IEEE ELECTRON DEVICE LETTERS, VOL. EDL-4, NO. 10, OCTOBER 1983

Creation and Termination of Substrate Deep Depletion in Thin Oxide MOS Capacitors by Charge Tunneling

M. S. LIANG, C. CHANG, Y. T. YEOW, C. HU, MEMBER, IEEE, AND R. W. BRODERSEN, FELLOW, IEEE

Abstract—**Deep depletion in both p-type and n-type substrates can be induced by minority carriers tunneling away from the substrate. When this occurs, tunneling current becomes saturated at the rate of carrier generation in the substrate, with the excess applied voltage dropped across the deep-depletion region. We present a quantitative model for this phenomenon based on balancing the tunneling current and the space-charge generation current. Conversely, the usual transient deep depletion in n-type substrate MOS capacitors can be terminated by tunneling-induced electron-hole pair generation, except for those with ultrathin oxides (<40 Å).**

THE SATURATION of gate tunneling current at high positive bias in p-type substrate MOS capacitors has been reported previously by several authors [1]-[3]. The phenomenological explanation was known due to the limitation of the substrate minority carrier generation. In this work, a more clear and quantitative treatment is presented. New phenomena of hole tunneling through ultrathin oxide as well as the termination of substrate deep depletion in n-type substrate MOS capacitor due to tunneling-induced pair generation will also be discussed.

The silicon-gate samples are arsenic-doped poly MOS capacitors fabricated on ⟨100⟩ p-type substrate with T_{ox} of ~100 Å and substrate resistivity of 5-10 Ω·cm and on ⟨100⟩ n-type substrate with T_{ox} of ~108 Å and doping of $1.3 \times 10^{17}/\text{cm}^3$, respectively. The metal gate samples are fabricated on ⟨100⟩ n-type substrate with T_{ox} of ~31 Å and substrate resistivity of 5-10 Ω·cm. The gate oxides are grown in dry oxygen ambient at 900°C followed by 10 min N_2 annealing at the same temperature. They are all sintered in forming gas at 400°C for 15 min after metallization. The thicknesses of the oxide are determined by ellipsometry and checked with low-frequency *CV* measurements.

We shall first examine the case in which electrons tunnel from the p-type substrate toward a positively biased gate. The Fowler-Nordheim (F-N) tunneling current J_g from the inversion layer of a p-type MOS capacitor is given by Weinberg [4] as

$$J_g = q \cdot \Theta \cdot N_{\text{inv}} \cdot Q(E_0) \tag{1}$$

where q is the electronic charge, E_0 is the lowest quantized energy level in the inversion well, Θ is the fraction of N_{inv} residing at energy level E_0, N_{inv} is the inversion carrier surface concentration, $Q(E_0)$ is the probability current of the electron at E_0. $Q(E_0)$ is also equal to $A \cdot E_{\text{ox}} \cdot e^{-B/E_{\text{ox}}}$, where E_{ox} is the oxide field; A, and B are constants.

Under steady-state conditions, this tunneling current must be supplied by the net electron-hole pair generation in the space charge region as given by (2). Bulk generation is assumed negligible in this study.

$$J_g = J_{\text{generation}} = q \cdot N_i \cdot (X_d - X_{d0})/2\tau_n \tag{2}$$

where τ_n is the minority carrier generation lifetime, X_d is the depletion width for an applied voltage V_g, and X_{d0} is the equilibrium depletion width of an inverted surface. Note that, at equilibrium, the generation rate is exactly balanced by the recombination rate (including the surface-effect components). Therefore, the net generation current J_n is zero at equilibrium, as expected from (2).

The applied voltage V_g is divided between the oxide voltage drop V_{ox} and semiconductor surface potential Ψ_s

$$V_g - \Phi_{ms} = \Psi_s + V_{\text{ox}} = \frac{0.5 q N_A X_d^2}{\epsilon_{\text{Si}}} + E_{\text{ox}} \cdot T_{\text{ox}} \tag{3}$$

where Φ_{ms} is metal–semiconductor work function difference, E_{ox} is the oxide field which is equal to $(q \cdot N_{\text{inv}} + q \cdot X_d \cdot N_A + Q_f)/\epsilon_{\text{ox}}$, and Q_f is the interface fixed charge.

The experimental points of a typical MOS capacitor tunneling current versus voltage relationship are plotted in Fig. 1. The theoretical J_g-V_g fit is obtained by solving (1), (2), and (3) using numerical iteration. This is drawn as the solid curve (curve *a*) in Fig. 1 for $T_{\text{ox}} = 100$ Å, $N_A = 2 \times 10^{15}/\text{cm}^3$, and $\tau_n = 0.5$ µS (the only fitting parameter). Our theoretical solution also indicates that a layer of constant inversion charges is always present at the Si-SiO$_2$ interface throughout the deep-depletion regime, as plotted in Fig. 1. For $V_g < 7$ V, the tunneling current is relatively small and can be adequately supplied by the generation current without noticeable increase in the width of the space-charge region and the capacitor can

Manuscript received May 20, 1983; revised July 29, 1983. This work was supported by the DARPA under Grant N00039-81-K-0251 and by JSEP under Grant F49620-79-C-0178.

The authors are with the Department of Electrical Engineering and Computer Science Electronic Research Laboratory, University of California, Berkeley, CA 94720.

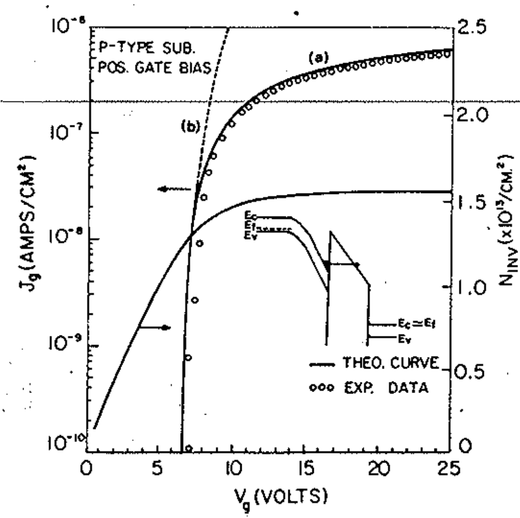

Fig. 1. Theoretical and experimental J_g-V_g plots of n+ gate-100-Å oxide p-type substrate MOS capacitor. A nearly constant inversion charge which is present throughout the deep-depletion regime is shown. The inset shows the energy band diagram of a MOS capacitor under the deep-depletion condition.

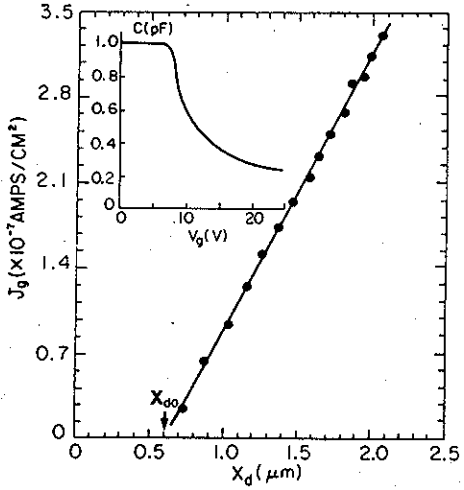

Fig. 2. Experimental plot of $J_g(V_g)$ versus depletion width $X_d(V_g)$. The inset shown experimental high-frequency CV plot for the same sample described in Fig. 1. Only the inversion and deep-depletion regions are represented in the CV curve ($V_T \sim 0.2$ V, $V_{FB} \sim -0.9$ V).

be considered to be in equilibrium. For $V_g > 7$ V, the depletion width has to expand in order to supply this tunneling current. Applied voltage in excess of the 7 V is dropped across the deep depletion region, and the J_g-V_g curve exhibits saturation. The energy band diagram depicting such tunneling process under deep depletion is shown in the inset of Fig. 1. The dashed curve (curve (*b*)) in Fig. 1 represents the ideal (F-N) curve of this capacitor (as though the generation current is infinite in supply) for comparison.

The change in the depletion width X_d with applied voltage can be monitored with the standard high-frequency CV measurement. This is shown in the inset of Fig. 2. It is clear that below $V_g \sim 7$ V the capacitance is equal to the equilibrium high-frequency capacitance in inversion. For $V_g > 7$ V, the depletion layer widens in order to maintain current continuity and hence the CV plot exhibits the deep-depletion characteristics.

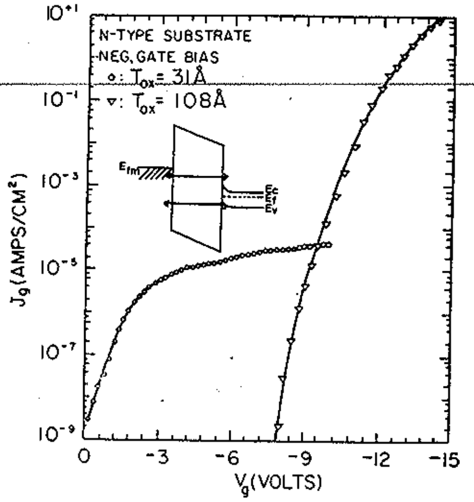

Fig. 3. Tunneling characteristics of a metal gate 31-Å and a poly-gate 108-Å oxide n-type substrate MOS capacitors with areas of 760 μm^2 and 575 μm^2, respectively. The saturation of tunneling current in 31-Å oxide indicates the tunneling of holes from the valence band edge of the substrate which is shown in the inset of the figure.

Equation (2) indicates that a plot of $J_g(V_g)$ from the I-V measurement versus $X_d(V_g)$ from the CV measurement should yield a straight line with a slope of $N_i/2\tau_n$. Such a plot is shown in Fig. 2. It is clear that the data points indeed fall on a straight line giving $\tau_n \sim 0.5$ μS. The theory indicates that N_{inv} and E_{ox} remain essentially constant in the deep-depletion region for $V_g > 7$ V, therefore Ψ_s is approximately equal to $(V_g - 7)$ V. Using this and one-sided abrupt junction relationship $N_A(X_d) = 2/q\epsilon_{Si}A^2 [d(1/C^2)/dV_g]$ the doping profile near the surface can be determined from the CV plot. The extracted dopant concentration of $2 \times 10^{15}/\text{cm}^3$ agrees with the 5-10-Ω·cm resistivity of the p-type substrate used.

The tunneling-induced deep-depletion phenomenon is also observed in ultrathin gate oxide (<40 Å), n-type substrate MOS capacitors under slow negative gate voltage ramp. The tunneling characteristics of a 31-Å oxide along with a 108-Å oxide MOS capacitors are shown in Fig. 3. In order for the substrate to become deep depleted under slow gate voltage ramp, holes must be removed from the surface inversion layer by the tunneling of holes from the substrate, as depicted in the inset of Fig. 3. Experimental results by Faraone *et al.*, using a p-channel MOSFET incorporating with a metal/tunnel-oxide/n-substrate device, indeed indicate significant hole tunneling currents in ultrathin oxides (<22.5 Å) at low gate voltage [5]. In our 31-Å oxide capacitor, the measured current is believed to be mostly electron tunneling current from the gate since the hole current (limited by the generation current) from the substrate is not expected to be more than 1×10^{-7} Å/cm^2. However, the amount of this hole current must have become comparable to the space-charge generation current to cause the substrate to deep-deplete and the current to saturate. In thicker oxide, such as the 108-Å case, the hole tunneling becomes practically zero for which the J_g is entirely due to the tunneling of electrons from the gate. Under such circumstances, equilibrium at the substrate surface will be maintained as long as the ramp rate is sufficiently slow and no deep depletion will occur. This is why the 108-Å tunneling

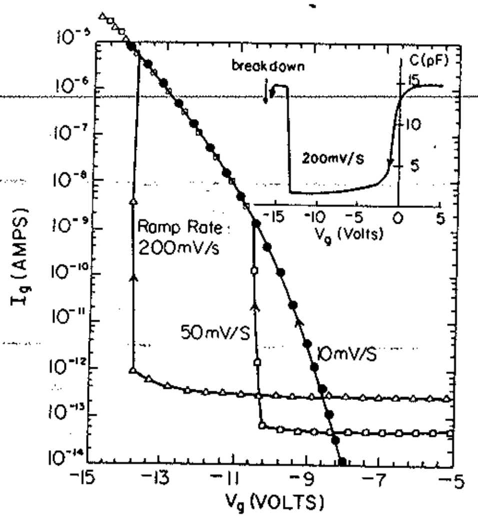

Fig. 4. Experimental I_g–V_g plots of negative gate biased n-type substrate MOS capacitor with area of 575 μm² for different ramp rates. The abrupt rises in I_g for 50 mV/s and 200 mV/s ramp rates demonstrate the terminations of transient deep depletion by tunneling-induced carrier generation. The inset shows the high-frequency *CV* plot of a capacitor with a area of 5000 μm² at 200 mV/s ramp rate.

curve in Fig. 3 follows the classical F-N relationship. In addition, as we shall explain later, pair productions will always occur in the substrate of thicker oxide capacitors due to impact ionizations by energetic electrons tunneling from the gate [6]. Hence the generation of holes in the substrate becomes further enhanced to keep the substrate free from any deep depletion. However, electron-hole pair productions usually do not occur in MOS capacitors with ultrafine oxides because the energy of the incident electrons is below the threshold energy for impact ionization [7].

A related but opposite phenomenon is also observed in thicker oxide, n-type substrate MOS capacitors under negative gate bias. As we know, substrate deep depletion can occur if the gate ramp rate is sufficently fast such that the substrate hole generation current can not keep up with the *CdV/dt* capacitor charging current. However, this transient deep depletion (created by fast ramping of V_g) can be terminated at a critical V_g depending on dV_g/dt. The corresponding I_g–V_g curves are shown in Fig. 4. The termination of deep depletion is not due to avalanche breakdown in the substrate since the avalanche breakdown voltage is expected to be in excess of 100 V and the values of C and I_g after the interruption indicates a total disappearance of the deep-depletion region beyond the critical voltage V_g. As this critical gate voltage is approached during the ramping, significant amount of electron-hole pairs is generated near the substrate surface by impact ionization of the energetic electrons tunneling from the gate. Indeed we have measured the quantum yield (number of electron-hole pairs generated per incident electron) of tunneling induced pair generation to be between 0 ~ 2 [7] depending on the thickness of the oxide. Electrons, being the majority carriers, will be swept into the bulk and the holes will drift to the surface inversion layer. This leads to a further build-up of the oxide voltage. Consequently, electron tunneling from the gate becomes further enhanced, resulting in a positive feedback process. Once the surface inversion layer is fully established the current will revert back to the usual F-N curve. The dependence of the critical V_g on the ramp rate is obviously related to the time required for building up certain amount of holes in the inversion layer. Therefore, the faster the ramp rate is, the higher the critical gate voltage will be.

The electron-hole pair generation by tunneling electrons is also evident in the high frequency *CV* curve. The inset in Fig. 4 shows that the capacitance jumps from the depletion capacitance to oxide capacitance C_{ox} at the critical gate voltage, as a result of the abundant supply of holes. One might note, however, that the termination of deep depletion should not be observed in inverted p-type substrate MOS capacitors because, in this case, the electrons are tunneling from the substrate and there are no electron-hole pair generations occurring in the substrate.

ACKNOWLEDGMENT

The authors would like to thank P. Ko and T. Y. Chiu for many useful discussions.

REFERENCES

[1] H. S. Lee, *Solid-State Electron.*, vol. 22, p. 385, 1979.
[2] P. M. Solomom, *Appl. Phys. Lett.*, vol. 30, p. 597, 1977.
[3] S. Davidoff, I. Kashat, and K. Klein, *Appl. Phys. Lett.*, vol. 34, p. 782, 1979.
[4] Z. A. Weinberg, *Solid-State Electron.*, vol. 20, p. 11, 1977.
[5] L. Faraone, F. L. Hsueh, and J. G. Simmons, presented at INFOS Conf., Eindhoven, The Netherlands, 1983.
[6] Z. A. Weinberg, W. C. Johnson, and M. A. Lampert, *Appl. Phys. Lett.*,
[7] C. Chang, M. S. Liang, C. Hu, and R. W. Brodersen, to be published.