

Investigation of ONO Layers for Application in One Time Programmable(OTP) SONOS Memories

H. Reisinger

Siemens Corporate Research and Development, Dep. ZFE T ME 1
Otto Hahn Ring 6, 81730 München, Germany

A new kind of SONOS memory is presented which is - in contrast to conventional SONOS EEPROM technology - not erasable but exhibits a data retention time which can be adjusted to be well above 1000 years.

1. INTRODUCTION

Two technologies are available for nonvolatile memories: The MNOS (or SNOS, SONOS, MONOS) and the floating gate (FLOTOX) technology¹). In both types of memories the gate dielectrics - i.e. ONO for SONOS- and $\approx 8\text{nm}$ SiO_2 for FLOTOX- memories - are heavily stressed while writing and erasing. So the much lower defect density of ONO which is 0.01cm^{-2} compared to 1cm^{-2} for SiO_2 clearly is an advantage of the SONOS technology. Another advantage is the fact that the charge in FLOTOX memories is stored on a conducting floating gate while in SONOS memories it is stored in a nonconductive Si_3N_4 layer. Leaky spots in the ONO films will cause only a local loss of charge while in the insulation of the floating gate they will cause a loss of charge all over the gate area.

The MNOS memory²) was invented nearly 30 years ago. Several hundred papers - dealing with MNOS physics and devices - led to today's SONOS memories with ONO thicknesses of about 10nm and programming voltages of below 10V^{3,4}). All data retention times that are reported, however, are hardly 10 years with the retention times of the modern thin SONOS memories even worse than the ones of the early thick MNOS memories.

This limited retention time is the major disadvantage of the SONOS technology. Its restricted application to military and space applications where the higher radiation hardness of the SONOS memories¹) is needed.

In this paper we want to show that - if reprogrammability is sacrificed - it is possible to increase the retention time by many orders of magnitude.

2. EXPERIMENTAL

The ONOs layers have thicknesses similar to the ones used in DRAMs. On (100) surfaces of $5\Omega\text{cm}$ n-type silicon bottom oxides from 30 to 50\AA were grown at 800°C in Ar-diluted O_2 . LPCVD Si_3N_4 layers were deposited at 700°C . The top oxides were grown by wet oxidation at 900°C with two different thicknesses 3.5 and 5nm corresponding to SiO_2 thicknesses of 150nm and to 300nm on silicon. Thicknesses were determined by Ellipsometry, capacitance measurements and TEM. MOS diodes with n^+ -doped polysilicon gates with areas from 10^{-4}cm^2 to 0.1cm^2 were used for the measurements. Minority carriers were generated by illumination with light.

3. RESULTS AND DISCUSSION

Fig.1 shows the write characteristics of two samples, one with thin the other with thick oxides. For the sample with

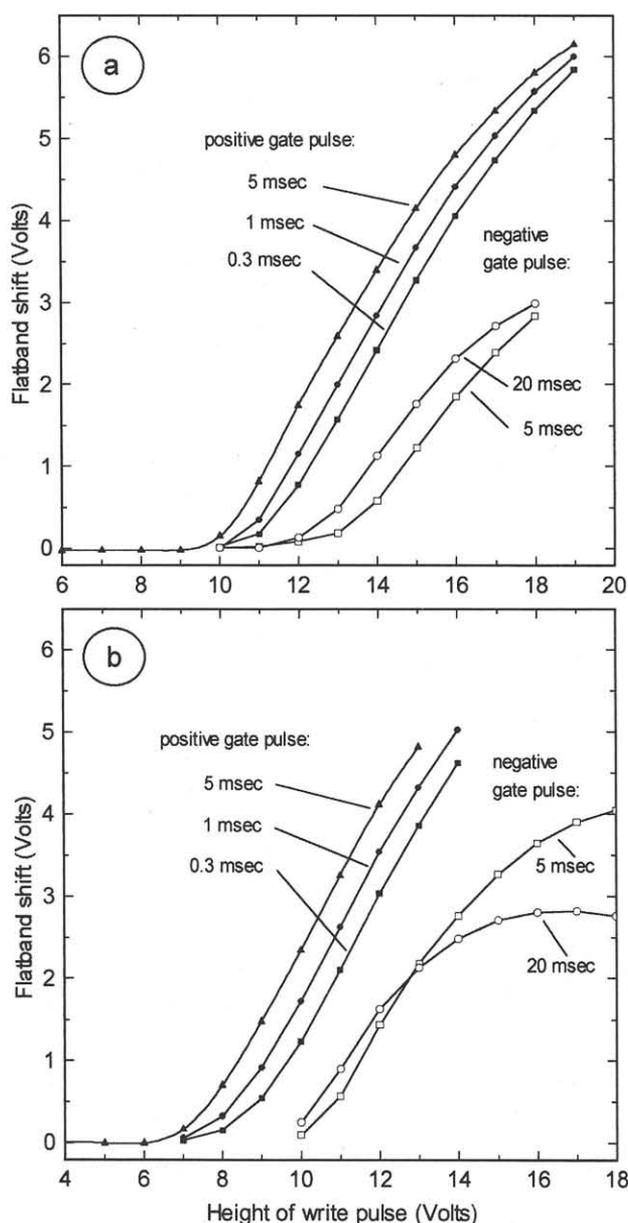


Fig.1: Write characteristics of 2 ONO films with thin and thick oxide films. (same thicknesses as in fig.3 a and b)

thick oxides a write up to a flatband shift of 2V would need a positive gate pulse of 1msec length and a height of

13V which is comparable to a conventional SONOS EEPROM. The sign of the charge trapped in the Si_3N_4 of SONOS EEPROMs, however, depends on the polarity of the gate pulse: Positive (write) pulses trap electrons,

The curves in Fig.1 for negative gate are shifted to higher voltages compared to positive gate. A part of this effect ($\approx 0.8\text{V}$) is due to the different doping levels of gate and substrate causing a workfunction difference of 0.4V ,

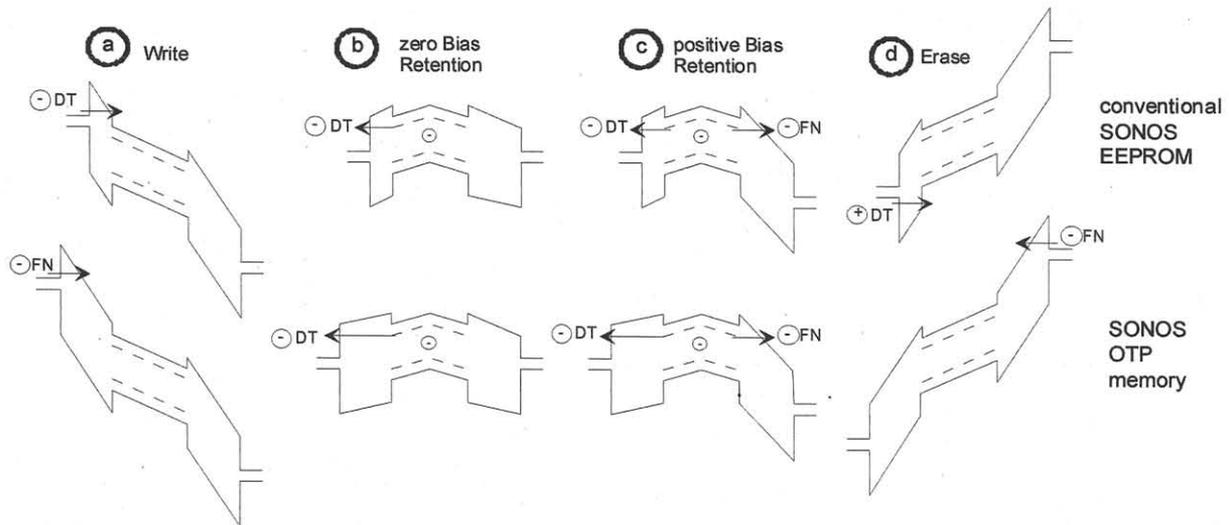


Fig.2: Energy band diagrams of the Si substrate, bottom SiO_2 , Si_3N_4 , top SiO_2 , Poly-Si gate (from left to right) structure for different gate biases. The dominant tunneling processes are shown for each diagram, DT and FN denote direct and Fowler Nordheim tunneling. Dashed lines mark trap levels in the Si_3N_4 . The trapped charge is drawn in the center of the Si_3N_4 for simplicity. Note that the oxide potential barrier is 1eV higher for holes than for electrons.

negative (erase) pulses detrapp the electrons and trap holes. In our ONOs for both positive as well as negative pulses the flatband shift is positive, so only electrons are trapped. The mechanism is illustrated in fig.2a and d.

In a SONOS EEPROM the bottom oxide is thin (20\AA) and the top oxide is thick ($>30\text{\AA}$). For both polarities and typical write/erase fields around 10MV/cm tunneling across the top oxide is in the Fowler Nordheim (FN) regime, across the bottomoxide in the direct tunneling regime for electrons and holes also. The tunneling probability for direct tunneling is much higher than the FN tunneling probability and is a strong function of the barrier (=oxide) thickness and the electric field while the FN current is only a function of the field and is independent of the oxide thickness. This asymetry causes the direct tunneling to be always dominant. The carrier transport - electrons during write, holes during erase - is always across the bottom oxide and never across the top oxide.

In the SONOS OTP memory both oxides are thicker than 30\AA . This means that for typical write/erase fields around 10MV/cm we are in the FN regime at both the bottom- and the top oxide. The ONO is electrically symmetric, so FN tunneling of electrons dominates the charge transport for both gate polarities. As a consequence we have a "write" for both polarities, an erasing is no longer possible. Theoretically there is the chance to erase even for thick oxides. For example with a 4nm bottom oxide and an 8nm top oxide one could have the direct tunneling across the bottom oxide and FN tunneling across the top oxide at a field of 5MV/cm . A simple estimation shows, however, that the small currents at fields that low would make times for erase pulses longer than 10^5 seconds which clearly is impracticable.

another part may be due to an asymmetric distribution of traps in the nitride. The bending and saturation of the curves for the large negative biases and the long pulse times are caused by positively charged traps in the oxide at the anode. These traps are known to be generated in the high field at the anode when the oxide is heavily stressed by FN current.

The initial slope of all the curves is approximately 1V/V . This is due to the fact that for each curve the field across the oxide at the injecting electrode - which decreases during the pulse due to the trapped charge and is determining the current - comes down to the same level for each pulse of a curve independent of the pulse height. The charge retention characteristics of ONOs are shown for two samples with thick (fig.3a) and thin (fig.3b) oxides. Conventional SONOS EEPROMs exhibit an almost complete closing of the memory window for SONOS EEPROMs for retention times of 10 years^{4,5}. Frequent reading of a memory cell may further bring down the retention time due to the applied gate voltage which may increase electric field (see fig.2b,c) and charge loss. In contrast there is still half of the charge left (at zero bias) for the sample in fig.3b allowing operation for at least another decade and nearly all the charge (99%) for the sample in fig.3a allowing operation for almost infinite time. Fig. 2b and c give a simplified illustration of the charge loss: In order to escape the charges trapped in the nitride have to tunnel from the trap position to the gate or to the substrate. The minimum tunnel distance is the oxide barrier. In a SONOS EEPROM the top (or blocking) oxide can be made as thick as needed but the bottom (or tunnel) oxide is always around 20\AA in order to keep erasability as pointed out above. The maximum bottom oxide thickness reported is

25Å⁴). As is easily seen in fig.2b for the retention state where the fields are low tunneling of charges sitting close to the SiO₂-Si₃N₄ interface is always in the regime of the direct tunneling. The tunneling current in this regime is an extremely strong function of the oxide thickness: Each Å more in oxide thickness brings half an order of magnitude decrease in current⁶).

For the OTP SONOS EPROMs this means that the thicknesses of bottom and top oxide can be tailored as needed for a given application. For a application where there is zero gate bias most of the time the most suitable

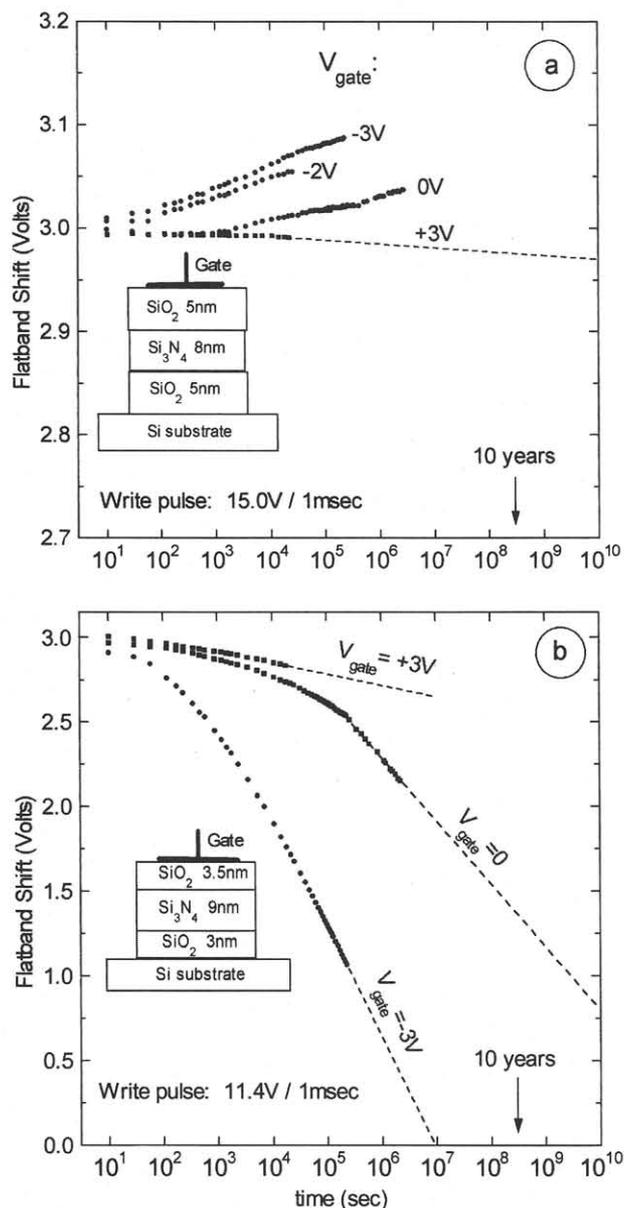


Fig.3: Unbiased and biased data retention of ONO films with different thicknesses of the oxide films.

Two examples for data retention in ONOs are given in fig. 3. For the symmetric ONO in fig. 3a with very thick oxides up to a time of 10⁷ sec there is no significant charge loss measurable. For zero and negative gate bias the flatband voltage rises with time. This indicates a redistribution of the trapped charge from its initial position close to the top oxide⁷) towards the center of the nitride. Negative gate biases speed up this movement. Only with a positive bias of 3V there is actually a slight charge loss. The charge loss of an asymmetric ONO with thinner oxides is shown in fig.3b. There is a significant charge loss across the bottom oxide for zero bias with a data retention well above 10 years, however. A negative bias increases the field across the bottom oxide thus increasing the charge loss. Positive bias decreases the field across the bottom oxide and so leads to a reduction of charge loss without drastically increasing the tunneling current across the (thicker) top oxide. The influence of bias on retention is the other way round if the asymmetry is inverted.

The data in fig. 3 were measured not at maximal operating temperature but at room temperature. Tunneling, however, is a process which is not thermally activated and so has only a slight temperature dependence. Actually measurements showed that the flatband shift of the ONOs after a 450°C/30min anneal is still 2V and for a complete discharge a 700°C/1h anneal is necessary.

4. CONCLUSION

By simply increasing the oxide thicknesses in SONOS memory structures the retention times can be made almost infinite. The oxide thicknesses also can be tailored to meet the requirements for a frequent reading of a memory cell, for example 10¹⁵ read cycles at 3V gate voltage. Erasing is no longer possible, however, with the thicker tunnel oxide. The voltage needed for writing is about 2 to 4V higher than the one needed for modern scaled SONOS EEPROMs.

References:

- 1) H.E. Maes, J. Witters and G. Groeseneken, Proc. 17th European Solid State Device Res. Conf. Bologna 1987, (1988) 157
- 2) H. Wegener, H. Pao, M O'Connell, and R.Olehiak, 1987 IEDM Tech. Digest
- 3) M.L. French, C. Chen, H. Sathianathan a. M.H. White, IEEE Trans. Compon. Packag. Manuf. Technol. A 17, no. 3, (1994) 390
- 4) T. Böhm, A. Nakamura, H. Aozasa, M. Yamagishi a. Y. Komatsu, Jpn. J. Appl. Phys. 35, (1996) 898
- 5) S. Minami and Y. Kamigaki, IEEE Trans. Electron Devices 38, no. 11, (1991) 2519
- 6) for a diagram see for example F.R. Libsch a. M.H. White, Solid-State Electronics 33 no.1, (1990) 105
- 7) A. Roy and M.H. White, IEEE Trans. Electron Devices 37, no. 16, (1990) 1504

ONO is symmetric, for an application where there is positive gate bias the top oxide has to be thicker than the bottom oxide (see fig.2c). Since the tunneling current is a strong function of thickness and field for a given bias the charge loss to one side always will be dominating, the one to the other side negligible.