Influence of the Poly Gate Depletion Effect on Programming EEPROM Cells

F. Schuler[a], P. Klein[b], K. Hoffmann[a], O. Kowarik[a]

[a]University of Bundeswehr Munich, ET4, D-85577 Neubiberg,
[b]Siemens AG, Balanstr. 73, D-81541 Munich

Abstract

To ensure a data retention time of e.g. 10 years a maximum allowed floating gate doping concentration can not be exceeded. The data retention time can be improved by reducing the floating gate doping concentration. This reduction causes a so far negligible depletion effect in the floating gate reducing the V_T-shift by approximately 1.5 Volts at a floating gate doping of $3 \cdot 10^{19}$ cm^{-3}. A trade-off will be discussed.

1. Introduction

It is known from literature (fig. 1) [1] that the data retention decreases drastically if the doping concentration of the floating gate (FG) N_{FG} is increased. Unfortunately a low doping concentration causes a non-negligible voltage drop across the FG. This effect is known by MOS transistors [2] and can be explained by the formation of a depletion region (DR) in the poly silicon. To our knowledge the poly gate depletion effect has not been considered so far by EEPROM models e.g. [3]. The voltage drop across the FG reduces the FN-programming efficiency and the V_T shift seriously.

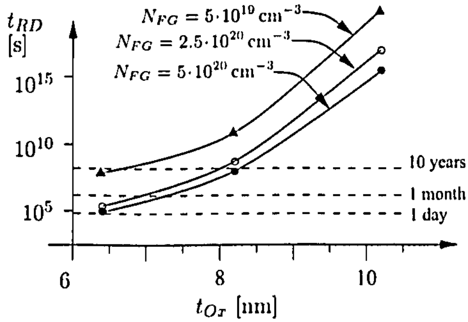

Figure 1: The read disturb time t_{RD} of EEPROM cells after 10^6 W/E stress cycles according to [1].

2. Measurement, Simulation and Discussion

Applying a positive voltage at the control gate (CG) depletion regions in the CG itself and in the FG will be created. The voltage drop across the depletion region depends on the doping levels. Because of the larger effective oxide thickness d_i^{eff} between the CG and the FG resulting from the ONO-structure the voltage drop in the CG can be neglected contrary to the one in the FG (fig. 2a).

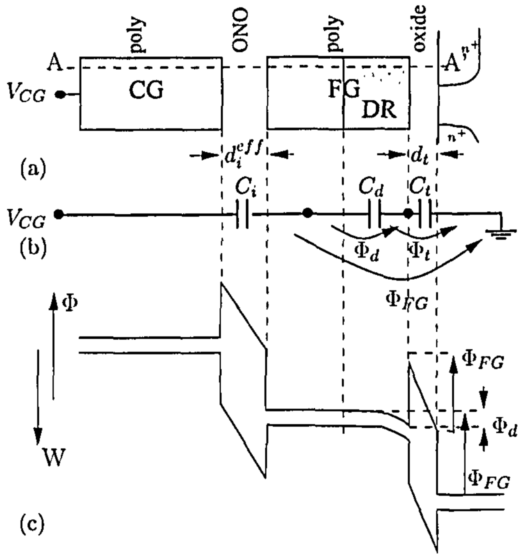

Figure 2: a) Cross-section of EEPROM, b) Equivalent circuit: cross-section A-A' c) Valence and conduction band

According to the subcircuit (fig. 2b) with the three connected capacitances a programming voltage of

$$\Phi_t = V_{CG}\frac{C_i'}{C_i' + C_t} \tag{1}$$

$$C_i' = \frac{C_i C_d}{C_i + C_d} \tag{2}$$

results.

The value of the depletion capacitance

$$C_d = A\sqrt{\frac{q\epsilon_0\epsilon_{poly}}{2\Phi_d}N_{FG}} \tag{3}$$

depends strongly on the doping concentration N_{FG} and the voltage Φ_d across it .

Measuring the capacitances of a test stucture with a contacted FG (fig. 3) the floating gate doping concentration N_{FG} near the interface can be determined. Evaluation of

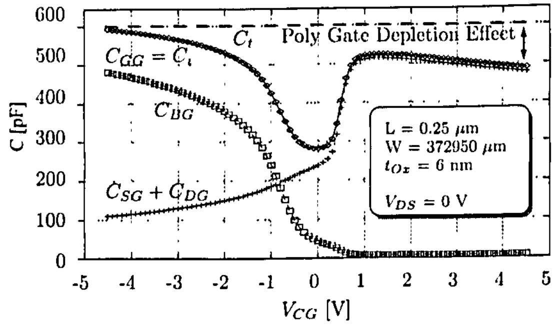

Figure 3: Capacitances behaviour of a quarter micron test structure demonstrating the influence of the poly gate depletion effect.

measured capacitances of quarter micron test structures results in an average poly gate doping concentration of $N_{FG} = 5 \cdot 10^{19}\text{cm}^{-3}$.

It is obviously that if the depletion capacitance C_d is not considered the calculated programming voltage will be larger than the expected one. Thus the real tunnel current

$$I_t = -B_1 \frac{\Phi_t}{d_t} \exp\left(-\frac{B_2}{\Phi_t / d_t}\right) \tag{4}$$

will be much smaller than the calculated one.

To further verify this effect device simulations (MEDICI) have been performed for an EEPROM cell with a tunnel oxide thickness of 6 nm and a variable doping concentration of the FG. The programming has been simulated by a CG voltage ramp of 0.2 ms, a maximal CG voltage of 12 V and a total programming time of 0.5 ms. The results are shown in fig. 4 and fig. 5: The lower the doping concentration the larger is the voltage drop Φ_d across the FG and the lower the resulting tunnel current.

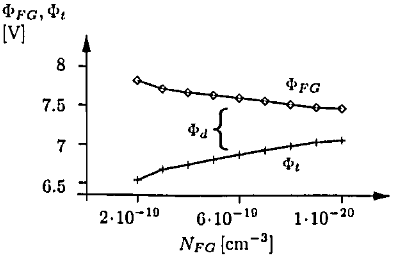

Figure 4: Programming voltage (Φ_t), floating gate voltage (Φ_{FG}) and voltage drop across depletion region (Φ_d) as functions of the poly gate doping concentration.

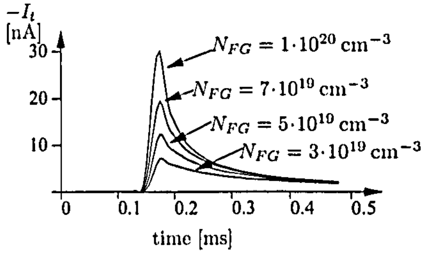

Figure 5: Tunnel current as a function of time for different poly gate doping concentrations.

According to the tunnel current performance the charge on the FG can be determined as a function of time and doping concentration as parameter (fig. 6). For a given ΔV_T of e.g. 1V the programming time as a function of doping concentration can be extracted (fig. 7). For a fixed programming time one can observe a linear dependence of the lost of the V_T-shift and the poly gate doping concentration (fig. 8).

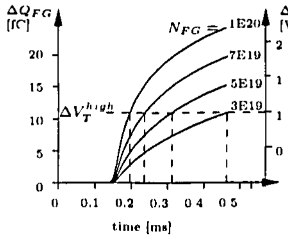

Figure 6: Charging of the FG and the resulting V_T-shift as a function of time and the poly gate doping concentration N_{FG} as parameter.

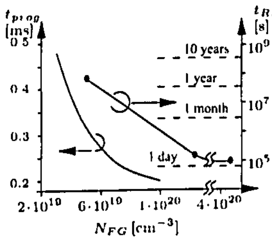

Figure 7: Programming time t_{prog} required to reach a V_T-shift of 1V and data disturb time t_{RD} according to [1] as functions of the poly gate doping N_{FG}.

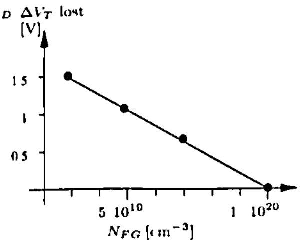

Figure 8: ΔV_T-lost for different poly gate doping concentrations and a programming time of 0.5 ms.

3. Conclusion

To ensure data retention of 10 years there is a maximum allowed poly gate doping concentration. It has been shown that the poly doping concentrations of typical quarter micron processes result in a ΔV_T-lost of more than $1V$ affected by the poly gate depletion effect. Therefore a compromise has to be choosen between required data retention and programming performance. This is particularly important if a high resolution as e.g. in multi value memory is required.

This work was sponsored by the DFG (Ho 1325/2-2).

References

[1] H. Watanabe et al., *Scaling of Tunnel Oxide Thickness for Flash EEPROMs Realizing Stress-Induced Leakage Current Reduction*, Symposium on VLSI Technology Digest of Technical Papers, p. 47, 1994

[2] N. D. Arora, R. Rios and C.-L. Huang, *Modeling the Polysilicon Depletion Effect and Its Impact on Submicrometer CMOS Circuit Performance*, Transactions on Electron Devices, vol. 42, No. 5, May 1995

[3] Lucy Chwan Liong and Po-ching Liu, *A Theoretical Model for the Current-Voltage Characteristics of a Floating-Gate EEPROM Cell*, Transactions on Electron Devices, vol. 40, no. 1, January 1993