International Journal of Modern Physics B
© World Scientific Publishing Company

World Scientific
www.worldscientific.com

ON THE RELATIONSHIP BETWEEN EFFECTIVE ELECTRON MOBILITY AND KINK EFFECT FOR SHORT-CHANNEL PD SOI NMOS DEVICES

MILIJA SARAJLIĆ*

IHTM-Center for Microelectronic Technologies and Single Crystal,
Njegoševa 12, 11000 Belgrade, Serbia,
milijas@nanosys.ihtm.bg.ac.yu

RIFAT RAMOVIĆ

School for Electric Engineering, Bul. Kr. Aleksandra 73a,
11000 Belgrade, Serbia,
ramovic@etf.bg.ac.yu

Received 27. September 2006
Revised 10. November 2006

Here is given a new approach for calculating triggering drain bias at the onset of the kink effect, kink voltage, V_{kink}, for PD SOI NMOS devices utilizing electron drift properties in the channel. This approach directly relates electron mobility in the channel to the kink effect and enables one to determine kink voltage knowing the device technology. It also gives possibility for calculating mobility from the kink voltage. Theory is compared to the previously published experimental results and based on this match behavior of the kink voltage for PD SOI NMOS components for various technology parameters is predicted. Explanation for the appearance of the kink in the volt regime below bandgap of silicon is also given. From this consideration design rules for PD SOI NMOS devices are derived in order to sooth the kink effect.

Keywords: SOI; SOI MOSFET; Kink Effect; Impact Ionization; Parasitic Bipolar Device.

1. Introduction

Silicon-on-insulator (SOI) CMOS offers a 20–35% performance gain over bulk CMOS.[1] Some of the recent applications of SOI are in high-end microprocessors and its upcoming uses are in low-power, radio-frequency (RF) CMOS and embedded DRAM (EDRAM), to name a few. As we move to the 0.1 μm generation and beyond, SOI is expected to be the technology of choice for system-on-a-chip applications which require high-performance CMOS, low-power, embedded memory, and bipolar devices.[1] The primary feature of MOS in SOI is that the local substrate ("body") of the device floats electrically, and therefore the substrate–source bias voltage, V_{BS}, is not fixed. This is especially prominent for Partially-Depleted (PD) SOI MOS devices where active layer is not totally depleted but it leaves an island of non-depleted region close to the buried oxide, Fig. 1.

* Corresponding author

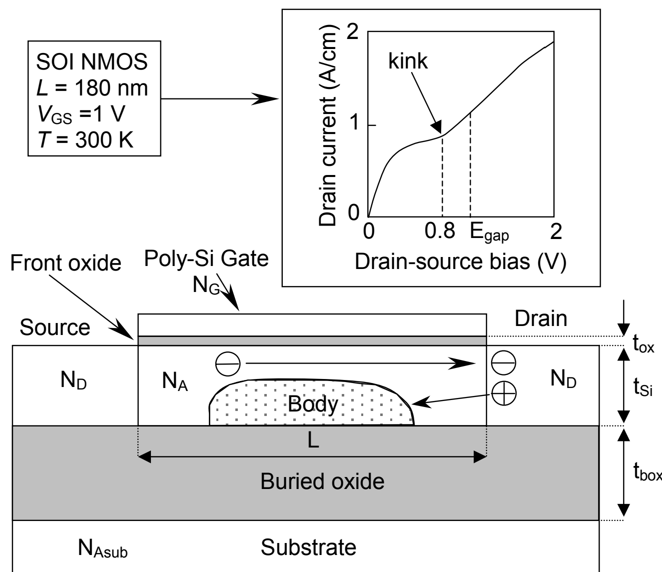


Fig. 1. Schematic representation of Partially Depleted SOI NMOS device. Mechanism of impact ionization and kink is sketched. At the inset, typical drain current characteristic is shown.[2]

As V_{BS} changes, the device threshold voltage, V_T, will change. This "instability" in V_T is what has made SOI device design very challenging. One manifestation of the threshold variation is the "kink effect," or increase in the output conductance of the device when drain-to-source bias, V_{DS}, comes near to 1 V, *i.e.* the band gap of silicon; however, the kink effect could appear even for the drain-to-source bias below the band gap of silicon, inset of the Fig. 1.[2] This is caused by the impact-ionization induced increase in V_{BS} with increasing V_{DS}, and the resulting reduction of V_T; when V_{DS} becomes large enough, impact ionization current (holes) flows to the non-depleted body, increasing the body charge and V_{BS}, resulting in a decrease in V_T. This is much more pronounced for SOI NMOS devices than for SOI PMOS devices because effective cross-section for impact ionization is much higher for electrons than for holes. For this reason, consideration in this work is restricted to NMOS devices although the same approach could be implemented for PMOS.

Numerical and semi-analytical modeling of the kink effect was already reported.[3, 4, 5]. However, there is a single article, to our knowledge, published so far, that deals with purely analytical modeling of the kink effect.[6] Although they take very profound physical properties into consideration, derived mathematical expressions are cumbersome and difficult for practical implementation. In this article it is showed that it is possible to make an analytical model of the kink effect for PD SOI NMOS devices based on electron drift properties in the channel. Kink effect and effective electron mobility in the channel

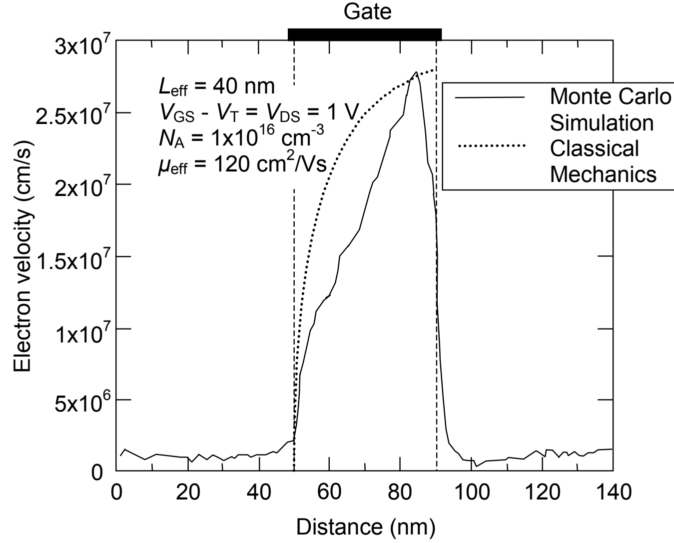


Fig. 2. Electron velocity vs. distance in the channel. Comparison between Monte Carlo simulation[6] and Classical Mechanics. Device parameters are: $L_{\mathrm{eff}} = 40$ nm, $V_{GS} - V_T = V_{DS} = 1$ V, $N_A = 1\times10^{16}$ cm^{-3}, estimated effective mobility $\mu_{\mathrm{eff}} = 120$ cm^2/Vs.

are directly related, and thus quite feasible formula is obtained. Previously published experimental data for the kink is fitted to the theory presented here introducing a single fitting parameter and from that point forward behavior of the kink for various technology parameters for PD SOI NMOS is predicted. Also, by this approach it is possible to explain the most striking effect of the nanoscaled SOI NMOS components and that is appearance of the kink for the drain bias below bandgap of silicon.[2] This theory is applicable for devices with effective gate length below 600 nm. The bottom channel length for application of this theory is appearance of the FD (Fully Depleted) devices. This approach is valid vice versa; it is possible to estimate electron mobility in the channel knowing the kink voltage, V_{kink}, *i.e.* triggering drain-to-source bias at the onset of the kink effect.

2. Theory

In this consideration classical mechanics is employed. The first question arises: is it possible to make proper model of electron drift in the channel by classical mechanics? Although quantum mechanics is used for description of the nanoscaled devices, it is still possible to apply classical mechanics in the consideration of the impact ionization effects. At the Fig. 2 electron velocity in the channel vs. distance as calculated by classical mechanics and Monte-Carlo method[7] is given. Monte-Carlo method is more accurate because it takes into account single collisions of the electron with phonons in the lattice.

As can be seen, Fig. 2, there is apparent difference between these two approaches for electron velocity along the channel, but maximum velocity is the same in both cases. Device parameters for the simulated component at the Fig. 2 are: front oxide thickness $t_{ox} = 1$ nm, active layer thickness $t_{Si} = 15$ nm, buried oxide thickness $t_{box} = 400$ nm, doping within drain/source regions $N_{SD} = 1\text{x}10^{20}$ cm^{-3}, active layer doping $N_A = 1\text{x}10^{16}$ cm^{-3}, substrate-to-source bias $V_{BS} = 0$ V, temperature $T = 300$ K, channel length $L = 40$ nm, gate-to-source bias above threshold and drain bias $V_{GS} - V_T = V_{DS} = 1$ V, estimated electron mobility in the channel is $\mu = 120$ cm^2/Vs. Constant longitudinal electric field and constant mobility along the channel is assumed in calculating classical velocity. What matters in estimating the kink effect is maximum electron velocity that the electron has at the drain interface, so regarding this effect it is convenient to model electron drift in the channel by the means of the classical mechanics.

It is interesting to calculate drift velocity that the electron has to attain in order to ionize silicon, *i.e.* to create electron-hole pair. Kinetic energy of a single electron is: $T = 1/2mv^2$, where m is the electron rest mass and v is the electron velocity. For impact ionization to occur, kinetic energy is to equal to the Si band-gap, $E_{gap} = 1.12$eV. If electron rest mass, $m = 9\text{x}10^{-31}$ kg is substituted, electron velocity of the impact ionization will be: $v_{ion} = 6.3\text{x}10^7$ cm/s, what is more than six times higher than electron saturation velocity for the silicon![8] For silicon saturation velocity, the majority of authors agree that it is below $1.0\text{x}10^7$ cm/s.[9] This means that all electrons causing impact ionization are high in velocity overshoot. To our knowledge, never before is this paradox reported in the literature. It is obvious that the "kink electrons", *i.e.* electrons that cause impact ionization, are really special species in compare to the rest of the "ordinary" electrons.

Fig. 3 depicts electron velocity vs. distance along the channel, as calculated by classical mechanics, for different channel lengths and for the same drain-to-source bias $V_{DS} = 1$V. Electron mobility is assumed to be equal to the maximum *i.e.* intrinsic mobility, $\mu = 1400$ cm^2/Vs. Velocity is derived from differential equation for the electron drift:

$$mv\frac{dv}{dx} = q\frac{V_{DS}}{L} - \frac{q}{\mu}v \quad (1)$$

where m is the electron rest mass, x is the distance along the channel, v is the electron drift velocity along the channel, q is the electron charge, V_{DS} is the drain-to-source bias, L is the channel length and μ is the electron mobility in the channel. Solution to this differential equation is:

$$v = \frac{\mu V_{DS}}{L}\left(1 + ProductLog\left(-e^{-1-\frac{qL}{m\mu^2 V_{DS}}x}\right)\right) \quad (2)$$

where $ProductLog(z)$ gives the principal solution for w in $z = we^w$.

As can be seen on Fig. 3 there are two distinct regions of electron drift. First region

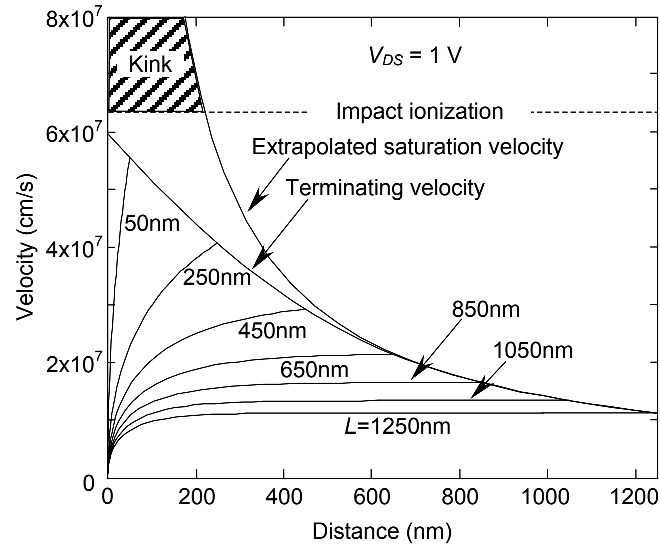


Fig. 3. Velocity in the channel vs. distance for different channel lenghts. Terminating velocity at the drain interface and extrapolated saturation velocity are also depicted. Shaded area represents the onset of the kink. Drain-to-soure bias is 1V.

belongs to the acceleration, where electron is accelerated up to the saturation velocity $v_{\text{sat}} = \mu V_{\text{DS}}/L$, and after that region it travels with constant velocity v_{sat}. On the Fig. 3 it is shown that for sufficiently short channel, electron will not attain saturation velocity at the point where it reach drain interface; this is pointed out by two different curves that are depicting extrapolated saturation velocity and terminating velocity at the drain interface. For channel lengths below 600 nm, electron drift velocity is extrapolated in order to find saturation velocity. Below 600 nm there is apparent difference between terminating velocity and extrapolated saturation velocity, Fig. 3. Key assumption is that electron is suffering no recombination in the acceleration region. In that manner, the same electrons that emerge at the source will immerge at the drain; this means that it is possible to model situation at the drain interface knowing situation at the source. This is not applicable for longer channel because electron will suffer multiple recombination and subsequent generations before it reaches drain interface. This assumption is reminiscent of the famous Shockley theory of the "lucky electron" where "lucky electron" will travel all the way through the channel without collisions with phonons, or it could even receive more phonons than it emits, so final velocity will be above the average. Here it is only expected for the electron not to suffer recombination during flight; but no ballistic electrons are taken into consideration. Taking all this into account this theory will be restricted only to the electrons where extrapolated saturation velocity and terminating

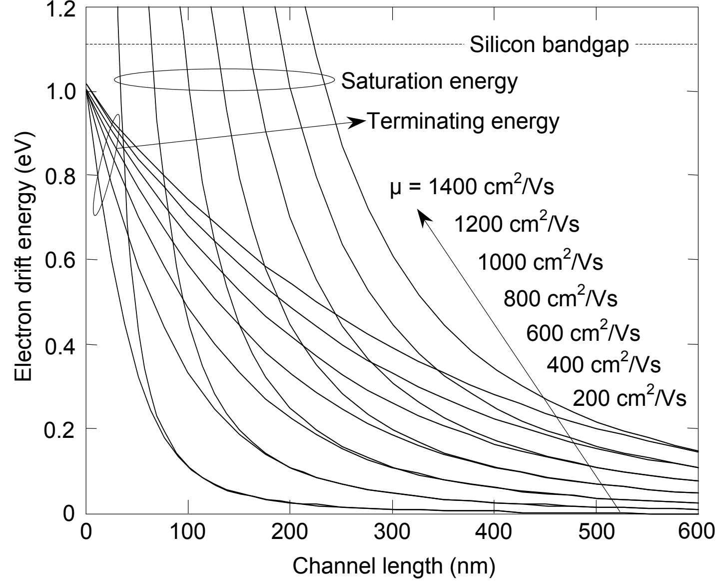


Fig. 4. Electron drift energy in the channel vs. distance for different mobility. Saturation energy and terminating energy are depicted. Drain-to-soure bias is 1V.

velocity do not equal. Fig. 4 depicts electron drift energy in the channel vs. channel length for various electron mobilities. Saturation electron energy and terminating electron energy are the same for longer channel, but for shorter channel they start to separate from each other. This separation occurs at shorter channel length as electron mobility is lowered. Theoretically, for the highest possible mobility, *i.e.* electron mobility for intrinsic silicon, $\mu = 1400\text{cm}^2/\text{Vs}$, separation between saturation velocity and terminating velocity occurs at L = 600nm, so this value of the channel length is declared to be the largest one for the application of this theory.

Now we have a possibility to formulate basic model. Impact ionization in the drain region will occur if saturation energy is equal to the silicon band-gap energy:

$$E_{gap} = \frac{1}{2} m v_{sat}^2 = \frac{1}{2} m \left(\frac{\mu V_{kink}}{L} \right)^2 \tag{3}$$

where V_{kink} is the drain-to-source bias at the onset of the kink effect. This comes from the observation that all electrons emerging from the source, no matter what initial velocity they have, are accelerated towards saturation velocity v_{sat}. It is straightforward from here to express the kink voltage:

$$V_{kink} = \frac{L}{\mu} \sqrt{\frac{2E_{gap}}{m}} \tag{4}$$

Very special issue in this approach is modeling of the electron mobility in the channel. There is no universal model that depends on geometry and technology parameters of SOI device. In this situation two successive stages in mobility modeling are employed. First, mobility for the doped Si with the same doping that active layer is made of is estimated. Semi-empirical formula is utilised:[10]

$$\mu = 200\frac{cm^2}{Vs}\left(1+\frac{2\times 10^{18}cm^{-3}}{N_A}\right) \tag{5}$$

where μ is the electron mobility of silicon depending only on doping concentration. This value will be used as maximum mobility in the channel of the SOI NMOS. In the second stage of the mobility modeling this expression is introduced:

$$\mu_{eff} = \frac{2\sqrt{b}\mu\left(E_s+\sqrt{b}\right)}{b+\left(E_s+\sqrt{b}\right)^2} \tag{6}$$

where μ_{eff} is the effective electron mobility in the channel, E_s is the transverse electric field in the front channel, V_{GS} is the gate-to-source bias, V_T is the threshold voltage, μ is the maximum mobility as calculated from (5), and b is the fitting parameter. This form of the mobility model is quite convenient because shape of the curve resembles typical mobility dependence on transverse surface electric field E_s, at the same time preserving point of maximum while fitting is done by adjusting parameter b. Transverse surface electric field E_s depends on the gate-to-source bias. Mobility dependence on the lateral electric field in the channel is not taken into account because that would prevent from obtaining sufficiently high electron velocity for the impact ionization to occur. In (6) it is envisioned that maximum mobility corresponds to the threshold bias V_T, based on the experimental observation from the Ref. 11. Therefore, it is required to find the threshold for a given device technology. Quasi-2D Model obtained through full depletion approximation[12] is utilized. Threshold formula read as:[12]

$$V_T = V_{FB} - 2\phi_f + \frac{t_{ox}}{t_{box}+\gamma t_{Si}}\left(-2\phi_f - V_{BS} + V_{FBsub} + \frac{qN_A t_{Si}\left(2t_{box}+\gamma t_{Si}\right)}{2\gamma\varepsilon_{Si}}\right)$$
$$-\frac{t_{box}+\gamma t_{Si}}{t_{box}+t_{ox}+\gamma t_{Si}}\frac{\left\lfloor 2(V_{BI}-2\phi_f)+V_{DS}\right\rfloor}{2ch(L/2\lambda)-2} \tag{7}$$

$$V_{FB} = \frac{kT}{q}ln\left(\frac{N_A N_G}{n_i^2}\right) \tag{8}$$

$$V_{BI} = \frac{kT}{q}ln\left(\frac{N_A N_{SD}}{n_i^2}\right) \tag{9}$$

$$\phi_f = \frac{kT}{q}ln\left(\frac{N_A}{n_i}\right) \tag{10}$$

$$\lambda = \sqrt{\frac{t_{Si} t_{box} t_{ox}}{\gamma(\gamma t_{Si} + t_{box} + t_{ox})}} \tag{11}$$

$$\gamma = \frac{\varepsilon_{Si}}{\varepsilon_{ox}} \tag{12}$$

$$V_{FBsub} = \frac{kT}{q} ln\left(\frac{N_{Asub}}{N_A}\right) \tag{13}$$

where V_{FB} is the flat-band voltage in the front channel, k is the Boltzmann constant, q is the electron charge, N_G is the doping of the polysilicon gate, n_i is the intrinsic carrier density, ϕ_f is the Fermi level, γ is a dimensionless constant, ε_{Si} is the dielectric permittivity of silicon, ε_{ox} is the dielectric permittivity of SiO_2, V_{FBsub} is the flat-band voltage in the back channel, V_{BS} is the substrate-to-source bias, N_{Asub} is the doping of the substrate, V_{BI} is the built-in potential between channel and source/drain regions, N_{SD} is the doping concentration in the source/drain regions and λ is the characteristic length.

Now it is possible to find relationship between gate-to-source bias, V_{GS}, and electric field in the front channel, E_s:[12]

$$E_s = \left(-\frac{\partial \phi}{\partial x}\right)_{x=0} = \gamma \frac{V_{GS} - V_{FB} - \phi_s}{t_{ox}} \tag{14}$$

where ϕ_s is the front channel potential. Eventually, expressions for ϕ_s and E_s are:[12]

$$\phi_s = V_{GS} - V_{FB} - \frac{t_{ox}}{t_{box} + \gamma t_{Si}}\left(-V_{BS} + V_{FBsub} + \frac{qN_A t_{Si}\left(2t_{box} + \gamma t_{Si}\right)}{2\gamma\varepsilon_{Si}}\right) \tag{15}$$

$$E_s = \frac{\gamma\left(\phi_s - V_{BS} + V_{FBsub} + \dfrac{qN_A t_{Si}\left(2t_{box} + \gamma t_{Si}\right)}{2\gamma\varepsilon_{Si}}\right)}{t_{box} + \gamma t_{Si}}. \tag{16}$$

3. Results and discussion

Fig. 5 depicts electron mobility in the channel vs. gate-to-source bias as calculated by (6) for different doping concentrations. Device parameters are: $t_{ox} = 5$ nm, $t_{Si} = 50$ nm, $t_{box} = 500$ nm. Mobility is lowered in the whole range of V_{GS} as active layer doping is increased. Maximum mobility corresponds to the threshold bias voltage.[11] By this mobility model it is possible to predict the onset of the kink effect. For the sake of model validation it is compared to the experimental results. Fig. 6 shows kink voltage vs. gate bias for the device with technology parameters: $t_{ox} = 4$ nm, $t_{Si} = 126$ nm, $t_{box} = 360$ nm, $N_A = 2\text{x}10^{18}\text{cm}^{-3}$, $L_{eff} = 70$ nm.[13] On the Fig. 6 experimental results from Shahidi et al.[13] and theoretical results from this work for SOI NMOS device are compared (dots). Also, electron mobility in the channel as implicitly given in (4) is calculated and compared to the theoretical curve obtained from (6). Theoretical curve is fitted to the experimental

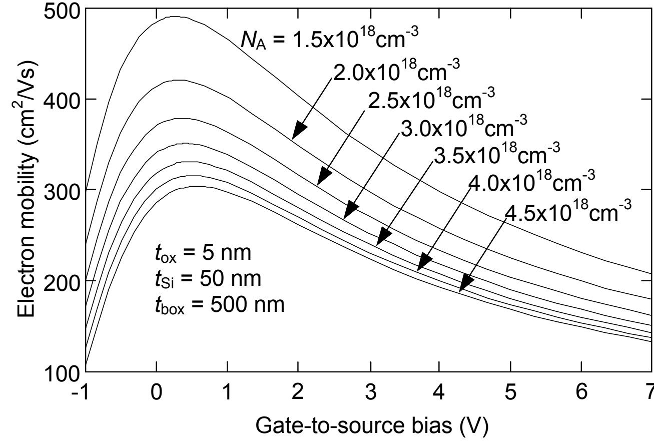


Fig. 5. Electron mobility in the channel vs. gate-to-source bias for different doping concentrations.

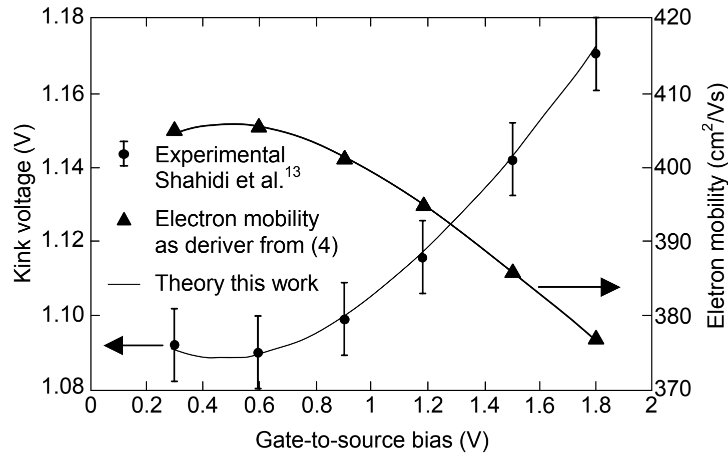
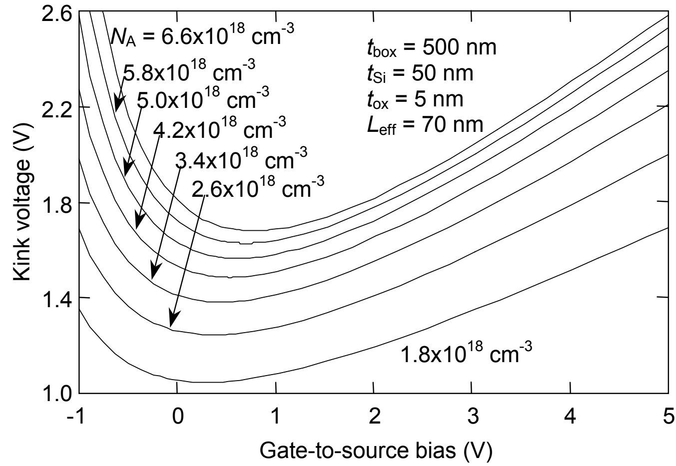
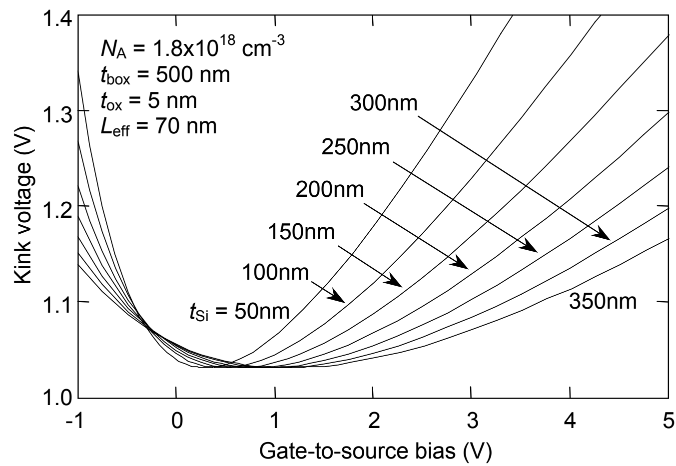


Fig. 6. Kink voltage vs. gate-to-source bias (dots) and corresponding electron mobility vs. gate-to-source bias (triangles). Comparison with experimental data from Shahidi et al.[13] Device parameters are: t_{ox} = 4 nm, t_{Si} = 126 nm, t_{box} = 360 nm, $N_A = 2x10^{18}cm^{-3}$, L_{eff} = 70 nm[13].

data through a single fitting parameter *b* in (6). By this fitting electron mobility in the channel is directly related to the corresponding kink voltage. Kink voltage is the lowest at the threshold, Fig. 6, reflecting the fact that electron mobility has maximum at the

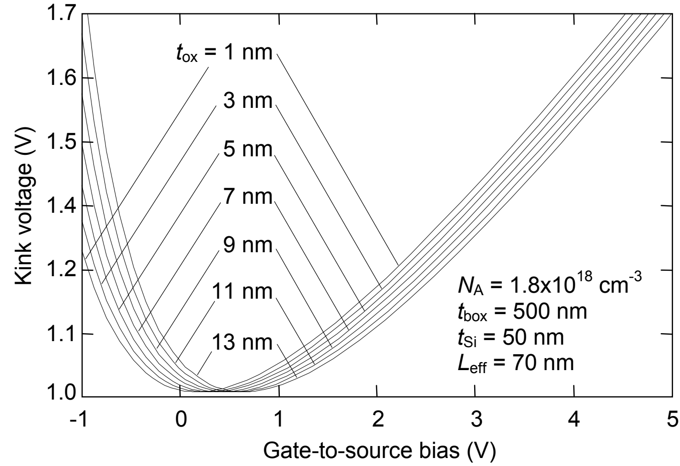


(a)

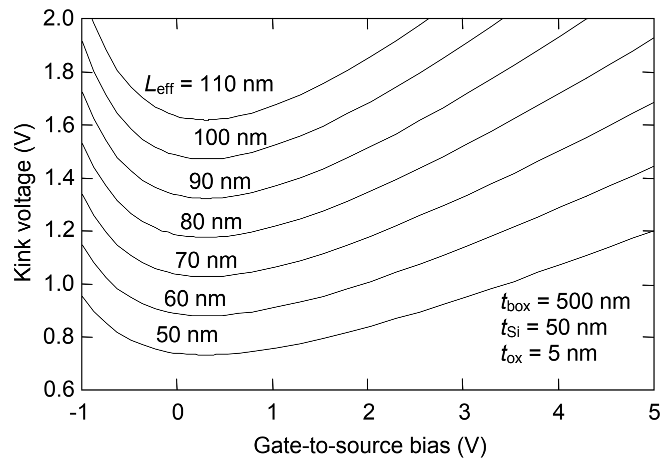


(b)

threshold, Fig. 5.[11] From this point forward behavior of the kink effect for various technology parameters for PD SOI NMOS devices is predicted, Fig. 7. Increase in doping concentration renders increase in the kink voltage in the whole range of the gate-to-source bias, due to decrease in electron mobility, Fig. 7(a). Device parameters are: $t_{ox} = 5$ nm, $t_{Si} = 50$ nm, $t_{box} = 500$ nm, $L_{eff} = 70$ nm. Variations in the active layer thickness will bring small changes in the kink voltage for the gate-to-source bias near threshold, Fig. 7(b), but it will bring remarkable decrease in the kink voltage with



(c)



(d)

Fig. 7. Kink voltage vs. gate-to-source bias depending on (a) the active layer doping concentration, (b) the thickness of the active layer, (c) the front oxide thickness, (d) the channel length.

increasing active layer thickness as we move eider below or above threshold. Channel doping density is here kept constant, $N_A = 1.8\text{x}10^{18}\text{cm}^{-3}$. Increasing front oxide thickness will increase kink voltage in the region below threshold, but it will decrease kink voltage in the region above threshold, Fig. 7(c). However, kink voltage change due to the

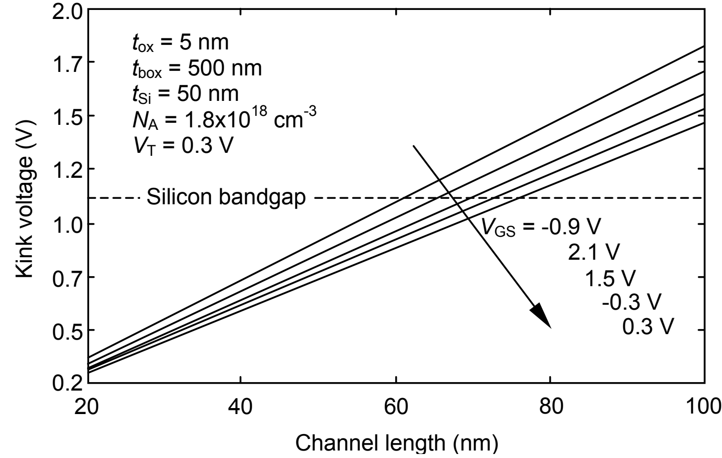


Fig. 8. Kink voltage vs. channel length depending on the gate-to-source bias. It is appealing to see appearance of the kink effect below silicon bandgap.

variations in the front oxide thickness is not so prominent. Channel doping, active layer thickness and buried oxide thickness are kept constant during this simulation. Decrease in the channel length will decrease kink voltage in the whole range of the gate-to-source bias thus making device more prone to the kink, Fig. 7(d). At Fig. 8 prediction for the kink voltage vs. channel length for the different gate-to-source bias is given. Device parameters are: t_{ox} = 5 nm, t_{Si} = 50 nm, t_{box} = 500 nm, N_A = 1.8x10^{18}cm^{-3}. The threshold voltage for this component is V_T = 0.3 V. As can be seen at the Fig. 8, the kink voltage exhibits linear behavior depending on the gate length. Lowest slope is for the gate-to-source bias equaling to the threshold voltage. Moving below or above threshold will increase the slope for the kink voltage dependence on channel length, Fig. 8. Most striking effect is appearance of the kink below of the bandgap of silicon. So far, proposed models in coping with this phenomenon include electron-phonon interaction, short-range and long-range electron-electron interaction and inversion layer quantization. Here it is shown that this approach gives quantitative results for the explanation of the effect.

In designing PD SOI NMOS devices, in order to avoid the kink effect, according to this work following rules should be obeyed: channel doping should be high thus increasing the kink voltage and allowing component more broad range of safe operation, Fig. 7(a). This is in contradiction to the design rules considering Short Channel Effects (SCE), which renders trade-off to the designer. Active layer should be thin as much as possible because that will significantly increase the kink voltage and render component more resistible to the kink, Fig. 7(b). Front oxide thickness is particular trade-off since decrease in its thickness will decrease the kink voltage when component is operating below threshold thus increasing leakage current in the off-state, but the same decrease in the front oxide thickness will increase the kink voltage when component is operating

above threshold thus making device more stable in the saturation regime, Fig. 7(c). Scaling gate length below 100 nm will give new remarkable effect *i.e.* kink voltage below silicon bandgap, Fig. 8. It is shown that simple theory developed here is able to explain this effect. Still, kink effect remains one of the most important problems in designing SOI based devices. In this article it is shown that device geometry and basic technology design could help in soothing this problem.

4. Conclusion

It is possible to directly relate electron mobility in the channel for PD SOI NMOS devices to the kink voltage *i.e.* triggering drain-to-source bias at the onset of the kink effect. This approach could be used in two manners: first, to give prediction for behavior of the kink voltage for various technology parameters; this can be utilized in modeling of PD SOI NMOS devices or circuitry simulations; second, it could be new method for measuring electron mobility in the front channel of these devices, *i.e.* knowing kink voltage one is able to derive corresponding electron mobility. This approach is capable of explaining appearance of the kink effect for the drain bias below silicon bandgap. It is important to notice that theory is applicable for the devices with channel lengths below 600 nm. Lower limit for the application of this theory is appearance of FD device.

Acknowledgments

This work has been partially supported by the Serbian Ministry of Science, Technologies and Development within the framework of the project IT.6151.B.

References

1. G. G. Shahidi, *IBM J. RES. & DEV.* VOL. 46, NO. 2/3 MARCH/MAY 2002.
2. Shahidi, et al. IBM DAMOCLES tutorial. Available at: www.research.ibm.com/DAMOCLES/html_file/segii.html
3. H. K. Yu, J. S. Lyu, S. W. Kang and C. K. Kim, *IEEE Trans. Electron Devices* 41, 726 (1994).
4. D. Suh and J. G. Fossum, *IEEE Trans. Electron Devices* 42, 728 (1995).
5. M. Youssef Hammad, Dieter K. Schroder, *IEEE Trans. Electron Devices* 48, 252 (2001).
6. S. S. CHEN and J. B. KUO, *Solid-State Electronics*, Vol. 41. No. 3, pp. 447-458, 1997.
7. S. Barraud, L. Clavelier, T. Ernst, *Solid State Electronics*, 49 (2005) 1090-1097.
8. Jacoboni C, Nava F, Canali C, Ottaviani, *Phys Rev B*, 1981, 24, 1014-26.
9. S. Takagi and A. Toriumi, *IEDMDig.*, 711-714(1992).
10. Siegfried Selberherr, *Analysis and Simulation of Semiconductor Devices*, pp. 88, Wien, Springer-Verlag, 1984.
11. J. R. Davis, A. E. Glaccum, K. Reeson, and P. L. F. Hemment, *IEEE Elec. Dev. Let.*,7(10), 570-572 (1986).
12. Milija Sarajlić, Rifat Ramović, *Proc. ETRAN,* Budva, Serbia and Montenegro, June , 2005.
13. G. Shahidi, J. D. Warnock, J. Comfort, S. Fischer, P. A. McFarland, A. Acovic, T. L. Chappell, B. A. Chappell, T. H. Ning, C. J. Anderson, R. H. Dennard, J. Y.-C. Sun, M. R. Polcari, B. Davari, *IBM J. RES. & DEV.* VOL. 39, NO. 1/2, JANUARY/MART 1995.