

Controlling STI-related parasitic conduction in 90 nm CMOS and below

E.Augendre, R.Rooyackers, D.Shamiryan, C.Ravit*, M.Jurczak and G.Badenes

IMEC vzw, Kapeldreef 75, 3001 Leuven, Belgium

*Philips Research Leuven, Kapeldreef 75, 3001 Leuven, Belgium

E-mail: augendre@imec.be

Abstract

In parallel to its continuous scaling, shallow trench isolation (STI) requires thorough optimisation with respect to its impact on device performance. In particular, the conduction occurring at the isolation edge of the device needs to be limited. Among the factors that influence the related control of threshold voltage and subthreshold current, this paper evaluates the impact of process parameters such as transistor architecture, trench profile (conventional or T-shape) and, for the first time, sacrificial oxidation strategy.

It is first proven that conventional STI architecture can provide the same gate oxide integrity and control of lateral conduction as T-shape STI. Second, it is shown that transistor scaling improves immunity to narrow channel effects. This is illustrated with an optimised conventional STI module showing hump-free operation and threshold voltage variation of 40 mV down to 0.11 μm wide nMOS transistors, making the approach suitable at least down to the 90 nm technology node.

1. Introduction

STI is the only lateral isolation scheme that meets the requirements of deep sub-micron technologies in terms of active area scaling and topography at gate level.

For each technology node, trench patterning and filling need adjustment towards the design rules. At the same time, significant optimisation efforts go to the tailoring of silicon and oxide profiles at the edge of active area in order to control both gate oxide integrity and device off-state current. In that respect, it is essential that the silicon corner is rounded and that the neighbouring trench oxide shows minimum recess, preventing the gate from wrapping around the silicon corner. These avoid field crowding that locally lowers the threshold voltage and results in a parasitic leakage path along the lateral transistor [1].

Silicon corner rounding is commonly achieved via high temperature oxidation of the trench sidewalls [2]. Preventing gate wrap-around is more problematic since significant oxide recess is likely to originate from all (isotropic) HF oxide etch steps that are carried out

between the post-planarisation nitride etch and gate stack deposition.

Several methods were proposed to limit gate wrap-around with STI, based on the covering of active area edge with a thick oxide (e.g. by forming a bird's beak [3] or a T-shaped trench [4]). This thick oxide delays the formation of the lateral oxide recess, but it also impacts effective channel width and can introduce non-uniformities due to the added pattern density dependence of T-shape formation. In this paper, we present for the first time the complementary approach that consists in drastically limiting the amount of critical HF oxide etch steps, with no negative draw-back on channel width or pattern uniformity.

Part 2 gives processing details, part 3 compares this approach to T-shape STI in terms of morphology, gate oxide integrity and transistor behaviour and part 4 demonstrates the interest of such an approach down to at least the 90 nm CMOS node.

2. Experimental

Processing started with the formation of active area stack (15 nm pad oxide and 150 nm nitride). After printing using 193 nm lithography, 400 nm, 325 nm and 250 nm-deep trenches were etched in the silicon.

A conventional straight STI profile is achieved by two-step etching: nitride etching is immediately followed by silicon trench etching with the same resist mask. In some of the 400 nm samples, an additional etch step was inserted after nitride etching to obtain T-shape STI. This step uses highly polymerising chemistry that intentionally creates polymers on the nitride sidewalls. These polymers are masking the subsequent trench etching, producing an offset between nitride and silicon sidewalls [4] (although in contrast to reference [4], polymerisation was here decoupled from silicon etching, resulting in an almost horizontal top silicon surface).

After stripping, all samples underwent the same high temperature sidewall oxidation that rounded silicon corners and trenches were filled with high-density plasma (HDP) oxide. Samples that had received polymerising showed a T-shape oxide profile, with HDP oxide covering the edges of active area. Next, all wafers received dummy-free planarisation [5] and nitride was etched away.

The reference process continued with the complete removal of the pad oxide and the growth of a fresh 20 nm sacrificial screening oxide, through which well implantation was carried out. For some wafers, the pad oxide was not etched but instead limited oxidation was performed to thicken it up into a 20 nm screening oxide and obtain the same doping profile. This approach significantly reduces the cumulated HF etch time between post-planarisation nitride etch and gate formation, which significantly decreases the resulting oxide recess at the edges of active area. This approach is presented in Figure 1 and will be referred to as “short HF”.

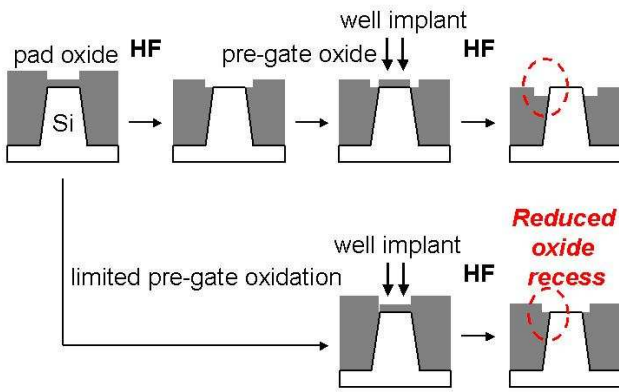


Figure 1. Principle of the “short HF” route. In contrast to the reference process (top), limiting HF dips reduces oxide recess (bottom).

The rest of the processing consists in 0.18, 0.13 and 0.09 μm nMOS flows with respective gate oxide thickness (t_{ox}) of 3.5, 2.0 and 1.5 nm, related thermal budgets and identical as-implanted channel profile-. Results will be shown on long channel ($L=10 \mu\text{m}$) nMOS transistors, since they are most sensitive to narrow-channel effects [6].

3. Influence of STI architecture

3.1. Isolation morphology

Table 1 presents metrology results on the offset in active area width relative to mask dimension (measured by scanning electron microscopy), the depth of the lateral oxide recess (evaluated from silicide sheet resistance) and the oxide thickness that was remaining in trenches when the processing reached gate formation (extracted from polysilicon-capped field capacitors).

It is first visible that for the same mask dimension, active area lines defined by T-shape STI are 34 nm wider than with a straight STI profile. Second, the oxide recess decreases significantly with short HF time, and it is slightly further reduced by T-shape profile. Finally looking at the STI oxide thickness at gate level, it is first visible that in contrast to 325 and 250 nm trenches,

400 nm trenches were filled slightly below target. Notwithstanding that, the benefit of the short HF is obvious: it consistently reduces oxide loss by about 15 nm.

Table 1. Offset in active area width relative to mask dimension (ΔW), oxide recess depth and trench oxide thickness at gate level

STI profile / depth / HF time	ΔW (σ)	Recess (σ)
Straight / 400 nm / long HF	-19 (5) nm	26 (0.2) nm
Straight / 400 nm / short HF	-19 (5) nm	13 (0.1) nm
Straight / 325 nm / short HF	-19 (5) nm	12 (0.1) nm
Straight / 250 nm / short HF	-19 (5) nm	12 (0.1) nm
T-shape / 400 nm / short HF	+15 (9) nm	9 (0.6) nm
STI profile / depth / HF time	STI oxide thickness (σ)	
Straight / 400 nm / long HF	364 (7) nm	
Straight / 400 nm / short HF	380 (8) nm	
Straight / 325 nm / short HF	322 (8) nm	
Straight / 250 nm / short HF	258 (8) nm	
T-shape / 400 nm / short HF	377 (9) nm	

3.2 Gate oxide integrity

Pre-gate sacrificial oxidation strategy can influence gate oxide integrity [7]. For this reason, this parameter was evaluated by measuring time-dependant dielectric breakdown at constant voltage on gate capacitors that overlap on STI, using a thin-oxide suited algorithm [8].

Figure 2 shows that the measured charge-to-breakdown is independent from capacitor perimeter and does not show any extrinsic failure, proving that the short HF approach has no impact on gate oxide integrity.

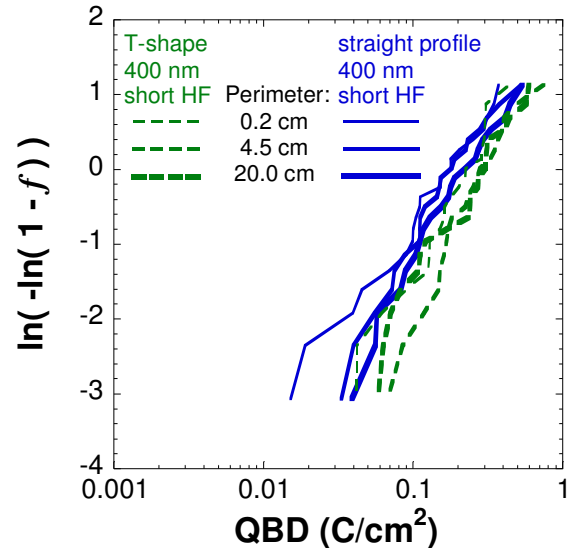


Figure 2. Charge-to-breakdown measured at $V_g=-4.8 \text{ V}$ on $2 \cdot 10^{-3} \text{ cm}^2$ capacitors with $t_{\text{ox}}=3.5 \text{ nm}$

3.3. Transistor performance

Figure 3 presents the biased-body transfer characteristics of $W/L=10/10\ \mu\text{m}$ nMOS transistors with $\text{tox}=3.5\ \text{nm}$. A hump is visible in the curve with long HF time. This hump is not present when short HF time was used, indicating that it can be attributed to gate wrap-around and that reducing HF time is sufficient to achieve hump-free performance.

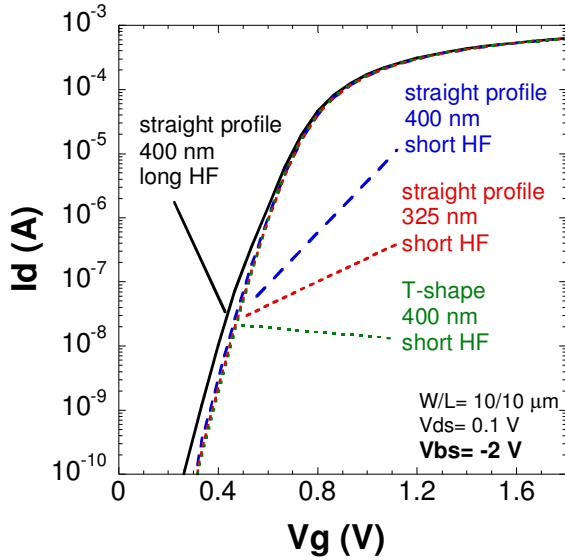


Figure 3. Transfer characteristics of nMOSFETs with $\text{tox}=3.5\ \text{nm}$

The dependence of linear threshold voltage (V_{th}) on actual transistor width is plotted in Figure 4 for nMOS transistors with $L=10\ \mu\text{m}$ and $\text{tox}=3.5\ \text{nm}$.

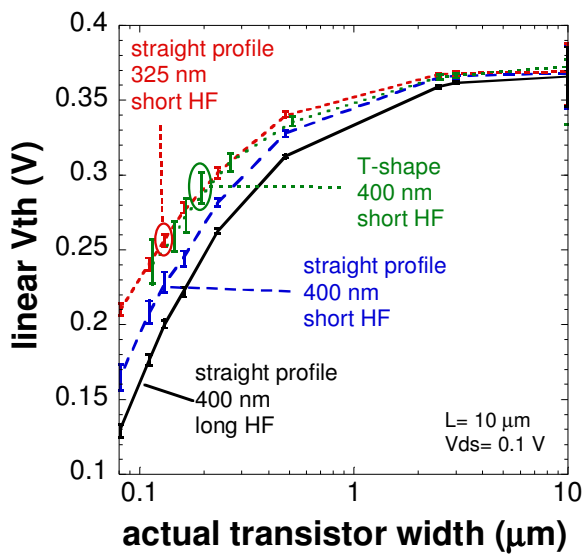


Figure 4. Dependence of V_{th} on actual transistor width for nMOS transistors with $\text{tox}=3.5\ \text{nm}$

Here again, the gate wrap-around reduction coming from short HF time is beneficial, improving the control of threshold voltage in narrow devices. This behaviour is yet improved equally by the conventional 325 nm-deep trenches (thanks to sufficient trench filling) and by T-shape STI, which is intrinsically more resistant to incomplete filling. However, higher V_{th} spread is observed for T-shape STI, probably caused by polymerisation/etch non-uniformity.

In summary straight-profile STI with well-controlled filling and short HF provide uniform hump and narrow channel effect control for a $0.18\ \mu\text{m}$ technology. Next paragraph examines its application to the following technology nodes.

4. Influence of device architecture

The biased-body transfer characteristics of $W/L=1/10\ \mu\text{m}$ nMOS transistors processed in with 0.18, 0.13 and $0.09\ \mu\text{m}$ technologies are presented in Figure 5. All characteristics are hump-free, indicating that the suppression of gate wrap-around with minimised HF time is also suitable for $0.13\ \mu\text{m}$ and $0.09\ \mu\text{m}$ technologies.

As a matter of fact, these technologies could be less susceptible to biased-body hump appearance since along with their lower long-channel V_{th} , they have a lower body factor. Demonstrating hump-free operation with $\text{tox}=3.5\ \text{nm}$ is a further guaranty of subthreshold conduction control.

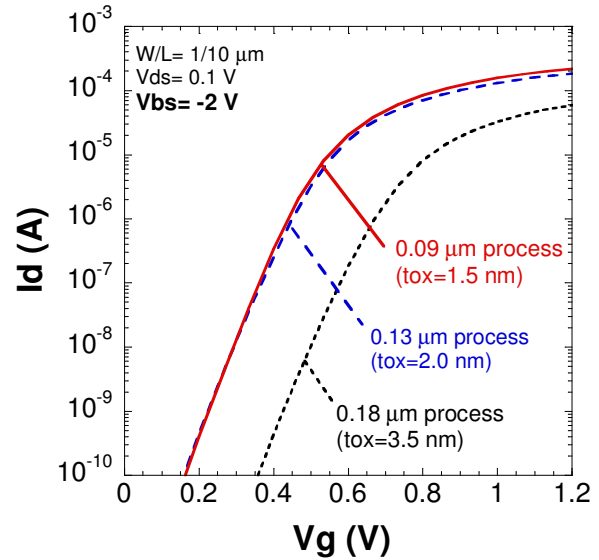


Figure 5. Transfer characteristics of nMOSFETs processed with 0.18, 0.13 and $0.09\ \mu\text{m}$ technologies

Figure 6 presents the dependence of V_{th} on transistor width for long nMOSFETs processed in different technologies, using the reduced-HF STI module. It is striking that with the same isolation architecture, the control of narrow channel devices is improved

significantly in scaled technologies. Using the 0.09 μm process, the threshold voltage difference between 10 μm - and 0.11 μm -wide transistors is reduced to 40 mV. This might be explained in terms of reduced boron redistribution [9] originating from curtailed thermal budget or from gate oxide nitridation.

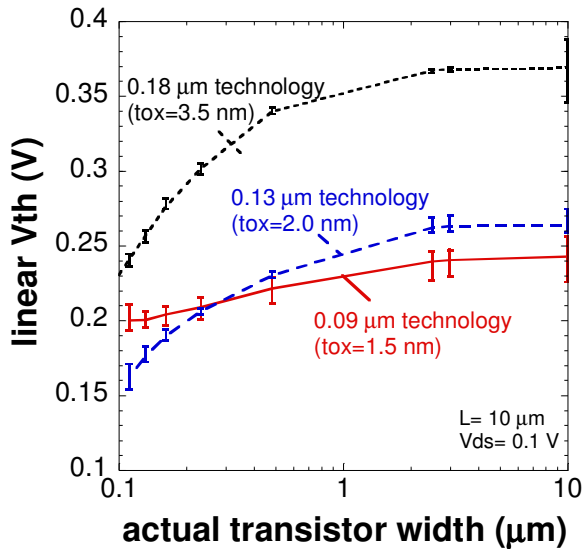


Figure 6. Dependence of V_{th} on actual transistor width for nMOS transistors of three technologies

5. Conclusions

The minimisation of pre-gate oxide etching steps by thickening the pad oxide and using it as sacrificial implant oxide has been presented. This approach allows minimising the lateral oxide recess that occurs at the edge of active area and causes gate wrap-around. It is also compatible with good gate oxide integrity as detected failures were intrinsic and perimeter-independent. Using the pad oxide as sacrificial implant oxide was compared to T-shape STI, which is the complementary route to minimise gate wrap-around. With a well-controlled trench filling, minimising HF time makes STI patterning simpler and results in a better process uniformity. In these conditions, it also achieves as good control of subthreshold conduction and of threshold voltage in narrow devices as T-shape STI. Finally, threshold voltage variation as small as 40 mV down to 0.11 μm wide nMOS transistors was demonstrated, making conventional STI with minimised oxide etch compatible at least down to the 90 nm CMOS technology node.

Acknowledgements

The authors wish to thank G. Van den Bosch for help in reliability characterisation, the SPT division of IMEC, H. Jin and W. Wang from Hua Hong for module development and the IMEC Pilot Line for processing.

This work has been partially funded by the European Union under the IST-1999-11599 HUNT project.

References

- [1] N. Shigyo *et al.*, "Three-dimensional analysis of subthreshold swing and transconductance for fully recessed oxide (trench) isolated $\frac{1}{4}$ - μm -width MOSFETs", *IEEE Trans. Electron Devices*, vol. 35, 1988, pp. 945-951
- [2] C.-P. Chang *et al.*, "Enabling STI for 0.1 μm technologies and beyond", *1999 Symp. on VLSI Tech. Digest of Technical Papers*, 1999, pp. 161-162
- [3] K. Horita *et al.*, "Advanced shallow trench isolation to suppress the inverse narrow channel effects for 0.24 μm pitch isolation and beyond", *2000 Symp. on VLSI Tech. Digest of Technical Papers*, 2000, pp. 178-179
- [4] B. H. Roh *et al.*, "Easily manufacturable shallow trench isolation for Gigabit dynamic random access memory", *Jpn. J. Appl. Phys. (A)*, vol. 35, 1996, pp. 4618-4623
- [5] G. Badenes *et al.*, "A new dummy-free shallow trench isolation concept for mixed signal applications", *J. Electrochem. Soc.*, vol. 147, 2000, pp. 3827-3832
- [6] F. Nouri *et al.*, "Narrow device issues in deep-submicron technologies – the influence of stress, TED and segregation on device performance", *Proceedings of ESSDERC'2000*, 2000, pp. 112-115
- [7] S. Deleonibus *et al.*, "Optimization of ultra thin gate oxide in a SILO/RTN isolation process for advanced ULSI", *J. Electrochem. Soc.*, vol. 41, 1994, pp. 2811-2820
- [8] P. J. Roussel *et al.*, "Accurate and robust noise-based trigger algorithm for soft breakdown detection in ultra thin dielectrics", *IEEE Trans. on Dev. and Mat. Reliab.*, vol. 1, 2001, pp. 120-127
- [9] A. Ono *et al.*, "TED control technology for suppression of INCE in 0.1 μm MOS devices", *IEDM Technical Digest of Technical Papers*, 1997, pp. 227-230