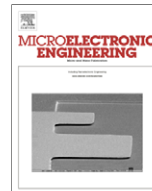
Microelectronic Engineering xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Microelectronic Engineering

journal homepage: www.elsevier.com/locate/mee



Impact of PDA temperature on electron trap energy and spatial distributions in SiO_2/Al_2O_3 stack as the IPD in Flash memory cells

X.F. Zheng [a], W.D. Zhang [a,*], B. Govoreanu [b], J.F. Zhang [a], J. Van Houdt [b]

[a] School of Engineering, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[b] IMEC, RDO/PT Division, Kapeldreef 75, B-3001 Leuven, Belgium

ARTICLE INFO

Article history:
Received 3 March 2009
Accepted 5 March 2009
Available online xxxx

Keywords:
Electron trap
Pulsed *C–V*
Energy distribution
Spatial distribution
High-κ dielectrics
Al_2O_3
IPD
Floating gate
Flash memory

ABSTRACT

SiO_xN_y/high-κ dielectric stack will soon replace the conventional SiO_xN_y-*based* dielectric stacks in the future generations of flash memory cells. Characterizing and reducing electron traps in the high-κ layer is an important task, as the large trap density may limit the memory retention via the trap-assisted tunneling. Since the post-deposition annealing (PDA) has great impact on the microstructure of high-κ dielectric, it is important to understand how PDA affects the properties of electron traps, such as the trap density, energy and spatial distributions. It is demonstrated in this paper that, by using a recently developed two-pulse *C–V* measurement technique, the energy and spatial distributions of electron traps throughout the SiO_2/high-κ stack can be characterized, and PDA temperatures have different impacts on traps at different energy levels and spatial locations.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

The scaling of the conventional SiO_xN_y-based tunnel and control dielectric layers in Flash memory technology is fast approaching its limits [1], as the increasing leakage current through thinner SiO_xN_y layers will result in a faster data loss rate. According to the ITRS Roadmap 2007 [2], this becomes the most pressing issue to be solved in floating gate Flash memory, in order to meet the charge retention and endurance requirements for future generations of Flash memory cells. Furthermore, a strong further reduction of the dielectric thickness between control and floating gates, i.e., the interpoly dielectric (IPD), is required to maintain the coupling ratio, starting from 45 to 40 nm technology generation for floating gate Flash devices [1,2]. Data retention, however, does not allow the thickness of SiO_xN_y to be further reduced.

Introduction of high-κ dielectrics as the IPD for floating gate Flash cells has been proposed as a potential solution [1,2]. Higher dielectric constant will increase the IPD capacitance without reducing its physical thickness, therefore help maintaining the coupling ratio and allow the cell size to continue down-scaling. However, the commercial application of high-κ layers in flash memory has been held back by their large electron trap density, which may limit the memory retention due to trap-assisted tunneling [1,2]. There is a pressing need for the characterization of electron traps in the high-κ layer, especially for characterizing their energy and spatial distributions, which are the essential information required for estimating the leakage current.

The existing techniques commonly used for dielectric characterization include various charge pumping (CP) [3,4], capacitance–voltage (CV) [5], conductance [6], transfer-characteristics (IV) measurements [7]. These techniques were developed mainly for characterizing the instability of MOS devices used in CMOS technologies and only have limited ability in determining the energy and spatial distribution of defects. Although these techniques allow the energy distribution of interface states to be measured within the bandgap of Si, they give little information on the energy distribution of defects in the bulk of dielectrics [4]. On the spatial distribution, some efforts [3,8,9] were made to probe into the dielectric, but none of these techniques can give a distribution throughout the dielectric. It has been demonstrated that our recently developed two-pulse *C–V* measurement technique [10] can overcome the major shortcomings of conventional measurement techniques and can probe the electron traps throughout the SiO_2/high-κ stack.

Reducing electron traps in the high-κ layer is important for the suppression of leakage current. It has been reported that PDA at different temperatures significantly changes the microstructure of high-κ layers, which may in turn affect the properties of electron traps [11]. It is possible that electron trap density can be reduced

* Corresponding author. Tel.: +44 151 231 2868; fax: +44 151 298 2624.
E-mail address: W.Zhang@ljmu.ac.uk (W.D. Zhang).

0167-9317/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mee.2009.03.066

ARTICLE IN PRESS

through optimizing post-deposition annealing (PDA) temperature. There is, however, little quantitative information available on how PDA affects electron traps, especially in terms of their energy and spatial distribution.

It is demonstrated in this paper that, by using the recently developed two-pulse *C–V* measurement technique, the energy and spatial distributions of electron traps throughout the SiO_2/high-κ can be obtained. It is observed that a decrease of PDA temperature from 1000 to 900 °C does not reduce trapping. A further reduction to 800 °C mainly reduces the electron traps at certain energy levels. PDA also has different impacts on electron traps at different spatial locations.

2. Devices

SiO_2/Al_2O_3 MOS capacitors were fabricated at IMEC using a process flow similar to that used for forming the Inter-poly Dielectric (IPD) stacks in floating gate flash memory devices. A 2 nm-thick High Temperature Oxide (HTO) was deposited as bottom layer on a *n*-type Si substrate without pn junctions. A 6 nm Al_2O_3 layer was then deposited by atomic layer chemical vapor deposition (ALCVD), followed by a conventional post deposition anneal (PDA). PDA was carried out in N_2 at 800, 900 or 1000 °C, covering the range of interest. Device fabrication was completed by deposition and etching of a TiN gate. The size of capacitors is 9×10^{-4} cm^2.

3. Measurement technique

The charge pumping technique relies on the measurement of a recombination current, which is proportional to frequency. Since detrapping time increases exponentially with the tunneling distance [12], it requires lower frequency in order to be able to probe traps situated farther from the Si substrate. The practical limitation of the frequency is typically around 100 Hz, giving a detrapping time less that 10 ms. This makes the charge pumping only suitable to probe the traps within 2 nm from the substrate [8]. The SiO_2/High-κ layers required in Flash memory cells are much thicker, usually in the range of 5–15 nm, so that the majority space is beyond the reach of even advanced CP technique.

The traditional CV/IV measurements have been used to probe charge loss at long time scales [12,13], but these measurements are too slow to characterise the defects near the interface because of fast detrapping [7,13]. To improve the measurement speed, the single pulse *I–V* technique was developed, but it did not give the energy and spatial distribution of defects [14].

In this work, we use the recently developed two-pulse *C–V* technique [10], as illustrated by Fig. 1. The waveform of the two gate pulses is shown in the inset of Fig. 1. The first pulse is used to charge the capacitor and the flat band voltage after electron trapping can be determined from its ramp-down trace. The capacitor is then discharged for a period of time, $T_{discharge}$, under a gate bias

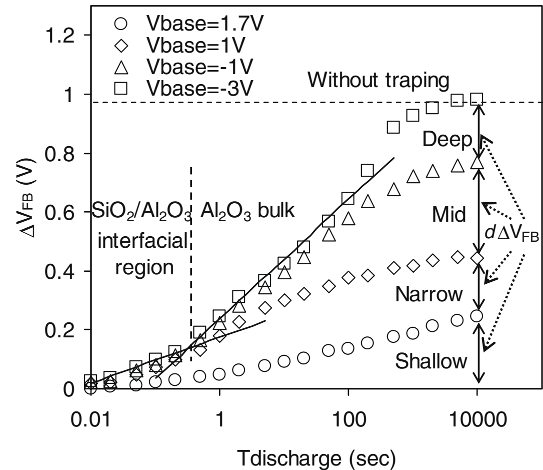


Fig. 2. The discharge induced ΔV_{FB} against discharge time at various V_{base}. The dashed horizontal line is the total trapping level.

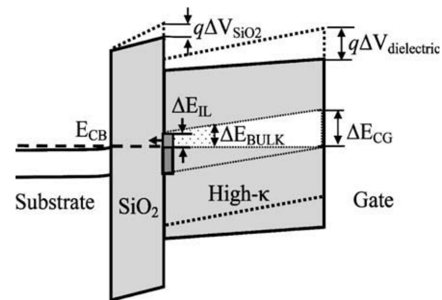


Fig. 3. Energy band diagram of the SiO_2/Al_2O_3 stack during discharge. $E_{IL} = q\Delta V_{SiO_2}$ and $E_{CG} = q\Delta V_{dielectric}$.

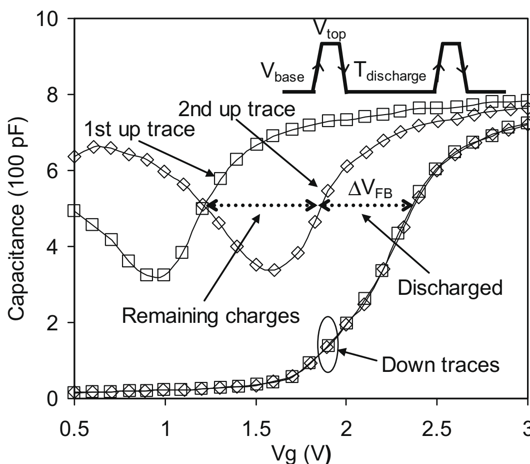


Fig. 1. Typical 2-pulse *C–V* test results. Inset shows the waveform. Test conditions are V_{top} = 5 V, T_{top} = 1 ms, $T_{discharge}$ = 10 s, and ramp slope = 10 kV/s.

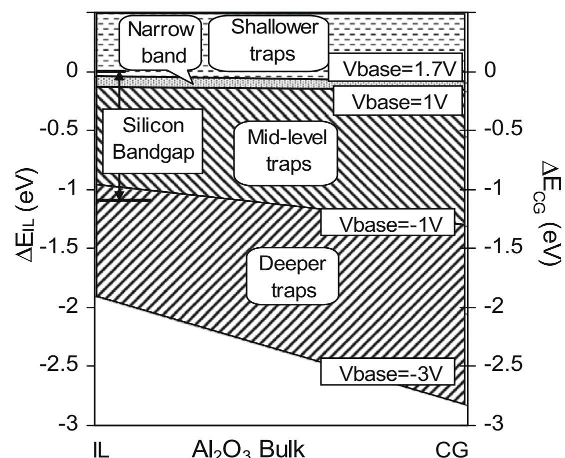


Fig. 4. Illustration of dischargeable energy bands with respect to the Si E_{CB} at the corresponding discharge bias.

Please cite this article in press as: X.F. Zheng et al., Microelectron. Eng. (2009), doi:10.1016/j.mee.2009.03.066

ARTICLE IN PRESS

of V_{base} until the beginning of the 2nd pulse. The amount of discharge is determined by the V_{FB} shift between the ramp-up trace of the 2nd pulse and the ramp-down trace of the 1st pulse. Since there is no limit on the maximum $T_{discharge}$, this provides an effective way to overcome the shortcomings of charge pumping techniques and to probe defects across the SiO_2/high-κ stack, as demonstrated in Ref. [10].

Fig. 2 shows the effect of $T_{discharge}$ and V_{base} on ΔV_{FB}. The change in the slope of ΔV_{FB} versus log($T_{discharge}$) at about 200 ms when moving into the high-κ layer confirms that trapping density in Al_2O_3 is higher than that in SiO_2. The excellent agreement between the discharge saturation after approximate 1000 s at $|V_{base} = -3\text{ V}|$ and the total trapping level confirms that 100% discharge was achieved. It is also shown that the Al_2O_3 layer dominates trapping. This demonstrates the capability of the 2-pulse C–V technique for probing traps throughout the entire stack.

Fig. 3 shows that only electron traps with energy level higher than the Si conduction band bottom, E_{CB}, can be discharged via tunneling, since there is no empty state in Si bandgap [4]. When the gate bias increases, the energy level of electron traps at the SiO_2/high-κ and gate interfaces will increase by $\Delta E_{IL} = q\Delta V_{SiO_2}$ and $\Delta E_{CG} = q\Delta V_{dielectric}$, respectively, and traps within this energy region become dischargeable. For a given V_{base}, ΔE_{IL} and ΔE_{CG} is determined from the model in Ref. [15]. Fig. 4 shows that by carefully selecting V_{base}, discharging within four energy regions in the Al_2O_3 bandgap can be studied: the shallower region mainly above Si E_{CB}; the narrow region near Si E_{CB} with a band width of about 0.1 eV; the mid-level region mainly corresponding to Si bandgap; and the deeper region mainly below Si valence band top. The differences in ΔV_{FB} at different V_{base} in Fig. 2, $d\Delta V_{FB}$, are caused by discharging traps within the corresponding region.

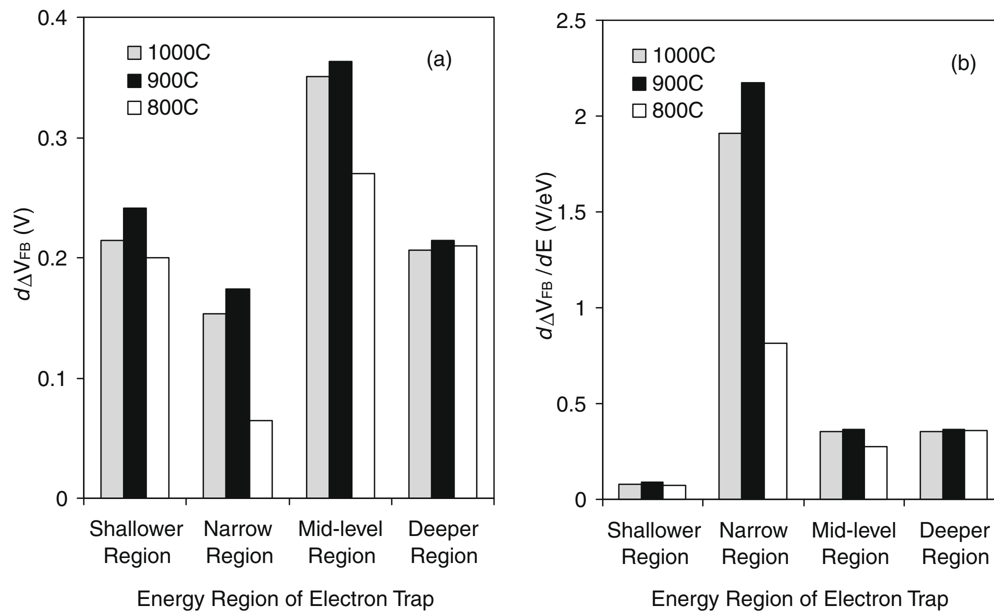


Fig. 5. (a) The total trapping in each region after PDA at different temperatures. (b) The trapping densities per eV in each region after PDA at different temperatures.

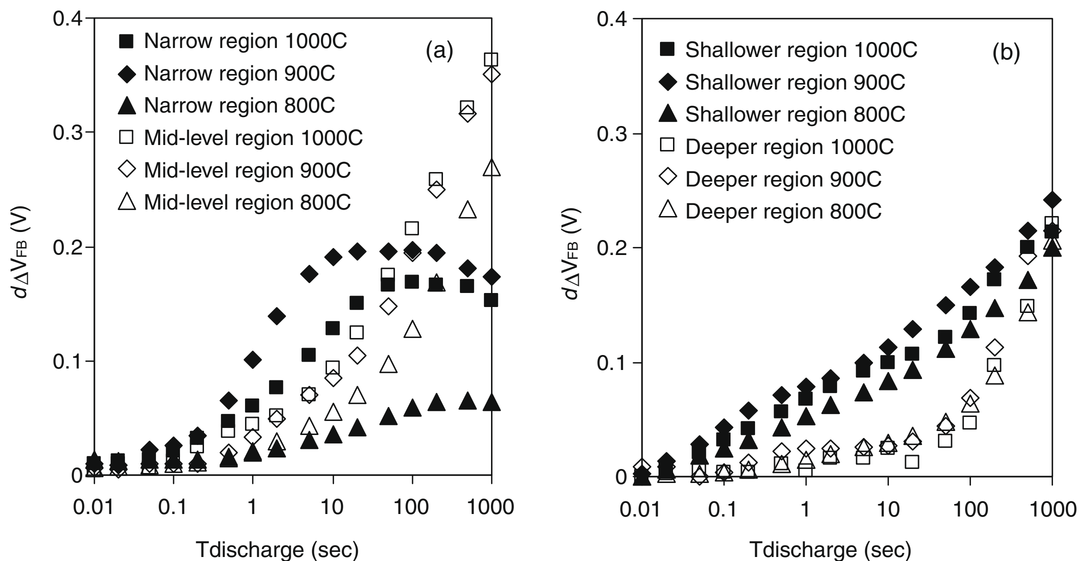


Fig. 6. Discharge against $T_{discharge}$ in each region after PDA at different temperatures. (a) Shows the narrow region and mid-level region. (b) Shows the shallower region and deeper region.

Please cite this article in press as: X.F. Zheng et al., Microelectron. Eng. (2009), doi:10.1016/j.mee.2009.03.066

4. Impact of pda temperature

Fig. 5a compares the total trapping density in each energy region after different PDA. Trapping in all regions actually increases slightly when PDA temperature reduces from 1000 to 900 °C, therefore a reduction of temperature in this range will not reduce trapping. A further decrease of temperature to 800 °C, however, leads to a clear decrease of trapping, as Al_2O_3 is crystallized at temperatures above 800 °C and crystallization introduces additional electron traps [11]. The largest reduction occurs in the narrow region near E_{CB}. The reduction is 67%, 27%, 17% and 2% in the narrow, mid-, shallower, and deeper regions, respectively. As a result, the PDA temperature mainly affects electron traps near E_{CB} and in the Si bandgap, where the O-vacancy is located [16]. It should also be noted that although $d\Delta V_{FB}$ is the lowest for the narrow region, the trap density per eV is actually the highest here due to its small band width of about 0.1 eV, as shown in Fig. 5b. Despite of the increased trap density, better retention has been observed in memory cells with crystallized Al_2O_3 IPD layers [5]. This may be attributed to its increased SiO_2/Al_2O_3 band offset from 2.3 to 2.6 eV [11], which reduces the leakage.

To assess the spatial location, Fig. 6a and b show the discharge against time. A longer time corresponds to discharge further away from the substrate [12]. In comparison with the 800 °C result, the sharp rise in the narrow region after 900 °C PDA only occurs at 1–10 s, indicating they are close to SiO_2/Al_2O_3 interface. Therefore the traps near E_{CB} generated by high PDA temperatures are mainly located in the transition region between SiO_2 and Al_2O_3. In contrast, higher PDA temperatures only induce a modest increase of trapping in the shallow and mid-level regions, and no increase in the deeper region. The slopes of $d\Delta V_{FB}$ versus $\log(T_{discharge})$ for these three regions do not decrease for longer $T_{discharge}$, indicating that the trapping towards the gate does not reduce. In fact, the slope for the deep region starts to increase from 100 s, indicating that these deep traps are likely situated near the gate interface.

5. Conclusions

Impacts of PDA temperature on electron trap energy and spatial distributions in SiO_2/Al_2O_3 have been investigated. A decrease of temperature from 1000 to 900 °C will not reduce trapping. A further reduction to 800 °C mainly reduces the electron traps near E_{CB} and in the bandgap of Si. When moving towards the gate, the traps near E_{CB} reduce sharply, while the deeper trap actually increases.

Acknowledgements

The authors are grateful for the kind support of their colleagues. Fruitful discussions with Mr. D. Ruiz Aguado, Dr. M. Rosmeulen, Dr. P. Blomme and Dr. R. Degraeve at IMEC are very much appreciated. The test sample used in this work was provided by IMEC. The work is sponsored by the EPSRC under Grant EP/C508793/2 and by the HEFCE PRF scheme.

References

[1] J. van Houdt, in: Proceedings of IRPS, 2005, pp. 234–239.
[2] International Technology Roadmap for Semiconductors, 2007, <http://public.itrs.net>.
[3] M.B. Zahid, R. Degraeve, L. Pantisano, J.F. Zhang, G. Groeseneken, in: Proceedings of IRPS, 2007, pp. 55–60.
[4] E. Cartier, B.P. Linder, V. Narayanan, V.K. Paruchuri, in: IEDM Tech. Dig., 2006, pp. 57–60.
[5] D. Wellekens, J. de Vos, J. van Houdt, K. van der Zanden, in: Proceedings of the NVSMW Workshop, 2008, pp. 12–15.
[6] M.J. Uren, J.H. Stathis, E. Cartier, J. Appl. Phys. (1996) 3915–3922.
[7] C.Z. Zhao, J.F. Zhang, M.B. Zahid, B. Govoreanu, G. Groeseneken, S. de Gendt, J. Appl. Phys. (2006) (Art. No. 093716).
[8] Y. Wang, V. Lee, K.P. Cheung, in: IEDM Tech. Dig., 2006, pp. 491–494.
[9] D. Heh, C.D. Young, G.A. Brown, P.Y. Hung, A. Diebold, E.M. Vogel, B. Joseph, G. Bersuker, IEEE Trans. Elect. Dev. (2007) 1338–1345.
[10] W.D. Zhang, B. Govoreanu, X.F. Zheng, D. Ruiz Aguado, M. Rosmeulen, P. Blomme, J.F. Zhang, J. van Houdt, IEEE Elect. Dev. Lett. (2008) 1043–1046.
[11] V.V. Afanas'ev, A. Stesmans, B.J. Mrstik, C. Zhao, Appl. Phys. Lett. 81 (2002) 1678–1689.
[12] J.F. Zhang, S. Taylor, W. Eccleston, J. Appl. Phys. (1992) 1429–1435.
[13] W.D. Zhang, J.F. Zhang, R. Degraeve, G. Groeseneken, IEEE Trans. Elec. Dev. (2002) 1868–1875.
[14] A. Kerber, E. Cartier, L. Pantisano, M. Rosmeulen, R. Degraeve, T. Kauerauf, G. Groeseneken, H.E. Maes, U. Schwalke, in: Proceedings of IRPS, 2003, pp. 41–45.
[15] B. Govoreanu, D. Wellekens, L. Haspeslagh, J. de Vos, J. van Houdt, in: IEDM Tech. Dig., 2006, pp. 206–209.
[16] J. Robertson, Rep. Prog. Phys. 69 (2006) 327.

Please cite this article in press as: X.F. Zheng et al., Microelectron. Eng. (2009), doi:10.1016/j.mee.2009.03.066