Extended Abstracts of the 1999 International Conference on Solid State Devices and Materials, Tokyo, 1999, pp. 136-137

Source/Drain Extension-to-Gate Overlap Scaling in Deep Sub-Micron MOSFETs

A. Inani, H. Deshpande, †S. Krishnan, †M.-R. Lin, and Jason C. S. Woo

Department of Electrical Engineering, University of California, Los Angeles, CA 90095, USA
Phone: 1-310-206-3279 Fax: 1-310-206-8495 E-mail: woo@ee.ucla.edu
†Technology Development Group, Advanced Micro Devices, 1 AMD Place, Sunnyvale, CA 94088, USA

1. Introduction

Continuous miniaturization of MOS transistors has enhanced their performance substantially. However, as the channel length of MOS transistors is reduced, short channel effects become worse and tend to degrade the electrical characteristics severely. It has long been recognized that ultra-shallow source/drain junctions may hold the key to keeping these adverse effects under control. In particular, the Source/Drain Extension (SDE) regions have gained a lot of attention. However, some of the devices fabricated in the past with such ultra-shallow junctions have shown lower than expected drive currents [1]. Hence, a lot of effort has been concentrated towards designing and optimizing such shallow SDE in the sub-100nm regime. This paper aims at further improving the understanding of SDE scaling issues.

2. Minimum Source/Drain Extension Overlap

It has been shown, with experimental data, that a minimum SDE-to-gate overlap (L_{ov}) is required to obtain a good drive current [2]. This has been deduced from the fact that the I_{dsat} seems to flatten out after a certain value of L_{ov}. It is believed that this occurs owing to some SDE-to-gate coupling. However, the basic principle and mechanism behind such a requirement is not clear. In this paper, we explore this phenomenon further and explain it in a simple, understandable way and also show that it is technology and processing dependent to a large extent.

Extensive device simulations were performed coupled with process simulations using SILVACO tools to study this effect [3]. The test structure used is similar to the one used in [2]. An offset spacer is used to obtain varying SDE-to-gate overlap. The metallurgical channel length (L_M) is fixed precisely for a particular set of simulations by varying the drawn channel length (L_D). The remaining physical parameters such as the oxide thickness and junction depths are fixed. The effect of the spreading resistance due to silicide is not taken into account. With these as the boundary conditions, device simulations were performed by varying the composition of the SDE doping as well as the channel doping profile in order to understand this effect.

3. Results and Discussion

In the first set of simulations, the L_M is precisely controlled at 70nm and the L_{ov} varied by varying the L_D. Fig. 1 shows the variation of I_{dsat} versus L_{ov} for two different channel doping profiles, namely, uniform doping and an optimized super steep retrograde doping [4]. It is clearly seen that the devices with the same L_M do not exhibit the same minimum requirement on L_{ov}. The slight difference in the magnitudes of currents also occurs because of the change in the channel doping. Fig. 2 shows the variation of I_{dsat} versus L_{ov} for the same uniform channel doping profile while changing the dose of the SDE. Again, it is clearly seen that the L_{ov} requirements vary with processing variation. In this case the difference in the magnitudes is also caused by the reduced R_{SD} in the heavily doped SDE. Finally, in Fig. 3, the effect of dopant species with different diffusion constants, e.g., Antimony and Phosphorus, is shown. It can be seen that the minimum overlap requirement varies with the dopant species. Note that in order to reduce the overlap requirement, a slow diffusing dopant is required. It also results in a higher current drive due to lower series resistance

Based on the above results, we propose that the requirement for a minimum SDE-to-gate overlap comes from the differential lateral diffusion of the dopants in the SDE depending on the processing details. This can be understood from the schematic in Fig. 4. During the lateral diffusion of the dopants in the SDE region, the tail end of the junction (near the channel; shaded here) does not remain as highly doped as the rest of it (called LDD here). It is the resistance of this region which is dominant in this phenomenon. The part of this LDD region which is under the gate is accumulated and its contribution to the overall R_{SD} is less. As shown in Fig. 4(a), under normal circumstances, only part of this region may be accumulated and hence the R_{SD} is not the minimum possible. The worst case (Fig. 4(b)) is when none of this region is under the gate. The optimum occurs when all of this region is under accumulation and hence the R_{SD} is minimum (Fig. 4(c)). This is the minimum required L_{ov}. Again, increasing the overlap further (Fig. 4(d)) does not really help since there is practically no further improvement in the R_{SD} and on the other hand, we end up increasing the overlap capacitance as well as the overall series resistance of the SDE. Fig.5 shows the dopant concentration in the SDE region for the two cases discussed in Fig. 2 above. It is seen here that the length of the SDE is different in the two cases and this leads to a difference in the low-doped tail in the two cases.

In order to further strengthen the above, we simulate a structure (inset of Fig. 6) with box SDE regions with a uniform doping of 1e20 cm^{-3}. Since there is no such low-doped tail region in the SDE of this structure, the minimum L_{ov} requirement does not exist in this case. This is clearly observed in Fig. 6. The L_M here is 40nm and the results are compared to a more realistic structure obtained using process simulations. The SDE in the box structure shows a much reduced R_{SD} and hence a correspondingly higher current.

4. Conclusions

It has been shown that a minimum SDE-to-gate overlap

is required in sub-100nm MOS devices in order to avoid current degradation. The overlap length, L_{ov}, is process dependent. The key factor dominant behind such a phenomenon is the resistance of the low-doped tail of the SDE region occurring near the channel.

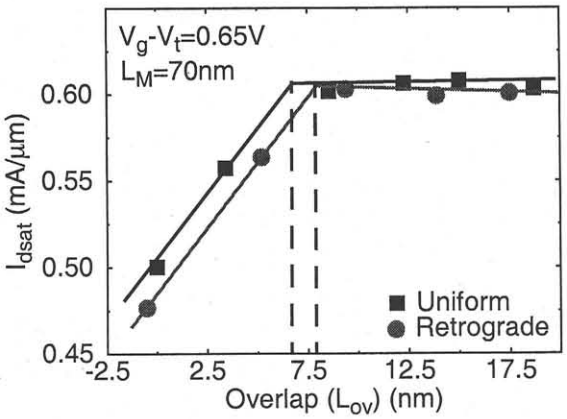

Fig. 1: I_{dsat} versus SDE-to-gate overlap (L_{ov}) for uniform channel and retrograde channel profiles. The SDE dose is kept constant. L_M=70nm.

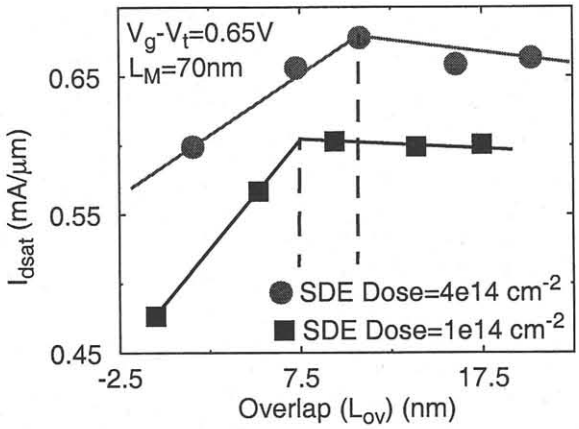

Fig. 2: I_{dsat} versus SDE-to-gate overlap (L_{ov}) for two different SDE dose. The same retrograde channel profile has been used in both cases. L_M=70nm.

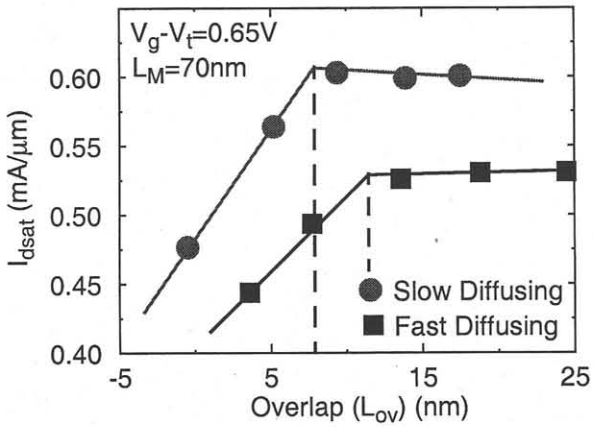

Fig. 3: I_{dsat} versus SDE-to-gate overlap (L_{ov}) for two different SDE implant species with different diffusion constants (e.g., Phosphorus and Antimony). The same retrograde channel profile has been used in both cases. and L_M=70nm.

References

[1] M. Takase et al., IEDM Tech. Dig., 1997 p475-478.

[2] S. Thompson et al., VLSI Symp., 1998 p132-133.

[3] SILVACO International, Santa Clara, CA 94054, USA.

[4] A. Inani et al., to be published in JJAP.

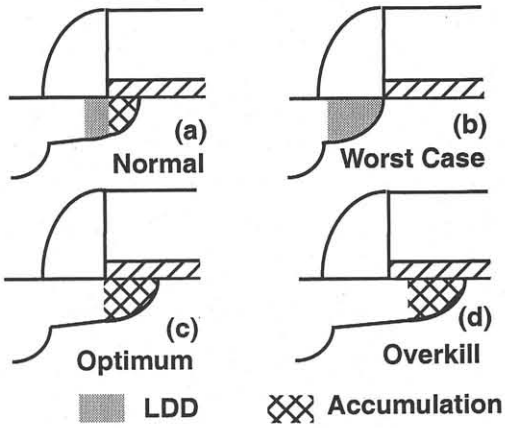

Fig. 4: The tail end of the SDE is lightly doped (LDD): (a) In the normal scenario, the LDD might be partially accumulated; (b) in the worst case, there is no overlap and hence no accumulation; (c) the optimum is when all of the LDD is accumulated, and (d) more overlap does not help anymore.

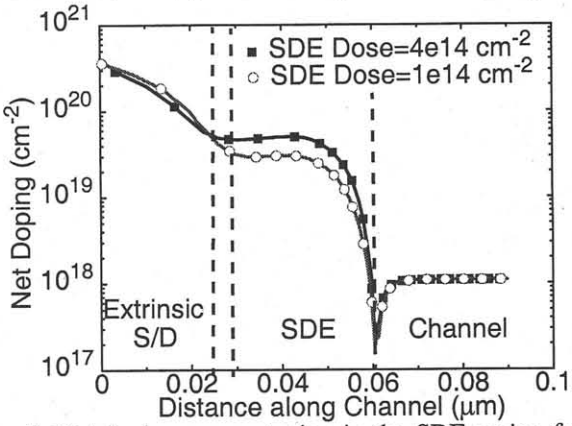

Fig. 5: Net doping concentration in the SDE region for two different SDE doses. The total length of the SDE differs in the two cases as shown by the dotted lines. The (low doped) tail end of the extension also varies in a similar fashion.

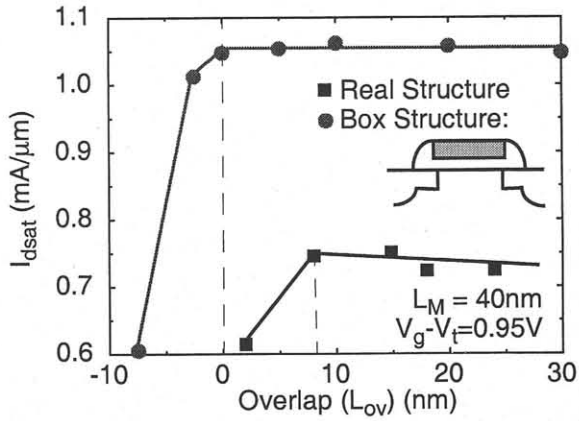

Fig. 6: I_{dsat} versus L_{ov} for a box structure, shown in inset, and for a real structure. The same retrograde channel profile has been used in both cases and L_M=40nm.