an ultimate MOSFET [5]. The value is higher than the projected V_{DD}'s for upcoming MOSFETs. However, the current prototypes of a Z-RAM cell operate now at 2.2 V [4]. Therefore, a decrease in supply voltage to 1.4 V is significant.
Simulated structures and models are described in the next section. Then results are presented and analyzed.

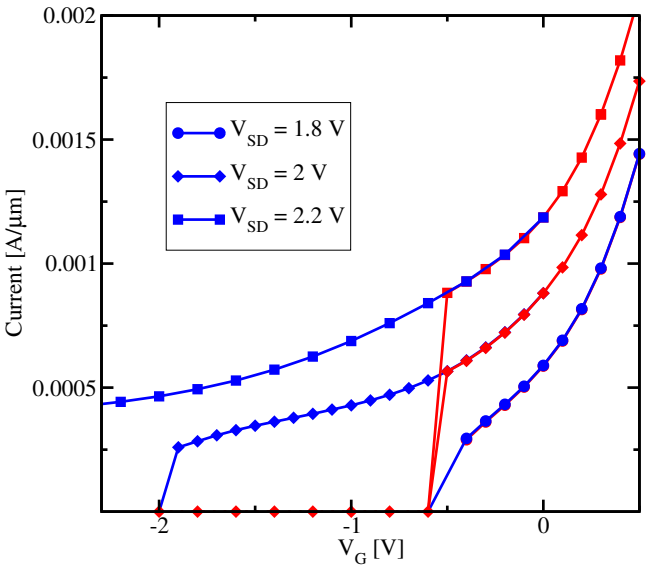

Figure 1: $I-V_G$ for a 50 nm double gate MOSFET with silicon body thickness 10 nm, for different source-drain voltagesV_{SD} . Hysteresis behavior is clearly observed for V_{SD} = 2.0 V, while for V_{SD} = 2.2 V the transition to the low current state is not observed even for $V_G \approx$ -2 V.

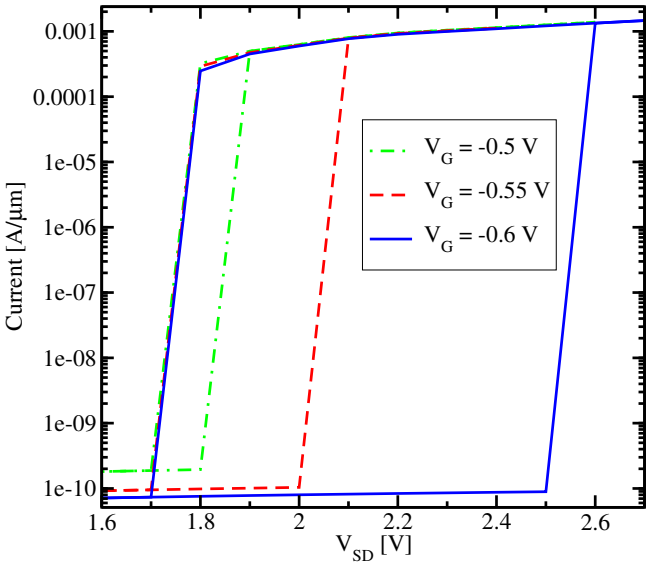

Figure 2: $I - V_{SD}$ in a logarithmic scale for a 50 nm double gate MOSFET with silicon body thickness 10 nm, for different gate voltagesV_G . Hysteresis behavior is clearly observed for V_G = -0.6 V and V_G = -0.55 V, while is has nearly extinct at V_G = -0.5 V.

STRUCTURES

For our analyses we have designed two double-gate structures. One structure has a gate length of 50 nm and a lightly doped ($N_A = 10^{15}$ cm^{-3}) Si body of 10 nm thickness. We have used metal gates with mid-gap work function and oxide with equivalent thickness of 2 nm. Source and drain extentions are heavily doped to $N_D = 10^{20}$ cm^{-3} in order to provide enough injected electrons. This structure corresponds to the current technology node [9].

The scaled double-gate structure has a gate length of 12.5 nm and a lightly doped Si body of 3 nm thickness. An oxide with an equivalent thickness of 1 nm is chosen, The source and drain extensions are heavily doped to $N_D = 10^{20}$cm^{-3}.

The analyses were performed with the MINIMOS-NT device simulator [10]. Impact ionization is essential to the functionality of a Z-RAM cell. Electron-hole recombination is very important as antagonistic mechanism. Similar parameters for impact ionization and for recombination are used for both structures. Band-to-band tunneling was also included with a standard model [10].

RESULTS

The results for current calculations as function of the gate voltage for 50 nm double-gate structure are shown in Fig. 1 and Fig. 2. At high positive gate voltages the current values do not depend on the gate voltage scan direction shown by left and right triangles.

For negative gate voltages the situation is completely different. In a forward scan direction for the gate voltage, the current stays low for both values of the source-drain voltage until a certain critical value is reached. This part of the $I-V_G$ corresponds to the subthreshold regime. Due to negligible DIBL in a 50 nm double-gate structure, the current dependence in the subthreshold regime is similar for both values of the source-drain current.

As soon as a critical current value is reached, the source-drain current rapidly increases by several orders of magnitude. The current keeps increasing for positive gate voltage values.

In a reverse gate voltage scan, the current first slowly decreases. For positive gate voltages, the current takes exactly the same values as for the forward scan. Therefore, the $I - V_G$ curve is completely reversible for both values of the source drain voltage, as already mentioned. However, the current value does not decrease sharply for moderately negative values of gate voltages, although it keeps slightly decreasing. The current values at the reverse scan remain several orders of magnitude higher than the values for the forward scan. The relatively large current value is maintained down to $V_G = -2$ V for V_{SD} = 2.0 V, where it abruptly decreases by several orders of magnitude. Thereby the MOSFET is turned back into the subthreshold regime, completing the hysteresis loop. Indeed, the previous current value cannot be reached just by inverting the scan direction. Instead, if one now increases the gate voltage, the current will follow the lower subthreshold branch until the critical

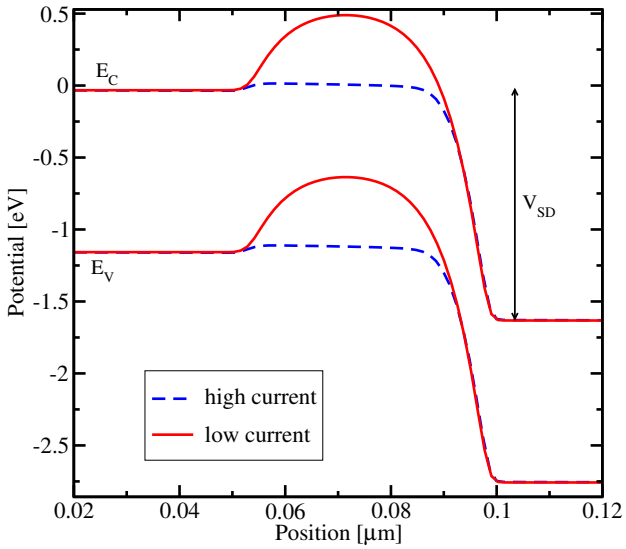

Figure 3: Cross section of the potential energy from the source to the drain along the middle of the channel in a 50 nm double gate MOSFET at V_G = -1 V and $V_{SD} = 2$ V. The potential barrier is high in the low current state, and the transistor is closed. In the high current state there is no barrier, and the transistor is opened.

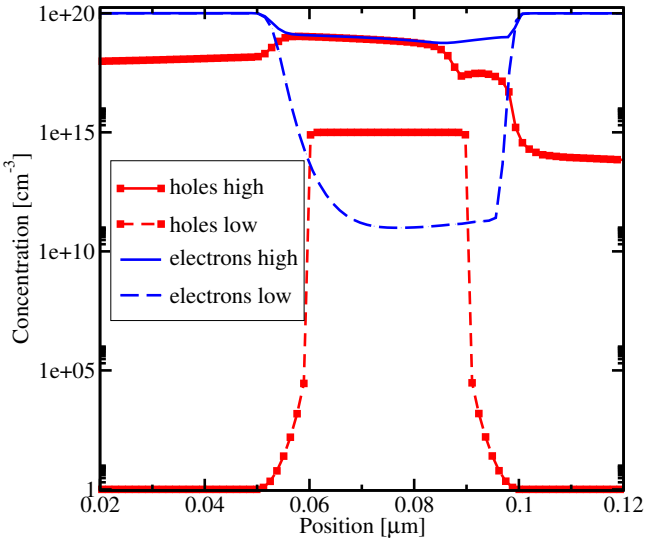

Figure 4: Cross section of the electron and hole concentrations along the middle of the channel in a 50 nm double gate MOSFET at V_G = -1 V and V_{SD} = 2 V for the two current states. The higher majority carrier concentration in the channel in the high current state is due to holes captured in the potential wells clearly seen in Fig. 3 under the gates.

current value is reached at $V_G = -0.6$ V. The point with relatively high current at $V_G = -2$ V can only be reached by inverting the gate voltage scan after the high current value was achieved at positive gate voltages. Interestingly, if we increase the source-drain voltage to $V_{SD} = 2.2$ V, the abrupt transition to the subthreshold regime during the reverse voltage scan cannot be observed for technically relevant negative gate voltage values. A similar behavior is observed when a parasitic bipolar transistor turns on in floating body SOI structures [11], which usually is considered as undesirable.

The two different current states corresponding to the same drain and gate voltages, seen also on $I-V_{SD}$ characteristics shown in Fig. 2 are essential for Z-RAM functionality [4]. We are now analyzing the physical reasons for these two different current states.

Fig. 3 displays the potential profile from the source to the drain electrode cut in the middle of the silicon body, for two different current states corresponding to the same source and drain voltage. In the low current state the potential has a large barrier under the gate preventing the current flowing from the source to the drain. On the contrary, in the high current state the potential is nearly flat in the source-gate region, and the transistor is opened. Such a difference in potential profiles is due to different charge distributions in the system shown in Fig. 4. In the state with low current the electron concentration in the channel is small and he majority carrier concentration (holes) is small too (Fig. 4). In the state with high current one naturally has a higher electron concentration in the channel.

More important, the hole concentration in the channel has also increased as shown in Fig. 4. The hole concentration in the channel is higher close to the gates (Fig. 5), in agreement with [2]. Holes are generated close to the drain region due to impact ionization. The electric field drives generated holes in the channel region where they accumulate. This accumulated positive charge pins down the conduction band to the potential in the source and opens the transistor.

Holes recombine with electrons primarily via the Shockley-Read-Hall mechanism. Excess holes visible in Fig. 4 flow into the source region. Their extra positive charge is compensated by additional electrons, resulting in a slightly higher electron concentration than the equilibrium concentration determined by $N_D = 10^{20}$ cm^{-3}. The nature of the two current states analyzed above allows to determine conditions, when the transitions between them occur. One important ingredient is impact ionization which is usually characterized by the multiplication factor $M > 1$. The positive feedback loop is activated when the collector current is larger than the base current. Because the hole base current is proportional to $M - 1$, the positive feedback corresponds to the condition [11]

$$\beta_F(M-1) > 1, \qquad (1)$$

where β_F is the common-emitter current gain. The increase of the drain current is triggered by the positive feedback, saturating when the transistor opens.

If we now reduce the gate voltage, an increasing number of holes must be stored under the gate to compensate

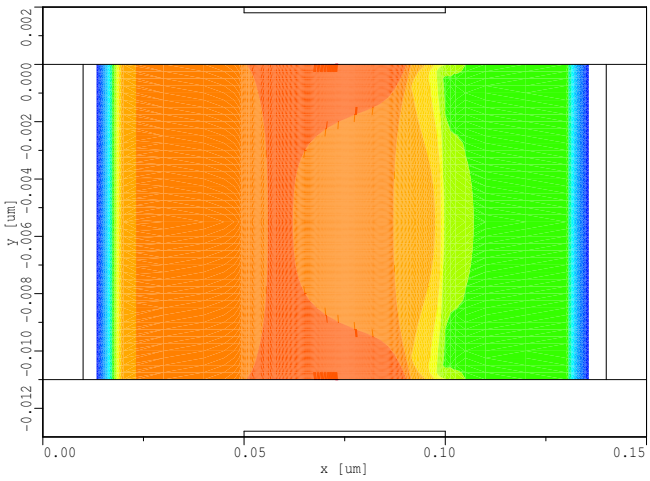

Figure 5: Contour plot in logarithmic scale of the hole distribution in the MOSFET channel for the high current state. Close to the gates the hole concentration is higher, in agreement with [4].

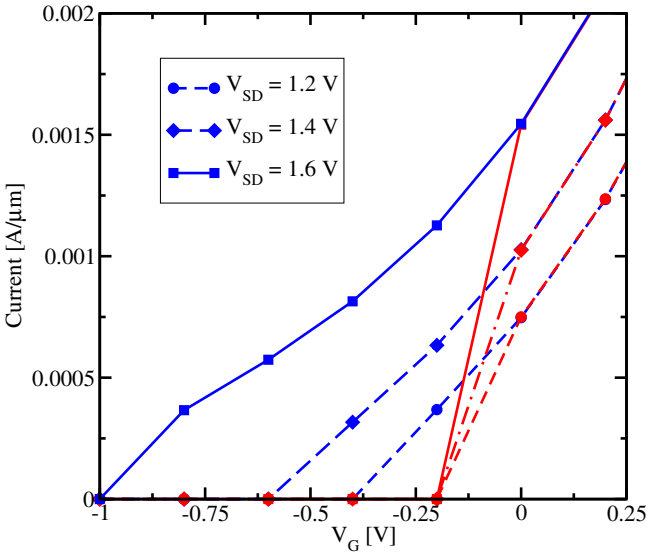

Figure 6: $I-V_G$ for a scaled double gate MOSFET with the gate length 12.5 nm and silicon body thickness 3 nm, for three source-drain voltage. Hysteresis behavior is clearly observed.

.

the gate voltage decrease and keep the transistor open. This results in an increased recombination rate that reduces β_F. At large negative gate voltages the condition (1) cannot be fulfilled. The positive feedback loop breaks, which results in a sudden decrease of the current. The number of generated holes drops. Their concentration under the gate reduces and they cannot screen the gate potential. This leads to a potential barrier increase which further reduces the current. The process stops when the transistor is closed.

The consideration above can explain the $I-V_G$ behavior at $V_{SD} = 2.0$ V for the reverse gate voltage scan. For the forward gate voltage scan the transistor stays in closed state for higher gate voltages than for the reverse scan. The reason is the absence of current in the closed state. However, due to an exponential current increase in the subthreshold regime at the gate voltage close to zero, the current reaches a critical value after which the positive feedback loop leading to transistor opening activates. The critical current values only slightly depend on drain voltage, due to a dependence of β_F on drain voltage.

We have demonstrated that a programming window, which is formed by the two current values and the two gate voltage values when switching appears, is sufficiently large for stable Z-RAM operation on 50 nm double-gate transistors. We now present simulations of a double-gate structure with 12.5 nm gate length. Results of $I-V_G$ calculations shown in Fig. 6 and Fig. 7 clearly display a hysteresis behavior similar to that observed for a 50 nm MOSFET. For all considered source-drain voltages the transition to the high current state appears at slightly negative gate voltages. For the reverse scan the transition to the low current state is observed at large negative gate voltages for $V_{SD} = 1.6$ V, while for $V_{SD} = 1.2$ V the hysteresis has nearly disappeared. For $V_{SD} = 1.4$ V, the transition to the low current state occurs at $V_G =$ - 0.6 V. This results in a relatively large programming window sufficient for successful Z-RAM cell operation. It is thus demonstrated that a Z-RAM cell built on a scaled double-gate MOSFET with 12.5 nm gate length preserves its functionality.

DISCUSSION

The $I-V_G$ behavior for a 12.5 nm gate length MOSFET looks analogous to the behavior of a 50 nm MOSFET. One difference between the results is that the current density for a thinner and shorter double-gate structure is nearly an order of magnitude smaller. However, this is not a substantial limitation, because the important criterion for Z-RAM functionality is the difference between the two values of current in the two different current states, which is still several orders of magnitude for a 12.5 nm double-gate structure.

Another important difference is that the supply voltages are 25-30% smaller for a Z-RAM based on a scaled MOSFET. The obtained substantial decrease in supply voltage is comparable with the anticipated decrease of V_{DD} for scaled logic devices. We should add, however, that the impact ionization model used in the simulations depends on the local field only. When the channel length is reduced, the local field in the channel at the drain end is expected to increase. Therefore, the local field impact ionization model can overestimate impact ionization. Another potential limitation of the applica-

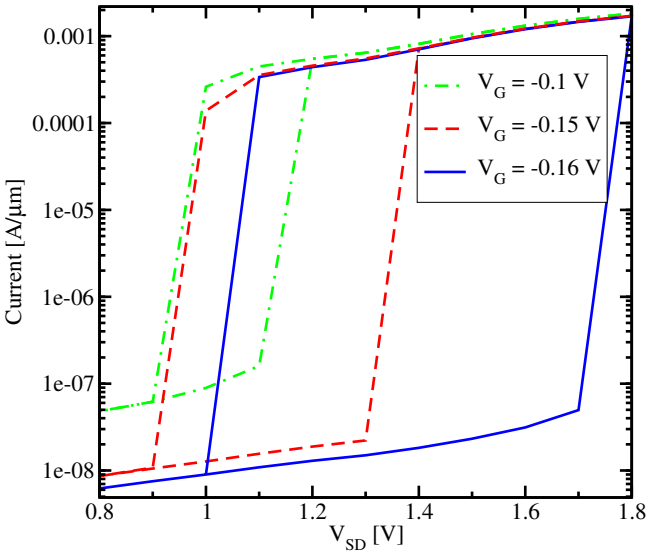

Figure 7: $I - V_{SD}$ in a logarithmic scale for a 12.5 nm double gate MOSFET with silicon body thickness 3 nm, for different gate voltages V_G. Hysteresis behavior is clearly observed.

bility of our approach is that in scaled devices transport becomes quasi-ballistic, and the impact ionization models used in Monte Carlo simulations of hot carrier transport [12] should become valid. These models are characterized by threshold energies above which impact ionization starts, with the lowest threshold of 1.2 eV. It was recently argued that due to the presence of energetic carriers in an injected distribution substantial impact ionization can be present even, when the source-drain voltage is smaller than the threshold [13]. However, because of the gap increase due to size quantization in a 3 nm silicon film, we believe that 1.2 eV is a fair estimate of a Z-RAM cell supply voltage, which is consistent with $V_{SD} = 1.4$ V obtained in our simulations.

CONCLUSION

We have shown that a Z-RAM cell built on a scaled double-gate MOSFET preserves its functionality by providing a wide voltage operating window with large current differences. We also predict a decrease in the supply drain-gate voltage to 1.2-1.4 V, which is about 25-30% smaller than in current prototypes.

REFERENCES

[1] S. Okhonin, M. Nagoga, J. Sallese, and P. Fazan, "A SOI capacitor-less 1T-DRAM concept," in *SOI Intl. Conf. Techn. Dig.*, 2001, pp. 153–154.

[2] D. Ban, U.E. Avci, D.L. Kencke, and P.L.D. Chang, "A scaled floating body cell (FBC) memory with high-k+metal gate on thin-silicon and thin-BOX for 16 nm technology node and beyond," in *Proc. 2008 VLSI Symposium*, 2008, pp. 92–93.

[3] M.G. Ertosun, H. Cho, P. Kapur, and K.C. Saraswat, "A nanoscale vertical double-gate single-transistor capacitorless DRAM," *IEEE Electron Device Lett.*, vol. 29, no. 6, pp. 615–617, 2008.

[4] S. Okhonin, M. Nagoga, E. Carman, R. Beffa, and E. Faraoni, "New generation of Z-RAM," in *IEDM Techn. Dig.*, 2007, pp. 925–928.

[5] K. K. Likharev, "Sub-20-nm electron devices," in *Advanced Semiconductor and Organic Nano-Techniques*, H. Morkoc, Ed. New York: Academic Press, 2003, pp. 239–302.

[6] Y. Naveh and K. K. Likharev, "Modeling of 10 nm-scale ballistic MOSFETs," *IEEE Electron Device Lett.*, vol. 21, no. 5, pp. 242–244, 2000.

[7] M. Lundstrom, "The ultimate MOSFET and the limits of miniaturization," in *Intl. Semiconductor Device Research Symposium Techn. Dig.*, 2007, pp. 1–1.

[8] V. A. Sverdlov, T. J. Walls, and K. K. Likharev, "Nanoscale silicon MOSFETs: A theoretical study," *IEEE Trans.Electron Devices*, vol. 50, no. 9, pp. 1926–1933, 2003.

[9] W. Xiong, C. R. Cleavelin, C.-H. Hsu, M. Ma, K. Schruefer, K. V. Armin, T. Schulz, I. Cayrefourcq, C. Mazure, P. Patruno, M. Kennard, K. Shin, X. Sun, T.-J. Liu, K. Cherkaoui, and J. Colinge, "Intrinsic advantages of SOI multiple-gate MOSFET (MuGFET) for low power applications," *ECS Transactions*, vol. 6, no. 4, pp. 59–69, 2007.

[10] *MINIMOS-NT 2.1 User's Guide*, Institut für Mikroelektronik, Technische Universität Wien, Austria, 2004.

[11] J.-P. Colinge, *Silicon on Insulator Technology: Materials to VLSI.* Boston: Kluwer Academic Publishers, 2004.

[12] M. V. Fischetti, S. E. Laux, and E. Crabbé, "Understanding hot-electron transport in silicon devices: Is there a shortcut?" *J.Appl.Phys.*, vol. 78, no. 2, pp. 1058–1085, 1995.

[13] J. Guo, M. Alam, and Y. Ouyang, "Subband gap impact ionization and excitation in carbon nanotube transistors," *J.Appl.Phys.*, vol. 101, pp. 064 311–1–064 311–5, 2007.