Received 9 August 2020; revised 6 September 2020; accepted 26 September 2020. Date of publication 22 October 2020; date of current version 28 January 2021. The review of this article was arranged by Editor T.-L. Ren.

Digital Object Identifier 10.1109/JEDS.2020.3032957

Understanding and Mitigating Stress Memorization Technique of Induced Layout Dependencies for NMOS HKMG Device

YING-FEI WANG [1], QING-CHUN ZHANG [2], PING LI [2], XIAO-JING SU [1], LI-SONG DONG [3], RUI CHEN [3], LI-BIN ZHANG [1], TIAN-YANG GAI [1], YA-JUAN SU [3], YA-YI WEI [1], AND TIAN CHUN YE [1]

1 University of Chinese Academy of Sciences, Beijing 100049, China
2 Semiconductor Manufacturing International Corporation, Shanghai 201203, China
3 Department of Integrated Circuit Advanced Process Center, Institute of Microelectronics, Chinese Academy of Sciences, Beijing 100029, China

CORRESPONDING AUTHOR: Y.-J. SU and Y.-Y. WEI (e-mail: suyajuan@ime.ac.cn; weiyayi@ime.ac.cn)

This work was supported in part by the National Nature Science Foundation of China under Contract 61804174.

ABSTRACT For the first time, this research addresses the notable layout proximity effects induced by stress memorization technique in planer high-k/Metal gate NMOS device systematically, including width effect, different shallow trench spacing effect, and length of diffusion effect. Based on the oxygen diffusion mechanism analysis of layout proximity effects in high-k/Metal gate NMOS device, an optimized process is proposed to suppress the layout dependency. The experiment result indicates that modified low temperature stress memorization technique process can suppress layout dependency efficiently without performance degradation of the devices.

INDEX TERMS Layout proximity effects, high-k HfO_2, stress memorization technique, Al diffusion.

I. INTRODUCTION

The effects of process-induced stress enhancements on transistor performance are becoming increasingly important to advanced CMOS technology. For the technology nodes above the 40 nm, stress memorization technique (SMT) by dislocation is widely used for performance boost in high-k metal-gate nMOSFET (including planer device and bulk FinFET) [1]–[3]. SMT process starts after spacer formation and source/drain (S/D) implantation, including stressor cap layer deposition and high temperature annealing. During the annealing step, stacking fault defects form in S/D regions due to the Si recrystallization growth, which causes the tensile strain in channel. The longitudinal tensile stress in channel can increase the electron mobility. However, several previously negligible physical effects are becoming increasingly important as a result of the high temperature annealing in the SMT process. At advanced integration nodes, layout proximity effects (LPEs) have become a non-negligible problem having impact on the performance of device [4]–[6]. Most of researches on SMT has focused on improving the mobility of nMOSFETs, while very few studies have been identified that have reported on LPEs stemming from stress process.

In this article, the notable LPEs induced by the SMT process in planer HKMG (high-k first/metal gate last) nMOSFETs are systematically studied, including the width effect (W) and the different shallow trench spacing (DSTS) effect.The DSTS includes DSTS-X (along the channel length direction) and DSTS-Y (along the channel width direction) [7]. The experiments used common LPE verification structures for the corresponding layout parameters. The layout dimension definitions are shown in Fig. 1. Electrical parameters of the nMOSFETs are extracted by wafer acceptance test (WAT). Moreover, a modified SMT process flow for suppressing LPE is proposed.

II. EXPERIMENTAL

To verify the efficiency of the proposed technique, the HKMG nMOSFETs with three different processes are fabricated in this article. Based on the regular HKMG process, three processes are developed as the regular SMT process, modified SMT process and without SMT process,

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

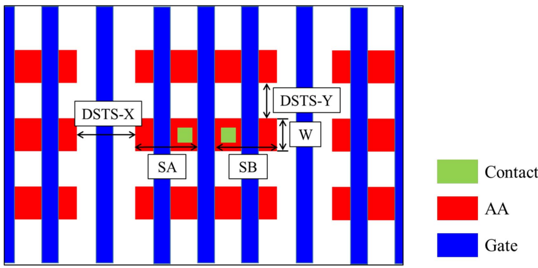

FIGURE 1. Sample layout of NMOS transistor with LOD (SA/SB), channel width (W) and different shallow trench spacing (DSTS-X/Y) dimension definition. Magnetization as a function of applied field. The active device is surrounded by dummy devices.

respectively. The simplified process flows are shown respectively in Fig. 2 (a), (b), (c). The devices went through the shallow trench isolation (STI)/well formation, gate dielectric (high-k), gate poly, spacer and S/D pre- amorphization implantation. For the regular SMT process, the high-tensile-stress film is then deposited. Standard spike anneal was then performed at temperatures ranging from 1000°C to 1100°C for stress memorization and S/D activation. The SMT film was subsequently removed. Fig. 2 (d) shows the cross-section schematic description of the nMOSFET structure with regular SMT process. Replacement metal (aluminum type) gate and backend processes were then performed to complete the device fabrication. For the modified SMT process, low temperature annealing at temperature 600°C to 700°C for 30 seconds was conducted after the stress film deposition. During low temperature annealing, the amorphized S/D regions occurs solid-phase epitaxy regrowth (SPER), which created edge dislocations. In addition, after the stress film removed, a supplementary spike annealing (> 1000°C) process was performed to active S/D regions. For comparison, the devices without SMT process were fabricated at the same time.

III. REGULAR SMT PROCESS LPES STUDY

For HKMG devices, the LPEs induced by the regular SMT process mainly include width effect and DSTS-Y effect. This section studies the effect of layout parameters on the saturated threshold voltage ($V_{T,sat}$) of nMOSFETs.

A. WIDTH EFFECT

For different channel widths, the $V_{T,sat}$ of nMOSFETs may change, which is called the width effect. Fig. 3 shows the dependency of $V_{T,sat}$ on channel width for nMOSFETs, with the other layout parameters fixed. As the reduction of channel width, the $V_{T,sat}$ of nMOSFETs drops from 220mV to 100mV, with reduction of 54% percentage.

B. DSTS EFFECT

The isolation between AA is achieved by STI technique. For different AA spacings along channel length (X) or channel width (Y) directions, electrical parameters may change, which is called DSTS-X effect or DSTS-Y effect. Noting that the LOD and other layout parameters are fixed with the DSTS-X or DSTS-Y varied.

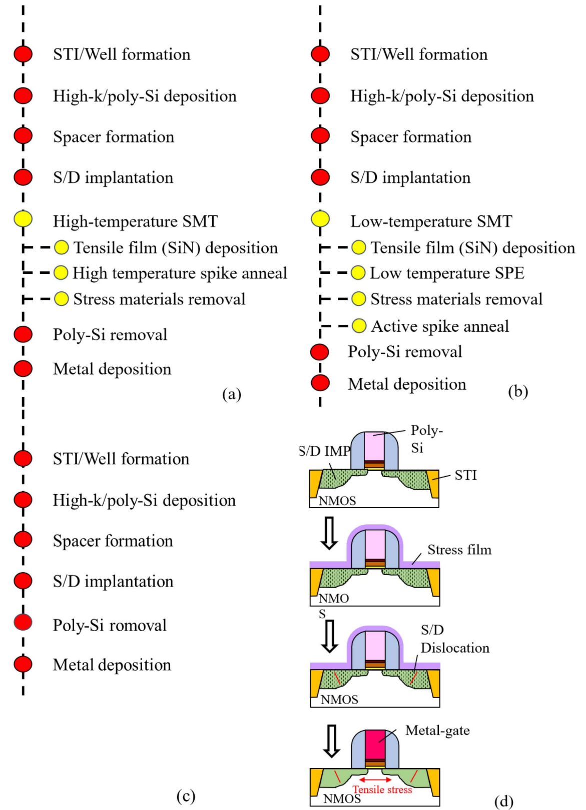

FIGURE 2. (a) The HKMG process with standard spike annealing regular SMT. (b) The HKMG process with low temperature annealing modified SMT. (c) The HKMG process without SMT. (d) Schematic description of the nMOSFET structure with process flow sequences for the SMT process.

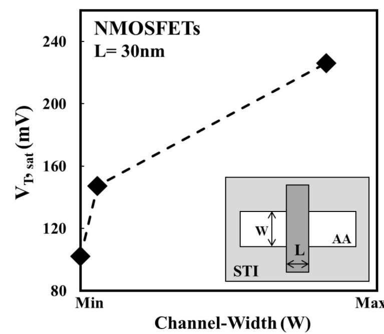

FIGURE 3. Dependence of $V_{T,sat}$ on channel width (W) for the SMT device at gate length(L) = 30nm. The inset of the Figure shows the layout geometry.

Fig. 4 shows the effect of DSTS-Y and DSYS-X on regular SMT induced $V_{T,sat}$ shift, referenced to DSTS-Y = 0.505um or DSTS-X = 0.08um. The results show that the $V_{T,sat}$ of nMOSFET is a highly relevant of STI width (on Y-direction). As the reduction of DSTS-Y, the $V_{T,sat}$ of nMOSFET increased. On the other hand, it is also observed that the $V_{T,sat}$ is independent of DSTS-X.

C. ANALYSIS AND HYPOTHESIS

According to Sections I and II, the $V_{T,sat}$ of nMOSFETs is regulated by channel width and DSTS-Y layout parameters.

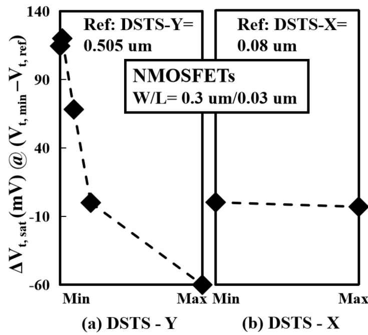

FIGURE 4. $V_{T,sat}$ dependency of (a) DSTS-Y and (b) DSTS-X effect for conventional SMT process nMOSFETs (SA/SB = 0.205 um).

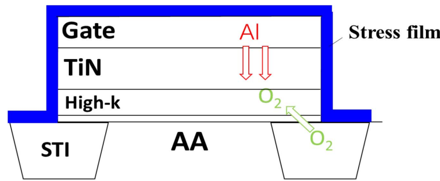

FIGURE 5. Schematic shows the diffusion of oxygen mechanism after the stress film deposited. The oxygen in STI diffuse into the high-k layer through the gate-STI overlap regions during the high temperature annealing step.

In addition, they have opposite effects on the $V_{T,sat}$ shift. In this article, a hypothesis was developed to explain the mechanism of the result. In the high-k first/metal gate last scheme, the gate stack is filled with Al, work function metals and some barrier layers [8]. For nMOSFETs, the work function metal is known to be Al-based metal and the migration of Al is the key element for n-type EWF. On the other hand, there is always a certain amount of oxygen in STI (High Density Plasma oxide) due to process problems. For regular SMT process, during the high temperature annealing with a stressor film covered, the oxygen enters the high-k layer through the gate/STI overlap portion [9], as shown in Fig. 5. The oxygen distribution in the high-k layer affects the diffusion of Al in metal gate. The decrease of the channel width and the increase of the DSTS-Y results in a rise of the oxygen concentration in the high-k layer. The migration of Al eventually caused the work function and the $V_{T,sat}$ shift of nMOSFETs. Since the gate/STI overlap only in the channel width direction, the phenomenon does not exist along the channel length direction.

IV. MODIFIED SMT PROCESS LPES STUDY

As mentioned above, the high temperature annealing step in the regular SMT process results in the diffusion of oxygen to the gate stack and aggravates the $V_{T,sat}$ shift of nMOSFETs. On the other hand, due to the LOD effect, the performance gain of the regular SMT process is reduced. This section investigated a modified SMT process that reduces the LPEs while improving the performance of nMOSFETs. For the modified SMT process, SPER takes the place of high temperature annealing (after the stress film deposition) and the oxygen diffusion is suppressed.

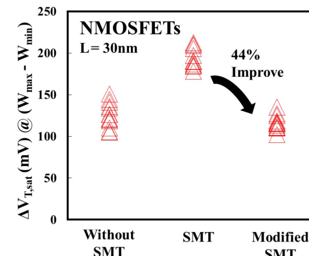

FIGURE 6. nMOSFETs $V_{T,sat}$ shift by channel width effect versus different process flows.

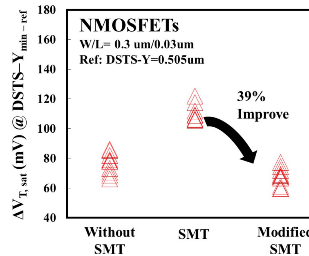

FIGURE 7. nMOSFETs $V_{T,sat}$ shift by DSTS-Y increase versus different process flows.

A. WIDTH AND DSTS-Y EFFECT COMPARISONS

Fig. 6 characterizes the impact of channel width on nMOSFETs $V_{T,sat}$ with different processes (without SMT, with regular SMT and with modified SMT). The width effect is characterized by the $\Delta V_{T,sat}$ $(= V_{T,sat}(W = max) - V_{T,sat}(W = min))$. It can be observed that the width effect of nMOSFET decreases by 44% with modified SMT using the low temperature annealing.

As stated above, the DSTS-X effect is not obvious for HKMG devices with the SMT process. Therefore, the DSTS-Y effect of nMOSFETs for different processes are studied. DSTS-Y effect results for the nMOSFETs with different processes are in Fig. 7, where regular SMT has the largest $V_{T,sat}$ shift, and modified SMT $V_{T,sat}$ shift was significantly reduced (about 39%). The DSTS-Y effect is characterized by the $\Delta V_{T,sat}$ $(= V_{T,sat}(DSTS-Y = min) - V_{T,sat}(DSTS-Y = ref)$, Ref: DSTS-Y $= 0.505$um).

B. LOD EFFECT IMPROVED AND PERFORMANCE COMPARISON

The layout parameter LOD (so called SA/SB) refers to the distance from the gate edge to the AA edge, as shown in Fig. 1. Different from the width effect and DSTS effect, the LOD effect is defined as the performance regression of nMOSFET with the reduction of SA/SB. Fig. 8 shows the LOD dependency for different process flows (with regular SMT and with modified SMT). The normalized nMOSFET performance is defined as the saturated drive current (I_{dsat}) at fixed off-current (I_{off}), reference to LOD=0.985um. For the regular SMT process, as the LOD decreases, the performance of nMOSFETs dropped by about 10%. The modified SMT process can effectively keep the LOD effect to about 6%. This is on account of the low temperature SPER, which can

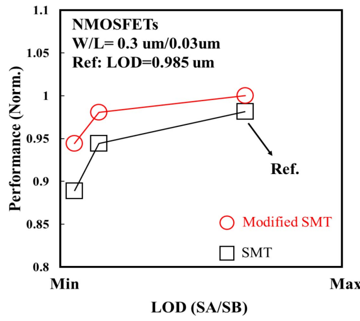

FIGURE 8. LOD (SA/SB) dependency for nMOSFETs normalized performance of regular SMT process and modified SMT process. The inset of this Figure shows the LOD layout geometry.

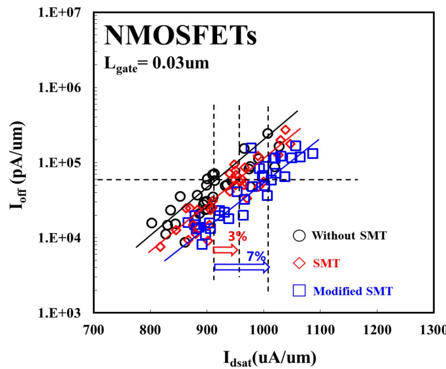

FIGURE 9. $I_{off} - I_{dsat}$ curves of nMOSFETs performance under different process flows. I_{dsat} enhances about 3% by the regular SMT process and 7% by modified SMT process at fixed I_{off}, respectively.

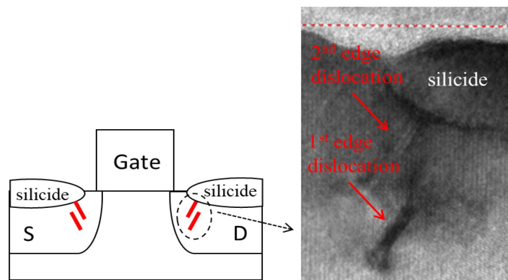

FIGURE 10. A schematic showing edge stacking fault defect inside S/D region typically appear in the NMOS with modified SMT process and the TEM image.

prevent the loss of S/D dislocations by controlling the rate of epitaxy regrowth.

Furthermore, the low temperature SPER step is more conducive generating edge dislocations in the S/D region for SMT effect enhancement. I_{dsat} vs I_{off} for nMOSFETs by different process flows are plotted in Fig. 9. I_{dsat} is improved in both SMT processes. The effect of performance improvement with regular SMT process is 3% and with modified SMT is 7%. The transmission electron microscope (TEM) image (Fig. 10) shows mask edge dislocation image when an SPER is applied, where two dislocation defects can be found in the S/D region. This multiple stacking faults are the sign of plastic deformation caused by SMT [10].

V. CONCLUSION

In summary, this work highlights the impact of the SMT process on layout proximity effects of HKMG nMOSFET, including width effects, DSTS-Y effects and LOD effect. Based on experimental results, the mechanisms of LPEs was investigated. Furthermore, a modified low temperature SMT process is proposed to reduce the above LPEs effectively. The modified SMT process shows high saturated drive current gain of ∼7%, which can be a feasible solution for sub 28nm HKMG nMOSFET performance boosting.

ACKNOWLEDGMENT

This research is supported by Key Laboratory of Microelectronics Devices & Integrated Technology, Institute of Microelectronics, Chinese Academy of Science, Beijing, China.

REFERENCES

[1] K.-Y. Lim *et al.*, "Novel stress-memorization-technology (SMT) for high electron mobility enhancement of gate last high-k/metal gate devices," in *IEDM Tech. Dig.*, 2010, pp. 229–232.

[2] Y. Liu *et al.*, "NFET effective work function improvement via stress memorization technique in replacement metal gate technology," in *VLSI Symp. Tech. Dig.*, 2013, pp. T198–T199.

[3] J. G. Min *et al.*, "The impact of dislocation on bulk -Si FinFET technologies: Physical modeling of strain relaxation and enhancement by dislocation," in *Proc. IEEE 13th Nanotechnol. Mater. Devices Conf. (NMDC)*, 2018, pp. 1–4.

[4] J. V. Faricelli, "Layout-dependent proximity effects in deep nanoscale CMOS," in *Proc. IEEE CICC*, 2010, pp. 1–8.

[5] T. Y. Wen *et al.*, "Fin bending mitigation and local layout effect alleviation in advanced FinFET technology through material engineering and metrology optimization," in *VLSI Symp. Tech. Dig.*, 2019, pp. T110–T111.

[6] F. Sato *et al.*, "Process and local layout effect interaction on a high performance planar 20nm CMOS," in *VLSI Symp. Tech. Dig.*, 2013, pp. T116–T117.

[7] C. Ndiaye, V. Huard, R. Bertholon, M. Rafik, X. Federspiel, and A. Bravaix, "Layout dependent effect: Impact on device performance and reliability in recent CMOS nodes," in *Proc. IEEE IIRW*, 2016, pp. 24–28.

[8] S. Hyun *et al.*, "Aggressively scaled high-k last metal gate stack with low variability for 20nm logic high performance and low power applications," in *VLSI Symp. Tech. Dig.*, 2011, pp. 32–33.

[9] M. S. Sivanaresh and N. R. Mohapatra, "Analysis and modeling of the narrow width effect in gate-first HKMG nMOS transistors," *IEEE Trans. Electron Devices*, vol. 62, no. 4, pp. 1085–1091, Apr. 2015.

[10] T.-M. Shen, S.-J. Wang, Z.-R. Xiao, C.-C. Wu, J. Wu, and C. H. Diaz, "Impact of SMT-induced edge dislocation positions to NFET performance," in *Proc. 73rd Annu. Device Res. Conf. (DRC)*, 2015, pp. 184–184.