Impact of Technology Parameters on Inverter-Delay of UTB-SOI CMOS

Thomas Schulz, Christian Pacha, and Lothar Risch

Infineon Technologies AG Corporate Research, Otto-Hahn-Ring 6, 81730 München, Germany
thomas.schulz@infineon.com

Abstract

Ultra-thin-body silicon-on-insulator (UTB-SOI) is one of the most promising candidates for future CMOS technologies with minimum feature sizes below 50nm [1]. In this paper we analyse the impact of different parameters such as doping profile, gate workfunction and local interconnects on the inverter propagation delay. For that purpose we have simulated fully-depleted (FD) SOI-MOSFETs with thin undoped silicon bodies using a coupled device and circuit simulation.

1. Introduction

An alternative to the classical scaling concept of bulk-MOSFETs with continuous increase in channel doping concentration and n^+/p^+ -polysilicon dual gate workfunction is the emerging concept using extreme thin and undoped silicon films with midgap gate materials to achieve the ITRS Roadmap specifications. In this concept the effective channel length is no longer a figure of merit for device performance because the metallurgical pn-junction vanishes. New parameters such as the gradient of the doping concentration, the silicon film thickness and the gate workfunction of new materials have now significant impact on device performance.

Due to the lack of accurate physics-based compact models for FD-UTB-SOI with undoped silicon bodies we investigate the impact of these technological parameters on the inverter delay using the MIXEDMODE simulator from Silvaco [2]. In such a circuit simulator the intrinsic devices are calculated by device simulation (ATLAS) and the SPICE input language is used for circuit specification.

2. Doping concentration gradient

The cross section of the device structure for the mixedmode simulation is given in **Fig. 1**. A comparison concerning doping profile and dc-performance between two devices with different doping concentration-gradients is shown in **Figs. 2, 3, and 4**. Note, that the proportion of channel length and silicon thickness L_G/t_{Si} is ten resulting in an eminent suppression of short channel effects. The doping concentrations are assumed to be constant in the source and drain regions and to

decay with a Gaussian distribution into the channel region. For comparison there are two curves shown in **Fig. 2** for 3 and 5 nm/decade with an exponential decrease, too. The assumed Gaussian profile is more flat at high concentrations and more steep for low concentrations compared to the average value of 3 and 5 nm/decade, respectively. The two different profile curves of arsenic crossing each other at a concentration of $1E19 \text{ cm}^{-3}$. This crossing point is also the beginning of the gate overlap from $x=50\text{nm}$ to $x=100\text{nm}$. The distance between the crossing points at the boron curve ($1E14 \text{ cm}^{-3}$) defines the effective channel length of the UTB-SOI devices with “undoped” body in the region where $N_{SD}(x) < N_b$.

The first result is that the difference between 10nm and 30nm in effective channel length of the two devices has no impact on the on-current, the difference is only 1-2%. Therefore, only the output characteristics of the nMOS and pMOS devices with 5 nm/decade are shown in **Fig. 4**. Even if the body doping concentration increases to $1E16 \text{ cm}^{-3}$ there is no influence on the device performance.

In contrast to this the steeper doping concentration is unfavourable for the off-current despite of the smaller gate overlap region. The devices with 3nm/decade have a ten times larger off-current at $|V_{DS}|=0.9\text{V}$ as shown in **Fig. 3**. The origin of these different behaviours is the band-to-band tunnelling current (GIDL, gate-induced drain leakage) which is higher at steeper doping profiles. The GIDL-current is the off-current limitation for UTB-SOI MOSFETs. Thus, in the following we concentrate on the second device with 5nm/decade (see Table 1). For mixedmode calculations we use the following models: Drift-Diffusion with band-to-band tunnelling but without impact ionisation.

Table 1. UTB-SOI device parameters.

Parameter	Value
Gate length L_G [nm]	50
Effective channel length L_{eff} [nm]	10
Spacer length L_{sp} [nm]	10
Equivalent front-oxide thickness t_{fox} [nm]	1.1
Si-body thickness t_{Si} [nm]	(10), 5
Back-oxide thickness t_{box} [nm]	100
Body doping N_b [cm^{-3}]	10^{14}
Source/drain doping N_{SD} [cm^{-3}]	10^{21}

3. Gate workfunction

New candidate materials for dual workfunction metal gates are under discussion (for nMOSFETs Tantalum nitride TaN or TaSi_xN_y and for pMOSFETs Ruthenium Ru or RuO₂). Optimised midgap gates are possible with Nitrogen implantation into Molybdenum Mo [3] or Ru-Ta alloys by varying the atomic concentration [4]. However, for these materials the value of the final workfunction depends on processing and underlying dielectric. Therefore, we investigate the impact of the supply voltage of the device as a function of the gate workfunction on the inverter delay. To evaluate the performance of a basic inverter circuit we use the simulation scheme illustrate in **Fig. 5**. The simulation scheme ensures a realistic input pulse forming by the first inverter and also considers the fan-out and wire loading by R_{load} and C_{load} . The inverter is optimised to switch at $V_{DD}/2$ by adjusting the channel width ratio W_p/W_n . (e.g. $2.46\mu m/1\mu m$ @ $V_{DD}=0.8V$, $\Phi_M=4.71eV$). The simulated average inverter delay $t_D=0.5(t_{D10}+t_{D01})$ is based on the delays of the 1-0 and the 0-1 transitions (**Fig. 6**).

Fig. 7 shows I_{ON} versus I_{OFF} curves of the nMOSFET for different gate workfunctions and supply voltages. The dashed lines mark the roadmap specifications for the different types of devices (HP=high performance, LOP=low operation power, LSTP=low standby power). The simulations results indicate that for the HP-device a wide range of different gate materials and supply voltages will satisfy the roadmap specifications. For the LOP-device the design window is much smaller. Only midgap gate materials in the range of 4.71 ± 0.1 eV can achieve on-currents exceeding $600\mu A/\mu m$ and off-currents below $300pA/\mu m$. The LSTP specification is not achievable for this 50nm-device with any workfunction combination because the GIDL-current limits the off-current to values above $100pA$.

Finally, the simulated average gate delay is shown in **Fig. 8** for three different supply voltages as a function of different dual-workfunction combinations. The inverter with gate workfunctions of 4.51 eV for the nMOSFET and 4.91 eV for the pMOSFET achieves 3.85 ps in delay time at $V_{DD}=1$ V. The calculated gate delay metric of the inverter is $\tau_{inv}=C_{OX}V_{DD}/I_{ON}=3.83$ ps with $C_{OX}=(W_n+W_p)L\epsilon_0\epsilon_{OX}/t_{OX}=5$ fF and $I_{ON}=0.5(I_{ON,n}+(W_p/W_n)I_{ON,p})=I_{ON,n}=1320\mu A$. The dimensions chosen are $L=50nm$, $W_n=1\mu m$, $W_p=2.21\mu m$, $t_{OX}=1.1nm$.

This gate delay metric is comparable with $\tau=3.70ps$ of the single HP-nMOSFET in the ITRS 1999 Roadmap but the new ITRS 2001 Roadmap demands gate delays of 1.2ps for HP-nMOSFETs. A rule of thumb says that an inverter with a fan-out of 4 is 14 times slower than the single nMOSFET. In this case the simulated FO4-inverter has to be faster than 16.8ps as shown in **Fig. 9**. Consequently the devices achieve also the delay specifications of the ITRS 2001 Roadmap.

4. Local Interconnects

A third technological innovation in UTB-SOI MOSFETs is a low-k intermetal dielectric (IMD). However, the integration significantly complicates moving from SiO₂ ($k=4$) to porous low-k materials ($k=2$) or to ideal airgaps ($k=1$). To assess the necessity of introducing low-k IMD for local interconnects we investigate three different scenarios (see **Fig. 10**) for the comparison of two MOSFETs with different channel widths ($W=50nm$, left column and $W=1\mu m$, right column). In the first case the interconnect of $1\mu m$ length has no neighbouring line, and in the second and third case there are one and three neighbouring lines, respectively (cases A, B and C correspond to $C_{WIRE}=0.01$; 0.06 and 0.12fF).

Fig. 11 depicts the impact of the channel width on the inverter delay t_D as a function of the local interconnect capacity C_{WIRE} . Consequently, wide driver MOSFETs are not sensitive to the capacitive surrounding. In contrast to this, for narrow MOSFETs such as used in low power logic gates the inverter delay is increasing dramatically. Nevertheless, the gate capacitance is still the dominating contribution of the inverter delay t_D as shown in **Fig. 9** for large fan-out.

5. Conclusions

The impact of various technological parameters on inverter delay of UTB-SOI devices with undoped bodies has been analysed by mixedmode simulations for different supply voltages. We find that with a steeper doping concentration gradient there is no influence on the on-current, but a significant off-current due to the increase of the GIDL-current occurs.

For high performance devices there is a wide design window with different gate workfunctions to achieve the roadmap specifications. Devices for low operation power circuits can meet the roadmap targets with midgap workfunctions but devices for low standby power applications suffer from the high GIDL-current.

The use of low-k materials instead of silicon dioxide for local interconnects can improve the speed of digital circuits up to 20% for circuit designs with condensed wiring and with minimum sized driver MOSFETs.

6. References

- [1] Y.-K. Choi, et. al., "Ultrathin-Body SOI MOSFET for Deep-Sub-Tenth Micron Era", *IEEE Electron Device Letters*, May 2000, pp. 254-255.
- [2] Silvaco International, *Atlas User's Manual*, Santa Clara, October 1996.
- [3] R. Lin, et. al., "An Adjustable Work Function Using Mo Gate for CMOS Devices", *IEEE Electron Device Letters*, January 2002, pp. 49-51.
- [4] H. Zhong, et. al., "Properties of Ru-Ta Alloys as Gate Electrodes For NMOS and PMOS Silicon Devices", *IEDM Tech. Dig.*, Dec. 2001, pp. 467-470.

Figure 1. Device structure for drift-diffusion and mixedmode simulation.

Figure 2. Doping profile along the cutline for two devices with different concentration-gradients.

Figure 3. Subthreshold I_D - V_{GS} characteristics of the 50nm UTB-SOI N- and P-MOSFETs.

Figure 4. Output I_D - V_{DS} characteristics of the 50nm UTB-SOI N- and P-MOSFETs (5nm/decade).

Figure 5. Simulation scheme for the inverter chain with the MOSFETs n1,2; p1,2 the input pulse forming at V_{IN} and the fan-out RC-load at V_{OUT_2} .

Figure 6. Transient and transfer characteristics of the inverter. The voltages V_{IN} , V_{OUT_1} and V_{OUT_2} correspond to the labelled circuit nodes in Figure 5.

Figure 7. I_{ON} - I_{OFF} curves of the nMOSFET for different gate workfunction combinations and supply voltages (HP, LOP, and LSTP = ITRS specs.).

Figure 8. Inverter delay t_D as a function of the gate workfunction for different supply voltages.

Figure 10. 3D-illustration for the comparison of two MOSFETs with different channel widths ($W=50nm$, left column and $W=1\mu m$, right column) for three scenarios. In the first case the interconnect (light green) has no neighbouring line and in the second and third case there are one and three neighbouring lines, respectively. ($C_{WIRE}=0.01; 0.06$ and $0.12fF$).

Figure 9. Impact of different supply voltages as a function of the fan-out on inverter delay t_D .

Figure 11. Inverter delay t_D as a function of the local parasitic wiring for different channel widths for FO1.