PHYSICAL REVIEW B, VOLUME 63, 233406

Electric-field-induced transport of protons in amorphous SiO_2

R. A. B. Devine
Center for High Technology Materials, University of New Mexico, 1313 Goddard SE, Albuquerque, New Mexico 87106

G. V. Herrera
Sandia National Laboratories, Mail Stop 1045, P.O. Box 5800 Albuquerque, New Mexico 87185-1405

(Received 5 September 2000; published 24 May 2001)

The temperature and time dependence of electric-field-induced motion of thermally generated protons in 20 nm amorphous SiO_2 films has been investigated. The short-time behavior is found to involve an activation energy of $\sim$0.38 eV and a hopping distance of $\sim$0.55 nm. The activation energy is considerably smaller than that found for radiation-generated proton motion ($\sim$0.82 eV). The hopping distance is consistent with transport involving motion between a bridging oxygen atom and a next-nearest-neighbor (or farther) bridging oxygen atom; this suggests that motion is primarily via cross-ring hopping.

DOI: 10.1103/PhysRevB.63.233406 PACS number(s): 66.30.Dn

Mobile protons have been generated in amorphous SiO_2 films following irradiation (if the films contain quantities of H in some form),[1–3] following room-temperature exposure[4] of preirradiated films to an atmosphere containing H_2, and following[5–8] annealing of sandwiched $Si/SiO_2/Si$ structures in N_2-H_2 mixtures at temperatures in the range 500–800 °C. In the last case, if the sandwiched oxide layers were of the thermally grown type, an initial anneal at high temperature (typically above 1000 °C) in an inert or weakly oxidizing atmosphere was necessary prior to the lower-temperature H_2 anneal. Buried oxide films[5,6] such as those manufactured by ion implantation (separation by implantation of oxygen, SIMOX) or by a process of bonding and etchback[9] (bonded and etchback silicon on insulator, BESOI or SMARTCUT®) generally exhibit the proton generation phenomenon during a hydrogen anneal without the >1000 °C high-temperature preconditioning.[8]

Interest from an applied standpoint in the electric-field-induced motion of the mobile protons in SiO_2 stems from concerns about the behavior of devices such as metal-oxide-semiconductor field-effect transistors (MOSFET's) subjected to irradiation. Radiation-induced protons may drift toward the semiconductor/oxide interface, where they generate interface states that can degrade the inversion-channel carrier mobility and result in threshold voltage variation.[1–3,10] On the positive side, thermally generated mobile protons appear to be stable and to drift in the presence of an electric field. Furthermore, they do not interact at the interface to create interface states. This suggests that MOSFET devices might be constructed with incorporated mobile charges[11] which transform them into transistors with an integrated memory function. When the protons are placed close to the semiconductor/oxide interface they induce an inversion layer which conducts (i.e., the transistor is in the ''on'' state), while when they are placed close to the gate electrode/oxide interface they have no effect and the transistor can be engineered to be ''off.'' In this case it is the retention time for which the protons remain ''blocked'' at the interfaces and the switching time required for the protons to move from one interface to another that will dictate the real feasibility of the integrated memory transistor as a commercially viable device.

From the above discussion it is clear that we must be concerned with the intimate details of the nature of the field-induced transport of protons in the amorphous $(a\text{-})SiO_2$ film of which the gate oxide of the MOSFET is comprised and the interaction of the protons at the interface, if any. We have constructed MOSFET devices that have been subsequently protonated thermally to induce the memory function. We have measured the transit time of the protons across the gate oxide in the presence of an applied field through the switching of the inversion-layer current and have studied how this transport time varies as a function of the device temperature. From these data and their comparison with available data on protons generated in SiO_2 by different methods, a clearer physical picture of the protons and the nature of their diffusion emerges.

MOSFET devices having gate oxide thicknesses of 20 nm were constructed in a manner described previously.[12] The starting Si wafer was p type so that the MOSFET's manufactured were n channel. Although devices having gate lengths (L) of 5 μm and widths (Z) of 100 μm were available on our wafers, the presence of substantial fixed oxide charge and reduced mobile charge[7] in the shorter-channel transistors led us to study components having 20 μm gate lengths. After manufacture of the basic devices, a three minute 1050 °C anneal was carried out in flowing N_2 gas both to activate the gate, source, and drain dopants and to create the precursor sites necessary for subsequent proton generation during a second anneal step. This anneal involved an atmosphere of $N_2/10\%\ H_2$ for 30 min at 700 °C followed by a fast pull to room temperature. The device characteristics were measured using a Hewlett-Packard 4145 transistor tester. In the first instance, the gate was polarized with a voltage V_G of $+6$ (-6) V with respect to the source contact for a period of 30 s and the gate voltage then swept to -6 ($+6$) V in a period of approximately 3 s (0.2 V steps). The source-drain current (I_{ds}) for a source-drain potential (V_{ds}) of 0.1 V was monitored during this sweep. A typical transistor hysteresis curve is shown in Fig. 1. In the down-sweep mode, the initial positive polarization step drove the protons to the oxide/semiconductor interface where they induced an inversion layer even in the absence of a potential on the gate electrode.

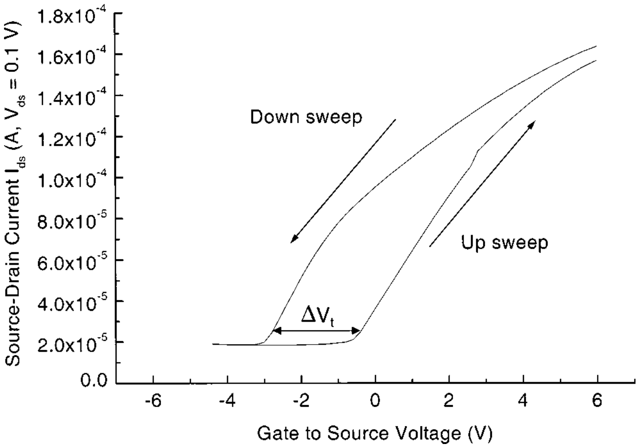

FIG. 1. Typical hysteresis curve of the source-drain current $I_{\rm ds}$ as a function of gate-to-source voltage V_G, for a MOSFET with 20 μm gate length, 100 μm gate width, and 20-nm-thick gate oxide after protonation at 700 °C. In the down-sweep mode the gate was first biased for 30 s at +6 V and then potential was swept to −6 V in less than 3 s. In the up-sweep mode, the gate was biased at −6 V for 30 s prior to sweeping to +6 V in less than 3 s. The hysteresis due to mobile protons is indicated by ΔV_t. The source-drain voltage $V_{\rm ds}$ was 0.1 V.

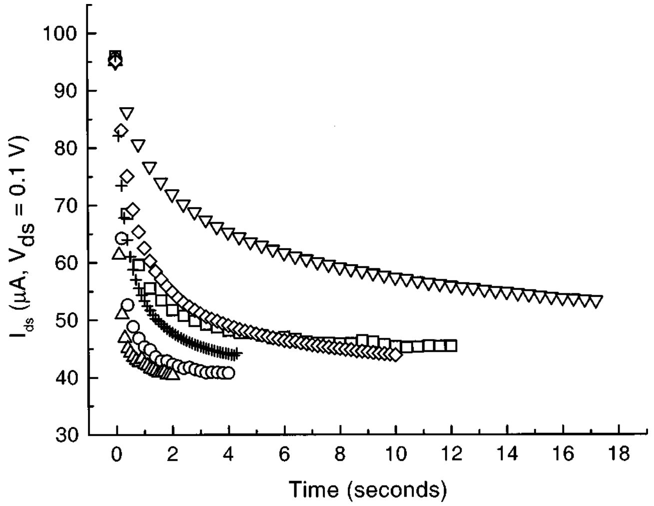

FIG. 2. Time dependence of the source-drain current in MOSFET's with 20 μm gate length, 100 μm gate width, and 20 nm gate oxide thickness for different pulsed gate voltages. The gate voltage was initially +6 V for more than 2 min. At 85 °C (□), −2; (○), −4; (△), −6 V; at 50 °C (▽) −2; (◇) −4; (+) −6 V. The source-drain voltage $V_{\rm ds}$ was 0.1 V in all cases.

Their presence at the insulator/substrate interface induced a negative threshold-voltage shift. In the up-sweep case, the application of a negative polarization prior to sweeping resulted in the protons being placed at the gate electrode/oxide interface where they played no role as far as the inversion layer was concerned. The hysteresis (ΔV_t) shown in Fig. 1 (~ -2 V) enables us to conclude that $\sim 2.2\times 10^{12}$ positive mobile charges per cm^2 were at the interface when the gate was polarized positively. Hysteresis as large as −5.5 V, has been obtained for similar transistors. Note that in Fig. 1 we observe a gate-voltage-independent leakage current ~20 μA. This leakage results from the architecture of the transistor/Si wafer substrate and its Ohmic nature was confirmed by varying the source-drain voltage. In a second set of measurements, a positive polarization was applied to the gate electrode for an extended period (>2 min) and then the source-drain current measured using a pulsed gate-source potential. The pulse potential applied for a time t_p was used to displace the protons away from the oxide/substrate interface and during the 0 V gate bias part of the pulse $I_{\rm ds}$ was measured. By accumulating a large number of pulses a plot of $I_{\rm ds}(t,V_G)$ was obtained. A similar pulsed technique was used in studies of SIMOX and SMARTCUT samples reported in Ref. 8. An example for applied voltages of −2, −4, and −6 V at 50 and 85 °C is shown in Fig. 2. We observe a rapid drop of $I_{\rm ds}(t)$ followed by a long-time ‘‘tail,’’ which characterizes charge transport in a dispersive medium.[13] Note that the minimum ‘‘off’’ current is not zero because even at 0 V gate bias with the protons at the gate electrode the transistor was still partially conducting (see the up-sweep curve of Fig. 1). Measurements of $I_{\rm ds}(t,V_G,T)$ were carried out for devices held at 23, 50, and 85 °C using a hot chuck facility installed on the micromanipulator stage used with the HP 4145 transistor tester. In the following analysis we will not use the ‘‘complete’’ $I_{\rm ds}(t,V_G,T)$ curves such as those shown in Fig. 2 but will concentrate on the initial-slope regime $-dI_{\rm ds}(t,V_G,T)/dt|_{t\to 0}$. Values of this parameter for different V_G and T are given in Table I.

In the small-$V_{\rm ds}$ regime in which we have operated $I_{\rm ds}$ can be expressed as[14]

$$I_{\rm ds}(t,V_G,T)=(Z/L)C_{\rm ox}\mu_n(V_G-V_t)V_{\rm ds}, \qquad (1)$$

where μ_n is the mobility of the inversion-channel carriers (here electrons) and $C_{\rm ox}$ is the capacity of the gate oxide. V_t is the threshold voltage, which we will express in terms of the volume density of charge in the gate oxide, $\rho(x,T,t)$, as[14]

$$V_t=-(1/C_{\rm ox})\left[(1/d_{\rm ox})\int x\rho(x,T,t)dx\right]. \qquad (2)$$

The integral limits are taken from $x=0$ to $d_{\rm ox}$ (where $d_{\rm ox}$ is the gate oxide thickness) measured from the gate electrode. Differentiation of Eq. (1) yields

$$dI_{\rm ds}(t,V_G,T)/dt|_{t\to 0}=-(Z/L)C_{\rm ox}\mu_n V_{\rm ds}dV_t/dt. \qquad (3)$$

TABLE I. Variation of the source-drain current with time, $dI_{\rm ds}(t,V_G,T)/dt|_{t\to 0}$, deduced from data such as those shown in Fig. 2 for different temperatures and gate voltages V_G. The source-drain voltage for all experiments was 0.1 V. The units for $dI_{\rm ds}(t,V_G,T)/dt|_{t\to 0}$ are μA s^{-1}.

	$dI_{\rm ds}(t,V_G,T)/dt\|_{t\to 0}$,		
Gate voltage (V)	30 °C	50 °C	85 °C
−2	5.15	21.6	68.5
−4	17.8	60.5	156.8
−6	36.8	139.1	337.4

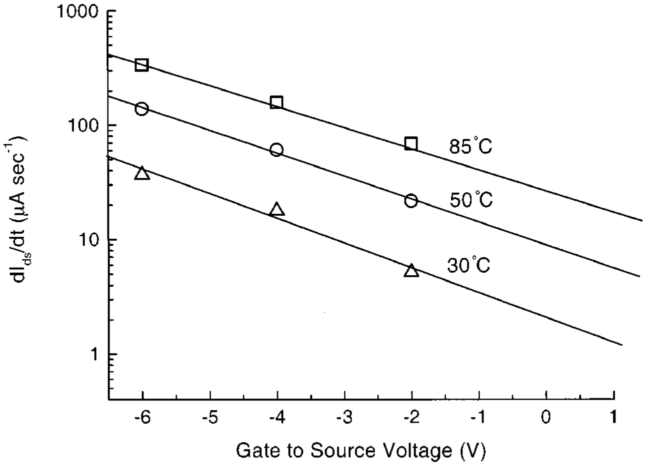

FIG. 3. Initial slope of the variation of the source-drain current with time, $dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}$, from Fig. 2 as a function of the pulse amplitude of the gate voltage (V_G) for various temperatures. The source-drain voltage was 0.1 V in all cases.

In consequence, measurement of $dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}$ provides direct information on the time variation of the proton distribution in the gate oxide under the influence of the applied gate voltage (and hence electric field). Assuming that for short times the sheet of proton charge of density Q_{H^+} composed of the protons moves with a velocity $v(T)$, then $dV_t/dt=Q_{\mathrm{H}^+}v(T)$ and

$$dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}=-(Z/L)\mu_n V_{\mathrm{ds}} Q_{\mathrm{H}^+}v(T). \qquad (4)$$

Clearly, the temperature dependence of $dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}$ and the electric field (or V_G) dependence, if any, will be contained in $v(T)$. For simplicity we will assume that $v(T)=R_{\mathrm{fi}}/\tau$ where R_{fi} is the intrinsic hopping distance and $1/\tau$ the hopping rate.

In the first instance, the proton in the $\mathrm{SiO_2}$ network bonds to an intertetrahedral bridging oxygen and can be treated as a small polaron. The jump rate for electric-field-assisted hopping is then[15]

$$R=1/\tau\propto[1/(4E_a k_B T)^{1/2}] \times\exp[-(4E_a-qER_{\mathrm{fi}})^2/(16E_a k_B T)], \qquad (5)$$

where k_B is Boltzmann's constant, E_a the activation energy for hopping, and q the elemental charge on the proton. Plotting the data from Table I in Fig. 3, we see that as a function of V_G the absolute value of the initial slope increases with voltage (equivalent to electric field) in a logarithmic manner. Taking the logarithm of Eq. (5) we see that the slope of the $\ln[dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}]$ versus V_G plot is proportional to $qR_{\mathrm{fi}}/2k_B T d_{\mathrm{ox}}$. From the experimental data we then deduce $0.53\leqslant R_{\mathrm{fi}}\leqslant 0.55$ nm. From Eqs. (4) and (5) we have

$$\ln[dI_{\mathrm{ds}}(t,V_G,T)/dt|_{t\rightarrow 0}] \propto\ln[1/(4E_a k_B T)^{1/2}]-E_a/k_B T+qER_{\mathrm{fi}}/2k_B T \qquad (6)$$

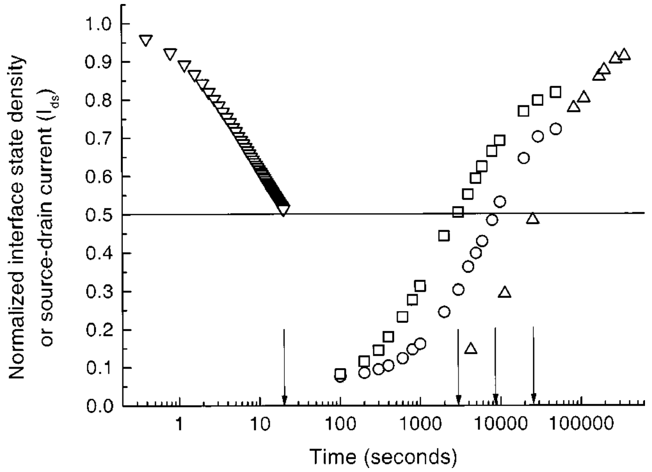

FIG. 4. Time dependence of the normalized source-drain current I_{ds} or interface state density as a function of time in MOSFET's subjected to similar electric fields. (▽) source-drain current due to thermally generated protons first driven to the 20 nm gate oxide/substrate interface then subjected to an electric field of ~ -1 MV cm^{-1}. (△) Interface states resulting from protons generated in 77 nm gate oxide by exposure of irradiated MOSFET's to a $\mathrm{H_2}$-containing atmosphere at room temperature. The electric field was +0.65 MV cm^{-1} (Ref. 4). (○) and (□) interface states generated in 47.8 nm gate oxide MOSFET's by protons created during ^{60}Co irradiation. The electric fields were +0.56 and +1.07 MV cm^{-1}, respectively (Ref. 3). The arrows indicate the time for a 50% variation, $t_{1/2}$.

so that, using the value of R_{fi} together with the data from Table I, we can estimate E_a. We obtain the value of $E_a \sim 0.38$ eV, which is significantly smaller than the value of 0.82 eV previously advanced[1] for the activation energy for the diffusion of radiation-generated protons in $\mathrm{SiO_2}$. We must stress that our temperature measurements were performed over a limited range of temperature (23–85 °C) so there will be imprecision in the determination of E_a. We do not, however, anticipate the error to be as large as the factor greater than 2 suggested by the numbers just cited.

The dispersive transport nature of charges (protons and holes) in the amorphous $\mathrm{SiO_2}$ network has been well established.[1–7] All of the experiments involving ''indirect'' detection of proton motion, which require, for example, a reaction leading to the appearance of interface states at the $\mathrm{SiO_2}$/substrate interface, suggest that proton motion through the oxide is relatively slow. In experiments where the proton motion is sensed ''directly'' via the effect of its charge, for example, on the threshold or flat band voltage and hence inversion-layer current I_{ds} [Eq. (3)], it appears orders of magnitude more rapid. This is illustrated in Fig. 4, where we show time-dependent data for radiation-induced and thermally generated proton motion. Care must be exercised in comparing the results from the two types of experiment shown since they do not sense the same quantity. For example, from Eq. (2) we see that the threshold voltage (and hence the source-drain current I_{ds}) due to mobile charges depends upon an integral that is the first moment of the charge distribution at any time. The interface state density, due to the arrival of protons at the interface and a charge conversion of interface states, measures the time dependence of $[1-\int\rho(x,T,t)dx]$ (where we assume the integral is normalized to unity at time $t=0$). The full significance is seen if we consider the case of charges moving as a sheet. If the sheet is initially at the gate electrode/oxide boundary, no interface states are measured at the substrate/oxide interface until the entire charge sheet has crossed the oxide and created them, whereas Eq. (2) suggests that a threshold voltage shift (source-drain current) will be detected immediately the weighted integral becomes finite. We therefore conclude that the ''indirect'' measurements actually sense an effective motion of the protons that is modulated and apparently slowed. This is because of the localized spatial nature of the proton–interface-state reaction necessary to show that the protons have arrived at the interface. Intuitively we might anticipate that this ''slowed'' motion would yield an effective, higher, activation energy for motion, which may explain why previous[1] activation energies (~0.82 eV) are substantially larger than the value ascertained in this work (~0.38 eV). We furthermore underline that in the present work we have obtained a value for the activation energy in the early stages of proton motion, whereas radiation-induced proton measurements involved the late stages of proton motion where the full influence of the dispersive nature would be felt.

Finally, we address the problem of the atomic nature of proton motion. In the SiO_2 network, the intratetrahedral O-O distance, or suggested shortest proton hop distance, is ~0.26 nm. In the present work we have determined that the fundamental hopping distance (R_{fi}) is at least twice the intrinsic O–nearest-neighbor-O distance. Calculations[16] assuming the shortest-distance hop suggest that the associated activation energy is $1\leqslant E_a\leqslant 1.2\,\mathrm{eV}$, a satisfactory range of values given the previously assumed activation energy ~0.82 eV. The significantly lower value of activation energy obtained in the present experiments and the larger hop distance cannot, then, be reconciled with first-nearest-neighbor hopping. An alternative explanation is required. The results would appear to indicate that a lower-energy diffusion path is available, perhaps involving O–next-nearest-neighbor-O distances (cross-ring hopping).[17] More detailed theoretical studies of this point are required.

The authors are extremely grateful to Professor David Emin of the Physics Department of the University of New Mexico for helpful discussions and comments. Part of this work was supported by the U.S. Department of Energy under Contract No. DE-AC04-94AL85000. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy.

---

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

[2] N. S. Saks and D. B. Brown, IEEE Trans. Nucl. Sci. **NS-36**, 1848 (1989).

[3] D. B. Brown and N. S. Saks, J. Appl. Phys. **70**, 3734 (1991).

[4] R. E. Stahlbush, A. H. Edwards, D. L. Griscom, and B. J. Mrstik, J. Appl. Phys. **73**, 658 (1993).

[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 (London) **386**, 587 (1997).

[6] N. F. M. Devine, J. Robertson, V. Girault, and R. A. B. Devine, Phys. Rev. B **61**, 15 565 (2000).

[7] R. A. B. Devine, K. Vanheusden, and G. V. Herrera, Appl. Phys. Lett. **77**, 579 (2000).

[8] R. E. Stahlbush, R. K. Lawrence, and H. L. Hughes, IEEE Trans. Nucl. Sci. **NS-45**, 2398 (1998).

[9] SMARTCUT is a registered trademark for a silicon-on-insulator wafer marketed by SOITEC, S.A., Parc Technologique des Fontaines, 38926 Crolles Cedex, France.

[10] P. J. McWhorter and P. S. Winokur, Appl. Phys. Lett. **48**, 133 (1986).

[11] W. L. Warren, K. J. R. Vanheusden, D. M. Fleetwood, and R. A. B. Devine, U.S. Patent No. 58305875 (1996).

[12] R. A. B. Devine, J. Appl. Phys. **88**, 2862 (2000).

[13] H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).

[14] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981), Chap. 8.

[15] D. Emin, Phys. Rev. B **43**, 11 720 (1991).

[16] H. Kurtz and S. P. Karna, IEEE Trans. Nucl. Sci. **NS-46**, 1574 (1999).

[17] H. Kurtz (private communication).