Gate polysilicon optimization for deep-submicron MOSFETs

J. Schmitz, H. P. Tuinhout, A. H. Montree,
Y. V. Ponomarev, P. A. Stolk and P. H. Woerlee

Philips Research Laboratories, Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

Abstract

The use of amorphously deposited silicon and fine-grained polysilicon as MOS gate material is discussed. A variety of deposition and anneal conditions was evaluated on MOS capacitors and transistors. Gate depletion and MOSFET matching have been studied as a function of deposition condition and gate activation temperature. It is shown that polysilicon gate material has better properties than α-Si for CMOS generations beyond 0.18 μm.

Introduction

Downscaling of CMOS requires a decreasing gate oxide thickness, as well as a reduction of the thermal budget. As a result, the fabrication of a polysilicon gate with ideal capacitor plate behaviour becomes increasingly difficult. Polysilicon gates show depletion at the gate oxide interface. Gate depletion is more pronounced with thinner gate dielectrics and with lower gate doping levels. It degrades the saturation current of a transistor and may increase process spread.

In this paper, we consider the scalability of polysilicon as a gate material. It is shown that with a tailoring of the process conditions, low gate depletion can be achieved on both NMOS and PMOS transistors with ultra-thin gate oxide, thus extending the applicability of the conventional approach beyond 0.18 μm CMOS.

Polysilicon deposition

In a modern CMOS process with reduced thermal budget, we found that the gate is not uniformly doped at the gate oxide interface [1]. Instead, the dopants are predominantly present around the grain boundaries, due to the high dopant diffusivity along these boundaries. The situation is sketched in Figure 1.

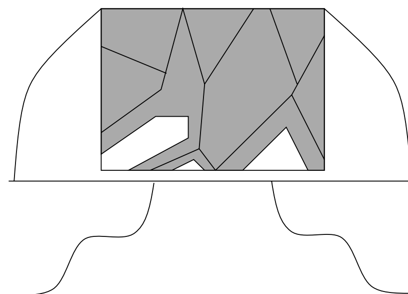

Figure 1: Inhomogeneously doped polysilicon gate.

As a result, the creation of the inversion layer does not start uniformly at one gate voltage, but some regions are inverted at a higher voltage. An effective solution to the problem of incomplete grain doping is to reduce the grain sizes in the polycrystalline gate [2]. We have achieved this using a standard LPCVD deposition of polysilicon (in a horizontal batch furnace) at a relatively low temperature. Figure 2 shows top-view SEM images of polysilicon layers after LPCVD deposition at various deposition temperatures.

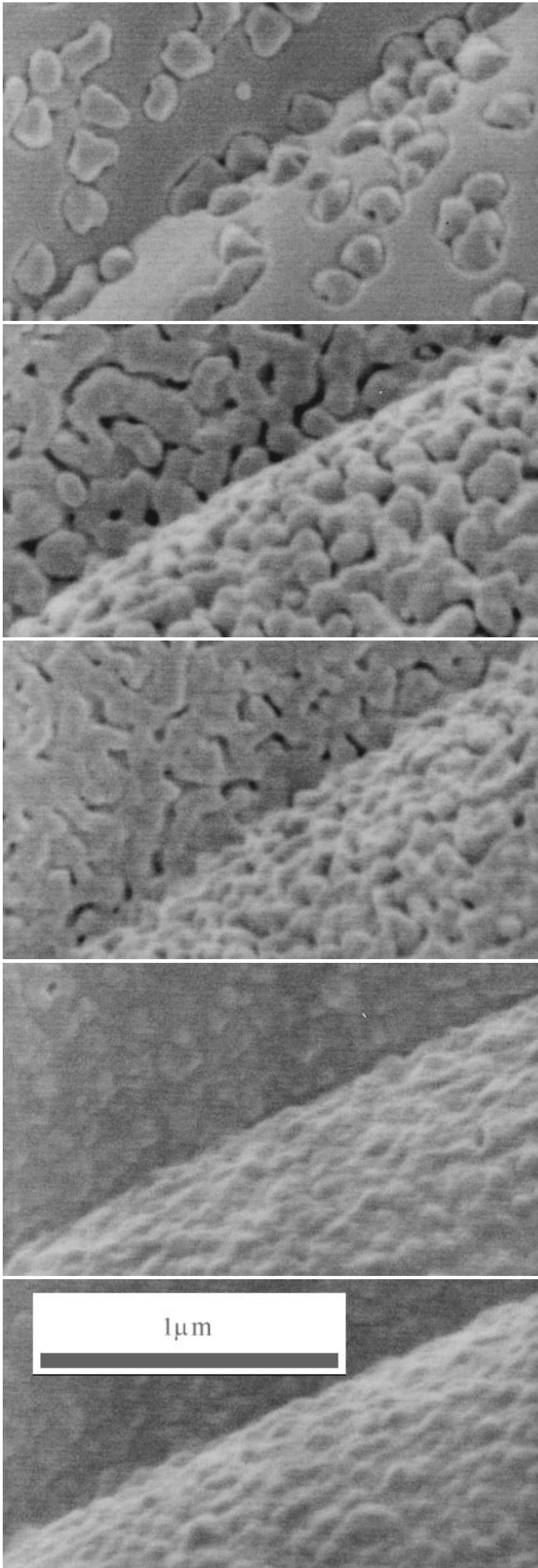

Figure 2: Top view SEM images of polysilicon deposited at 590°C, 595°C, 600°C, 605°C and 610°C.

Samples were made in the range of 580°C–640°C with 5°C steps. The polysilicon thickness in all these depositions is 150 nm; the layers were deposited on wafers covered with 350 nm thermal oxide. The average grain sizes in the polycrystalline varieties are less than 100 nm in all cases; for the recrystallised amorphous silicon, the average grain size was found to be around 150 nm, while there are occurrences of grains as large as 500 nm.

There is a clear transition from amorphous to polycrystalline occurring in the 580–610°C regime, as visible in Figure 2 and confirmed with XRD. The transition phase starts with the occurrence of crystals in an amorphous bulk; then, so-called hemispherical grains (HSG) are formed at 600°C; and at a slightly higher temperature, fine-grained polysilicon is formed, with a columnar structure (as observed with TEM). At temperatures too close to the HSG phase, voids occur. These are seen in the SEM-pictures as dark spots.

Above 600°C, the grain size slowly increases with increasing temperature - it also depends on the total layer thickness. The roughness of these layers is more or less constant. This is quantified using Atomic Force Microscopy and depicted in Figure 3. The polysilicon deposited at 610°C has the smallest average grain size among the samples without voids. Figure 2 illustrates that a temperature control significantly better than ±5°C is required to reproduce this material.

Device fabrication

NMOS and PMOS capacitors and transistors were fabricated with 150 nm α-Si and 610°C-poly-Si gate material. The gate oxide thickness was 3.1 nm (determined with ellipsometry); the doping dosages of these gates were 2.5×10^{15} (boron) and 4×10^{15}

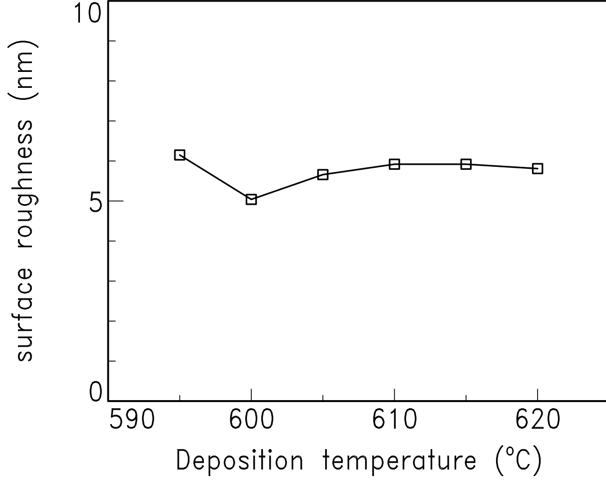

Figure 3: Surface roughness (RMS) of 150 nm fine-grained polysilicon layers, measured with AFM.

(arsenic). We restricted the experiments to self-aligned gate/source/drain doping; no gate predoping was used.

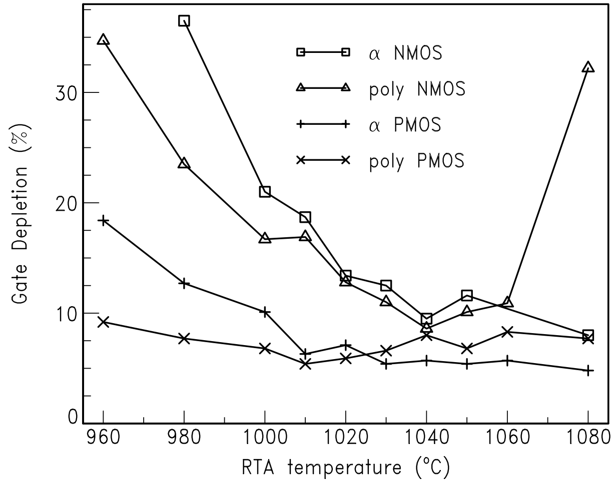

Figure 4: Gate depletion for NMOS and PMOS capacitors fabricated with α-Si or fine-grained poly-Si, and with various gate activation temperatures (20 seconds anneal).

Gate depletion data are shown in Figure 4. The gate depletion G is defined as

$$G = \frac{C_{acc} - C_{inv}}{C_{acc}} \times 100\%, \qquad (1)$$

where C_{acc} and C_{inv} are measured at $\pm$2V. The figure indicates that at a fixed thermal budget between 960°C and 1050°C, the gate activation is always better in the fine-grained polysilicon gates. This translates into a lower minimum thermal budget requirement for gate activation.

Interestingly, the gate depletion increases again at very high activation temperatures; particularly in the fine-grained NMOS gate. This is interpreted as a result of arsenic segregation towards Si-boundaries, or even arsenic evaporation from the top of the gate to the ambient.

In the studied temperature range, we did not observe boron penetration at all, with the α-Si nor the fine-grained polysilicon samples. This is illustrated in Figure 5. The slow shift in threshold voltage in this wide temperature range is a combined effect of the channel dopant diffusion, gate workfunction change and gate depletion change.

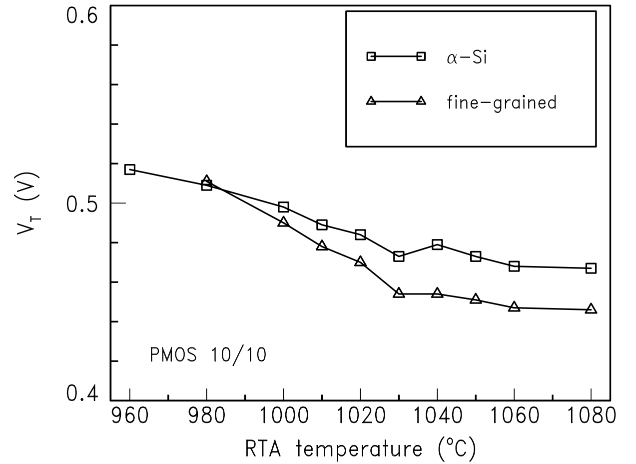

Figure 5: Threshold voltage of PMOS transistors as a function of gate activation temperature (20 seconds anneal). There is no indication of boron penetration.

MOSFET matching

Statistical differences between closely spaced, supposedly identical transistors have been characterized. This so-called matching measurement is a means to identify small differences in transistors on a microscopic scale. In earlier work [1], we have shown that local gate depletion as

sketched in Figure 1 results in increased mismatches between transistors. We found that particularly the V_T mismatch standard deviation for large transistors quantifies the gate material quality effectively.

A striking difference between α-Si and fine-grained polysilicon is seen in the NMOS transistor mismatch (see Figure 6). The transistor matching is very good for the fine-grained polysilicon gates; it is close to target already at a gate activation temperature of 1000°C! This is unexpectedly good in view of the large gate depletion of these transistors (see Figure 4).

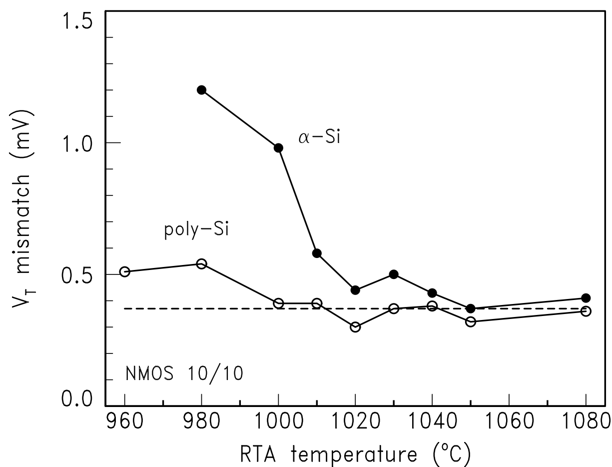

Figure 6: NMOS matching behaviour as a function of gate activation temperature (20 seconds anneal), for α-Si and poly-Si gates. The dotted line indicates the target matching value for this process.

No major matching differences were observed between the two gate materials on PMOS transistors. This is shown in Figure 7. Only at very high temperatures, the α-Si variety appears slightly better. The transistor matching is close to the theoretical limit (determined by statistical variations in the number of dopant atoms in the channel), both for the NMOS and for the PMOS transistors. Such good matching performance allows compact design of analog circuitry.

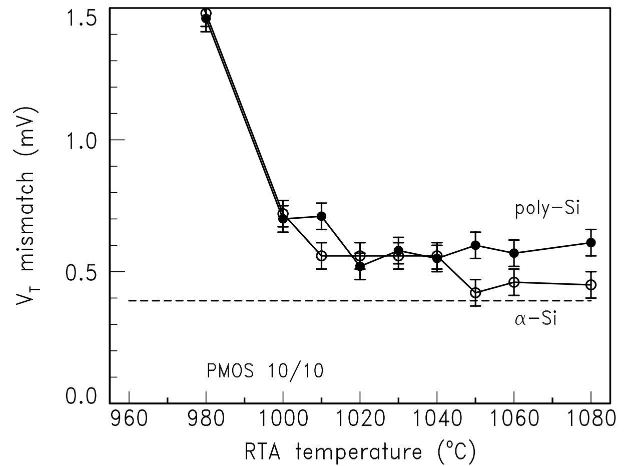

Figure 7: PMOS matching behaviour, as in Figure 6.

Conclusions

The use of fine-grained polysilicon with sub-100 nm grain size is shown to result in improved gate activation and improved MOSFET matching performance at low thermal budget. The material is shown to form better capacitors and transistors in both flavours. These results can be understood as a result of a better dopant redistribution in the gate due to the improved density of grain boundaries.

Using this fine-grained polysilicon, a reduction of the thermal budget is feasible which allows formation of shallower junctions and steeper retrograde wells; both being very effective in the improvement of the short channel behaviour of the transistors.

We would like to thank J. B. F. aan de Stegge for his deposition work, and W. B. de Boer for fruitful discussions. This work was partially supported by the EC within the ULTRA project (Esprit E23806).

References

[1] H. P. Tuinhout et al., IEDM ’97 p.631.
[2] S. Shimizu et al., VLSI’97 p.107.