SAND99-0531J

A nonvolatile MOSFET memory device based on mobile protons in SiO_2 thin films[§]

K. Vanheusden[a]*, W.L. Warren[b], R.A.B. Devine[c], D.M. Fleetwood[b], B.L. Draper[b], J.R. Schwank[b]

[a]US Air Force Research Laboratory, 3550 Aberdeen Ave., SE, Kirtland AFB, NM 87117-5776 USA

[b]Sandia National Laboratories, Albuquerque NM 87185, USA

[c]France Télécom-CNET, B.P. 98, 38243 Meylan, France

RECEIVED
MAR 15 1999
OSTI

Abstract

It is shown how mobile H^+ ions can be generated thermally inside the oxide layer of $Si/SiO_2/Si$ structures. The technique involves only standard silicon processing steps: the nonvolatile field effect transistor (NVFET) is based on a standard MOSFET with thermally grown SiO_2 capped with a poly-silicon layer. The capped thermal oxide receives an anneal at ~1100 °C that enables the incorporation of the mobile protons into the gate oxide. The introduction of the protons is achieved by a subsequent 500-800 °C anneal in a hydrogen-containing ambient, such as forming gas (N_2 : H_2 95:5). The mobile protons are stable and entrapped inside the oxide layer, and unlike alkali ions, their space-charge distribution can be controlled and rapidly rearranged at room temperature by an applied electric field. Using this principle, a standard MOS transistor can be converted into a nonvolatile memory transistor that can be switched between "normally on" and "normally off". Switching speed, retention, endurance, and radiation tolerance data are presented showing that this non-volatile memory technology can be competitive with existing Si-based non-volatile memory technologies such as the floating gate technologies (e.g. Flash memory).

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

1. Introduction

In microelectronics technology, annealing in a H_2 containing atmosphere at temperatures < 450 °C is commonly used to passivate electrically active interface traps [1]. However, the role of hydrogen in the Si/SiO_2 system is far more complex as indicated by the vast research effort in this field. The incorporation of atomic or molecular hydrogen into the technologically relevant Si/SiO_2 system is almost unavoidable during device processing. Hydrogen is present at the Si/SiO_2 interface, terminating dangling Si bonds (P_b centers [2]), and in the SiO_2 film, either bonded to the network (Si-OH) or in molecular form at interstitial sites (H_2, H_2O).

Recently, we have begun to explore the effects of annealing at temperatures > 450 °C, and have found that hydrogen can generate both fixed and mobile positive charges in thin film amorphous SiO_2. The fixed positive species is remarkably similar to oxidation induced fixed charge [3] and have recently been attributed to near-interfacial over-coordinated oxygen sites induced by interaction with hydrogen [4]. The mobile species, which will be discussed in this work, have only been observed to date in $Si/SiO_2/Si$ structures and are identified as mobile protons imprisoned inside the oxide of these layered structures [5]. These protons are rearrangable at room temperature by applying a bias and are immobile in the absence of a bias, which makes this phenomenon potentially useful as the basis for a new generation of nonvolatile memory devices [5].

Nonvolatile memory technologies are expanding in response to the fast growing demand for low-power applications such as portable consumer electronics. This memory market is currently dominated by floating gate technologies, using hot-carrier injection or tunneling to store charge. We propose a device based on mobile protons located in the gate dielectric and show that it may offer a promising and inexpensive alternative to existing nonvolatile memory technologies. The device is based on a

stacked Si/SiO_2/Si structure, subjected to an anneal at ~ 1100 °C in an inert ambient. Mobile protons (Q_H+) can be generated in the dielectric of such structures by subjecting them to a forming-gas anneal at temperatures ≥ 400 °C. The protons are found to be stable, even under field and thermal stressing and irradiation, i.e., they do not easily escape the dielectric or trap electrons [6]. These observations have been confirmed by other workers [7].

There are several reports in the literature of mobile protons in SiO_2, which have been generated electrically or by irradiation in specific types of SiO_2 [8,9]. However, the reaction kinetics that generate the mobile protons during annealing in the materials described above are still largely unexplored. To obtain more information about the atomic reactions that govern the generation of the observed mobile protons in the SiO_2, we have performed a study of the chemical kinetics by looking at different forming-gas anneal conditions and anneal sequences. Based on our observations we suggest a model of dynamic equilibrium, which assumes a reversible interfacial reaction with a temperature dependent reaction balance.

2. Experimental Procedures

2.1 Samples

Although the stacked single-crystalline Si / SiO_2 / poly-crystalline Si (*c*-Si/SiO_2/*poly*-Si) structure is the preferred material for commercial application in a memory device, we also examined *c*-Si/SiO_2/*c*-Si structures (also called silicon-on-insulator or SOI) to gain insights into the mechanisms involved. The *c*-Si/SiO_2/*poly*-Si samples were formed by using a 11-nm dry thermal SiO_2 (p-type substrate) grown at Texas Instruments, and a 40-nm dry thermal SiO_2 (p-type substrate) grown at Sandia's Microelectronics Development Laboratory. These thermal oxides were then capped with a poly-crystalline Si layer followed by annealing at 1100 °C in Ar + 1% O_2 for 30 min. In the class of

SOI materials separation by implantation of oxygen (SIMOX) material was used. It is formed by implanting a *p*-type Si(100) substrate with 190-keV O^+ ions to a dose of 1.8×10^{18} cm^{-2} followed by a subsequent anneal for 6 hours at 1325 °C in Ar + 1% O_2. This process resulted in a 200-nm mono-crystalline Si layer on top of a 400-nm buried amorphous SiO_2 layer. Another set of SOI samples was made by first growing a 1-μm thermal oxide, then covering this with a 330-nm thick layer of polycrystalline Si, which was subsequently zone-melt recrystallized (ZMR). Unibond® material was also investigated [10]. This SOI material is formed by implanting hydrogen (~ 6×10^{17} cm^{-2}) into a wafer, below a thermally grown SiO_2 layer (300 nm thick), followed by bonding this wafer to another wafer. This bonding involves a low temperature anneal at 400 to 600 °C to split the first wafer at the boundary defined by the implant, followed by a high temperature anneal at 1100 °C to strengthen the bonding interface. Note that one common feature of all the material used in this study is that they all were subjected to a high-temperature treatment in the range from 1100 to1325 °C. It turns out that this anneal is crucial on these types of thermal oxides to the observation of generation of mobile protons during the hydrogen anneals performed later at much lower temperatures, as discussed below.

Square or rectangular capacitor structures (c-Si/SiO_2/c-Si or c-Si/SiO_2/poly-Si, areas 0.16-10.30 mm^2) were obtained by wet etching the top Si film in a HNO_3-CH_3COOH-HF mixture using a mask, leaving behind small isolated islands of the top Si layer. The size of the capacitors facilitates lateral diffusion of hydrogen into the buried oxide during the forming gas anneal. Lateral diffusion of hydrogen into the buried oxide is crucial because the amount of reactive hydrogen reaching the buried oxide by diffusion through the top *c*-Si layer appears to be negligible; this effect may be due to the low solubility of hydrogen in Si [11]. Mobile protons (deuterons) are generated in the gate oxide of these capacitor structures by annealing in a flow of forming gas [Ar:H_2(D_2); 95:5 (by volume, 99.999% pure)]. The use of a quartz tube furnace with a calibrated (trapezoidal) temperature profile allowed the

samples to be cooled at a given linear rate by pulling the sample out of the furnace at a constant speed. Very short anneal times (~10 s) could be achieved by rapid insertion and fast pulling of the samples from the preheated furnace. For the radiation study, devices containing mobile protons were subsequently irradiated using a 10-keV x-ray source at a dose rate of 4 krad(SiO_2)/s. Samples were irradiated without applied bias (floating gate) at substrate temperatures between 77 and 500 K.

2.2 Analysis Techniques

The net charge within the oxide layers was derived from capacitance midgap shifts at room temperature. To determine the midgap voltage shifts, 1 MHz capacitance-voltage (C-V) measurements were performed. Because the protons are mobile at room temperature and move in the presence of an applied electric field, a reversible midgap voltage shift (hysteresis) is observed which can be directly linked to the density of mobile protons, Q_H+, in the gate oxide [5]. In the case of SOI, the top-Si midgap voltage shifts were also obtained from current-voltage (I-V) curves measured using the point-contact transistor method or pseudo-MOSFET (Ψ-MOSFET) technique [12] on the same capacitor structures. This simple test "device" is based on the upside-down MOS structure, taking advantage of the SOI configuration. The buried oxide plays the role of the gate oxide and the top Si layer represents the transistor body. Two metal probe tips are placed on the top Si layer to form the source and drain point contacts (V_D = 0.2 V), while the gate voltage (V_G) is applied to the back of the Si substrate.

3. Results

3.1 Detection and Identification of Mobile Charge

Figure 1 (a) shows the hysteresis in the Ψ-MOSFET *I-V* curves on SIMOX after it received a 550 °C forming-gas anneal. Similar features were observed upon annealing at different temperatures in the range from 500 to 800 °C in forming-gas. Curve 1 was recorded with a *decreasing* gate (substrate) bias (from positive to negative) after the bias was kept constant at the initial + 40 V for 5 min. Curve 2 was subsequently recorded using the *opposite* gate voltage sweep direction (from negative to positive) after holding the gate bias at the initial - 70 V for 5 min. The hysteresis is apparent as the original *I-V* curve (curve 1) was not retraced by reversing the voltage sweep direction (curve 2). Curve 3 was recorded after curve 2 using the same procedure described for curve 1, showing the reversibility of the process. The reversible voltage shift, ΔV, in the *I-V* plots is caused by electric field induced migration of mobile positive charge from one Si/ SiO_2 interface to the other (areal number density ~ 2×10^{12} cm^{-2}). these properties were not observed after annealing in Ar, showing that the presence of hydrogen in the anneal ambient triggers the reversible shifts.

Figure 1 (b) shows the hysteresis of the *C-V* curves on SIMOX buried oxide after it received a 550 °C anneal. Following the forming gas anneal the top Si was removed in KOH solution, and a metal gate was formed on the exposed oxide. The three curves were measured using the same sequences as in Fig. 1 (a). Again, the negative voltage shift, ΔV, in the *C-V* plot is caused by mobile positive charges in the buried SiO_2. However, since the semiconductor body is now the substrate Si (as illustrated in the insets), ΔV will now be maximum if the positive charges are located near the SiO_2/substrate-Si interface, i.e., the opposite interface as compared to the *I-V* measurements. Curve 1 reveals the presence of positive charge near the substrate-Si/SiO_2 interface. The overlay of curves 2 and 3 shows that, once the charged species have been drifted towards the metal-gate/SiO_2 interface

under the negative gate bias, reversing the gate bias polarity in the metal-gate/SiO_2/Si substrate capacitor does not induce a reversible shift, but leads to a permanent change.

The next set of experiments is designed to provide more information about the nature of the mobile species by studying the isotope mass effect on charge motion. In Fig. 2 we show the results of an experiment using the Ψ-MOSFET structure in which a ZMR sample was annealed in H_2 at 600 °C. A positive gate (substrate) bias (+60 V) was applied to 'drive' any mobile positive charges to the SiO_2/top–Si interface where they were sufficient to induce an inversion layer in the Si and thus give rise to a source-drain current (I_D) of 45 μA. The gate bias was then switched to –10 V so that the mobile charges were drawn back towards the gate and the variation of I_D with time monitored. In Fig. 2 we plot the initial rate of decay of the normalized drain current ($\propto \partial I_D / \partial t$) versus inverse temperature obtained from a series of switched bias measurements as described above, carried out on Ψ-MOSFETs on samples annealed in gas containing either H_2 or D_2. As can be seen in Fig. 2, the normalized decay decreases exponentially with decreasing temperature. The slope of both curves is the same but a small offset is apparent.

3.2 Performance of Prototype Nonvolatile Memory Devices

It is easy to see how the reversible flat band voltage shift observed in this work could be utilized in a nonvolatile FET (NVFET) memory device. An n-channel transistor can be changed to 'normally on' or 'normally off' by applying a positive or negative gate (substrate) bias which will drift the protons to the top-Si/SiO_2 or substrate-Si/SiO_2 interface. For a memory device this can be interpreted as writing the device to a bit state '1' or '0', respectively.

For memory devices, a short write time is desirable. The write time for the NVFET device is the time it takes for the protons to move from one interface to the opposite interface. The ionic transport

through a SiO_2 film is a complex function of the ionic species involved, time, temperature, gate voltage, and oxide layer thickness. A detailed analysis of this problem is beyond the scope of this work. As a crude approximation of the write time, we will use the inverse of the initial rate of decay of the normalized drain current $[(\partial I_D/\partial t)/I_0]^{-1}$. Figure 3 (a) and (b) plot the initial rate of decay vs. the applied oxide field for a fixed oxide thickness, and vs. the thickness of the buried oxide layer for a fixed gate voltage, respectively. It follows from Fig. 3 that the speed of a memory device will be approximately proportional to V_G^2 and d^{-3}. For test devices fabricated using a poly-Si-capped 40-nm thermal oxide substrate, a write time of about 80 ms was attained (write/erase voltage $V_G = \pm 4$ V). Write times as fast as 1 ms can be extrapolated from the fitted curve in Fig. 3 (b) for 10-nm thermal oxides.

The most important requirement for a memory device is that it retains the stored information over a specified time and temperature range. Retention tests were performed by drifting the protons to the top-Si/SiO_2 or substrate-Si/SiO_2 interface (corresponding to the two possible memory states: high and low zero-bias drain current), followed by heating the device to 200 °C at floating gate for extended times. The applied drain voltage was 0.2 V which is too low to disturb our experiments. The results of this experiment are shown in Fig. 4. No instabilities in the initial (programmed) drain current could be observed for times up to 25 h. Room temperature experiments show retention times of at least one order of magnitude larger. Furthermore, fatigue experiments performed on a protonated c-Si/12-nm thermal SiO_2/poly-Si capacitor structure demonstrate that these devices easily endure over 10^6 write-erase cycles without any degradation. This invariance means that the loss of mobile protons over time is negligible. Fatigue resistance over 10^6 cycles is greater than the specifications of most commercial nonvolatile memories.

Radiation tolerance or hardness is another quality of non-volatile memory technologies that would be of considerable interest for certain applications. The data in Fig. 5 show the results of a cumulative

radiation dose experiment for poly-Si capped thermal oxide capacitors. The samples were first annealed in forming gas at 600 °C to generate mobile protons inside the thermal oxide layer, and subsequently irradiated at various temperatures under floating gate conditions. It can be seen that the density of mobile protons drops from ~ 9×10^{11} cm^{-2} to ~ $4 - 7 \times 10^{11}$ cm^{-2} after a dose of 10 Mrad(SiO_2), depending on the substrate temperature. Comparing our experimental data with trapped-hole annihilation data, obtained under similar experimental conditions, shows that the effective rate of annihilation is about 200 times slower for mobile protons as compared to trapped holes. By comparing the effective cross section to that obtained for electron trapping by trapped holes [13], an effective capture cross section $\sigma_{H+} \cong \sim 2 \times 10^{-15}$ cm^2 can be obtained for radiation induced electron trapping by mobile protons in SiO_2.

3.3 Diffusion and Reaction Kinetics

By comparing the rate of mobile proton buildup during forming-gas annealing for 0.16 mm^2 vs. 10.30 mm^2 samples, we found that the buildup is much faster in the smaller devices. For a 0.16 mm^2 device annealed at 600 °C, Q_H+ reaches about 60 % of its equilibrium state value after 250 s.

Figure 6 shows how Q_H+ varies as a function of the cool-down conditions after performing a 600 °C forming gas anneal for 20 minutes. The anneal time was chosen to be sufficiently long to allow Q_H+ to reach equilibrium at 600 °C. The inset visualizes a few typical thermal cycles. The general conclusion from Fig. 6 is that the largest Q_H+ is obtained by quenching the sample from 600 °C to room temperature (fast pull). Cooling more slowly yields fewer mobile protons in the buried oxide. It is interesting to note that this effect is stronger in the 11-nm thermal oxide capped with poly-Si as compared to the SOI buried oxide.

To further explore the pull rate effect on Q_H+, samples were first annealed in forming gas at 600 °C for 30 min and then fast pulled to maximize Q_H+. Following this treatment the samples were then annealed again at a lower temperature in the same ambient for 30 min, and then again quenched to room temperature. The results of this experiment are summarized in Fig. 7, showing that the second anneal at a lower temperature reduces the initial Q_H+, but only for temperatures ≥ 400 °C. In fact, both the pull rate and the double anneal experiments demonstrate that the temperature of the final anneal affects the Q_H+.

4. Discussion

4.1 Identification and Transport Kinetics of Mobile Protons in SiO_2 films

The difference in properties between the Ψ-MOSFET and the MOS capacitor (Figs. 1 a and b) is evidence that the charged species can escape the SiO_2 dielectric through the metal gate, but are unable to exit through the Si/SiO_2 interface (barrier). It follows from the sequence in Fig. 1 (b) that escape occurs through the metal gate under a negative gate bias, which proves that the mobile species involved are positively charged. The much larger solubility of hydrogen in metals such as Al or Hg as compared to Si [11], together with the crucial role of hydrogen in the anneal ambient, and the relatively rapid motion of the positive charge at room temperature as compared to Na^+ motion [14] collectively provide evidence that the mobile charges are protons.

For the small V_D used in our experiment (0.2 V) the Ψ-MOSFET is operating in the so-called linear region so that I_D is related to the threshold voltage, V_t, by [1]

$$I_D = A\left[(V_G - V_t)V_D - (0.5 + B)\,V_D^2\right] \qquad (1)$$

where the constants A and B engulf an ensemble of transistor related parameters. The threshold voltage in the presence of oxide charge is linearly related to the mid-gap voltage shift, ($V_t \propto \Delta V_{mg}$), so that changes in I_D are induced by changes in the oxide charge contribution to the threshold voltage: $\partial I_D / \partial V_t = \partial I_D / \partial \Delta V_{mg}$. The switched gate bias data shown in the inset of Fig. 2 result from the field assisted drift of positive charge away from the region of the top-Si/ SiO_2 interface. If we wish to calculate the exact time dependence of I_D we must model the field-assisted diffusion of the positive charge and calculate ΔV_{mg} for each instant in time. This problem is non-trivial since it has been demonstrated that the diffusion of charge in *a*-SiO_2 is dispersive [15]. Consequently, even if the initial charge profile is a sheet of charge at the interface, the spatial profile of the charge distribution evolves with time. It is not an easy task to deduce a simple analytical form for the time dependence of ΔV_{mg} (and hence I_D). If we consider the decrease in I_D for short times we can, to a first approximation, assume that the movement of the sheet of charge close to the Si/SiO_2 interface is non-dispersive. We take the very simplified example in which the sheet of charge moves with an effective "velocity" proportional to μ / τ_0 where $\mu(E)$ is proportional to the average hopping distance per hop of the charge (a hopping like diffusion is assumed [15]). This distance is dependent upon the electric field, E, in the oxide. The intersite hopping time, τ_0, depends upon the particle mass, m, the absolute temperature, T, and E. Under these assumptions we can write:

$$\partial \Delta V_{mg} / \partial t = -Q_H^{+}(\mu/\tau_0) / \varepsilon \varepsilon_0 \text{ Å}, \qquad (2)$$

where Q_H+ is the initial areal charge density at the Si/SiO_2 interface. We therefore expect the variations of I_D with time for very short times to be

$$\partial I_D / \partial t \propto \partial \Delta V_{mg} / \partial t \propto 1/\tau_0. \qquad (3)$$

Now $\tau_0 \propto \sqrt{m}\, exp(E_a/kT)$ [16], where E_a is the activation energy for hopping. Plotting $\Delta I_D\, T/I_0\, \Delta T$ as a function of 1/T in Fig. 2, we derive a value for E_a of 0.8 eV (the slope of the curve). Note further from Fig. 2 that this value is independent of the mass difference between H and D, and that there is a constant offset between the two curves corresponding to a factor of $\sqrt{2}$. All these observations are in agreement with the theoretical model elaborated here which is based on proton drift in SiO_2. These finding provide further evidence that the mobile species are indeed protons.

The excellent retention reported in Fig. 4 appears to defy basic diffusion laws. However, under floating gate conditions the mobile protons evidently stay near their programmed interface due to the large electrostatic attraction between the protons in the dielectric, and their negative image charge in the Si.

The conclusion from Fig. 5 that the mobile proton has a much smaller capture radius for electrons as compared to the trapped hole in SiO_2 can be understood qualitatively when one compares the ionization energy (difference between the charged and neutral state) as obtained from *ab-initio* theoretical calculations. If one assumes the model where the mobile H^+ ions are temporarily bound to bridging oxygen atoms inside the SiO_2 network, an ionization energy of about 3 eV is obtained from first-principles quantum-mechanical calculations [17]. This is much less than the ~ 8 eV calculated using the same formalism for the *E'* center [18] – a typical trapped hole center in SiO_2. This energy difference becomes obvious if one realizes that the Si_2–O–H^+ configuration is quite stable in SiO_2 [16] and is therefore less likely to trap an electron as compared to a classical coulomb-attractive defect center. These conclusions are very interesting from a technological standpoint because it implies that a memory device based on mobile protons has the potential of being not only nonvolatile, but also extremely radiation tolerant.

4.2 Reaction Model for the Proton Generation Process

We suggest that the faster proton buildup in smaller devices as compared to large area devices as observed here is due to the fact that the reactive hydrogen species, which trigger the proton generating reaction, enter the oxide from the edges of the capacitor and diffuse laterally along the oxide layer [19]. Hence, the amount of the reactive species reaching the oxide by diffusion through the Si overlayer is negligible for these samples and the time and temperature ranges explored in this study. This conclusion appears to be in contradiction with the conclusion from an earlier work [20] were it was argued that H^0 diffuses easily through the top Si layer at 600 °C. A possible clue to this difference in the conclusions may be that the relevant diffusing species are different (H^0 in Ref. [20] vs. H_2 or H^+ in this study). However, if we assume lateral diffusion into the buried oxide for the experiment in Fig. 1, we calculate approximately the time it takes for the diffusing species to reach its equilibrium concentration at 600 °C in a 400×400 μm^2 capacitor. For H_2 this time is ~ 300 s, and for H^0 ~ 40 s (diffusion coefficients taken from Ref [21]). Comparing these times with the characteristic time for Q_H+ buildup reported in this work leads us to conclude that the rate of the reaction(s) that generates the mobile protons may well be diffusion limited [22].

The observations in Figs. 6 and 7 allow us to develop a better insight into the dominant chemical reaction responsible for the variations in Q_H+. The rate equation for the proton reaction can be written as:

$$\frac{d[H^+]}{dt} = + k_f [H^0][X] - k_r [H^+][X] \qquad (4)$$

where X is the reactive site, suggested to be a bridging oxygen at the interface [4]. The forward and reverse rate constants, k_f and k_r, are assumed to obey the Arrhenius relation: $k = k_0 \exp(-E/kT)$. In order to explain the observations in Figs. 6 and 7, we assume that the forward reaction in Eq. 4 dominates ($k_f > k_r$) around 600 °C, while the reverse reaction gains weight with decreasing temperature and becomes competitive ($k_f < k_r$) at ~ 400 °C. Both reactions become negligible at temperatures below 400 °C, as discussed in Ref. 19. In terms of the activation energies, E_f and E_r, and the pre-exponential factors, k_{0f} and k_{0r}, for the forward and reverse reaction, respectively, this means that $E_f > E_r$ and $k_{0f} > k_{0r}$ should both be valid in the studied temperature domain. The assumption $E_f > E_r$ is rather straightforward for thermally activated reactions. The suggested difference between k_{0f} and k_{0r} is less obvious. The pre-exponential factors k_{0f} and k_{0r} are directly proportional to the fraction of time the reactive site X is occupied by the potential reactant (H^0 for k_{0f}, and H^+ for k_{0r}). Because of the presence of positive fixed charge near the interfaces, which is being generated along with Q_H+ during the hydrogen anneal process [4], the mobile protons will be repelled from the interface during the anneal. Hence, as near interfacial fixed charge builds up, H^+ is less likely to reach the reactive sites than H^0 and the assumption $k_{0f} > k_{0r}$ becomes reasonable.

6. Conclusions

We have provided compelling evidence that annealing of $Si/SiO_2/Si$ structures in forming gas above 400 °C can introduce mobile protons into the SiO_2 layer. It is shown that these protons are stable and confined within the dielectric layer, and that their space-charge distribution is easily rearrangable by applying a gate bias. By analyzing speed, retention, endurance, and radiation tolerance of nonvolatile memory elements based on this novel phenomenon, we have shown that this memory

technology has the potential of being competitive with existing Si-based non-volatile memory technologies such as Flash. Whether this potential is realized depends on addressing manufacturability issues associated with integrating protonic memory elements into a full CMOS process flow [23]. A diffusion and reaction kinetics study suggests that the proton generation reaction is a diffusion-limited reaction with a temperature dependent reaction balance.

Acknowledgment

The authors would like to thank Cliff Renschler, Mike Knoll, Jeff Bullington, and Zach Lemnios for stimulating discussions.

References

[§]The part of this work performed at US Air Force Research Laboratory was supported under contract PE62601, Project 8809. The part of this work performed at Sandia National Laboratories (SNL) was supported by the Defense Advanced Research Projects Agency, and by the Department of Energy (DOE). SNL is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for DOE under contract DE-AC04-94AL85000.

*electronic mail: karel@unm.edu

[1] S.M. Sze, in Physics of Semiconductor Devices (Wiley, New York, NY, 1981), ch. 8.

[2] K. L. Brower, Phys. Rev. B **42**, 3444 (1990).

[3] B.E. Deal, M. Sklar, A.S. Grove, and E.H.Snow, J. Electrochem. Soc. **114**, 266 (1967).

[4] W.L. Warren, K. Vanheusden, J.R. Schwank, D.M. Fleetwood, P.S. Winokur, and R.A.B. Devine, Appl. Phys. Lett. **68**, 2993 (1996).

[5] K. Vanheusden, W.L. Warren, R.A.B. Devine, D.M. Fleetwood, J.R. Schwank, M.R. Shaneyfelt, P.S. Winokur, and Z.J. Lemnios, Nature **386**, 587(1997).

[6] K. Vanheusden, W.L. Warren, R.A.B. Devine, J.R. Schwank, D.M. Fleetwood, R.G. Polcawich, S.P. Karna, and R.D. Pugh, IEEE Trans. Nucl. Sci. **44**, p. 2087 (1997).

[7] R.K. Lawrence, H.L. Hughes, R.E. Stahlbush, D.I. Ma, and M.E. Twigg, in Proceedings of the IEEE International SOI Conference (IEEE Catalog No. 97CH36069, Piscataway, NJ, 1997), p.142.

[8] N. Lifshitz and G. Smolinsky, Appl Phys. Lett. **55**, 408 (1989).

[9] See, e.g., P.S. Winokur, H.E. Boesch, J.M. McGarrity, and F.B. McLean, IEEE Trans. Nucl. Sci. **24**, 2113 (1977).

[10] A.J. Auberton-Herve, T. Barge, F. Metral, M. Bruel, B. Aspar, C. Maleville, H. Moriceau, T. Poumeyrol, in Amorphous and Crystalline Insulating Thin Films, Vol. **446**, edited by W.L. Warren, R.A.B. Devine, M. Matsumura, S. Cristoloveanu, Y. Homma, J. Kanicki (Materials Research Society, Pittsburgh, PA, 1997), p. 177.

[11] B. L. Sharma, in Diffusion in Semiconductors (Trans. Tech. Pub., Germany, 1970), p. 87.

[12] S. Cristoloveanu and S. S. Li, in Electrical characterization of silicon-on-insulator materials and devices (Kluwer Academic Publishers, Boston, MA, 1995) p.104.

[13] D.M. Fleetwood and J.H. Scofield, Phys. Rev. Lett. **64**, 579 (1990).

[14] S.R. Hofstein, IEEE Trans. Elect. Dev. **14**, 749 (1967).

[15] F. B. McLean, IEEE Trans. Nucl. Sci. **27**, 1651 (1980).

[16] N. S. Saks and R. W. Rendell, IEEE Trans. Nucl. Sci. **39**, 2220 (1992).

[17] A.M. Ferreira, S.P. Karna, C.P. Brothers, R.D. Pugh, B.K. Singaraju, K. Vanheusden, W.L. Warren, and R.A.B. Devine, in Amorphous and Crystalline Insulating Thin Films, Vol. **446**, edited by W. L. Warren, R.A.B. Devine, M. Matsumura, S. Cristoloveanu, Y. Homma, J. Kanicki (Materials Research Society, Pittsburgh, PA, 1997), p. 247.

[18] S.P. Karna, R.D. Pugh, C.P. Brothers, W. Shedd, and B.K. Singaraju, , in Proceedings of the 1997 IEEE International SOI Conference, (IEEE Catalog No. 97CH36069, Piscataway, NJ, 1997), p. 32.

[19] K. Vanheusden, W. L. Warren, D. M. Fleetwood, J. R. Schwank, M. R. Shaneyfelt, B. L. Draper, P. S. Winokur, R. A. B. Devine, L. B. Archer, G. A. Brown, and R. M. Wallace, Appl. Phys. Lett. **73**, 674 (1998).

[20] S. M. Myers, G. A. Brown, A. G. Revesz, and H. L. Hughes, J. Appl. Phys. **73**, 2196 (1993).

[21] D.L. Griscom, J. Appl. Phys. **58**, 2524 (1985).

[22] K. Vanheusden, S.P. Karna, R.D. Pugh, W.L. Warren, D.M. Fleetwood, R.A.B. Devine, and

A.H. Edwards, Appl. Phys. Lett. **72**, 28 (1998).

[23] D.M. Fleetwood, W.L. Warren,. K. Vanheusden, R.A.B. Devine, M.R. Shaneyfelt, B.L. Draper, J.R. Schwank, T.L. Meisenheimer, P.S. Winokur, and M.G. Knoll, in Proceedings of the 1998 IEEE International Nonvolatile Memory Technology Conference (IEEE Catalog No. 98EX141, Piscataway, NJ, 1998), p. 91.

Figure Captions

FIG 1. Voltage shifts of *I-V* (a) and *C-V* (b) curves on SIMOX (*c*-Si/SiO_2/*c*-Si structure) after it received a 550 °C forming gas anneal. *I-V* curves were measured using the Ψ-MOSFET technique, and *C-V* curves at 1 MHz using standard MOS capacitors, as shown in the insets.

FIG. 2. Arrhenius plot of the initial rate of drain current decay vs. inverse temperature. Values were obtained using normalized drain current curves like the one shown in Fig. 2. The thermal activation energy E_a for ionic motion is obtained from the slope of a linear least squares fit to the data points (see Ref. 5).

FIG. 3. Initial rate of decay of the point-contact FET drain current obtained from curves such as the one shown in the inset of Fig. 2. All data were recorded at room temperature as a function of applied gate voltage for a fixed buried oxide thickness of 1 μm (a), and as a function of buried oxide thickness (b) for a fixed applied gate voltage of −10 V (see Ref. 5).

FIG. 4. Retention behavior for memory test devices fabricated using point-contact FETs on ZMR substrate material. The devices were hydrogenated at 600 °C.

FIG. 5. Density of mobile protons as a function of cumulative radiation dose for poly-Si capped thermal oxide capacitors. Irradiations were performed at different substrate temperatures, while all measurements were done at room temperatures.

FIG. 6. Mobile proton density vs. cooling rate in Unibond SOI and 11-nm thermal SiO_2 substrates. All samples were annealed in forming-gas at 600 °C for 20-min, and subsequently cooled at various rates in unaltered ambient (see Ref. 19).

FIG. 7. Normalized proton density in Unibond vs. anneal temperature. The open symbols represent data taken after a single anneal. The closed symbols refer to samples first annealed in forming gas at 600 °C, followed by a second anneal at a lower temperature (see Ref. 19).

FIG. 1(a)

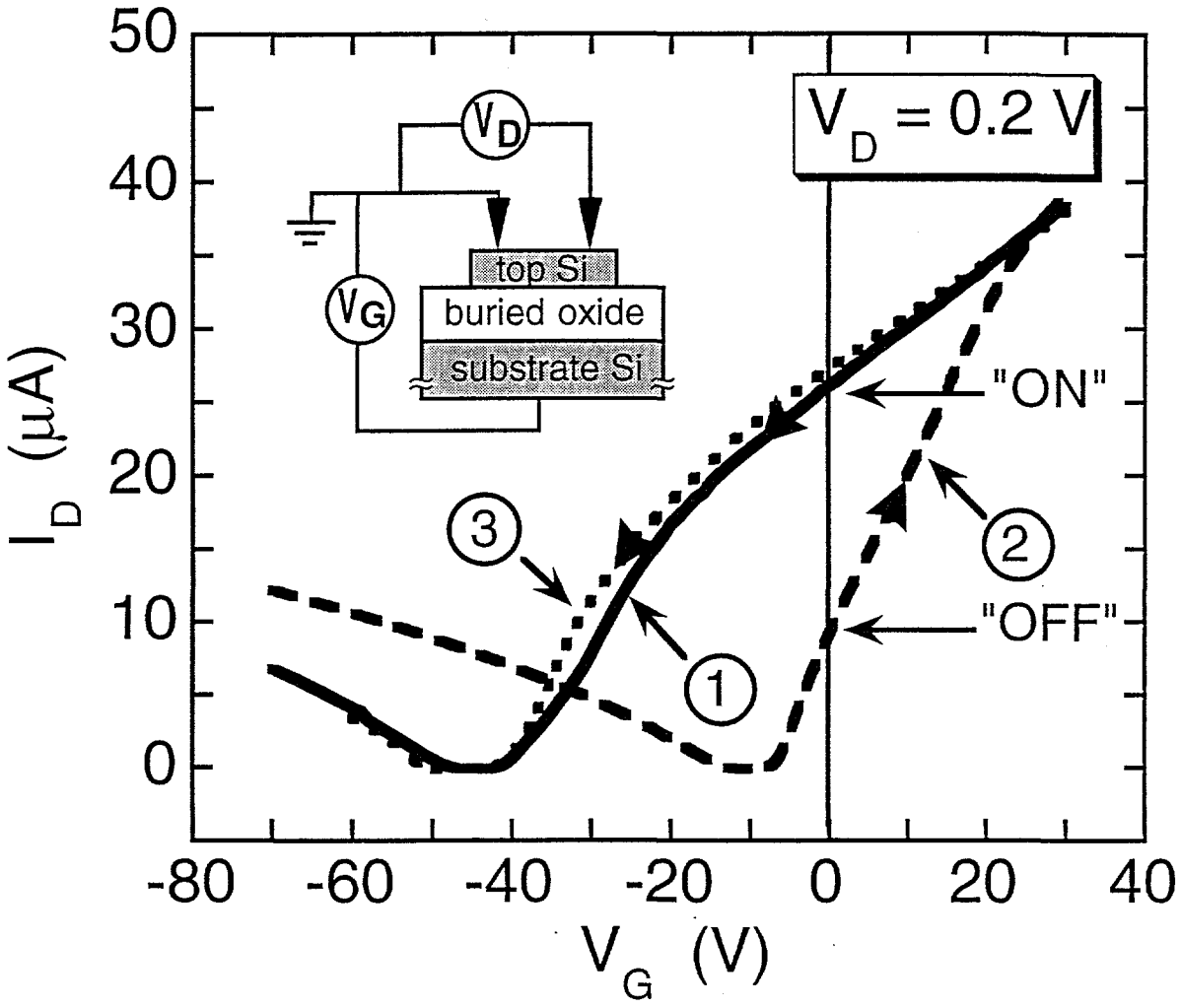

50
40
30
20
10
0
I_D (μA)
-80
-60
-40
-20
0
20
40
V_G (V)
V_D = 0.2 V
V_D
V_G
top Si
buried oxide
substrate Si
"ON"
"OFF"
1
2
3

FIG. 1 (b)

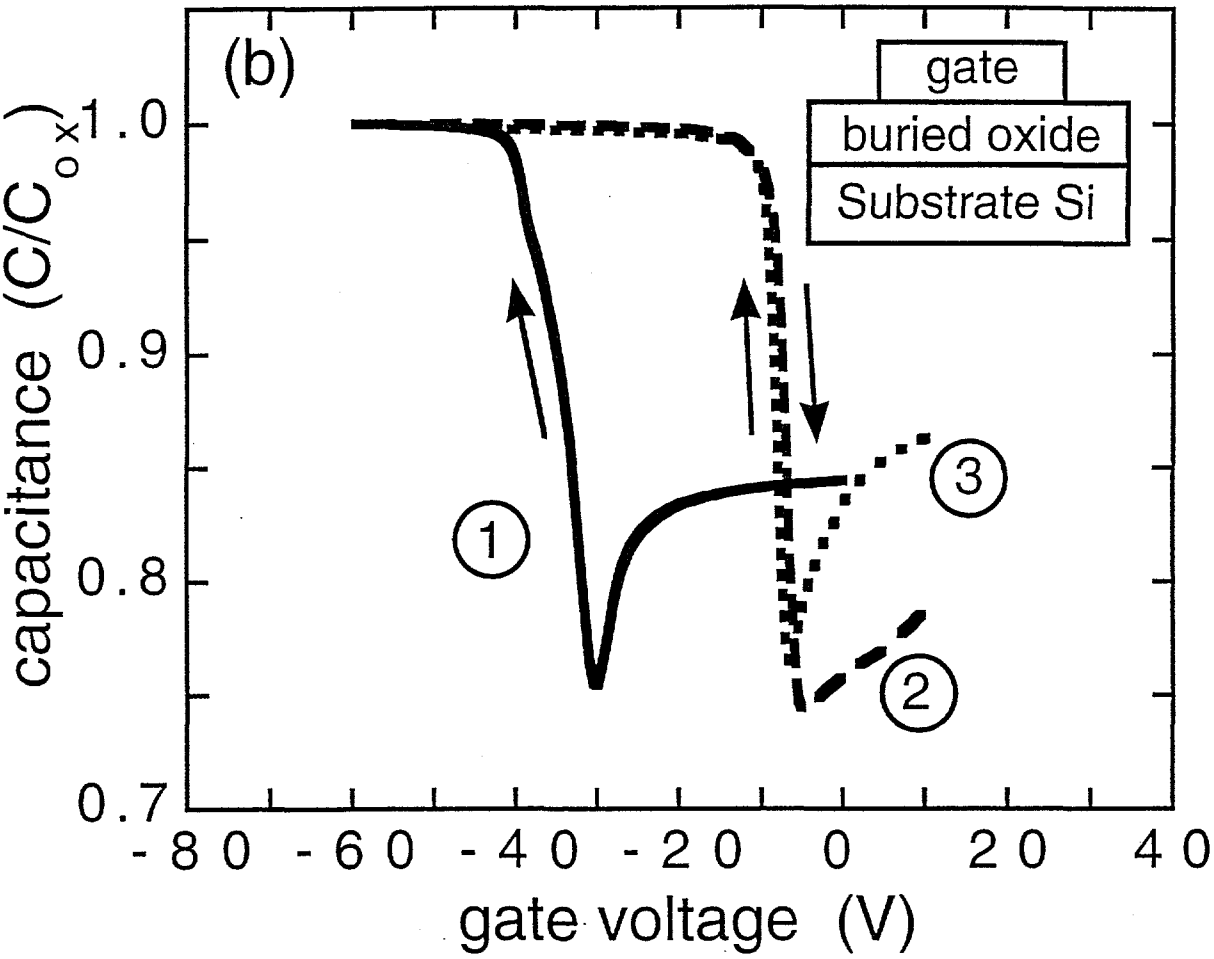

(b)
gate
buried oxide
Substrate Si
capacitance (C/C$_{ox}$)
1.0
0.9
0.8
0.7
-80
-60
-40
-20
0
20
40
gate voltage (V)
1
2
3

FIG. 2

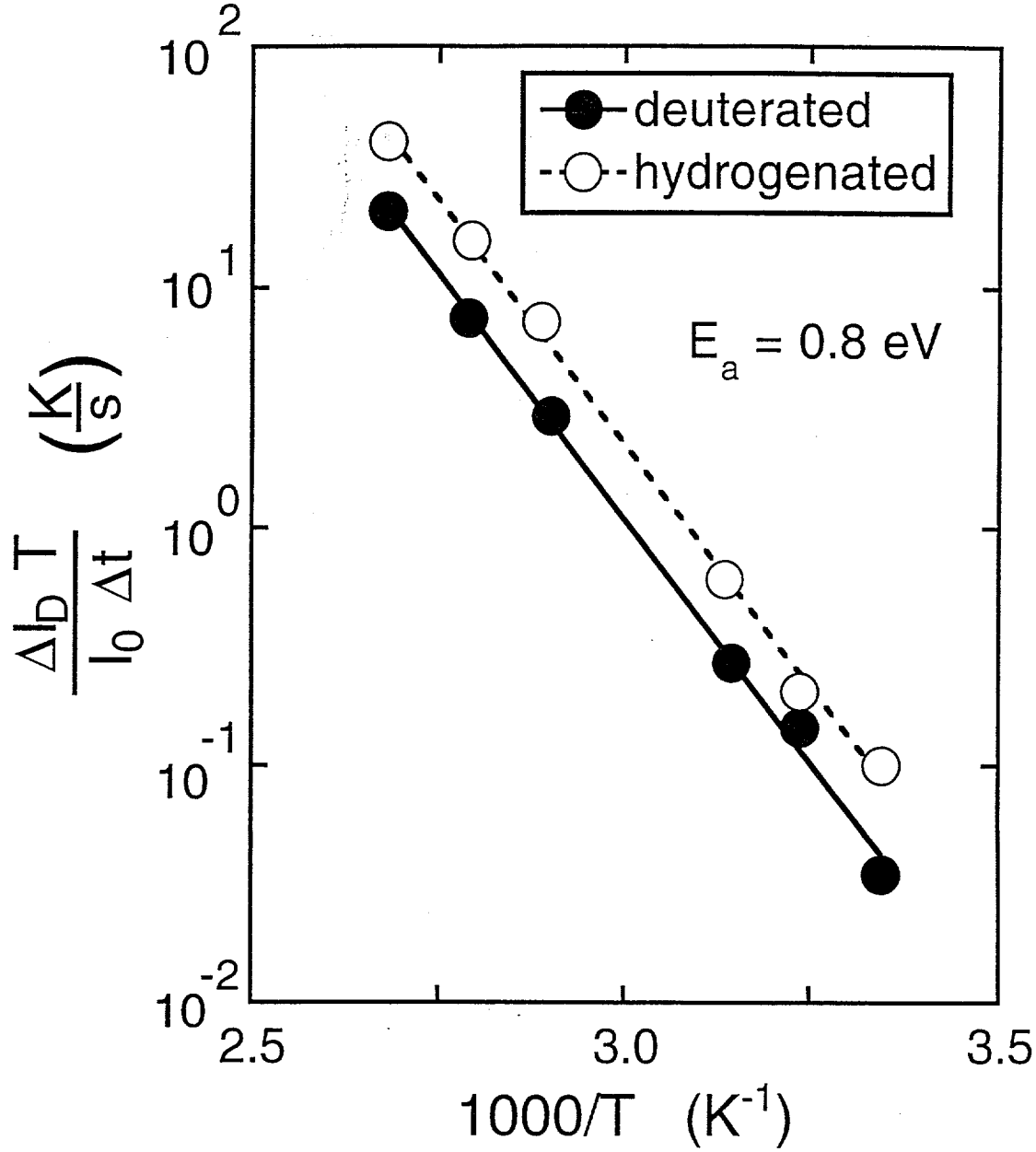

deuterated
hydrogenated
E_a = 0.8 eV
$\frac{\Delta I_D}{I_0} \frac{T}{\Delta t}$ $\left(\frac{K}{s}\right)$
10^2
10^1
10^0
10^{-1}
10^{-2}
2.5
3.0
3.5
1000/T (K^{-1})

FIG. 3(a)

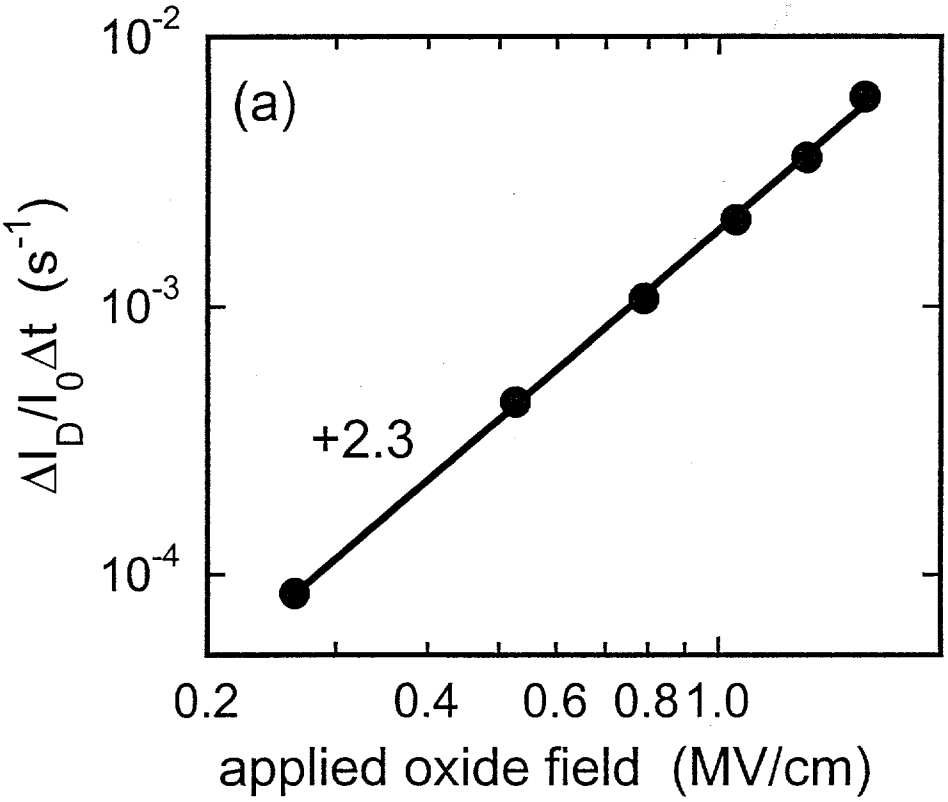

(a)
+2.3
10^{-2}
10^{-3}
10^{-4}
$\Delta I_D/I_0\Delta t$ (s^{-1})
0.2
0.4
0.6
0.8
1.0
applied oxide field (MV/cm)

FIG. 3(b)

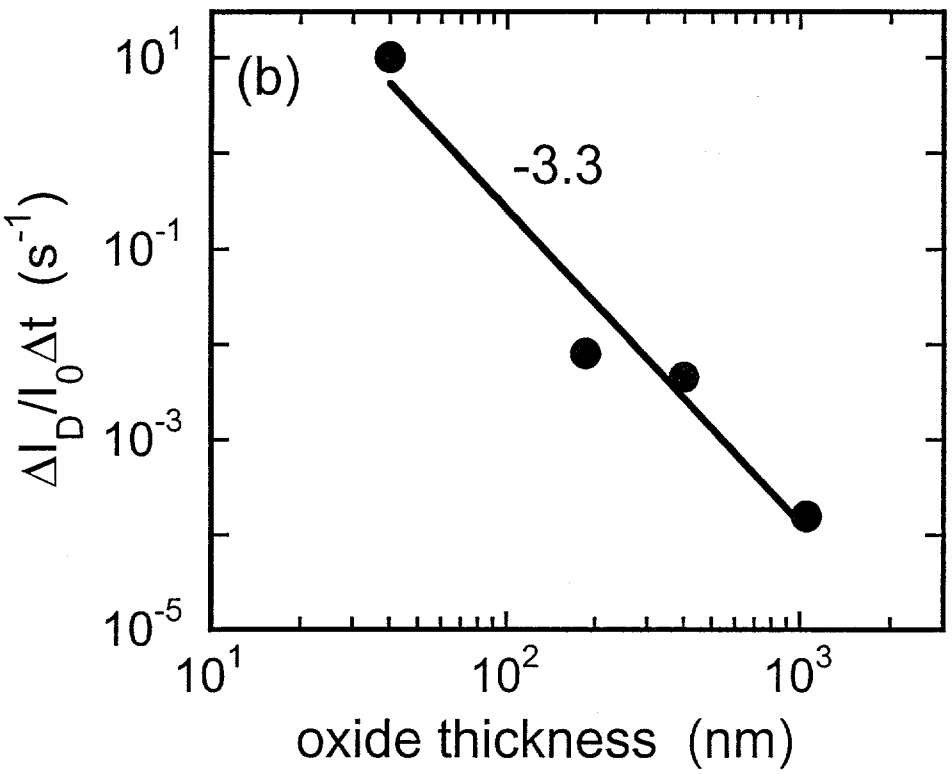

(b)
-3.3
10^{1}
10^{-1}
10^{-3}
10^{-5}
$\Delta I_D/I_0\Delta t$ (s^{-1})
10^{1}
10^{2}
10^{3}
oxide thickness (nm)

FIG. 4

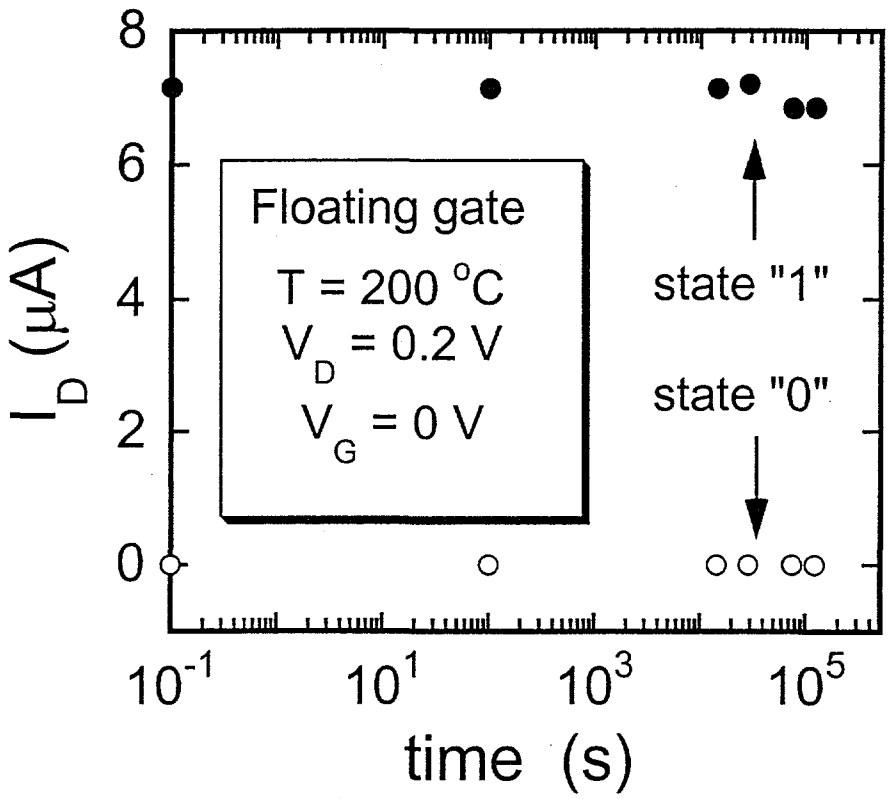

8
6
4
2
0
I_D (μA)
Floating gate
T = 200 °C
V_D = 0.2 V
V_G = 0 V
state "1"
state "0"
10^{-1}
10^{1}
10^{3}
10^{5}
time (s)

FIG. 5

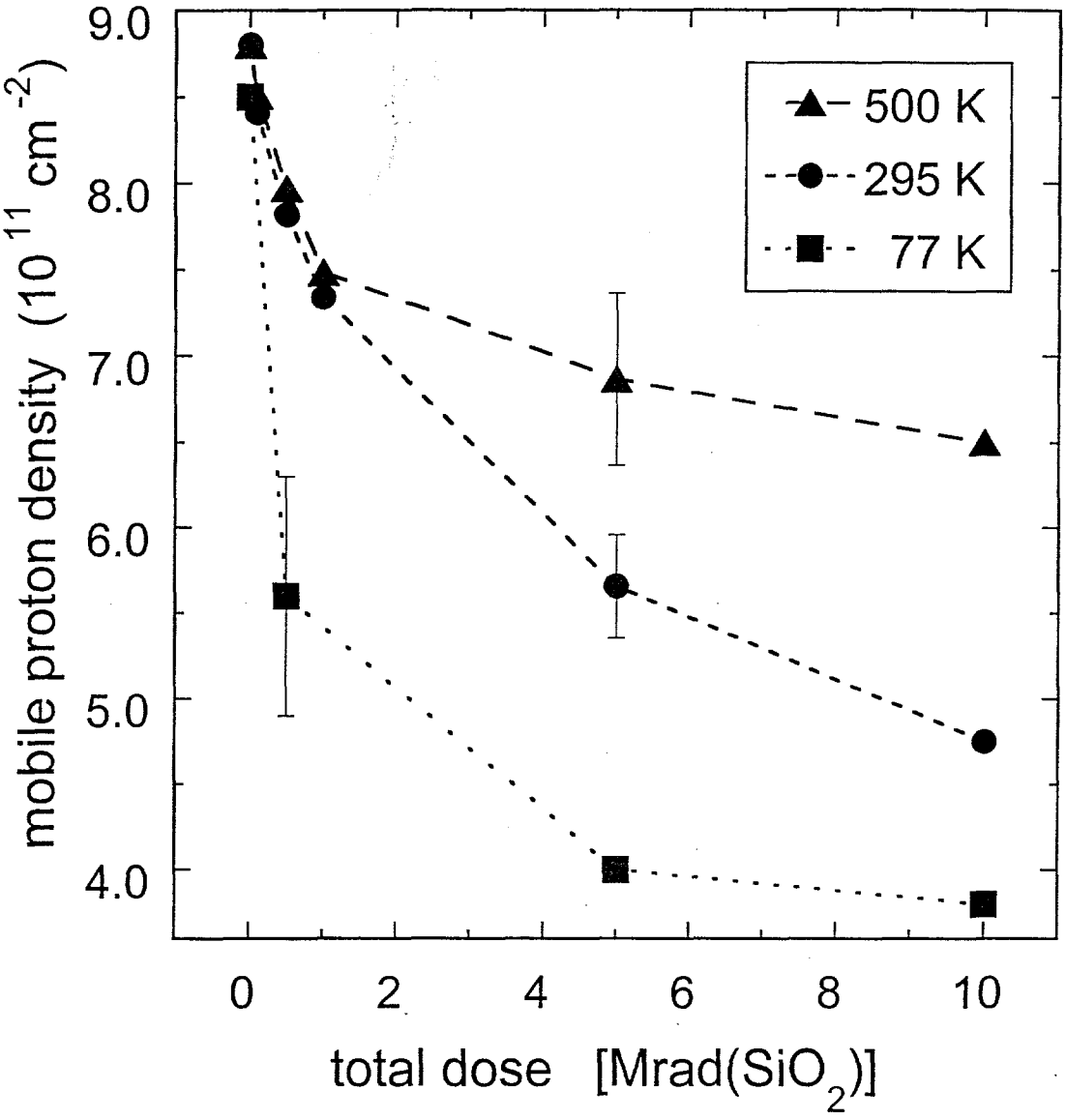

9.0
8.0
7.0
6.0
5.0
4.0
mobile proton density (10^{11} cm^{-2})
500 K
295 K
77 K
0
2
4
6
8
10
total dose [Mrad(SiO_2)]

FIG. 6

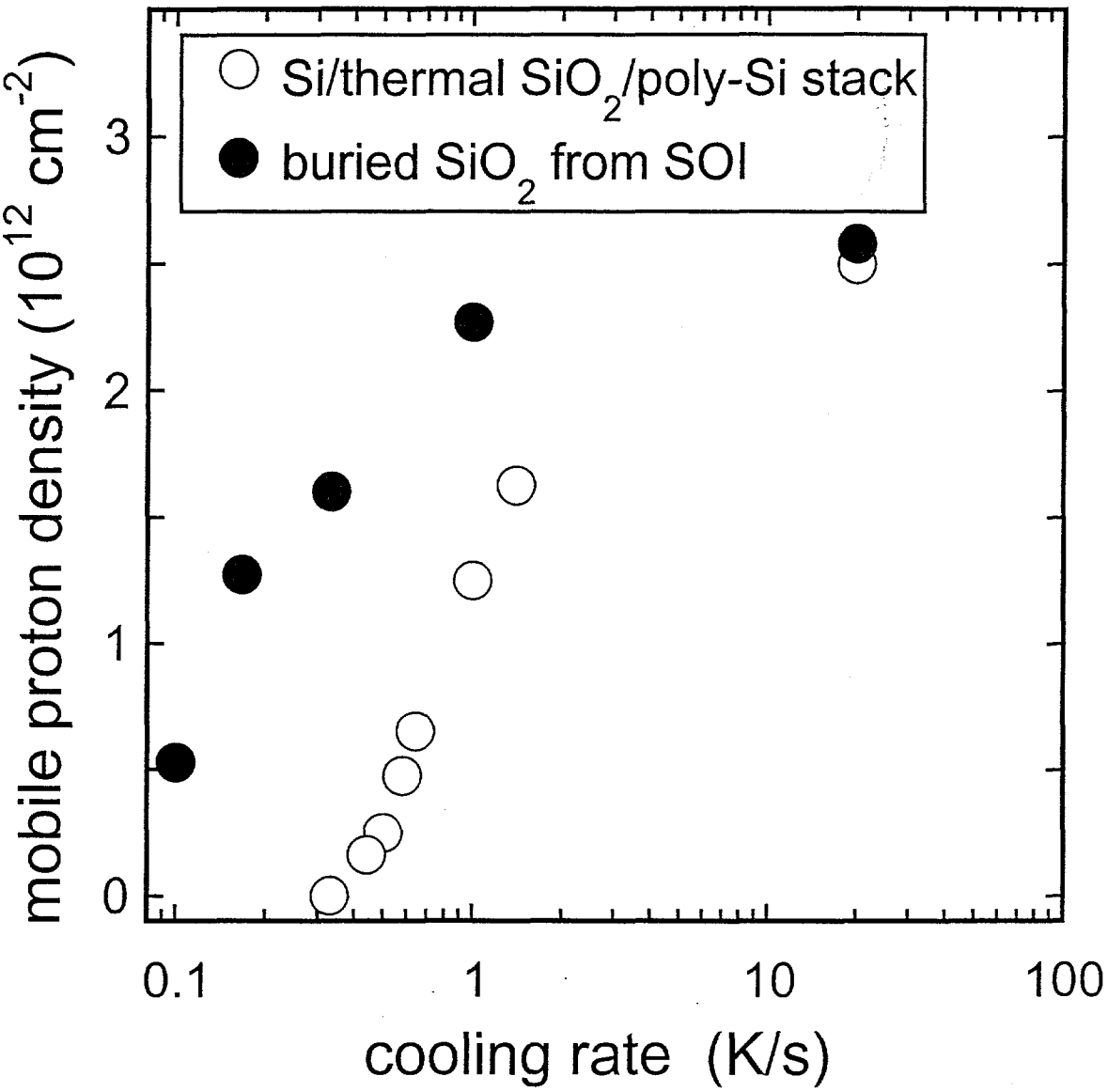
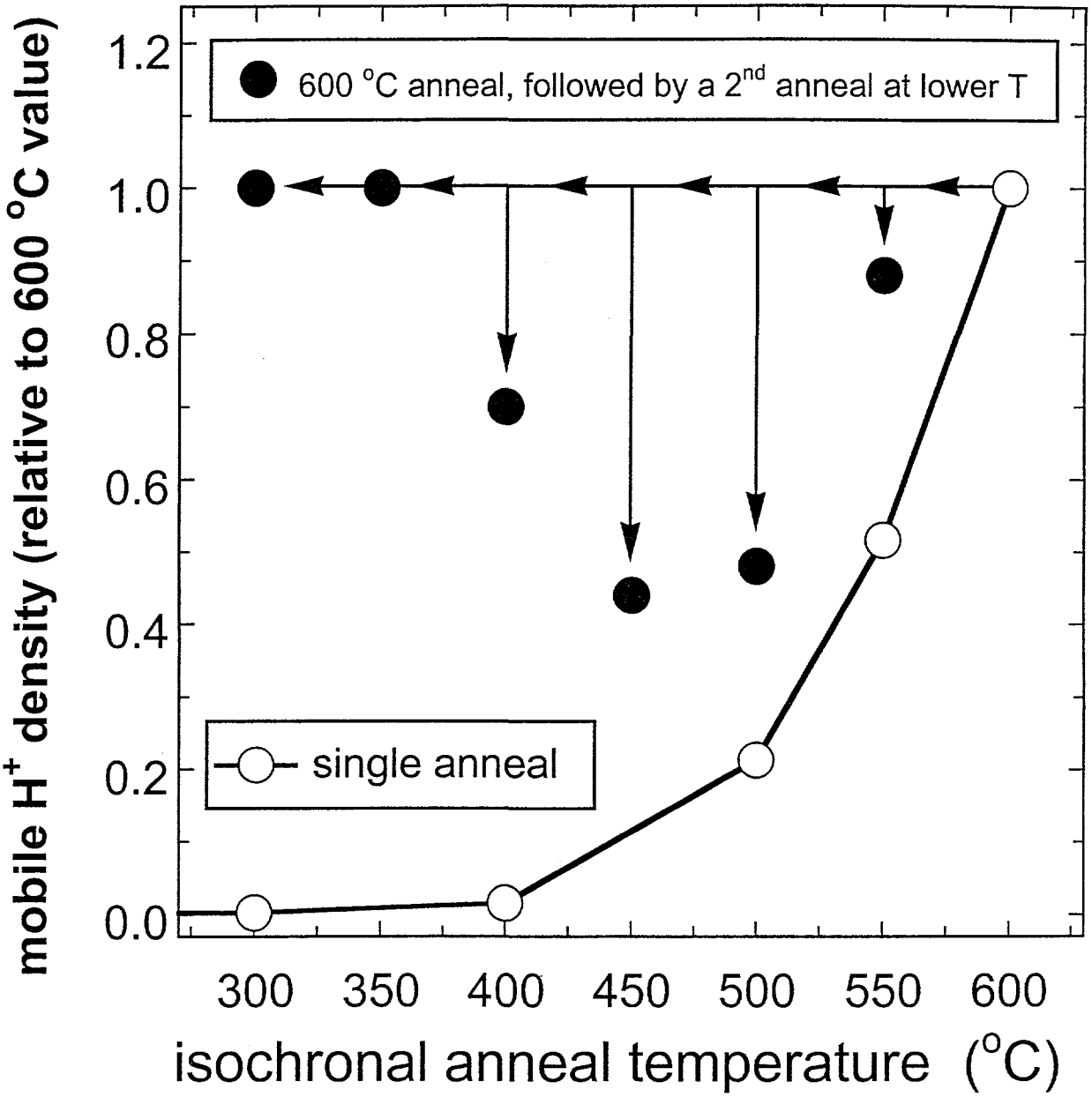

Si/thermal SiO2/poly-Si stack
buried SiO2 from SOI
mobile proton density (10^12 cm^-2)
0
1
2
3
0.1
1
10
100
cooling rate (K/s)

FIG. 7


600 °C anneal, followed by a 2nd anneal at lower T
single anneal
mobile H+ density (relative to 600 °C value)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
300
350
400
450
500
550
600
isochronal anneal temperature (°C)