Ultra-thin Silicon Dioxide Leakage Current and Scaling Limit

K.F. Schuegraf, C.C. King, and C. Hu

Department of Electrical Engineering and Computer Science

University of California, Berkeley, CA 94720

Abstract – Modifications are made to Fowler-Nordheim tunneling current analysis to accurately model the measured conduction characteristics of insulator layers thinner than 6 nm. The most significant is direct tunneling for which a closed form expression is introduced. Polysilicon depletion and electron wave interference are also considered. 4 nm is found to be a practical limit for SiO_2 scaling due to direct tunneling leakage almost independent of power supply voltage.

Introduction

The success of silicon dioxide as a gate insulator in MIS structures is partly due to its rather wide forbidden band and high energy barriers at electrode interfaces. Consequently silicon dioxide approaches the ideal case of electrode limited conduction, i.e. Fowler-Nordheim (FN) tunneling, in contrast to films of silicon nitride or tantalum pentoxide which are characterized by bulk-limited mechanisms, such as Frenkel-Poole emission. Excellent agreement with FN Theory for oxide films in excess of 6 nm establishes this claim. For thinner oxide films, FN tunneling is no longer adequate for accurate modeling of oxide leakage currents. Yet, leakage current may limit the scaling of MOSFET gate oxide, not to mention the EEPROM tunneling oxide.

Polysilicon Gate Depletion

Test capacitors of varying area and oxide thickness were studied. Thickness measured by CV and ellipsometry showed uniformity across each wafer of ±1 Å. The standard FN expression

$$J = A\ E_{ox}^2\, e^{-\frac{B}{E_{ox}}}. \qquad (1)$$

$$B = \frac{8\pi\sqrt{2m_{ox}}}{3hq}\ \phi_b^{3/2}. \qquad (2)$$

cannot fit the measured J-V in Fig. 1 if one takes E_{ox} as V_g/x_{ox} even when B is treated as a fitting parameter. Good fit was obtained by including polysilicon depletion [1] even though the polysilicon was in-situ doped with phosphorus as opposed to implanted with arsenic.

$$E_{ox}x_{ox} = V_{ox} = V_g - V_{poly} \qquad (3)$$

where $V_{poly} = \frac{\varepsilon_{ox}^2\, E_{ox}^2}{2q\varepsilon_{Si}N_{poly}}$ until V_{poly} is pinned at 1.12 V due to the saturation of band bending in strong inversion. Of course, (3) must be modified for other polarities and doping types. [2]

Direct Tunneling Model

Fig. 2 illustrates the difference between direct and FN tunneling. Although direct tunneling has been analyzed with Airy functions and numerical integration before [3], we have derived a closed form expression with the use of the WKB and other approximations. The current density is unchanged from (1) for $V_{ox} > \phi_b$ and increased by the following factor for $V_{ox} < \phi_b$

$$\frac{1}{[1 - (\frac{\phi_b - V_{ox}}{\phi_b})^{1/2}]^2}\, e^{\frac{B(\frac{\phi_b - V_{ox}}{\phi_b})^{3/2}}{E_{ox}}}. \qquad (4)$$

Fig. 3 shows that the seemingly insignificant change from triangular to trapezoidal barrier caused a dramatic change in J-V slope and can increase the leakage current by orders of magnitude below 3.5 V. Fig. 4 shows the good fit between the closed form model and the data.

Electron Wave Interference

Fig. 5 illustrates the physical basis for this model. The tunneling electron sees the FN triangular potential, but upon emission into the oxide conduction band, it establishes a partial standing wave between reflection-inducing potential discontinuities. The WKB transmissivity becomes modified by a multiplying factor

$$\frac{(1 + R)^2}{(1 - R)^2 + 4\,R \sin^2 \frac{\left(\frac{B(\frac{V_{ox} + E_{tr} - \phi_b}{\phi_b})^{3/2}}{E_{ox}} - \frac{\pi}{2}\right)}{2}}. \qquad (5)$$

R accounts for both wave reflection at the oxide silicon interface and wave attenuation, phenomenologically modeled by an oxide mean free path, $\lambda_{scat} = 1.5$ nm. A previous scattering model [4] reported $\lambda_{scat} = 3.2$ nm, suggesting, however, that this bulk value may degrade near the oxide interface. Fig. 6 compares data with this model. Fig. 7 presents an excellent fit between this model, varying only x_{ox}, for the range 3 to 6 nm. The subtle change in shape of the measured characteristic between the 4.1 and 4.3 nm oxides can be attributed to interference. This interference undulation in quantitative agreement with theory may therefore be used to monitor dielectric thickness uniformity and interface flatness to within at least 0.2 nm.

The absence of quantum interference in the FN regime for large x_{ox} can be explained as increased carrier scattering in the oxide so that resonant waves are attenuated and single pass theory suffices. Moreover, while under the barrier, e.g. in direct tunneling, electrons do not demonstrate the interference effect because the electron wave number is imaginary.

Voltage Limit due to Oxide Leakage

The high electric fields present in scaled devices mandate voltage scaling, in order to ensure oxide reliability. Presently, TDDB lifetimes dictate maximum operating voltage[5]; subsequently, this may be limited by gate leakage. Defining V_{max} as the voltage necessary to induce a 0.1pA/μ^2 gate leakage current, leads to the thickness dependent voltage scaling requirement given in Fig. 8. The exceedingly rapid drop of acceptable voltage below 4 nm suggests that 4 nm is a practical limit for oxide scaling in ULSI applications. Furthermore, Fig. 8 establishes the convergence of the intrinsic TDDB and gate leakage criteria and raises the possibility that gate leakage will set the ultimate limit to oxide scaling at 4 nm.

Acknowledgements

This research is supported by AMD and Signetics under MICRO, JSEP under contract F49620-87-0041 and ISTO/SDIO administered by ONR under contract N00014-85-K-0603.

References

[1] S.J. Wang, I.C. Chen, and H.L. Tigelaar, "Effects of Poly Depletion on the Estimate of Thin Dielectric Lifetime," *IEEE Electron Device Letters*, pp.617-619, 1991.

[2] E. Rosenbaum, R. Moazzami, and C. Hu, "Implications of Waveform and Thickness Dependence of SiO_2 Breakdown on Accelerated Testing," *Int. Symposium VLSI TSA*, pp.214-218, 1991.

[3] K.H. Gundlach, "Zur Berechnung des Tunnelstroms durch eine Trapezfoermige Potentialstufe," *Solid State Electronics*, pp.949-957, 1966.

[4] C.N. Berglund, and R.J. Powell, "Photoinjection into SiO_2: Electron Scattering in the Image Force Potential Well," *Journal Applied Physics*, pp.573-579, 1971.

[5] R. Moazzami, and C. Hu, "Projecting Gate Oxide Reliability and Optimizing Reliability Screens," *IEEE Transactions Electron Devices*, pp.1643-1650, 1985.

92CH3172-4/92/0000-0018$3.00 © 1992 IEEE

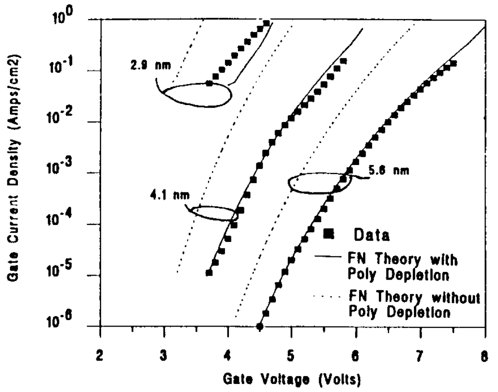

Fig. 1. Illustrates the necessity of including poly depletion correction in ultra-thin gate dielectrics.

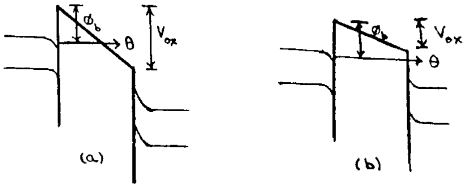

Fig. 2. (a) Fowler-Nordheim tunneling is associated with triangular barrier. (b) Direct tunneling is associated with a trapezoidal barrier, i.e. when $V_{ox} < \phi_b$.

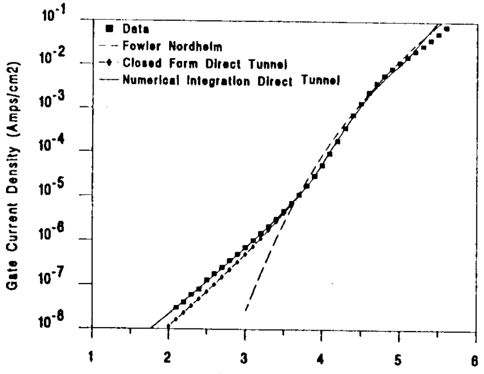

Fig. 3. A closed form tunneling expression for direct tunneling agrees well with data and more exact theory.

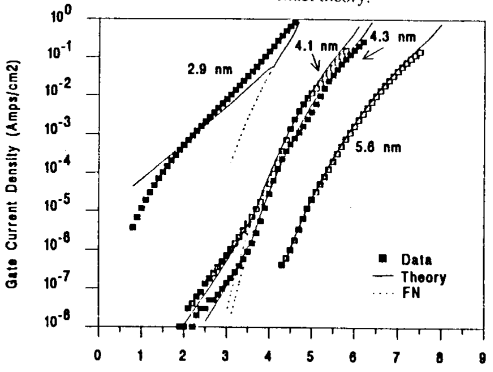

Fig. 4. Closed-form Tunneling Expression correctly shows the large direct tunneling leakage current below 3.5 V.

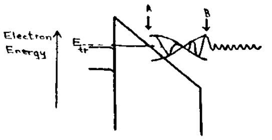

Fig. 5. Interference of waves due to multiple reflection at A and B can cause undulations in the J-V characteristics. E_{tr} denotes transverse energy available to tunneling electron.

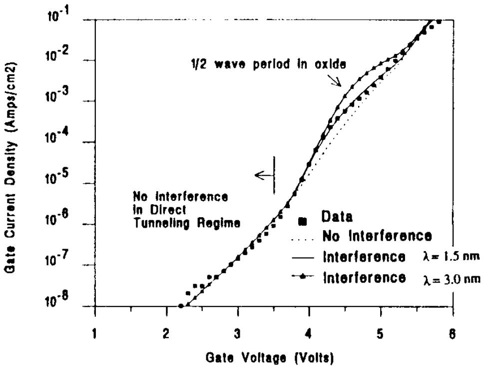

Fig. 6. Establishes improvement of the resonance model over closed-form expression for x_{ox} = 4.3nm

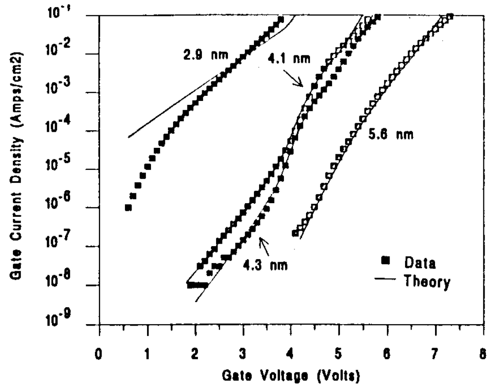

Fig. 7. Fit using numerical integration of resonance term, varying only the oxide thickness.

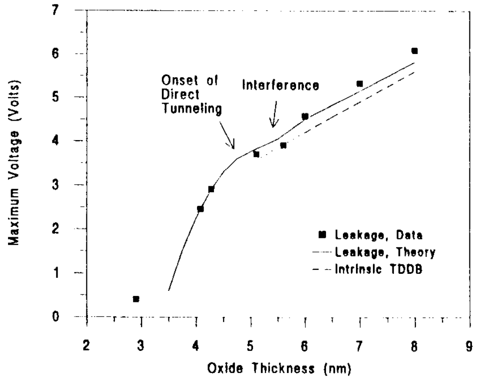

Fig. 8. Maximum acceptable voltages from standpoint of intrinsic TDDB and Gate Leakage. TDDB Data is extrapolated to ultra-thin region, where its validity awaits further study.