Vol 16 No 10, October 2007
1009-1963/2007/16(10)/3104-04
Chinese Physics
© 2007 Chin. Phys. Soc.
and IOP Publishing Ltd

Effect of STI-induced mechanical stress on leakage current in deep submicron CMOS devices

Li Rui(李　睿)[a)b)c)†], Yu Liu-Jiang(俞柳江)[c)],
Dong Ye-Min(董业民)[c)], and Wang Ching-Dong(王庆东)[c)]

[a)] *Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai* 200050, *China*
[b)] *Graduate School of Chinese Academy of Sciences, Beijing* 100049, *China*
[c)] *Grace Semiconductor Manufacturing Corporation, Shanghai* 201203, *China*

(Received 12 December 2006; revised manuscript received 10 January 2007)

The shallow trench isolation (STI) induced mechanical stress significantly affects the CMOS device off-state leakage behaviour. In this paper, we designed two types of devices to investigate this effect, and all leakage components, including sub-threshold leakage ($I_{\rm sub}$), gate-induced-drain-leakage ($I_{\rm GIDL}$), gate edge-direct-tunnelling leakage ($I_{\rm EDT}$) and band-to-band-tunnelling leakage ($I_{\rm BTBT}$) were analysed. For NMOS, $I_{\rm sub}$ can be reduced due to the mechanical stress induced higher boron concentration in well region. However, the GIDL component increases simultaneously as a result of the high well concentration induced drain-to-well depletion layer narrowing as well as the shrinkage of the energy gap. For PMOS, the only mechanical stress effect on leakage current is the energy gap narrowing induced GIDL increase.

Keywords: CMOS, shallow trench isolation stress, leakage, gate-induced drain leakage
PACC: 7340G, 7300, 7360J, 7340Q

1. Introduction

The control of the off-state drain leakage ($I_{\rm off}$) is one of the main challenges for MOSFETs scaling.[1] High leakage current in deep submicron regime is becoming a significant contributor to power dissipation of CMOS circuits as threshold voltage ($V_{\rm th}$), channel length, and gate oxide thickness ($T_{\rm ox}$) are reduced. For deep submicron technology, the $I_{\rm off}$ increases with changes in device scales[2]: 1) the sub-threshold leakage ($I_{\rm sub}$) increases exponentially with the reduction of $V_{\rm th}$; 2) the gate induced drain leakage ($I_{\rm GIDL}$) and gate edge direct tunnelling leakage ($I_{\rm EDT}$) increase as $T_{\rm ox}$ becomes thinner; 3) the band-to-band tunnelling leakage ($I_{\rm BTBT}$) at drain-to-bulk junction also increases due to the increasing source-drain-extension (SDE) and pocket concentration. Among these, $I_{\rm GIDL}$ has become one of the greatest limiting factors on advanced semiconductor devices.[3]

Additionally, with the continuous downscaling of CMOS technology, shallow trench isolation (STI) induced mechanical stress increases with the shrinkage of device active area. Many processing steps individually or collectively contribute to the development of STI stress, such as the STI formation, thermal oxidation processes after STI formation, etc.[4] STI stress will lead to strained region in silicon, and thus changes the silicon band-gap, the diffusivity of impurities in silicon, the mobility of both electrons and holes, and so on. These phenomena can cause non-negligible impacts on device performance. As a result, devices with the same gate length ($L_{\rm g}$) and width ($W_{\rm g}$) but different layout patterns and/or surroundings exhibit inconsistencies in characteristics, such as threshold voltage, trans-conductance ($g_{\rm m}$), saturation drain current ($I_{\rm dsat}$), etc. This problem has significant impact on precise analogue circuits.

In the present research, we have designed two devices with different active overhangs to explore the STI-induced mechanical stress effect on CMOS leakage current. All leakage components contributing to the off-state current, including $I_{\rm sub}$, $I_{\rm GIDL}$, $I_{\rm BTBT}$, and $I_{\rm EDT}$ were analysed. The results were then verified by TSUPREM-4 stress simulation. The 2D numerical process simulator TSUPREM-4 simultaneously solves force balance equations while taking into

†E-mail: rayli@gsmcthw.com

account the volume expansion during oxidation, thermal mismatch between materials, intrinsic strain in deposited layers and surface tension.[5]

2. Experiment

As shown in Fig.1, the only difference between the two devices is the active overhang (L_a): 0.4 μm for A-device and 5 μm for B-device. Other parameters, including gate width ($W_g = 10\,\mu$m), gate length ($L_g = 0.15\,\mu$m) and contact-to-gate spacing (0.13 μm) are identical for A- and B-device.

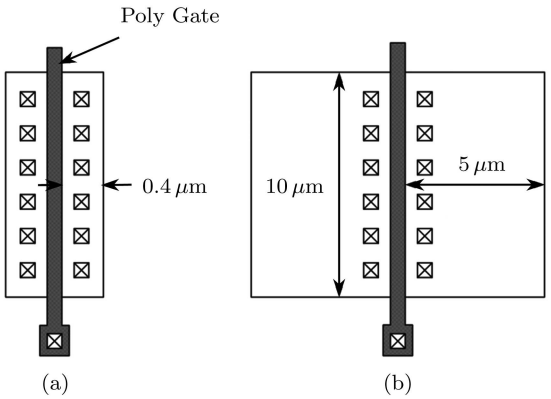


Fig.1. Layout of the two devices: (a) A-device: $L_a = 0.4\,\mu$m and (b) B-device: $L_a = 5\,\mu$m.

The devices were fabricated on (100) p-type 8-12 ohm.cm silicon substrate with a standard 0.15 μm self-aligned-silicide (SALICIDE) CMOS process. After laying out of active region, the two-step silicon etching is applied to form STI with top-corner-rounding (TCR). A 15 nm oxide liner was grown at the trench sidewall and high density plasma (HDP) CVD dielectric films were used to fill the trenches. The device channels were doped with boron and arsenic for NMOS and PMOS, respectively, before the growth of gate oxide. The thickness of gate oxide was 2.5 nm.

3. Results and discussion

3.1. Layout dependent device characteristics

Table 1 shows a comparison of the key MOSFET parameters of each two devices of NMOS and PMOS, including I_{off} and its four leakage components, as well as V_{th} and I_{dsat}. It is found that the difference in active overhang has a significant influence on device characteristics. The A-device with narrow active overhang has 13.9% and 3.0% higher off-state leakage than B-device for NMOS and PMOS, respectively, in spite of the same gate length and width. Moreover, the V_{th} and I_{dsat} dependence on layout pattern are not consistent between NMOS and PMOS. For NMOS, the V_{th} of A-device is 45 mV higher than that of B-device. While for PMOS, the V_{th} of both devices are identical. The layout influences on I_{dsat} have opposite trends in NMOS and PMOS, which can be explained by electron mobility enhancement and hole mobility degradation due to compressive stress along channel direction.[6] Gate voltage dependencies of g_m in NMOS and PMOS are shown in Fig.2. For NMOS, the gate voltage at maximum g_m (g_{m_max}) of A-device is 40 mV higher and g_{m_max} is 5% lower than B-device. For PMOS, the g_{m_max} positions are the same for both devices, and g_{m_max} of A-device is 15% higher than that of B-device.

Table 1. Comparison of key parameters of NMOS and PMOS devices.

Parameters	NMOS			PMOS		
	A-device	B-device	Diff /%	A-device	B-device	Diff/%
I_{off}/(pA/ μm)	2.86	2.51	13.9	1.02	0.99	3.0
I_{sub}/(pA/ μm)	0.29	0.78	−63.4	0.65	0.66	−1.8
I_{GIDL}/(pA/ μm)	2.41	1.57	53.2	0.16	0.13	22.4
I_{BTBT}/(pA/ μm)	0.17	0.16	3.8	0.20	0.19	5.2
I_{EDT}/(pA/ μm)	<0.01	<0.01	/	<0.01	<0.01	/
V_{th}/V	0.635	0.590	7.6	0.600	0.600	0.0
I_{dsat}/(μA/μm)	370	420	−11.9	189	174	8.6

To understand the mechanism of off-state leakage dependence on layout, the four leakage components were extracted and compared respectively. For NMOS, the major contributor is GIDL leakage, which is more than 60% of the total drain leakage in both structures. Moreover, the magnitude of I_{GIDL} in A-device is 53% higher than that in B-device, which results in higher I_{off} in A-device. As for other leakage

components, $I_{\rm EDT}$ is negligible, while $I_{\rm sub}$ and $I_{\rm BTBT}$ still need to be taken into consideration. The $I_{\rm sub}$ of A-device is far smaller than that of B-device, while their $I_{\rm BTBT}$ are comparable. However, PMOS characteristic is much different from NMOS. $I_{\rm sub}$ are dominating components and almost identical in both structures. The magnitude of $I_{\rm GIDL}$ in A-device is 22% higher than that in B-device, but contributes little to $I_{\rm off}$. The other two components, $I_{\rm BTBT}$ and $I_{\rm EDT}$, are similar as NMOS.

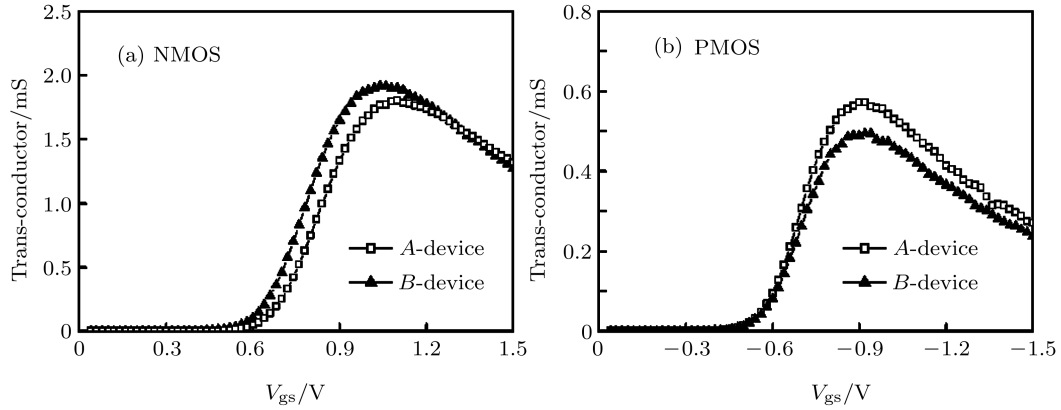


Fig.2. Comparison of transconductance curves of A- and B-device: (a) NMOS and (b) PMOS.

3.2. Analysis and mechanical stress simulation

We'll focus on the comparison of $I_{\rm sub}$ and $I_{\rm GIDL}$, while ignoring $I_{\rm EDT}$ and $I_{\rm BTBT}$ due to their negligible magnitude or inconspicuous difference. It is well known that $I_{\rm sub}$ is determined by MOSFET $V_{\rm th}$, and increases exponentially with decreasing $V_{\rm th}$. And $V_{\rm th}$ relates to effective gate oxide thickness (EOT), charge density at gate oxide interface ($Q_{\rm it}/Q_{\rm ox}$), and substrate dopant concentration ($N_{\rm b}$).[7] To investigate the mechanism of $V_{\rm th}$ behaviour, we used TSUPREM-4 to simulate the dopant concentration in MOSFET well region, while taking the stress effect into account. Figures 3(a) and 3(b) show the boron concentration in p-well and arsenic in n-well for both A- and B-device. It can be seen that boron diffusion is retarded in A-structure NMOS, which will lead to higher boron concentration near the surface and thus higher $V_{\rm th}$. However, arsenic concentrations are almost identical in both structures of PMOS, and won't cause $V_{\rm th}$ difference. The stress distributions in active area along channel direction were simulated for both devices with different active overhangs and are shown in Fig.4. It reveals that the compressive stress originates from STI edge and is attenuated to the centre of active region. As a result, the magnitude of stress under the channel region increases with decreasing $L_{\rm a}$. It has been reported that compressive stress will reduce the diffusion constant of boron.[8] However, the diffusivity of arsenic seems to be immune to mechanical stress. Based on the above results, we conclude that different $I_{\rm sub}$ behaviours originate from STI-stress induced impurity concentration variation.

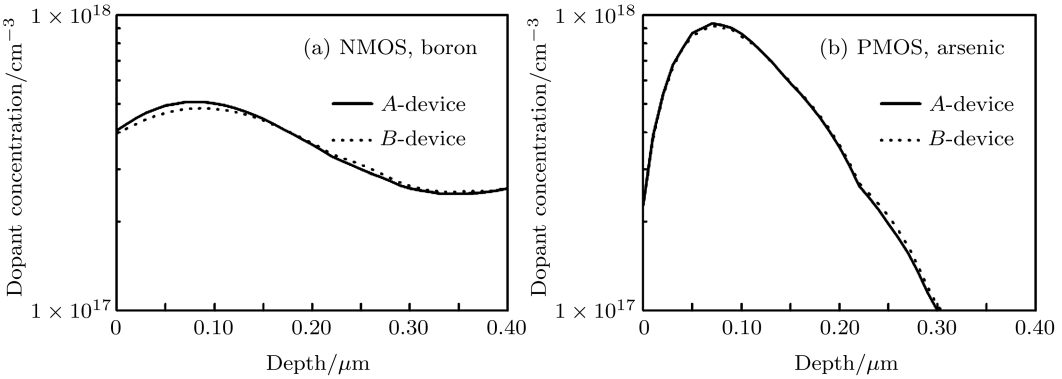


Fig.3. Simulated distribution of (a) boron and (b) arsenic dopant concentration in A- and B-structure NMOS and PMOS well regions, taking the stress effect into account.

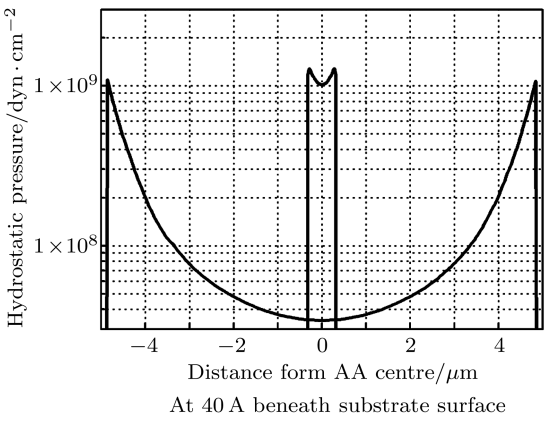


Fig.4. Simulated stress distributions along channel length direction of both *A*- and *B*-device.

Another main contributor of $I_{\rm off}$ is $I_{\rm GIDL}$, which is the band-to-band tunnelling leakage occurring at the interface of gate-to-drain overlap region. The mechanism of GIDL is rather complicated, and relevant to many factors such as drain and well dopant distribution, the density and distribution of interface trap at gate-to-drain region, silicon energy gap ($E_{\rm g}$), etc. For PMOS, the two devices did not show any difference in drain and well dopant, as well as $D_{\rm it}$, due to their threshold voltages being the same. As a result, the higher GIDL in *A*-device is believed to originate from the mechanical stress induced $E_{\rm g}$ narrowing.[9] When the silicon crystal is deformed by external compressive stress, the distances between silicon atoms are changed and thus the $E_{\rm g}$ becomes narrowed. This is consistent with the result of device simulation.[10] For NMOS, we have already known the retardance of boron diffusion in the channel region. Thus, higher boron concentration near the surface causes the drain-to-well depletion layer to be narrower. The narrowing of the depletion layer near the surface leads to an increase of local electric field, and hence higher GIDL. Meanwhile, the stress induced energy gap change will also make contribution in NMOS. Therefore, the GIDL dependence on mechanical stress in NMOS is more sensitive than in PMOS.

In recent years, the strained silicon technology has been widely used in advanced VLSI manufacturing.[11] Tensile and compressive stresses are introduced to enhance electron and hole mobility, respectively. From the engineering point of view, tensile stress in NMOS will also be helpful to decrease the total $I_{\rm off}$. However, compressive stress in PMOS will result in higher leakage based above analysis. Thus, a trade-off should be considered between on-state current increase and off-state leakage decrease in PMOS.

4. Summary

In summary, two styles of devices were fabricated to investigate the STI process induced mechanical stress effect on off-state leakage current. For *A*-structure NMOS with small active overhang, $I_{\rm sub}$ is reduced due to mechanical-stress-induced higher boron concentration in well region. However, the GIDL component increases simultaneously as a result of the narrowing of high well concentration induced drain-to-well depletion layer as well as the shrinkage of energy gap. Therefore, the total $I_{\rm off}$ of *A*-device is larger because GIDL is the main contributor of $I_{\rm off}$. While for PMOS, the only effect of mechanical stress on leakage current is the energy gap narrowing induced GIDL increase. And this effect should be considered when attempting to take advantage of the compressive stress to enhance hole mobility.

References

[1] Roy K, Mukhopadhyay S and Mahmoodi-Meimand H 2003 *Proc. of the IEEE* **91** 305

[2] Lin Y S, Wu C C, Chang C S, Yang R P, Chen W M, Liaw J J and Diaz C H 2002 *IEEE Trans. Electron Devices* **49** 1034

[3] Chen H F, Hao Y, Ma X H, Zhang J C, Li K, Cao Y R, Zhang J F and Zhou P J 2006 *Chin. Phys.* **15** 645

[4] Senez V, Hoffmann T, Robilliart E, Bouche G, Jaouen H, Lunenborg M and Carnevale G 2001 *IEDM Tech. Dig.* 831

[5] Synopsys TSUPREM-4 User Guide, Synopsys Inc. 2004

[6] Bianchi R A, Bouche G and Roux-dit-Buisson O 2002 *IEDM Tech. Dig.* 117

[7] Sze S M 1981*Physics of Semiconductor Devices* 2nd Ed. (New York: Wiley)

[8] Kuo P, Hoyt J L, Gibbons J F, Turner J E and Lefforge D 1995 *Appl. Phys. Lett.* **66** 580

[9] Egley J L and Chidambarrao D 1993 *Solid-State Electronics* **36** 1653

[10] Yang W, Qin G, Shao X, Yu Z and Tian L 2005 *IEEE Conference on Electron Devices and Solid-State Circuits* 769

[11] Tyagi S, Auth C, Bai P, Curello G, Deshpande H, Gannavaram S, Golonzka O, Heussner R, James R, Kenyon C, Lee S-H, Lindert N, Liu M, Nagisetty R, Natarajan S, Parker C, Sebastian J, Sell B, Sivakumar S, St Amour A and Tone K 2005 *IEDM Tech. Dig.* 245