Simulation of Semiconductor Processes and Devices 2016
Edited by E. Bär, J. Lorenz, and P. Pichler

Band-to-Band Tunneling Off-State Leakage in Ge Fins and Nanowires: Effect of Quantum Confinement

G. Eneman, A.S. Verhulst, A. De Keersgieter, A. Mocuta, N. Collaert, A. Thean
Imec
Kapeldreef 75, 3001 Heverlee, Belgium
E-mail: Geert.Eneman@imec.be

L. Smith, V. Moroz
Synopsys
Mountain View, CA, USA

***Abstract*— This simulation work studies whether band-to-band-tunneling leakage in short-channel germanium FinFETs and nanowires can be mitigated by the band gap widening resulting from quantum confinement. Through a combination of drift-diffusion and coupled Poisson-Schrödinger simulations, two possible solutions are investigated: can the BTBT rate be lowered sufficiently? Secondly, can the tunnel path be cut off by band gap widening? Our results indicate that by exploiting the band gap widening in narrow devices, gate-induced drain leakage due to band-to-band tunneling in Ge FinFETs and nanowires can meet the high-performance and low operating power off-state leakage specifications of CMOS technology. However, the low standby power off-state target seems out of reach.**

***Keywords*—germanium; MOSFETs; leakage; band-to-band tunneling**

I. Introduction

For CMOS digital logic applications, alternative channel materials like germanium and $In_xGa_{1-x}As$ are considered as replacements for silicon. Their main selling point is the excellent mobility, in particular the electron and hole mobility for germanium, and electron mobility for $In_xGa_{1-x}As$. Recent publications show promising performance for Ge P-type devices [1,2] and $In_xGa_{1-x}As$ NFETs [3,4]. For Ge NFETs, despite the challenge to achieve a high-quality gate passivation and low-resistive source-drain contacts, significant progress is reported for the channel mobility [5,6].

Besides the promise of excellent performance, future FETs with germanium channels also need to meet the specifications with respect to off-state leakage. Here lies another challenge for Ge, as its small indirect (0.66 eV) and direct (0.82 eV) band gap [7] will lead to significantly increased leakage currents. In particular, Gate-Induced Drain Leakage (GIDL), which is mainly caused by Band-To-Band Tunneling (BTBT) has been identified as a potential showstopper [8-10]. Fig. 1 compares the BTBT generation rate in bulk germanium and silicon. At the same electric field, both direct tunneling (between the Γ points of the conduction and valence bands) and indirect tunneling (between the L-valleys of the conduction band and the top of the valence band) in germanium are significantly higher than indirect BTBT in silicon.

The purpose of this work is to study in how far the band gap widening in narrow fins and scaled nano-wires can be exploited to mitigate BTBT in the drain junctions of germanium-channel devices. As can be seen from Fig. 1, both direct and indirect tunneling in Ge are high and need to be considered. Two separate possible solutions are considered: First, band gap widening lowers the BTBT generation rates (Sec. IIA). Secondly, the BTBT tunnel path may be cut off in a short FET due to the larger band gap in a narrow channel (Sec. IIB).

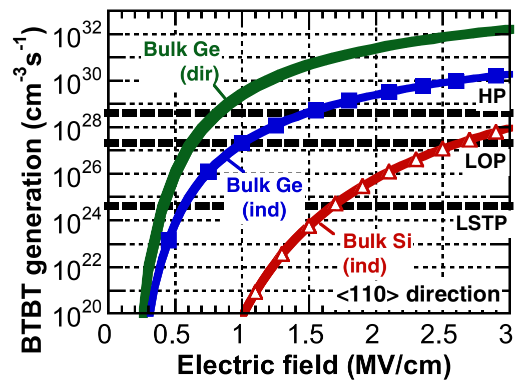

Figure 1. Direct ('dir') and indirect ('ind') band-to-band tunneling generation rate of bulk Ge and Si as a function of the electric field, based on [11]. For silicon, indirect BTBT is dominant while direct tunneling is negligible. At the same electric field, both indirect and direct tunneling in germanium are significantly higher leakage than in silicon. The 'HP', 'LOP' and 'LSTP' levels correspond approximately to a fin/wire off-state current of 100 nA/μm, 5 nA/μm and 10 pA/μm, as is detailed further in the text.

In this work, the electric fields are simulated in off-state for fins and wires with 14 nm gate lengths, using Sentaurus-Device [7]. The band structure of Ge in the cross-section of a FinFET and nanowire is calculated using Sentaurus Band-Structure [7], a 2-dimensional coupled Poisson-Schrödinger simulator that uses a 6-band k.p solver for the valence bands, and effective-mass approximation with non-parabolicity correction for the conduction bands.

II. Results

A. BTBT Rates in Fins and Wires

In a first step, the band structure is calculated for 30 nm-tall Ge FinFETs with varying widths, and rounded Ge wires with varying diameters. Fig. 2 shows the indirect and direct band gap of relaxed and stressed Ge FinFETs and wires. Especially for fin widths and wire diameters below 10 nm, significant band gap widening is observed. This is counteracted by the channel stress, that leads to lower band gaps.

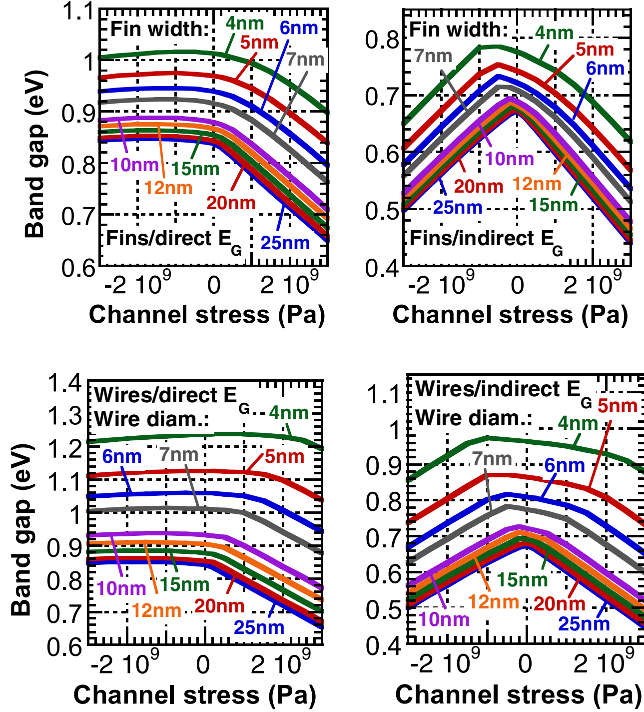

Figure 2. Direct (left) and indirect (right) band gap of Ge-channel FinFETs (top) and nanowires (bottom) with <110> channels, as a function of longitudinal (i.e. <110>) stress, for various fin widths and wire diameters. For fins, the fin height is 30nm. Tensile stresses have positive value, compressive stresses are negative.

Secondly, the subband energies and curvature masses are used to calculate the BTBT rate using a formulation for direct tunneling in confined systems. For bulk materials, Kane's formalism leads to the following generation rate G for BTBT in a direction x [12, 13]:

$$G = \frac{9}{\pi^2}\frac{g(qE)^2}{36\pi\hbar^2 E_G^{1/2}}\frac{m_{r,y}^{1/2} m_{r,z}^{1/2}}{m_{r,x}^{1/2}} \cdot \exp\left(-\frac{\pi m_{r,x}^{1/2} E_G^{3/2}}{2\hbar q E}\right) \quad [\#/(\mathrm{cm}^3\mathrm{s})] \tag{1}$$

In the above formula, g is the spin degeneracy, q is the electron charge, E is the magnitude of the electric field, E_G is the direct band gap, and $m_{r,a}$ is the reduced mass in a direction a, given by $m_{r,a}^{-1}=m_{lh,a}^{-1}+m_{e,a}^{-1}$. $m_{lh,a}$ and $m_{e,a}$ are the light-hole, resp. the electron curvature masses in direction a.

For systems that are confined in the y and z directions in a structure with dimensions L_y and L_z, the following generation rate is used for BTBT in the transport direction x [13]:

$$G = \sum_{\substack{Sub-\\bands}} \frac{9}{\pi^2}\frac{g\, qE\pi}{18\hbar}\frac{1}{L_y}\frac{1}{L_z} \cdot \exp\left(-\frac{\pi m_{r,x}^{1/2}}{2\hbar q E \cdot E_{G,0}^{1/2}}\left(E_{G,0}+\Delta E_G\right)^2\right) \; [\#/(\mathrm{cm}^3\mathrm{s})] \tag{2}$$

$E_{G,0}$ is the band gap of the unconfined, bulk material, while ΔE_G corresponds to the band gap widening due to confinement. The summation over the various subbands is done such that the n'th conduction subband is coupled to the 2n'th valence subband.

The calculated direct BTBT generation rate G for Ge-channel FinFETs and wires is shown in Fig. 3. The generation rates for fins and wires have been calculated using (2) and the band structure resulting from Sentaurus Band-Structure simulations. The generation rates for bulk Ge and Si use (1) and parameters taken from [10]. Fig. 3 also includes estimations of the generation rates leading to an off-state current of 100 nA/µm ('HP'), 5 nA/µm ('LOP') and 10 pA/µm ('LSTP') in FinFETs, by assuming that the main BTBT in these devices takes place between the drain and channel region, over a distance that roughly corresponds to the spacer width W_S: $I_{OFF,BTBT} \sim q.G.F_W.F_H.W_S/(F_W+2F_H)$. The fin width F_W and the fin height F_H are taken constant as 7 and 30 nm, respectively. For W_S, 5 nm is assumed.

As confinement leads to larger band gaps, the generation rate decreases for thinner fins and wires. As can be seen from Fig. 3, it is impossible for Ge to reach the generation rate of bulk silicon by using confinement. For FinFETs targeting the 7 nm node and working at a supply voltage of 0.6 V, typical electric fields in the drain-to-channel junction are found to be in the order of 1 MV/cm (not shown). Fig. 3 indicates that, to meet the off-state specifications ('HP', 'LOP' and 'LSTP' lines, Fig. 3) at an electric field of 1 MV/cm, very narrow fins or very thin wires are required. Alternatively, the electric fields need to be reduced in Ge devices. This can be done by graded or asymmetrical junctions, however this may come at the expense of increased parasitic series resistance and hence performance.

B. Tunnel Paths

Tunneling from the drain to the channel region can be cut off when there is no tunnel path available. This happens when the band gap E_G is larger than the drain-to-channel barrier in off-state (Fig. 4): $E_G > E_{G,CutOff} = |E_{C,Drain} - E_{C,Channel}|$ where $E_{C,Drain}$ and $E_{C,Channel}$ correspond to the highest resp. lowest positions of the conduction band in the drain and channel region. To first order, the supply voltage determines the energy level in the drain ($E_{C,Drain}$), while $E_{C,Channel}$ depends on the target subthreshold off-state leakage. As a consequence, the height of the drain-to-channel barrier decreases for lower supply voltages and for higher target off-state subthreshold currents.

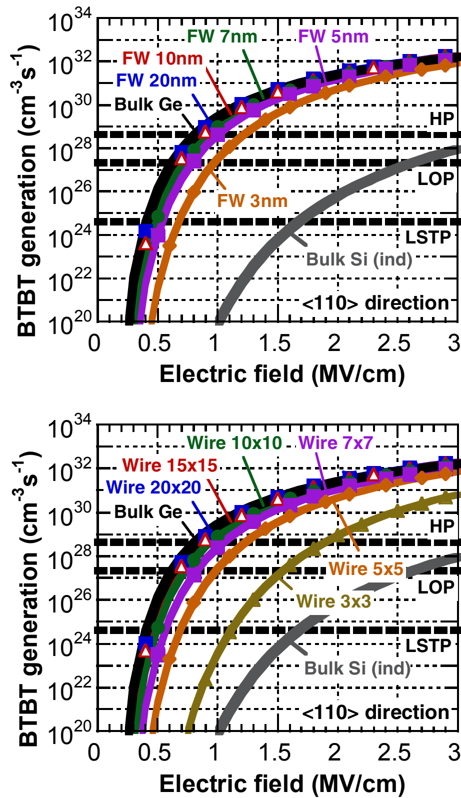

Figure 3. Direct BTBT generation rate of relaxed Ge FinFETs for different fin widths F_W (top) and nano-wires for various diameters (bottom, in nm). The 'HP', 'LOP' and 'LSTP' levels correspond approximately to a fin/wire off-state current of 100 nA/µm, 5 nA/µm and 10 pA/µm, respectively. Fin height is 30 nm.

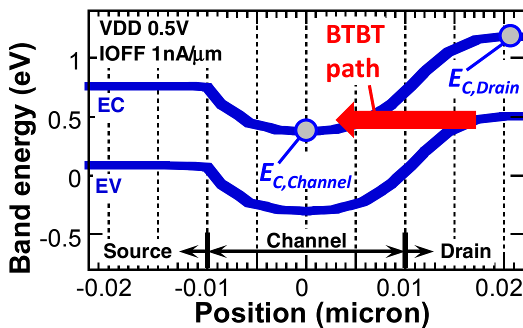

Figure 4. Simulated band structure of a Ge pFinFET in off-state, along a cutline running from the source to the drain, in the center of the FinFET channel. The drain voltage is -0.5 V, and the gate voltage is such that the subthreshold off-state leakage is 1 nA/µm. The BTBT GIDL is not included in this off-current. Whether a BTBT path is available for GIDL depends on the height of the drain-to-channel barrier ($E_{C,Drain} - E_{C,Channel}$).

Sentaurus-Device simulations were done to calculate $E_{G,CutOff}$ for a variety of supply voltages and subthreshold off-state targets. Comparison of $E_{G,CutOff}$ with simulations of the confined band gap (Fig. 2) leads to the supply voltage requirements of Fig. 5 and Fig. 6, that show the maximum supply voltage to cut off direct resp. indirect BTBT leakage paths for relaxed and strained Ge-channel fins and nanowires.

For the strained fins and wires, a compressive stress (2 GPa) from source to drain is assumed, which is preferred for p-type Ge devices with [110]-oriented channels. As the direct band gap is rather insensitive to compressive stress (Fig. 2, bottom), the maximum supply voltage for direct BTBT is insensitive to channel stress (Fig. 5). The indirect band gap reduces for compressive stress (Fig. 2, top), and hence the maximum supply voltage to cut off indirect BTBT is lower for strained devices (Fig. 6).

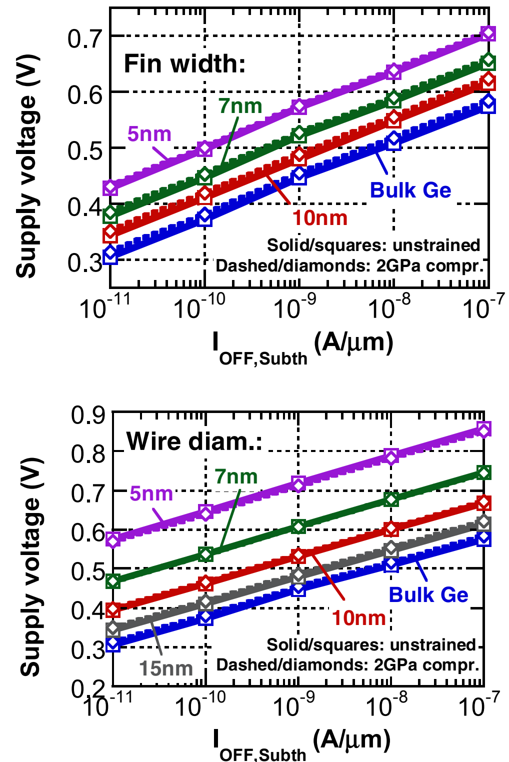

Figure 5. Maximal supply voltages to cut off *direct* BTBT in relaxed and strained bulk Ge and Ge fins and nanowires with <110> channels, as function of target subthreshold off-state (in A/µm). Stress is parallel to the channel, i.e. in <110> direction. (Top) FinFETs with fin height 30nm and various fin widths. (Bottom) Nanowires with various diameters.

Assuming a supply voltage of 0.6 V, Fig. 5 shows that for an I_{OFF} target of 100 nA/µm, the direct tunneling BTBT path is cut off for fins with a fin width below 10 nm and wires with a diameter of 15 nm and below. For an I_{OFF} of 5nA/µm, fin widths below 5 nm or wire diameters below 10 nm are required. For an I_{OFF} of 10 pA/µm, the direct tunnel path cannot be cut off for Ge fins, but it can likely be cut off for wires with a diameter below 5 nm.

As Ge's indirect band gap is smaller than the direct gap, it is more difficult to cut off indirect BTBT than direct BTBT, as indicated by the smaller maximal supply voltages for indirect BTBT (Fig. 6). For fins, it is impossible to cut off indirect BTBT by confinement at a supply voltage V_{DD} of 0.6 V (Fig. 6, top). For wires (Fig. 6, bottom), cutting off indirect BTBT at a V_{DD} of 0.6 V is only possible for unstrained wires with a

diameter of 5 nm and below, and a relaxed I_{OFF} target of 100 nA/µm.

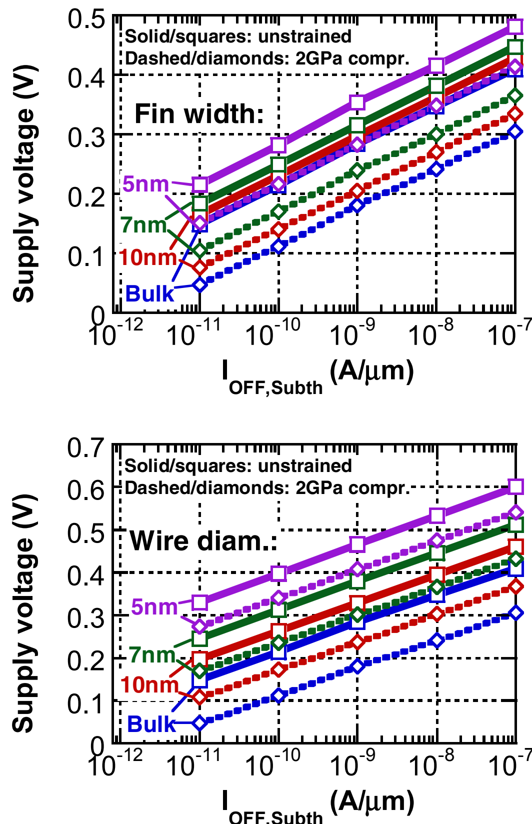

Figure 6. Maximal supply voltages to cut off *indirect* BTBT in relaxed and strained bulk Ge and Ge fins and nanowires with <110> channels, as function of target subthreshold off-state (in A/µm). Stress is parallel to the channel, i.e. in <110> direction. (Top) FinFETs with fin height 30nm and various fin widths. (Bottom) Nanowires with various diameters.

III. Conclusions

The purpose of this paper is to estimate whether GIDL caused by BTBT in Ge-channel fins and wires can be suppressed enough to meet the off-state leakage targets for digital CMOS technologies. As both the direct and indirect band gap of Ge are small, direct and indirect BTBT need to be considered. Band gap widening due to quantum confinement in scaled devices can mitigate the GIDL issue in two ways: 1. by lowering the BTBT generation rate to acceptable levels, or alternatively, 2. by cutting off the BTBT paths in the drain-to-channel junction. The conclusions in the following paragraphs are made assuming a supply voltage V_{DD} of 0.6 V, a junction design that leads to about 1 MV/cm electric fields in the drain junction, and limited channel stress (below 1 GPa).

GIDL can be reduced to meet a high-performance I_{OFF} target of 100 nA/µm for fins with a fin width below 10 nm, and wires with less than 15 nm diameter: for this I_{OFF} target, *indirect* tunneling is sufficiently low at 1 MV/cm even for bulk Ge (Fig. 1). The *direct* BTBT can be lowered to an acceptable level for fins with a fin width below 10 nm (Fig. 3, top). For wires with a diameter below 15 nm, no direct tunnel path is available. (Fig. 5, bottom).

To target an I_{OFF} of 5 nA/µm, fin widths of 5 nm and lower, and wire diameters below 10 nm are sufficient to cut off the *direct* BTBT path (Fig. 5, top and bottom). For the indirect component, it is likely that the band gap widening at these dimensions will reduce the leakage level sufficiently, as the indirect leakage at an electric field of 1 MV/cm already roughly meets the 5 nA/µm target even for bulk germanium (Fig. 1).

An off-state current of 10 pA/µm seems out of reach for Ge FinFETs, as for direct BTBT, neither the current level can be lowered sufficiently (Fig. 3, top), or the tunnel path can be cut off (Fig 5, top) for reasonable fin widths. For wires, the direct tunnel path can be cut off for diameters below 5 nm (Fig. 5, bottom). However, whether the indirect tunneling leakage level can be sufficiently reduced is unclear.

Besides exploiting the effects of quantum confinement to reduce leakage, which is the purpose of this work, several other optimizations can be considered to reduce BTBT leakage in Ge-channel devices. The most obvious options, reducing the electric fields that generate BTBT by lowering the supply voltage or re-designing the doping levels in the junctions, may be undesired as they will degrade the performance and drive current of the technology. Other options may be more promising, such as band gap engineering of the drains by growing a larger-band gap material in the drains, and could be subject of further study. Finally, the GIDL studied here is only one leakage mechanism that is a concern for Ge CMOS technologies. Experimental data suggests that defect-generated mechanisms such as perimeter junction leakage at the semiconductor/isolation interface may provide another roadblock for a low-leakage Ge technology [10], and therefore should also be considered.

References

[1] Y. Yeo et al., IEDM Tech. Dig., p. 28 (2015).
[2] J. Mitard et al., IEDM Tech. Dig., p. 418 (2014).
[3] A. Vardi et al, IEDM Tech. Dig., p.807 (2015).
[4] X. Zhou et al., VLSI Tech Dig., p. 166 (2016).
[5] H. Wu et al., IEDM Tech. Dig., p. 16 (2015).
[6] H. Arimura et al., IEDM Tech. Dig., p. 588 (2015).
[7] Sentaurus Device and Sentaurus Band Structure Reference Manual, I-2013.12 (2013).
[8] B. Obradovic et al., IEEE EDL, vol. 35(4), p. 473 (2014).
[9] T. Krishnamohan, SISPAD Proc, p. 97 (2008).
[10] G. Eneman et al., IEEE TED 56(12), p. 3115 (2009).
[11] K.-H. Kao et al, IEEE TED 59(2), p. 292 (2012)
[12] E.O. Kane, J. Phys. Chem. Solids 12, p. 181 (1959)
[13] A.S. Verhulst et al., to be published.