Suppression of Gate-Induced Drain Leakage by Optimization of Junction Profiles in 22 nm and 32 nm SOI nFETs

Andreas Schenk [a,*] ,

[a] *Integrated Systems Laboratory, ETH Zurich, Gloriastrasse 35, CH-8092, Switzerland*

Abstract

It is shown by TCAD simulations how the gate-induced drain leakage which dominates the OFF-current in 22 nm double-gate and 32 nm single-gate SOI nFETs with high-K gate stacks, can be minimized by proper variations of the junction profiles. Based on a microscopic, non-local model of band-to-band tunneling, transfer characteristics are computed after systematic changes in source/drain doping, body thickness, and HfO_2 layer thickness. This is done under the constraint of a minimal degradation of the ON-current. Variations which lead to the best compromise are highlighted. The obtained results suggest the alignment of the lateral doping profile with the gate corners to be the decisive factor. It is found that the ITRS target for 22 nm DG LSTP devices (10 pA/μm) can be met with acceptable degradation of the ON-current.

Key words: TCAD, simulation, gate induced drain leakage, non-local band-to-band tunneling, high-K gate stacks, DGSOI, SGSOI, 22 nm technology node

1. Introduction

The OFF-current of short-channel metal-oxide-silicon field effect transistors (MOSFETs) is increasingly dominated by band- to-band tunneling (BTBT) which causes "gate-induced drain leakage" (GIDL). Reasons that have been identified for bulk and partially depleted silicon-on-insulator (SOI) FETs are the heavily doped pn-junction between drain and substrate, or between drain and partially depleted (PD) body, and the use of heavy counter doping (halo implants) to suppress short-channel effects (see [1] and references therein). BTBT causes increased power dissipation in bulk FETs and unwanted threshold voltage shifts in PD SOI FETs due to the floating-body effect (kink effect), which can lead to extra sub-threshold leakage [1]. The focus in this paper is on fully-depleted (FD) SOI FETs with ultra-thin body where the kink effect is absent [2] and where the location of the BTBT rate can be different from that in bulk and PD SOI FETs.

To meet the $I_{\rm off}$ requirement defined by the ITRS, it has to be understood how the GIDL can be minimized without degrading the ON- current too much. Test vehicles for the present TCAD analysis are template devices of the 22 nm and 32 nm technology nodes with high-K gate stacks that had been designed in the EU project PULLNANO [3]. Section 2 describes the geometry and doping of these devices in detail. It also gives a comparison between direct gate tunneling (GT) and BTBT. In Section 3 the BTBT model is outlined and its parameters are validated against experimental data of Ref. [1]. The technology variations to minimize the GIDL are presented in Section 4. Finally, Section 5 provides the conclusion.

* Corresponding author. Tel.: +41 44 632 6689; Fax: +41 44 632 1194.

Email address: schenk@iis.ee.ethz.ch (Andreas Schenk).

2. Device description

2.1. *22 nm DGSOI nFET*

The starting structure of the 22 nm DGSOI nFET was defined in the EU project PULLNANO [3] as symmetrical double-gate architecture (see Fig. 1). The EOT is 1.11 nm obtained with a 0.7 nm interfacial oxide layer and 2.4 nm of HfO_2. The gate length is 22 nm, the body thickness 10 nm, and 4.8 V are chosen for the gate work function. The body is unstrained with ⟨100⟩ orientation and has a constant boron concentration of $1.2\times10^{15}\,\mathrm{cm}^{-3}$. The source/drain extensions are 60 nm long and the contacts are placed vertically at the ends of them. Reasonable junction profiles were defined to meet the OFF-current ITRS specification for 22 nm DG LSTP devices (10 pA/μm) [4], assuming thermionic emission as the only leakage mechanism. Structure and doping information were translated into Sentaurus-Device [5] input files by the PULLNANO consortium [3]. When direct GT and nonlocal BTBT are taken into account, the picture changes drastically (see Fig. 2). Whereas the high-K gate stack drops GT well below the ITRS limit at $V_{GS}=0$ V, BTBT increases the OFF-current by about 3 orders of magnitude. The assessment of the smallness of GT is based on the calibrated parameter set [3]: $\epsilon_{HfO2}=23$, $\epsilon_{SiO2}=3.9$, $\chi_{HfO2}=2.05$ eV, $m_{HfO2}=0.11\,m_0$, $\chi_{SiO2}=0.9$ eV, $m_{SiO2}=0.5\,m_0$. These parameters were found by extensive comparisons to experimental data of MOS capacitors and long-channel MOSFETs with similar gate stacks in [6]. An interesting feature visible in Fig. 2 is the coupling between GT and BTBT currents. BTBT generated holes lower the gate tunneling barrier via the floating-body effect. However, this increase of GT induced by BTBT is too small to reach the ITRS limit.

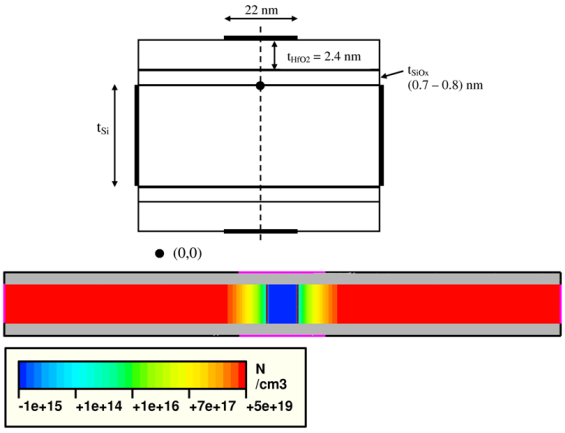

Fig. 1. Geometry and doping of the 22 nm DGSOI nFET.

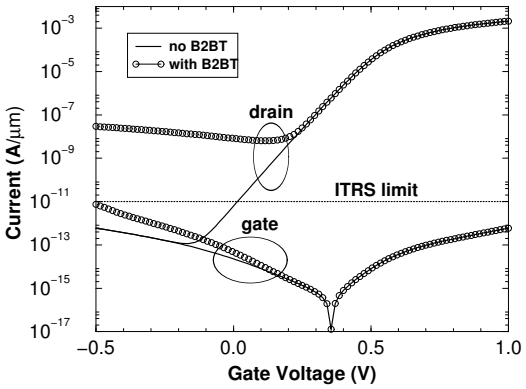

Fig. 2. Transfer and gate current characteristics of the 22 nm DGSOI nFET showing the suppression of gate leakage by the high-K gate stack, but a strong GIDL due to BTBT.

2.2. *32 nm SGSOI nFET*

The starting structure of the 32 nm SGSOI nFET [3] is shown in Fig. 3. Here, the EOT is 1.19 nm obtained with a 0.8 nm interfacial oxide layer and 2.3 nm of HfO_2. The gate length is 32 nm, the body thickness 7 nm, and the BOX thickness 20 nm. The p-type doping of the latter was adjusted to $1\times10^{18}\,\mathrm{cm}^{-3}$. The body is unstrained with ⟨100⟩ orientation and has a constant boron concentration of $1.2\times10^{15}\,\mathrm{cm}^{-3}$. Source and drain regions are elevated by 10 nm and planar source/drain contacts are placed on top of them 36 nm from the center of the channel. It is assumed that the spacers fully consist of Si_3N_4 with a relative dielectric constant of 7.5. Technologically relevant junction profiles were defined to meet an OFF-current target of 200 pA/μm, assuming thermionic emission as the only leakage mechanism. This relaxed OFF-current specification allows for a mid-gap work function of 4.6 V which was used throughout this study. With GT and BTBT turned on, the minimum drain leakage is about 5 times higher than 200 pA/μm and the gate current is as low as in the case of the 22 nm DGSOI nFET. As the chosen 200 pA/μm limit is just a concession to allow for a mid-gap work function, it is worthwhile to optimize also the 32 nm SGSOI for a low GIDL.

3. GIDL Model

Phonon-assisted BTBT is a crucial leakage mechanism in strong electric fields of MOS structures. It must be turned on in TCAD simulations, if the field exceeds approximately 8×10^5 V/cm. The non-local BTBT rate is modeled according to the microscopic theory of [7,8] implemented in [5]. The TCAD version of the model has the compact form

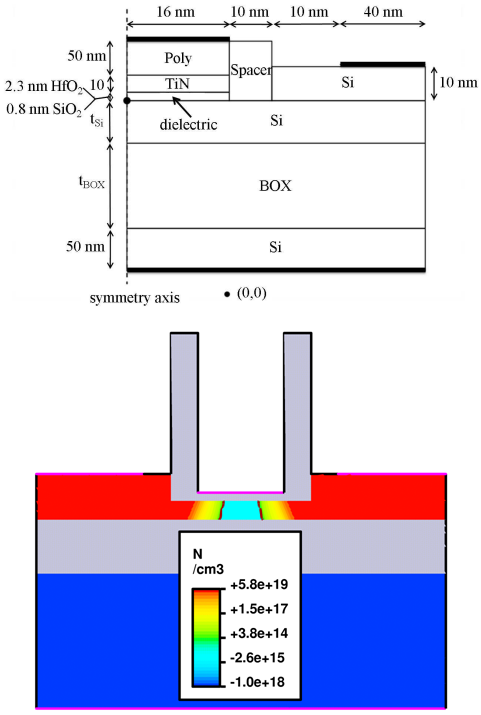

Fig. 3. Geometry and doping of the 32 nm SGSOI nFET.

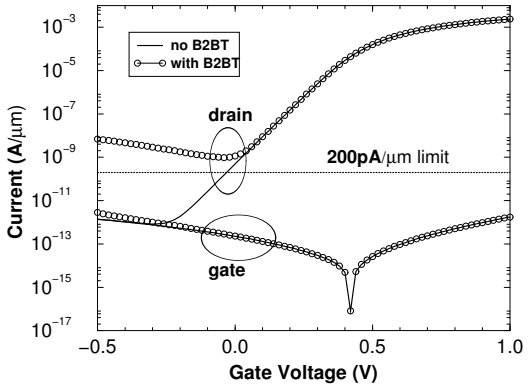

Fig. 4. Transfer and gate current characteristics of the 32 nm SGSOI nFET.

$$R_{\rm net}^{\rm BTBT} = A\,F^{7/2}\frac{\tilde{n}\,\tilde{p}-n_{\rm i,eff}^2}{(\tilde{n}+n_{\rm i,eff})\,(\tilde{p}+n_{\rm i,eff})}\times \qquad (1)$$
$$\left[\frac{(F_{\rm c}^{\mp})^{-\frac{3}{2}}\,\exp\left(-\frac{F_{\rm c}^{\mp}}{F}\right)}{\exp\left(\frac{\hbar\omega}{k_{\rm B}T}\right)-1}+\frac{(F_{\rm c}^{\pm})^{-\frac{3}{2}}\,\exp\left(-\frac{F_{\rm c}^{\pm}}{F}\right)}{1-\exp\left(-\frac{\hbar\omega}{k_{\rm B}T}\right)}\right].$$

$F_{\rm c}^{\pm}$ are "critical" field strengths

$$F_{\rm c}^{\pm} = B\,(E_{\rm g,eff}\pm\hbar\omega)^{3/2}\ , \qquad (2)$$

where the upper sign refers to tunneling generation and the lower sign to recombination. The quantity $\hbar\omega$ denotes the energy of the transverse acoustic (TA) phonon which yields the necessary momentum for the indirect transition. The height of the potential barrier for BTBT is given by the effective band gap $E_{\rm g,eff}$. As strong electric fields usually occur in heavily doped regions, the band gap is strongly modified by band gap narrowing [9] and DOS tails [10,11]. Therefore, in BTBT simulations this is the first parameter that has to be adjusted. The pre-factor B contains the *orientation-dependent* reduced effective mass. Although the full version of the BTBT model considers tunneling from/into all six conduction band valleys along a given field direction, this field direction varies in a MOSFET and even depends on the applied voltages. Hence, B has to be adjusted, too. The energy of the involved TA phonon was determined in early experiments with Esaki diodes by Logan and Chynoweth [12]. It turns out that the rate (1) is rather sensitive to the choice of $\hbar\omega$, hence it can be taken as additional fitting parameter.

Non-locality of BTBT is crucial, as it e.g. prevents tunneling in regions where no final states are available due to band bending. In a MOSFET this usually happens close to the gate oxide interface, i.e. in a region where the electric field F in the semiconductor becomes maximal. In Figs. 6 and 7 these regions are denoted as "dark space". In the model, non-locality is accounted for in a simple way by modified densities

$$\tilde{n} = n\left(\frac{n_{\rm i,eff}}{N_{\rm C}}\right)^{\frac{|\nabla E_{\rm F,n}|}{F}}\ ,\quad \tilde{p} = p\left(\frac{n_{\rm i,eff}}{N_{\rm V}}\right)^{\frac{|\nabla E_{\rm F,p}|}{F}} \qquad (3)$$

that enter the driving term and account for the splitting of the quasi-Fermi levels between the classical turning points. The BTBT model was validated using IV and CV data of special Si pn-junction diodes and their SIMS profiles, provided by P. M. Solomon, IBM Yorktown Heights [1]. From

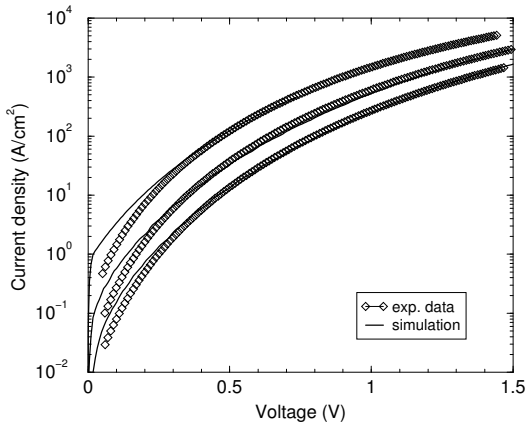

Fig. 5. Measured and simulated reverse-bias gate currents of pn-junction diodes with different junction profiles. Experimental data correspond to deep boron implantation dosages of 3.5, 4.0, and $6.3\times10^{14}\,\mathrm{cm}^{-2}$ [1] (from bottom to top).

the forward-bias branches of these samples, a series resistance of $21.5\,\Omega$ was extracted which also influences the currents at $V_G > 1\,V$ in the reverse-bias branches. This value is in good agreement with the measured series resistance. It effectively takes into account band structure details as well as intrinsic resistive effects like electron-hole scattering, the actual screening of the impurity scattering under degenerate conditions, and the actual activated doping. There is no unique fit of the four parameters A, B, $E_{g,\mathrm{eff}}$, and $\hbar\omega$. Changing just one of the default values provided in [5] already results in a fit comparable to that in Fig. 5. Using CV, reverse- and forward-bias IV, and various implantation dosages, the following set was determined: $A = 5\times10^{20}\,(\mathrm{cm\,s})^{-1}\mathrm{V}^{-2}$, $B = 2.04\times10^{7}\,\mathrm{V/cm}$, $E_{g,\mathrm{eff}} = 1.009\,\mathrm{eV}$, and $\hbar\omega = 30\,\mathrm{meV}$.

4. Technology variations for GIDL minimization

4.1. *22 nm DGSOI nFET*

4.1.1. *Steepness and size of the source/drain doping profiles*

Increasing the decay length of the lateral Gaussian donor profile as much as possible results in shallower junctions, but as expected, the effect on the OFF-current is practically zero, because the BTBT rate is *not* located in the pn-junction (see Fig. 6). The ON-current slightly degrades (2%). The location of the BTBT rate beneath the gate corners is a general feature of SOI MOSFETs with moderately

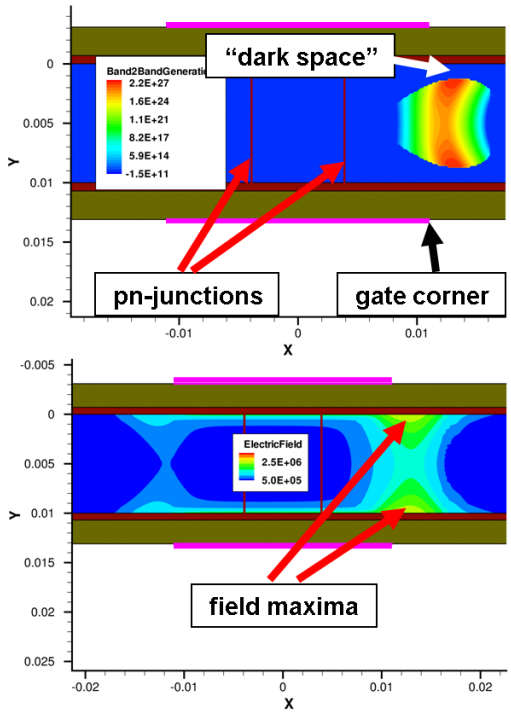

Fig. 6. Upper: BTBT rate location close to the gate corners in the 22 nm DGSOI nFET. Lower: Spatial distribution of the electric field.

doped body. Only with a heavily doped body, the maximum of the rate would occur at the metallurgical junction.

Decreasing the plateau level of the donor profile in the S/D regions also results in shallower junctions which now slightly shift outwards (see Fig. 8). A moderate decrease yields only a moderate reduction of the OFF-leakage, but a significant degradation of the ON-current. A strong reduction of the OFF-current can only be achieved with a strong decrease of the heavy S/D doping, at the price of a strongly reduced ON-current due to the much lower S/D conductivity.

4.1.2. *Gate overlap*

If the Gaussian donor profiles are shifted outwards, the gate overlap becomes smaller and the sub- threshold swing improves markedly due to the stronger source-drain potential barrier. The OFF-current reduction is proportional to the size of this shift as shown in Fig. 9. The ITRS limit is almost reached with "PeakPos"=24nm which means that the decay of the profile starts 24 nm from the center of the device. The pn-junctions are then almost

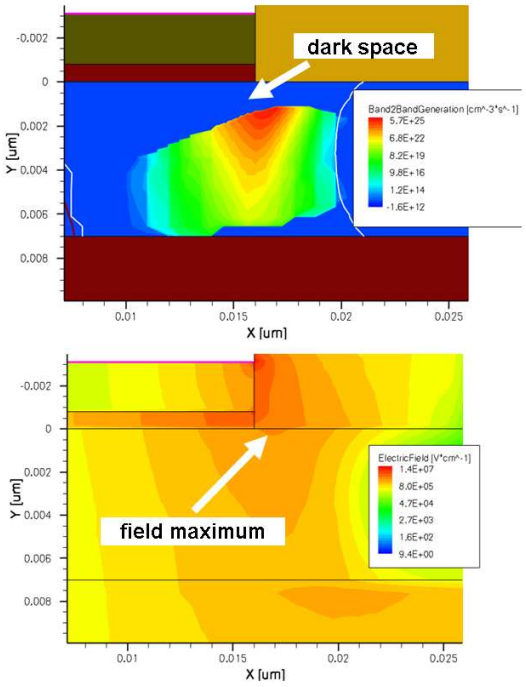

Fig. 7. Upper: BTBT rate location close to the gate corner in the 32 nm SGSOI nFET. Lower: Spatial distribution of the electric field.

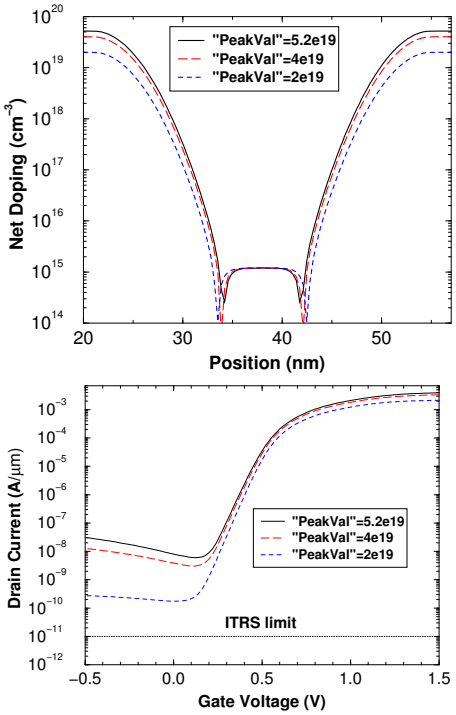

Fig. 8. Upper: Lowering of the plateau level of the S/D dopings. Lower: Corresponding transfer characteristics at $V_{DS} = 1\,V$.

at the position of the gate corners (almost zero gate overlap). The ON-current decreases by only 6%.

4.1.3. *Body thickness and HfO_2 layer thickness*

The voltage drop in the body and the effect of the "dark space" of the BTBT rate can be influenced by a change of the body thickness. However, decreasing the body thickness by 2 nm leads to a small reduction of the OFF-current only as shown in Fig. 10. At the same time, the ON-current degrades because of the lower mobility.

A larger voltage drop across the gate oxide should reduce the voltage drop beneath the gate corners and thus reduce the BTBT rate. However, increasing the HfO_2 layer thickness from 2.4 nm to 3.0 nm (which is a reasonable amount) produces only a very small effect (see Fig. 10). Hence, neither t_{Si} nor t_{HfO2} are effective in reducing the BTBT OFF-current.

4.2. *32 nm SGSOI nFET*

As can be seen in Fig. 7, the maximum of the electric field in the 32 nm SGSOI transistor in the OFF-state does *not* occur in the pn- junction either, but right to the drain-side gate corner (largest voltage

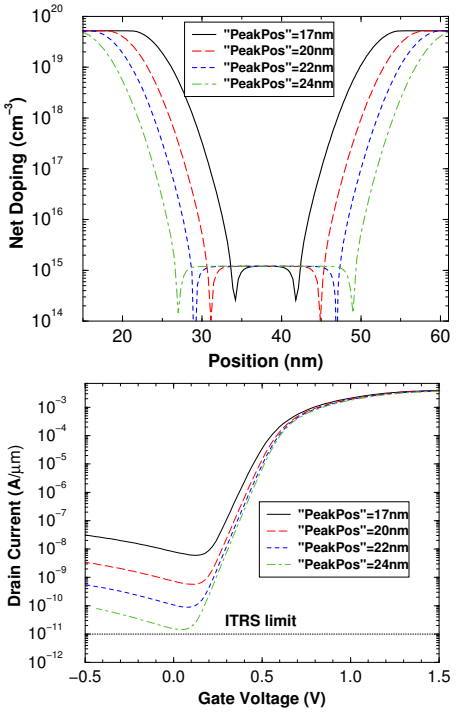

Fig. 9. Upper: Lateral shifts of the S/D doping profiles. Lower: Corresponding transfer characteristics at $V_{DS} = 1\,V$.

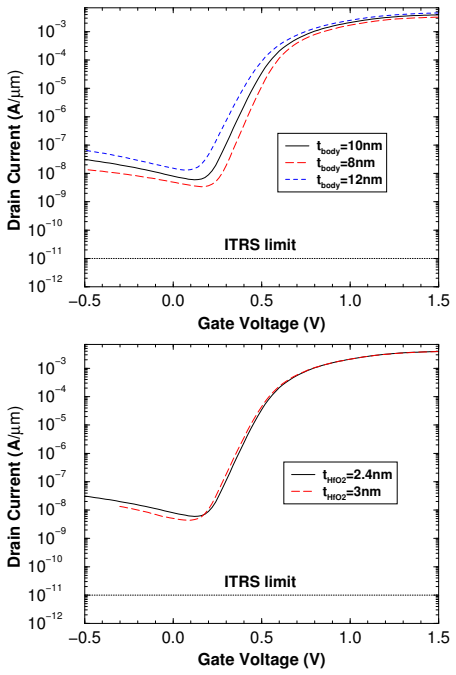

Fig. 10. Upper: Transfer characteristics at $V_{DS} = 1\,V$ for different Si body thicknesses. Lower: Transfer characteristics at $V_{DS} = 1\,V$ for different HfO_2 layer thicknesses.

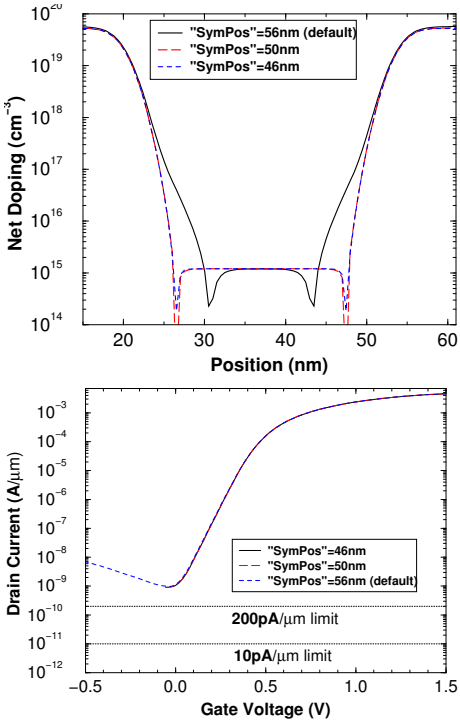

Fig. 11. Upper: Lateral profiles with different "SymPos" parameters for the LDD implants. Lower: Corresponding transfer characteristics at $V_{DS} = 1\,V$.

drop). The consequence is that the BTBT rate is also concentrated right to the gate corner and *not* at the metallurgical pn-junction. Therefore, as in the case of the 22 nm DGSOI FET, one cannot expect to change the BTBT rate much by changing the doping in the vicinity of the pn-junction. The doping profile was designed as combination of two error function profiles, one for the heavy S/D doping and the other for the LDD extensions [3]. The extensions produce the pn-junction which is displaced from the maximum of the BTBT rate by about 10 nm. Therefore, one anticipates significant effects only due to variations of the heavy S/D profiles. In fact, if the LDD extension profile is shifted outwards by more than 5 nm, it submerges with the heavy S/D profile which has no effect on the leakage current, as shown in Fig. 11.

4.2.1. *Lateral position of the heavy S/D doping profiles*

The heavy S/D doping profiles can be shifted outwards by increasing the "SymPos" parameter in the Error function (see Fig. 12). This has the effect that the heavy doping branch is no longer situated under the gate corner, where electric field and BTBT rate have their maxima. The $10\,pA/\mu m$ limit can be almost reached with "SymPos" = 66 nm, where the gate overlap with the heavy S/D profile has disappeared. Due to the worse capacitive coupling the threshold voltage shifts by $\sim 0.1\,V$, which is advantageous to meet the $10\,pA/\mu m$ target. However, the ON-current degrades at the same time by about 40%.

4.2.2. *S/D doping level*

The doping level in the S/D regions can be decreased by a smaller "MaxVal" parameter (see Fig. 13). A moderate decrease yields a moderate improvement for the OFF-leakage only, but also a moderate degradation of the ON-current. A stronger reduction of the OFF- current can only be achieved by decreasing the S/D doping significantly, at the price of a strongly reduced ON-current (lower S/D conductivity). For "MaxVal" $= 2 \times 10^{19}\,cm^{-3}$ the ON-current decreases by about 30%.

4.2.3. *Body thickness and HfO_2 layer thickness*

Similar changes of body thickness and HfO_2 layer thickness as in the case of the 22 nm DGSOI nFET were done for the 32 nm SGSOI nFET. If the body

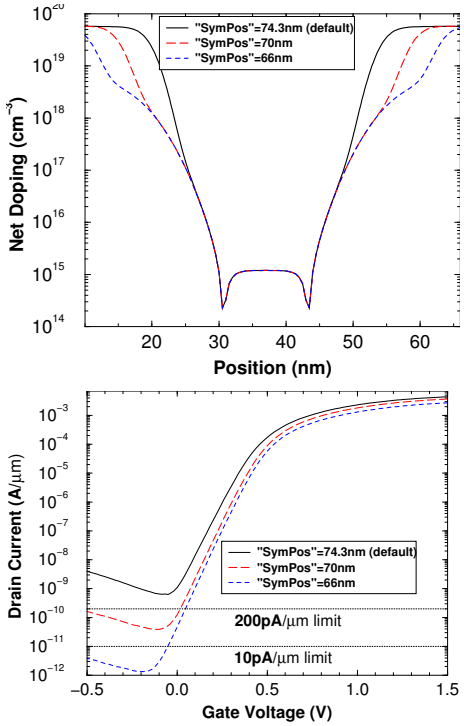

Fig. 12. Upper: Lateral profiles with different "SymPos" parameters for the heavy S/D implants. Lower: Corresponding transfer characteristics at $V_{DS} = 1\,V$.

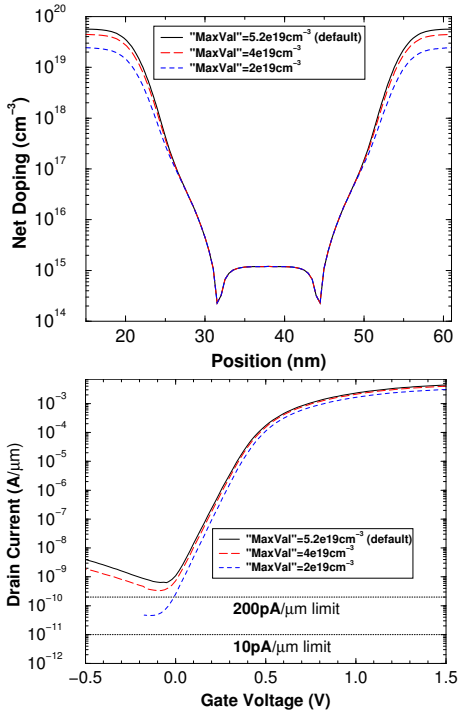

Fig. 13. Upper: Lateral doping profiles with different "MaxVal" parameters that change the S/D doping level. Lower: Corresponding transfer characteristics at $V_{DS} = 1\,V$.

thickness is decreased by 2 nm, the 200 pA/μm limit is almost reached as shown in Fig. 14. The ON-current degradation is 12%. However, quantum-mechanical confinement starts at a body thickness of 5 nm, which further reduces the mobility. This was not taken into account in the simulation.

Increasing the HfO_2 layer thickness from 2.3 nm to 3.0 nm produces only a small effect (see Fig. 14) and is thus not effective in reducing the BTBT OFF-current in accordance with the case of the 22 nm DGSOI nFET.

5. Conclusion

The 22 nm and 32 nm template FETs of this study exhibit a BTBT induced OFF-current which is about 2 - 3 orders larger than the 10 pA/μm target. The optimization of doping profile and geometry of the 22 nm DGSOI nFET shows that a reduction of the OFF-current has to be paid with a degradation of the ON-current. In trying to find the best compromise it turned out that the only efficient measure is a shift of the lateral doping profiles such that the pn-junctions become very close

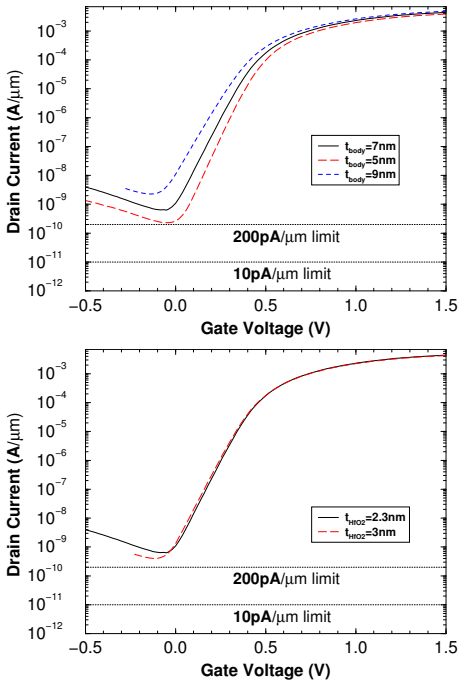

Fig. 14. Upper: Transfer characteristics at $V_{DS} = 1\,V$ for different Si body thicknesses. Lower: Transfer characteristics at $V_{DS} = 1\,V$ for different HfO_2 layer thicknesses.

to the gate corners. This brings I_{off} down to almost the ITRS limit while decreasing the ON-current by a few % only. All other variations studied (junction steepness, doping level, body thickness, gate oxide thickness) have either a minor effect or are linked with an unacceptable degradation of I_{on}. Also in the 32 nm SGSOI nFET the BTBT rate has its maximum under the drain-side gate corner and not in the pn-junction. With a work function of 4.6 eV it is impossible to find variations of the extension and S/D doping profiles such that I_{off} becomes lower than the 10 pA/μm target without unacceptable degradations of I_{on}. A possible remedy is given by a careful combination of changes of work function, S/D profile, and gate oxide thickness. All these findings are strongly related to the SOI architecture with lowly doped body, where the maximum of the BTBT rate is not located at the metallurgical junction, but close to the gate corner.

6. Acknowledgment

The author is grateful for the financial support by the European projects PULLNANO (IST-4-026828) and NANOSIL (IST-216171) He also thanks Dr. P. M. Solomon, IBM Yorktown Heights, for providing experimental data and for many valuable discussions.

References

[1] P. M. Solomon, J. Jopling, D. J. Frank, C. D'Emic, O. Dokumaci, P. Ronsheim, and W. E. Haensch, "Universal tunneling behavior in technologically relevant P/N junction diodes", JAP 95(10), 5800-5812, 2004.

[2] A. Schenk, "Scalability Study of Floating Body Memory Cells", accepted for SISPAD 2009, San Diego, California.

[3] EU-IST-4-026828-PULLNANO, http://www.pullnano.eu.

[4] http://www.itrs.net/reports.html.

[5] Synopsys Inc, Sentaurus Device User Guide, version Z-2007.03, Mountain View, California, (2007).

[6] S. Monaghan, P. K. Hurley, K. Cherkaoui, M. A. Negara, and A. Schenk, "Determination of electron effective mass and electron affinity in HfO_2 using MOS and MOSFET structures", Solid-State Electronics 53(4), 438 - 444 (2008).

[7] A. Schenk, "Rigorous Theory and Simplified Model of the Band-to-Band Tunneling in Silicon", Solid-State Electronics, vol. 36 (1), pp. 19-34 (1993).

[8] A. Schenk, "Physical Models for Semiconductor Device Simulation", Festkörperprobleme (Advances in Solid State Physics), vol. 36, pp. 245-263 (1996).

[9] A. Schenk, "Finite-Temperature Full-Random Phase Approximation Model of Band Gap Narrowing for Silicon Device Simulation", J. Appl. Phys. 84 (7), 3684-95 (1998)

[10] P. P. Altermatt, A. Schenk, and G. Heiser, "A simulation model for the density of states and for incomplete ionization in crystalline silicon. I. Establishing the model in Si:P", J. Appl. Phys., vol. 100 (11), 113714 (2006)

[11] P. P. Altermatt, A. Schenk, B. Schmithüsen, and G. Heiser, "A simulation model for the density of states and for incomplete ionization in crystalline silicon. II. Investigation of Si:As and Si:B and usage in device simulation", J. Appl. Phys., vol. 100 (11), 113715 (2006)

[12] R. A. Logan and A. G. Chynoweth, "Effect of Degenerate Semiconductor Band Structure on Current-Voltage Characteristics of Silicon Tunnel Diodes", Phys. Rev. **131**(1), 89 - 95 (1963)