Unexpected mobility degradation for very short devices : A new challenge for CMOS scaling

Antoine Cros[1,3], Krunoslav Romanjek[2], Dominique Fleury[1], Samuel Harrison[2], Robin Cerutti[1], Philippe Coronel[1], Benjamin Dumont[1], Arnaud Pouydebasque[2], Romain Wacquez[1], Blandine Duriez[2], Romain Gwoziecki[4], Frederic Boeuf[1], Hugues Brut[1], Gérard Ghibaudo[3], Thomas Skotnicki[1]

antoine.cros@st.com, phone : +33 4 38 92 29 26, fax : +33 4 38 92 29 53

[1]STMicroelectronics, [2]NXP Semiconductors, 850 rue Jean Monnet, 38926 Crolles, France,

[3]IMEP, [4]CEA-LETI, Minatec, Parvis Louis Néel, 38016 Grenoble, France

Abstract

A new mobility degradation specific to short channel MOSFETs is studied and elucidated. Pocket implants/dopants pile-up, interface states/oxide charges, remote Coulomb scattering or ballisticity are insufficient to explain this degradation. The role of non-Coulombian (neutral) defects, which can be healed by increasing the annealing temperature, is evidenced.

Introduction

To extend the CMOS scaling to the future technological nodes, emphasis has been particularly put in the last years in transport improvement. Indeed, enhancing the carrier mobility allows a gain in Ion, with no loss neither in leakage current nor in load capacitance, thus improving greatly the C.V/I ratio.

Mechanical stress approach has already demonstrated Ion [1] and mobility gains both for high gate length and short gate length transistors [2]. But mobility can also be improved by effective field reduction; indeed, channel doping increase – especially for short gate transistors with pocket implants – which is necessary for short channel effects control, increases the effective field and decreases the mobility. So, thin film fully depleted single gate (SG) or double gate (DG) transistors, which allow low channel doping thanks to their intrinsic short channel effects (SCE) control, should also benefit from a mobility gain [3]. However, if SOI fully depleted (FD) and moreover DG type transistors have demonstrated high static performances, no actual gain was achieved thanks to channel doping suppression for short gates, both in literature (Fig. 1) and from our experiments (Fig. 2), and recent papers propose to re-introduce dopants to relax the dimensional constraints (T_{Si}) [4].

As a mobility lowering was previously observed on SG [2,5], SOI-FD [6] and DG [7,8] transistors, we aim at studying in depth this unexpected limitation and understand it origins.

Experimental protocol

State-of-the-art devices competitive with the best-published results were used for this study: bulk transistors from pre-industrial technology (12Å and 17 Å oxide thickness, with gate length down to 45nm) and planar GAAs from SON type technology (doped channel [9]: 20Å oxide and undoped channel [10]: 22Å oxide, Tsi=15nm).

We focus in low field mobility, although short channel transport is also ruled by drift velocity. It has been shown that these 2 parameters are well correlated [11], and thus mobility can be considered as a good indicator of the transport quality. Y-function technique [12] (Fig. 3), which allows to suppress the effect of series resistances and to perform statistical extraction on wafers, was used; the key point in short channel mobility computation being the extraction of the effective (electrical) length, performed by gate-to-channel split C(V) measurements (Figs 4 & 5) [13]. The low field mobility, i.e. μ_0 parameter is computed from the extracted gain (β) parameter ($\beta=\mu_0.C_{ox}.W/L_{Eff}$).

The consistency of the method was checked with short channel split C(V) (Fig. 6) and magnetoresistance measurements [14,15], independent of the L_{Eff} knowledge.

Ambient temperature results

Comparing the mobility for low L_{Eff} for heavily doped ($6.10^{18}cm^{-3}$) and undoped (channel made by epitaxy) GAA transistors with equivalent Tsi (Fig. 7), it can be noticed that undoped channel shows a high mobility – as expected – for long (1µm) transistors, but for short gate lengths doped and undoped channels are equivalent. This well explains the lack of performance increase for undoped DG devices. As this phenomenon was observed on other SG or DG type architectures in literature, the root cause of this issue was investigated.

The mobility in bulk transistors is degrading for short gate lengths, for NMOS and PMOS (Fig. 8), mainly under 100nm electrical length. It was first attributed to pocket implants [5] or channel dopants pile-up [7], but comparing transistors with or without pockets (Fig. 9) shows that this effect is not predominant.

A cause for this mobility degradation could be interface or oxide charges located near the extensions (Fig. 10). The experimental degradation can be reproduced quantitatively (Fig. 11) provided high density of interface charges ($5x10^{11}cm^{-2}$) are added near source and drain, but a reverse short channel effect is also induced, which is not observed experimentally, neither on SG nor on DG without pockets.

The effect of the remote Coulomb scattering from extensions (analogous to polygate-to-channel one) was also simulated (not shown), but it only explains a worst case μ_0 degradation of

10% at L_{Eff}=30nm, on undoped DG transistor.
More fundamentally, a μ decrease with L is also predicted because of nearly ballistic transport [16] – but in our case the measured mobilities are far from the ballistic limit (Fig. 12).

Low temperature experiments & interpretation

In order to separate scattering mechanisms, low temperature mobility extractions were performed and fitted to a simple empirical model [17] (Figs. 13 & 14). Those measurements reveal the predominance of a non-Coulombian (and thus independent of temperature) mechanism on the mobility variation with the gate length. This scattering mechanism, of negligible impact on large gate lengths (L_G>150nm, Fig. 15), has a very significant impact on the mobility for very short devices (Figs. 16 & 17) even at ambient temperature, whereas the dominant mechanism stays phonon scattering at this temperature. This can be linked to the presence of neutral defects in silicon or at the interface near source and drain (≈50nm), explaining the gate length dependence (Fig. 18). The damaged zones merge for L_G<100nm, and thus the effect becomes more dramatic .
These defects can be crystalline defects, possibly induced by extensions implants; though, this degrading mechanism is not actually dependent of extensions implant energy (Fig. 19). Investigating the possibilities to repair defects, the effect of the activation anneal temperature was experimented. 1080°C spike anneal instead of 1050°C improves mobility on NMOS (+25% for short L) without any effect on the V_{Th} (Fig. 20), evidencing neutral defects healing. Mobility is even more improved on PMOS (+60%) but with a V_{Th} shift, implying also Coulombian defects (interface states) healing (Fig. 21).

Conclusion

The mobility degradation with the length reduction, which particularly impacts the transistors with gate length under 100nm, was observed and deeply investigated on both bulk and double gate transistors. This degradation is not attributable only to ballisticity, remote Coulomb scattering, pockets/dopants pile-up or interface states/oxide charges, but mainly to non-Coulombian scattering associated to process-induced defects located near the source and drain junctions. Mobility can be improved by annealing temperature increase; this introduces a new trade-off between junction shallowness/abruptness and transport quality.

Aknowledgements

This work was supported by the Crolles2 Alliance (STMicroelectronics, NXP Semiconductors & Freescale Semiconductor), and partially by the NANOCMOS & PULLNANO projects.

References

[1] T. Ghani et al, "A 90nm high volume manufacturing logic technology featuring novel 45nm gate length strained silicon CMOS transistors", IEDM technical digest, p.11.6, 2003
[2] François Andrieu et al, "Experimental and comparative investigation of low and high field transport in substrate- and process-induced strained nanoscaled MOSFETs", Proc. Symposium on VLSI technology, p.176-177, 2005
[3] H.-S. P. Wong , D. J. Franck & P. M. Solomon., "Device design considerations for double-gate, ground-plane, and single-gated ultra-thin SOI MOSFET's at the 25 nm channel length generation", IEDM Technical digest, pp 407-410, 1998
[4] Y. Shiho, D. Burnett, M. Orlowski and J. Mogab , "Moderately Doped Channel Multiple-FinFET for Logic Applications", IEDM technical digest, p. 39.3 , 2005
[5] K. Romanjek, F. Andrieu, T. Ernst & G. Ghibaudo, "Characterization of the effective mobility by split C(V) technique in sub 0.1 μm Si and SiGe PMOSFETs" Solid State Electronics, vol. 49, p. 721-6, 2005
[6] C.Gallon et al, "New magnetoresistance method for mobility extraction in scaled fully-depleted SOI devices", SOI conf. proc., p. 153-155, 2004
[7] A.Cros et al, "Experimental mobility study of DG-GAA MOSFETs down to 40nm gate length", EuroSOI Conference proceedings, 2005
[8] J. Widiez, T. Poiroux, M. Vinet, B. Previtali, M. Mouis, S. Deleonibus, « Experimental low field transport investigation in sub-0.1μm ultra-thin SOI single and double gate MOSFETs », Proceedings Silicon Workshop on Nanoelectronics, 2005
[9] S. Harrison et al, "Highly Performant Double Gate MOSFET realized with SON process", IEDM Technical Digest, p.18.6, 2003
[10] R. Cerutti et al, "New Design Adapted Planar Double Gate Process for performant low standby power application Proceedings Silicon Workshop on Nanoelectronics, p.12-13, 2005
[11] A. Lochtefeld, D. A. Antoniadis, "Investigating the relationship between electron mobility and velocity in deeply scaled NMOS via mechanical stress", IEEE Electron Device Letters, v. 22, n. 12, p 591-3, Dec. 2001
[12] G. Ghibaudo, "New method for the extraction of MOSFET parameters", Electronics Letters, n. 24, p. 544, 1988
[13] K. Romanjek, F. Andrieu, T. Ernst, G. Ghibaudo, "Improved split C-V method for effective mobility extraction in sub-0.1-μm Si MOSFETs" EDL, p.583-5, 2004
[14] W.Chaisantikulwat et al, "Magnetoresistance Mobility Measurements in Gate-All-Around SON MOSFETs", Euro-SOI proceedings, 2005
[15] Y.M.Meziani et al, "Magnetoresistance mobility measurement in sub 0.1μm Si MOSFETs", ESSDERC proceedings 2004
[16] M. S. Shur, "Low Ballistic Mobility in Submicron HEMTs", IEEE electron devices letters, vol. 23, n. 9, 2002
[17] K. Romanjek et al, WOLTE proceedings, p. 201-208, 2004
[18] F. L.Yang et al, "25nm CMOS Omega Fets", IEDM technical digest, 2002
[19] B. S. Doyle et al, "High performance fully-depleted tri-gate CMOS transistors", IEEE Electron Device Letters, v. 24, n. 4, p 263-5, Apr. 2003
[20] J. Kedzierski et al, "Metal-gate FinFET and fully-depleted SOI devices using total gate silicidation", IEDM technical digest, p. 247-50, 2002
[21] C. Jahan et al., "ΩFETs transistors with TiN metal gate and HfO2 down to 10nm", VLSI symposium proceedings, p. 112-13, 2005
[22] Lee et al, "A novel sub-50 nm multi-bridge-channel MOSFET (MBCFET) with extremely high performance" VLSI symp. Proc., p. 200-1, 2004
[23] Lee et al, "Sub-25nm single-metal gate CMOS multi-bridge-channel MOSFET (MBCFET) for high performance an low power application" VLSI symposium proceedings, p. 154-5, 2005
[24] L. Mathew et al, "Inverted T channel FET (ITFET) – Fabrication and Characteristics of Vertical-Horizontal, Thin Body, Multi-Gate, Multi-Orientation Devices, ITFET SRAM Bit-cell operation : A Novel Technology for 45nm and Beyond CMOS", IEDM technical digest, p. 30.2, 2005
[25] S. D. Suk et al, "High Performance 5nm radius Twin Silicon Nanowire MOSFET (TSNWFET) : Fabrication on Bulk Si Wafer, Characteristics, and Reliability", IEDM technical digest, p. 30.3, 2005
[26] A. Kaneko et al, "Sidewall Transfer Process and Selective Gate Sidewall Spacer Formation Technology for Sub-15nm FinFET with Elevated Source/Drain Extension", IEDM technical digest, p. 34.6, 2005

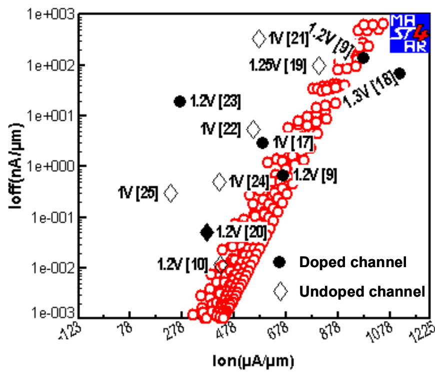


Figure 1 : MASTAR literature review for best published multi-gate NMOS and best trend cloud at Vdd=1.2V (doped channel). Vdd is indicated for each point. Currents normalized from interface conduction width. Circles : heavily doped channel; Diamonds : undoped channel transistors.

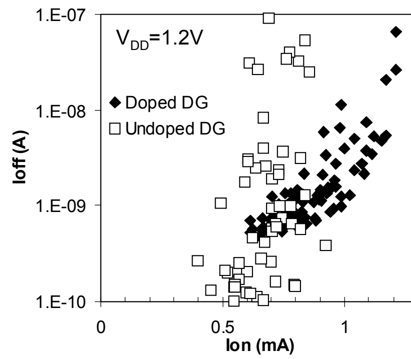


Figure 2 : Ioff(Ion) plots for GAA transistors with variable Lg down to 40nm, undoped and doped channel, T_{Si}=15nm. Normalized with mask width. Undoped channel doesn't show improved trade-off in spite of the reduced effective field.

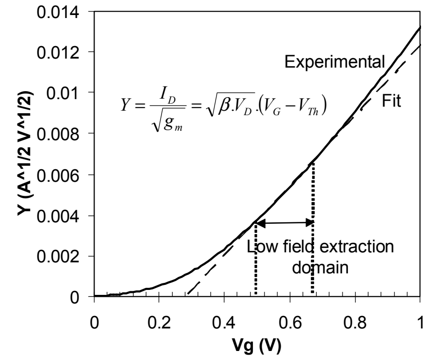


Figure 3 : Parameters extraction by Y function method. The Y function eliminates θ_1 and thus the effect of series resistances.

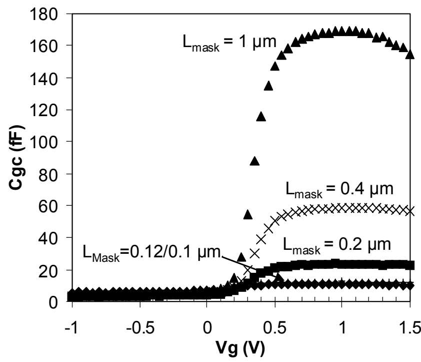


Figure 4 : Gate-to-channel capacitance measurements for 12Å oxide nmos bulk transistors, 5 mask lengths from 1µm to 100nm.

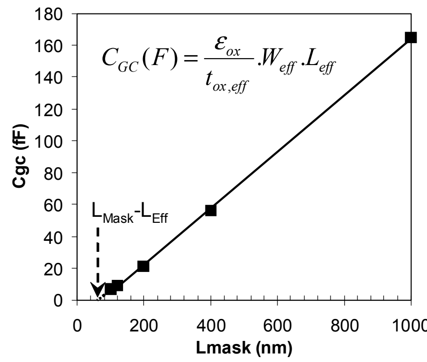


Figure 5 : Gate-to-channel maximum values function of mask length (minus parasitic capacitance). Linear regression interception with x-axis gives L_{Mask}-L_{Eff}.

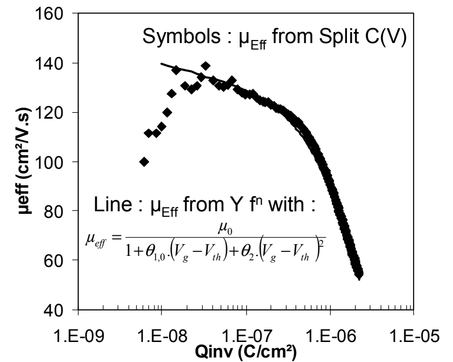


Figure 6 : Comparison of effective mobility extracted by split C(V) and from parameters extracted by Y function method.

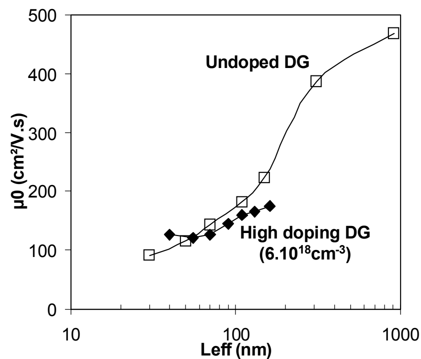


Figure 7 : Low field mobility compared for highly doped and undoped DG NMOSFETs

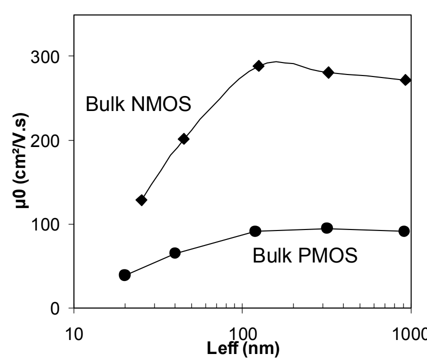


Figure 8 : Low field mobility for bulk MOSFETs (12A oxide)

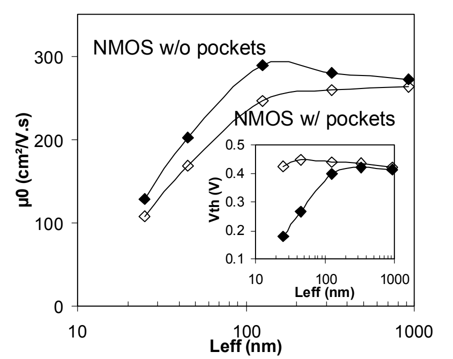


Figure 9 : Low field mobility for NMOSFETs (12A oxide) with or without pocket implants. Insert : V_{Th}

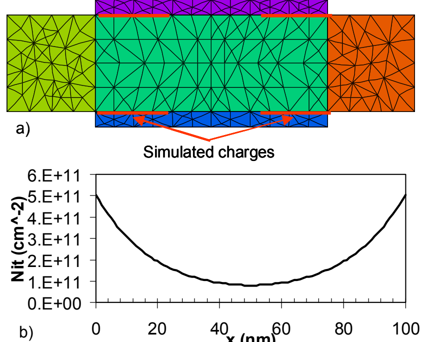


Figure 10 : Simulated undoped DG NMOSFET with interface states variable in the channel position (near source and drain junction) :

$$Nit = 5.10^{11}.\exp(-x/2.10^6) + 5.10^{11}.\exp(x - L/2.10^6)$$

a): simulated device, b) interface states distrib.

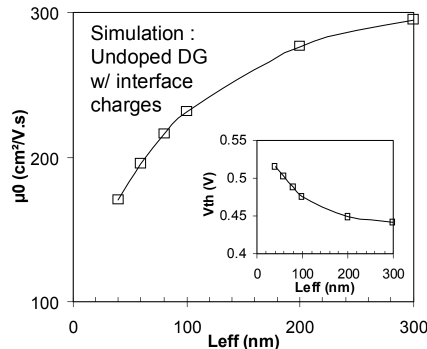


Figure 11 : FEM simulation results : low field mobility extracted from the $I_D(V_G)$ curves. Insert: V_{Th}.

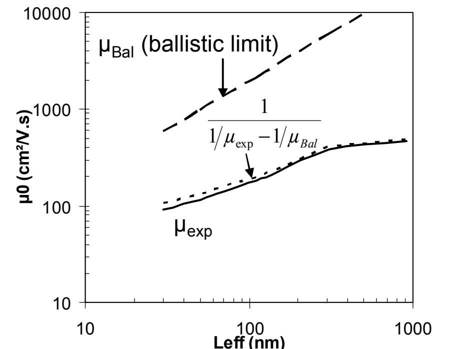


Figure 12 : Comparison of the maximum possible mobility with the ballistic limit from [17] and the experimental data from n-channel UDG.

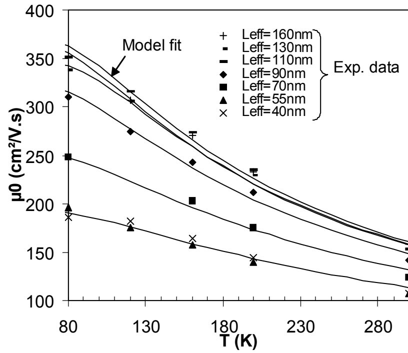
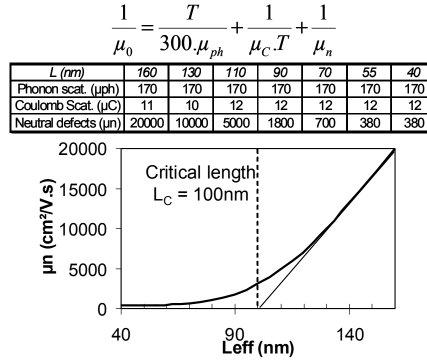


Figure 13 : Experimental data and model fit for low field mobility variation with T, doped DG.



$$\frac{1}{\mu_0} = \frac{T}{300.\mu_{ph}} + \frac{1}{\mu_C.T} + \frac{1}{\mu_n}$$

L (nm)	*160*	*130*	*110*	*90*	*70*	*55*	*40*
Phonon scat. (µph)	170	170	170	170	170	170	170
Coulomb Scat. (µC)	11	10	12	12	12	12	12
Neutral defects (µn)	20000	10000	5000	1800	700	380	380

Figure 14 : The mobility model allows to separate phonon (μ_{ph}), Coulombian scattering (μ_C) and non-Coulombian scattering (μ_n) thanks to their temperature dependence. Critical length Lc of 100nm is found for the non-Coulombian scattering parameter, suggesting extension of the defects of 50nm from junctions.

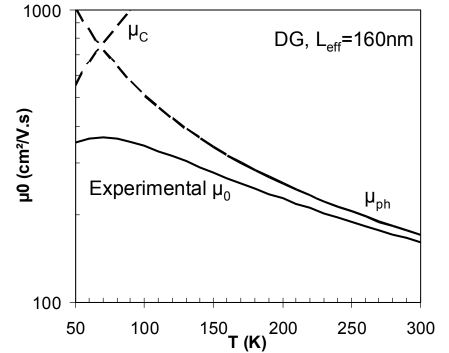


Figure 15 : Mobility limiting factors for doped DG transistor, L_{Eff}=160nm. Main limiting factor at ambient temperature is phonon scattering, while coulomb scattering plays a role only at low temperature.

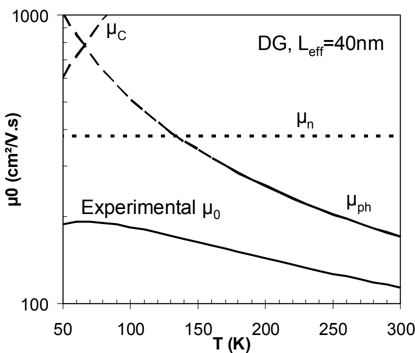


Figure 16 : Mobility limiting factors for doped DG transistor, L_{Eff}=40nm. Main limiting factor at ambient temperature stays phonon scattering, but neutral scattering plays a significant role.

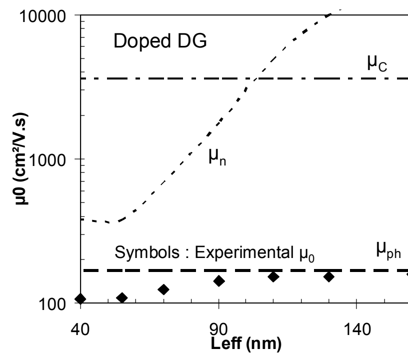


Figure 17 : Mobility limiting factors variation with the gate length. Phonon and coulomb scattering stay constant, but neutral scattering impact increases for short devices.

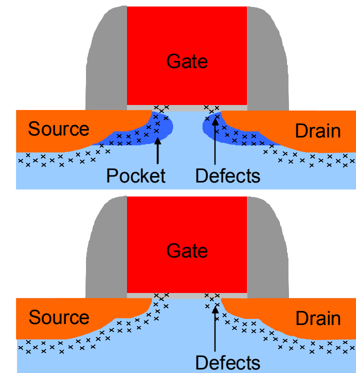


Figure 18 : Principle scheme of the localisation of the neutral defects. Pockets can be superimposed with the same localisation on the channel, but neutral defects are predominant electrically.

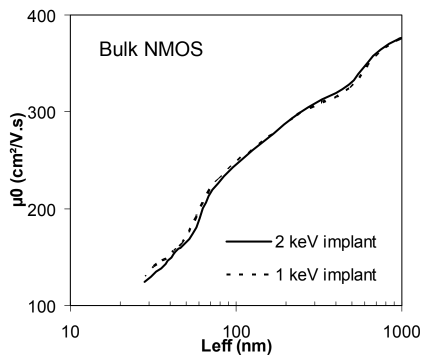


Figure 19 : Low field mobility for bulk NMOSFETs (17A oxide) with 2 different S/D extension implant energies: no significant difference is found.

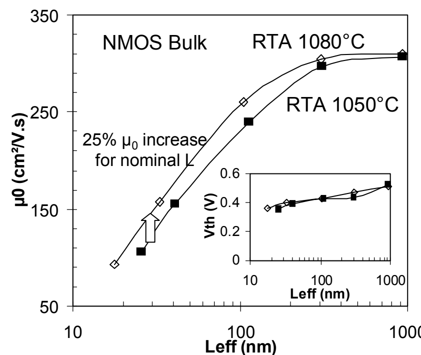


Figure 20 : Comparison of mobility variation with the effective length for 2 different spike anneal temperatures, bulk NMOS (insert : V_{Th}). Mobility recover without Vth variation (due to reverse short channel effect) indicates neutral defects healing.

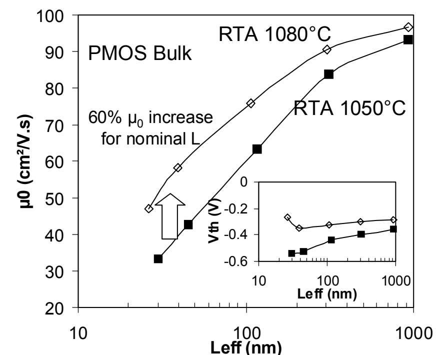


Figure 21 : Comparison of mobility variation with the effective length for 2 different spike anneal temperatures, bulk PMOS (insert : V_{Th}). Mobility recover with reverse short channel effect indicates interface states and neutral defects healing.