Spike annealing of boron-implanted polycrystalline-silicon on thin SiO₂

A. T. Fiory a)  
Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974

K. K. Bourdelle and P. K. Roy  
Bell Laboratories, Lucent Technologies, Orlando, Florida 32819

(Received 29 August 2000; accepted for publication 14 December 2000)

Spike thermal annealing is examined for electrical activation of B implants into 100 nm Si films deposited over 1.5 to 2.4 nm thermally grown SiO₂. These structures simulate gate stacks in advanced p-type metal–oxide–Si (PMOS) devices. Spike anneals, at minimized thermal budget, are shown to yield higher carrier concentrations in PMOS polycrystalline-silicon (poly-Si), as compared to conventional rapid thermal annealing. The activation energy for B diffusion through SiO₂ is found to be 3.71 to 3.83 eV and near that previously reported for furnace anneals. Boron penetration appears unaffected by photoexcitation from heating lamps. © 2001 American Institute of Physics. [DOI: 10.1063/1.1348307]

Boron penetration through the gate oxide (GOX) during activation annealing of B implanted polycrystalline-silicon (poly-silicon) gates 1–3 is a major concern in controlling threshold voltage and reliability 4 of high performance p-type metal–oxide–Si (PMOS) devices. High carrier concentration in polycrystalline-silicon at the GOX interface, the desired process goal, utilizes the rapid grain boundary assisted diffusion of B from the implanted region. 1 Boron diffusivity through the GOX, which is sought to be minimized, has been reported to become faster with decreasing SiO₂ film thickness. 5 Moreover, it was noted that rapid thermal annealing (RTA) leads to a higher B diffusivity in the oxide, when compared to furnace anneals. 6 The purpose of the present work is to study boron penetration through thin SiO₂ films and the electrical activation of implanted B by varying RTA temperature and cycle time as well as conditions of lamp irradiation. Short cycle annealing at high temperature, i.e., “spike annealing,” is used to reduce B penetration through the GOX while also increasing electrically active B concentration in the polycrystalline-silicon.

Gate quality SiO₂ films of thickness dOX of 1.5, 2.0, or 2.4 nm (measured by ellipsometry) were grown in a vertical furnace on 200 nm p/p⁺ epitaxial Si wafers. 100 nm amorphous Si films were deposited on SiO₂ films by chemical vapor deposition and implanted with 11B⁺ at 5 keV and 3 × 10¹⁵ cm⁻² dose. Square (12×12 mm²) samples were cleaved from the wafers, placed over an 11.5 mm aperture in a 200 mm epitaxial holder wafer and annealed in N₂ with 0.1%O₂ admixture to suppress B out diffusion. The wafer was heated from both sides with incandescent lamps in an AG Heatpulse RTA system. Samples received a 10 min final anneal at 425 °C in 10%H₂ forming gas. Results are derived from measurements of metal–oxide–semiconductor (MOS) capacitance–voltage and Hall van der Pauw transport.

The heavily doped Si substrate absorbs lamp radiation that is incident on the back sides of the wafer and sample. In several anneals, a second epitaxial cover wafer, supported 2 nm above the holder wafer, was used to shield the top of the implanted sample from lamp irradiation. Photoexcitation of the implanted polycrystalline-silicon is thus prevented directly by the cover and indirectly by rapid carrier recombination in the heavily doped substrate. Covered and uncovered samples annealed with similar temperature versus time profiles were found to produce equivalent results for carrier activation in the polycrystalline-silicon and B penetration through the oxide.

After annealing, MOS capacitor and van der Pauw mesa patterns were fabricated by masking the polycrystalline-silicon and wet-chemical etching. Surface oxide was removed with hydrogen–fluoride and electrical connections were made with In contact pads. MOS capacitor C–V was measured from the displacement current under triangular wave form gate-voltage sweeps. Carrier concentration in the polycrystalline-silicon at the oxide interface, pPOLY, was determined by comparing the accumulation capacitance under negative bias with a quantum C–V calculation. 7 The flat band gate voltage, VFB, was determined by Berglund’s method, i.e., fitting the Si surface potential from flat band wave form gate-voltage sweeps. Carrier concentration in crystal Si near the interfaces with the GOX, pSI, and mid-gap interface trap density ( ≤ 2 × 10¹⁰ cm⁻² eV⁻¹).

VFB = WPS + COX⁻¹(qNB – QEFF),

where COX is the oxide capacitance per unit area and q is the elementary charge. To a first approximation, WPS = q⁻¹kBT ln(pPOLY/pSI).

For pPOLY below about 10¹⁹ cm⁻³ Fermi level pinning at grain boundaries 9 depresses WPS, while for degenerate dop-

---

a)Electronic mail: atf@lucent.com
ing, above about $2 \times 10^{20} \text{cm}^{-3}$, the Fermi level in the polycrystalline-silicon is pinned near the intrinsic valence band edge.\(^{10}\)

The dependence of $V_{FB}$ on temperature for spike anneals is shown in Fig. 1 for the three $d_{OX}$. The anneals used a maximum 120°C/s heating rate, lamp turn off at approximately 20°C below peak temperature, and maximum radiative cooling at 80°C/s. The dashed curve is $W_{PS}$ from Eq. (2). Three temperature regions are seen in Fig. 1. At low temperatures, where the carrier concentration is below about $3 \times 10^{19} \text{cm}^{-3}$, grain boundary Fermi level pinning reduces $V_{FB}$ below the dashed curve. The onset of B penetration causes $V_{FB}$ to rise above the dashed curve at high temperature. Optimal PMOS device processing lies in the inflection region, where $V_{FB}$ is close $W_{PS}$, at temperature between 1050 and 1100°C.

The region of the onset of B penetration was studied by varying temperature and time of annealing of 2.4 nm oxide. The highest values are obtained by the spike method. Hall van der Pauw measurements of the polycrystalline-silicon show monotonic increases with temperature for average conductivity, from 418 to 739 $\Omega^{-1}\text{cm}^{-1}$, and average carrier density (Hall factor $r = 0.72$), from $8.9 \times 10^{19}$ to $1.6 \times 10^{20} \text{cm}^{-3}$. The highest values are obtained by the spike method. The average carrier mobility in the polycrystalline-silicon is about $29 \text{cm}^2\text{V}^{-1}\text{s}^{-1}$ and varies less than ±2% among the anneals. Mobility is lower than for uniformly doped crystal Si, 47 to 50 $\text{cm}^2\text{V}^{-1}\text{s}^{-1}$, owing to grain boundary defects.\(^{9}\) Figure 2 shows the temperature dependence of the carrier density in the polycrystalline-silicon at the oxide interface extracted from the $C-V$ analysis, with the effective annealing times along the right hand scale. The results show that B transport to the oxide interface is sufficiently rapid to allow maximum activation via the spike method.

An activation energy for the B penetration through the GOX, $E_A$, was deduced from the results for spike anneals and for anneals with extended dwell times at peak temperature. The latter used a heating rate of 50°C/s and soak times of 1, 5, 20, 80, and 200 s. A set of annealing temperatures, $T$, was selected for each oxide thickness by annealing samples from the same wafer to provide $V_{FB} \approx 0.5 \text{V}$. At this value of $V_{FB}$ the B penetration term in Eq. (1) is comparable to $W_{PS}$.

We use a two interface model,\(^{11}\) where the B concentration in the polycrystalline-silicon, $C_{POLY}$, is assumed constant and the interface distribution coefficient, $m_{Si}$, assumed the same for poly and crystal Si, to obtain B penetration dependence on time $t$:

$$N_B = 2C_{POLY}L_{OX} \left[ m_{Si} + (D_{OX}/D_{Si})^{1/2} \right]^{-1} \text{ierrorc}(y),$$

where $D_{OX}$ and $D_{Si}$ are time independent B diffusivities in the oxide and crystal Si, respectively, $L_{OX} = (4D_{OX})^{1/2}$ is a diffusion length in the oxide, $y = d_{OX}/L_{OX}$, and ierfc is the integral of the complementary error function. (Higher order terms in the general expression\(^{11}\) are less than 0.1% for $y > 0.8$). In Eq. (3) we use $C_{POLY} = p_{POLY}$, assuming that the B diffusion source is dominated by electrically active B in the adjacent polycrystalline-silicon grains. The value for $m_{Si}$ was set at 0.55.\(^{12}\)

Using Eqs. (1)–(3) an effective $D_{OX}$ was computed for each anneal (contribution from fixed charge term was assumed to be negligible). Heating and cooling transients are included in the definition of time, $t = t_0 + \Sigma_i \exp \left[ k_B^{-1}E_A(T_i^{-1} - T_0^{-1}) \right]$, where $T_i$ are the recorded temperatures, $t_0$ is the sampling time interval, and activation energy, $E_A = 3.8 \text{eV}$, selected to be self-consistent with the findings for the effective $D_{OX}$. Times computed for the spike anneals are 1.3 to 1.4 s; times for the soak anneals are 1.8 to 2.0 s longer than nominal dwell times at constant maximum temperature.

The temperature dependence of the effective $D_{OX}$ is shown on an Arrhenius scale in Fig. 3. While the effective diffusivity appears to increase with decreasing film thickness, the fitted slopes are nearly parallel, implying similar activation energies: the fits yield $E_A = 3.83 \pm 0.03$, 3.79 ± 0.02, and 3.71 ± 0.02 eV for $d_{OX} = 1.5, 2.0$, and 2.4 nm, respectively. The result for an optically shielded 5 s 1090°C anneal for $d_{OX} = 2.4 \text{nm}$ is shown by the star symbol, which falls on the line for the optically exposed samples. The 2.4 nm oxide was selected for comparison because it contains the largest fraction of low defect, bulk-like SiO$_2$. The 5 s anneal time allows comparison at high temperature with relatively small systematic differences in temperature transients (the covered wafer cooling more slowly).
B diffusion in SiO\(_2\) is a controversial subject with much scatter among published data. B diffusivity was reported to depend not only on GOX thickness but also on B concentration in polycrystalline-silicon and SiO\(_2\) and upon time.\(^1\)\(^{12}\) It is worth mentioning that the condition of fixed \(V_{\text{FB}}\) adopted in this work causes the diffusion length \(L_{\text{OX}}\) to be same to within \(\pm 10\%\) for all anneals. This analysis for the activation energy is possibly less influenced by systematic variations arising from uncertainties in \(C_{\text{POLY}}, d_{\text{OX}},\) and \(D_{\text{OX}}\).

While the increase in \(D_{\text{OX}}\) with decreasing \(d_{\text{OX}}\) in Fig. 3 is in qualitative agreement with data in Ref. 5, the present results do not support an associated decrease in \(E_A\) as proposed therein. The data for \(D_{\text{OX}}\) are close to previous results obtained in MOS structures after rapid thermal anneals.\(^4\)\(^6\) Diffusivities from RTA are 1–2 orders of magnitude larger than the values obtained for furnace anneals in MOS structures\(^12\)\(^{13}\) and in bulk silica.\(^14\) A possible explanation is the time dependence observed for B diffusivity in SiO\(_2\).\(^12\)

Our data on activation energy are significantly larger than \(E_A = 3.07 \text{ eV}\) previously deduced for RTA processing\(^6\) and are rather in line with \(E_A = 3.82 \text{ eV}\) for furnace anneal.\(^13\) The disagreement might be explained by the limited amount of RTA data available for analysis in Ref. 6.

In conclusion, the spike annealing method is shown to be consistent with the goal of managing minimum B penetration through SiO\(_2\) and maximum carrier activation in polycrystalline-silicon used in advanced PMOS gate structures. Boron diffusion through thin SiO\(_2\) films depends on the oxide thickness and is characterized by an activation energy in the range from \(3.71\) to \(3.83 \text{ eV}\). Shielding the lamp radiation has negligible influence on either the B penetration or the electrical activation of the polycrystalline-silicon.

The authors acknowledge stimulating discussions with H.-H. Vuong, K. A. Ellis, G. L. Timp, and C. S. Rafferty.