

APPLIED PHYSICS REVIEWS

Ultrathin (<4 nm) SiO₂ and Si–O–N gate dielectric layers for silicon microelectronics: Understanding the processing, structure, and physical and electrical limits

M. L. Green^{a)}

Agere Systems, Murray Hill, New Jersey 07974

E. P. Gusev

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

R. Degraeve

IMEC, Leuven 3001, Belgium

E. L. Garfunkel

Rutgers University, Piscataway, New Jersey 08854

(Received 1 March 2001; accepted for publication 21 May 2001)

The outstanding properties of SiO₂, which include high resistivity, excellent dielectric strength, a large band gap, a high melting point, and a native, low defect density interface with Si, are in large part responsible for enabling the microelectronics revolution. The Si/SiO₂ interface, which forms the heart of the modern metal–oxide–semiconductor field effect transistor, the building block of the integrated circuit, is arguably the world's most economically and technologically important materials interface. This article summarizes recent progress and current scientific understanding of ultrathin (<4 nm) SiO₂ and Si–O–N (silicon oxynitride) gate dielectrics on Si based devices. We will emphasize an understanding of the limits of these gate dielectrics, i.e., how their continuously shrinking thickness, dictated by integrated circuit device scaling, results in physical and electrical property changes that impose limits on their usefulness. We observe, in conclusion, that although Si microelectronic devices will be manufactured with SiO₂ and Si–O–N for the foreseeable future, continued scaling of integrated circuit devices, essentially the continued adherence to Moore's law, will necessitate the introduction of an alternate gate dielectric once the SiO₂ gate dielectric thickness approaches ~1.2 nm. It is hoped that this article will prove useful to members of the silicon microelectronics community, newcomers to the gate dielectrics field, practitioners in allied fields, and graduate students. Parts of this article have been adapted from earlier articles by the authors [L. Feldman, E. P. Gusev, and E. Garfunkel, in *Fundamental Aspects of Ultrathin Dielectrics on Si-based Devices*, edited by E. Garfunkel, E. P. Gusev, and A. Y. Vul' (Kluwer, Dordrecht, 1998), p. 1 [Ref. 1]; E. P. Gusev, H. C. Lu, E. Garfunkel, T. Gustafsson, and M. Green, *IBM J. Res. Dev.* **43**, 265 (1999) [Ref. 2]; R. Degraeve, B. Kaczer, and G. Groeseneken, *Microelectron. Reliab.* **39**, 1445 (1999) [Ref. 3]. © 2001 American Institute of Physics. [DOI: 10.1063/1.1385803]

TABLE OF CONTENTS

I. INTRODUCTION AND OVERVIEW.....	2058	E. Scanning probe microscopy.....	2073
A. SiO ₂ enabled the microelectronics revolution...	2058	III. ELECTRICAL CHARACTERIZATION	
B. Fundamental limits of SiO ₂ : Is the end in sight?.....	2060	METHODS AND PROPERTIES.....	2074
II. PHYSICAL CHARACTERIZATION METHODS FOR ULTRATHIN DIELECTRICS.....	2062	A. Characterization methods.....	2075
A. Optical methods.....	2062	1. C–V measurements.....	2075
B. X-ray based methods.....	2065	2. Gate tunnel current.....	2076
C. Ion beam analysis.....	2067	3. Charge pumping.....	2077
D. Electron microscopy.....	2072	B. Oxide degradation during electrical stress....	2078
		1. Interface trap creation.....	2078
		2. Oxide charge trapping.....	2078
		3. Hole fluence.....	2078
		4. Neutral electron trap generation.....	2079
		5. Stress-induced leakage current.....	2080
		6. Discussion of trap generation mechanisms..	2081

^{a)}Electronic mail: mlg@agere.com

C. Oxide breakdown. 2082

 1. Breakdown modeling. 2082

 2. Soft breakdown. 2083

 3. Breakdown acceleration models. 2084

 4. Temperature dependence of breakdown. 2085

 5. Oxide reliability predictions. 2086

IV. FABRICATION TECHNIQUES FOR ULTRATHIN SILICON OXIDE AND OXYNITRIDES. 2086

 A. Surface preparation. 2086

 B. Fabrication of ultrathin oxide and oxynitrides. 2086

 1. Thermal oxidation and oxynitridation. 2086

 2. Chemical deposition. 2088

 3. Physical deposition. 2090

 C. Postoxidation processing and annealing. 2090

 D. Hydrogen/deuterium processing. 2091

V. THE Si/SiO₂ SYSTEM. 2093

 A. The initial stages of oxygen interaction with silicon surfaces. 2093

 1. The passive oxidation regime. 2094

 2. The active oxidation regime. 2098

 3. The transition regime. 2099

 B. Growth mechanisms of ultrathin oxides, beyond the Deal–Grove model. 2100

VI. THE Si/SiO_xN_y SYSTEM. 2104

 A. Oxynitride properties. 2104

 1. Thermodynamics of the Si–O–N system. 2104

 2. Physical properties. 2105

 3. Diffusion barrier properties of oxynitride layers. 2105

 B. Oxynitridation processes. 2106

 1. NO processing. 2106

 2. N₂O processing. 2107

 a. Gas phase N₂O decomposition at high temperatures. 2107

 b. N incorporation and removal during N₂O oxynitridation. 2108

 3. Nitridation in NH₃. 2109

 4. Nitridation in N₂. 2109

VII. THE POST-SiO₂ ERA: ALTERNATE GATE DIELECTRICS. 2109

 A. Si₃N₄. 2110

 B. Alternate (higher) dielectric constant materials. 2110

TABLE I. Selected properties of SiO₂ gate dielectric layers.

Native to silicon (SiO ₂ is the only stable oxide phase on Si)
Low interfacial (Si/SiO ₂) defect density ($\sim 10^{10}$ eV ⁻¹ cm ⁻² , after H ₂ passivation)
Melting point=1713 °C
Energy gap=9 eV
Resistivity $\sim 10^{15}$ Ω cm
Dielectric strength $\sim 1 \times 10^7$ V/cm
Dielectric constant=3.9

I. INTRODUCTION AND OVERVIEW

A. SiO₂ enabled the microelectronics revolution

Nature has endowed the silicon microelectronics industry with a wonderful material, SiO₂, as is shown in Table I. SiO₂ is native to Si, and with it forms a low defect density interface. It also has high resistivity, excellent dielectric strength, a large band gap, and a high melting point. These properties of SiO₂ are in large part responsible for enabling the microelectronics revolution. Indeed, other semiconductors such as Ge or GaAs were not selected as the semiconducting material of choice, mainly due to their lack of a stable native oxide and a low defect density interface. The metal–oxide–semiconductor field effect transistor (MOSFET), Fig. 1, is the building block of the integrated circuit. The Si/SiO₂ interface, which forms the heart of the MOSFET gate structure, is arguably the worlds most economically and technologically important materials interface. The ease of fabrication of SiO₂ gate dielectrics and the well passivated Si/SiO₂ interface that results have made this possible. SiO₂ has been and continues to be the gate dielectric par excellence for the MOSFET. Figure 2 is a transmission electron photomicrograph of an actual submicron MOSFET, showing the SiO₂ gate dielectric as well as the Si/SiO₂ interface.

In spite of its many attributes, however, SiO₂ suffers from a relatively low dielectric constant (κ =dielectric constant, or permittivity, relative to air=3.9). Since high gate dielectric capacitance is necessary to produce the required drive currents for submicron devices,^{1,4} and further since capacitance is inversely proportional to gate dielectric thickness, the SiO₂ layers have of necessity been scaled to ever thinner dimensions, as is shown in Fig. 3. This gives rise to a number of problems, including impurity penetration through the SiO₂, enhanced scattering of carriers in the

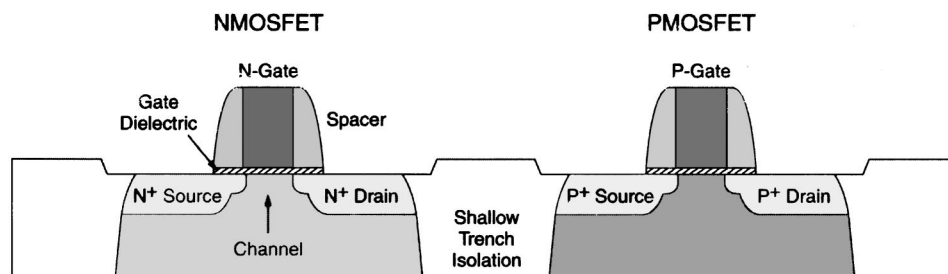


FIG. 1. Schematic illustration of a submicron (channel length) CMOSFET (complementary metal–oxide–semiconductor field effect transistor) (courtesy of C. P. Chang, Agere Systems).

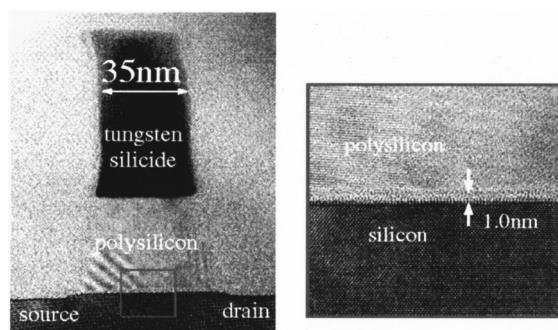
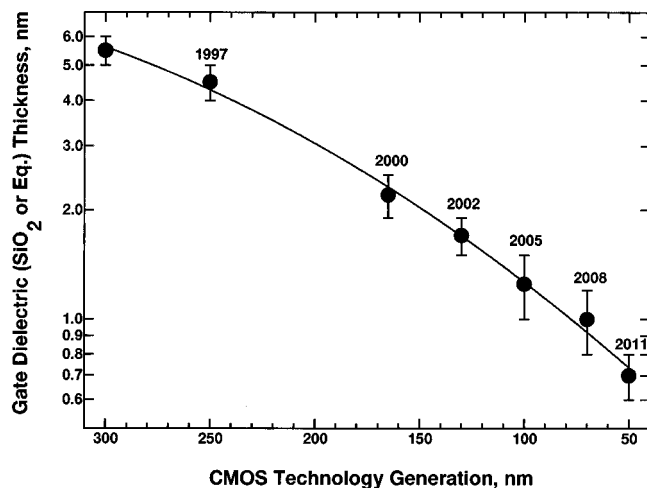


FIG. 2. Cross-section transmission electron photomicrographs of a 35 nm (channel length) transistor, and a detailed view of its 1.0 nm SiO₂ gate dielectric and Si/SiO₂ interface [from Timp *et al.* (Ref. 15)].

channel, possible reliability degradation, high gate leakage current, and the need to grow ultrathin and uniform SiO₂ layers.

This article summarizes recent progress and current scientific understanding of ultrathin (<4 nm) SiO₂ and Si-O-N gate dielectrics on silicon based devices. (Si-O-N will refer to oxynitride dielectrics for which $N/(N+O) \leq 0.5$, and most often much less than that.) We will emphasize an understanding of the limits of these gate dielectrics, i.e., how their continuously shrinking thickness, dictated by integrated circuit device scaling, results in physical and electrical property changes that impose limits on their usefulness. We will also discuss the near-future need for alternate gate dielectric materials such as Si₃N₄ and other metal oxides and nitrides, known as “high κ ” materials.

In a continuous drive to increase integrated circuit performance through shrinkage of the circuit elements, the dimensions of MOSFETs and other devices have been scaled since the advent of integrated circuits about 40 years ago, according to a trend known as Moore’s law.^{5–8} Moore’s law describes the exponential growth of chip complexity due to decreasing minimum feature size, accompanied by concurrent improvements in circuit speed, memory capacity, and cost per bit. SiO₂ gate dielectrics have decreased in thickness



from hundreds of nanometers (nm) 40 years ago to less than 2 nm today, to maintain the high drive current and gate capacitance required of scaled MOSFETs. Further, as can be seen in Fig. 3, SiO₂ thickness continues to shrink. Many ultrasmall transistors have been reported, with SiO₂ layers as thin as 0.8 nm.^{9–15} The International Technology Roadmap for Semiconductors,¹⁶ excerpted in Table II, shows that SiO₂ gate dielectrics of 1 nm or less will be required within 10 years. [It will become obvious while reading this article that SiO₂ layers thinner than ~1.2 nm may not have the insulating properties required of a gate dielectric. Therefore alternate gate dielectric materials, having “equivalent oxide thickness” less than 1.2 nm (for example), may be used. Equivalent oxide thickness, $t_{ox}(eq)$, is the thickness of the SiO₂ layer ($\kappa=3.9$) having the same capacitance as a given thickness of an alternate dielectric layer, t_{diel} ($\kappa=\kappa_{diel}$). Equivalent oxide thickness is given by the relationship: $t_{ox}(eq) = t_{diel}(3.9/\kappa_{diel})$.] At these thicknesses, the Si/SiO₂ interface becomes a more critical, as well as limiting, part of the gate dielectric. A 1 nm SiO₂ layer is mostly interface, with little if any bulk character. It contains about five layers of Si atoms, at least two of which reside at the interface.¹⁷ Given its prominence, much of this article will focus on the physical and electrical properties of the Si/SiO₂ interface.

Due to its commercial relevance, the Si/SiO₂ system has received an enormous amount of scientific attention. It is daunting to count the number of scientific papers: using an INSPEC Database we found 36 708 references (since 1969) devoted to this system. [We intersected the set “silicon” with the set “(SiO or SiO₂ or SiO_x)” and then subtracted the set “quartz.”] Only 2% of these references are cited in this article. Several excellent books^{18–23} and review papers on dielectric selection,²⁴ atomic scale interactions between Si and O,^{25–27} oxidation of Si,^{28–30} SiO₂ structure,³¹ interface structure and defects,^{32–34} reliability,³⁵ metrology,³⁶ Si-O-N,^{2,37,38} and general growth, structure and properties^{39,40} have been published. However, some basic scientific issues at the forefront of the field remain unresolved. Among these issues are an understanding of the exact diffusion mechanisms and incorporation reactions of oxidizing and nitriding species in SiO₂, an atomistic understanding of the initial stages of oxidation, the role of postoxidation processing on physical and electrical properties, the bonding structure at and near the Si/SiO₂ interface, the relationship between local bonding/chemistry and electrical defects, and the failure mechanisms in ultrathin dielectrics. All of these topics will be addressed in this article.

It is amusing and instructive to learn that not only is SiO₂ enabling to microelectronics, but also to some forms of life itself. Many forms of animal and plant life have cell membranes and exoskeletons composed of pure, crystalline (opaline) SiO₂.⁴¹ This should not be too surprising, since Si and O are the two most abundant elements in the earth’s crust. In particular, one celled animals called diatoms fashion their cell membrane via self-assembly (molecule by molecule), from SiO₂ dissolved in H₂O, at ambient temperature.⁴² They exist in thousands of symmetric mor-

TABLE II. MOSFET technology timetable, adapted from the International Technology Roadmap for Semiconductors (ITRS) (Ref. 16).

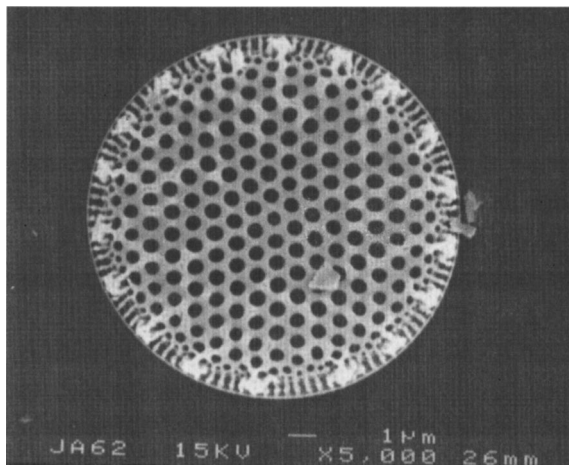
First year of production	1999	2001	2002	2005	2008	2011	2014
DRAM generation	1G	1G	4G	16G	64G	256G	1T
Minimum feature size, nm	180	160	130	100	70	50	35
Equivalent oxide thickness, nm	1.9–2.5	1.5–1.9	1.5–1.9	1.0–1.5	0.8–1.2	0.6–0.8	0.5–0.6

serve SiO₂ features as small as 100 nm. It is humbling to discover that diatoms evolved with such “scaled dimensions” 400 million years ago, and suggests that in the future, SiO₂, as well as other electronic materials structures, might be self-assembled by biomimetic processes. In fact, such structures have already been achieved.^{43,44}

B. Fundamental limits of SiO₂: Is the end in sight?

The use of ultrathin SiO₂ gate dielectrics gives rise to a number of problems, including high gate leakage current, reduced drive current, reliability degradation, B penetration, and the need to grow ultrathin and uniform SiO₂ layers. We may ask when any of these effects will fundamentally limit the usefulness of SiO₂ as a gate dielectric.⁴⁵

Due to the large band gap of SiO₂, ~9 eV, and the low density of traps and defects in the bulk of the material, the carrier current passing through the dielectric layer is normally very low. For ultrathin films this is no longer the case. When the physical thickness between the gate electrode and doped Si substrate becomes thinner than ~3 nm, direct tunneling through the dielectric barrier dominates leakage current.^{11,40,46–48} According to fundamental quantum-mechanical laws, the tunneling current increases exponentially with decreasing oxide thickness. Gate leakage currents measured on 35 nm transistors fabricated using advanced wafer preparation, cleaning, and oxidation procedures¹⁴ are shown in Fig. 5. The leakage current is seen to increase by one order of magnitude for each 0.2 nm thickness decrease.



Assuming a maximum allowable gate current density of 1 A/cm² for desktop computer applications, and 10⁻³ A/cm² for portable applications, minimum acceptable SiO₂ thicknesses (physical) would be approximately 1.3 and 1.9 nm, respectively.

Recent electron energy loss spectroscopy (EELS) experiments on ultrathin Si/SiO₂ interfaces, in a scanning transmission electron microscope (STEM), support these findings.¹⁷ Oxygen profiles across the interface were obtained and are shown in Fig. 6 for SiO₂ films 1.0 and 1.8 nm thick. The profiles consist of bulk-like regions and interfacial regions, and the interfacial regions are thought to be due to interfacial states. Based on this analysis, it has been calculated that the minimum oxide thickness, before leakage current becomes overwhelming, is 1.2 nm. This comes from the fact that a satisfactory SiO₂ tunnel barrier is formed when it is equal in thickness to six times the decay length of the interfacial states, about 6×0.12 nm≅0.7 nm, plus a 0.5 nm contribution from interfacial roughness.

Reduced drive (drain) current has been reported in small transistors with ultrathin gate dielectrics.¹⁰ Figure 7 shows that drain current increases with decreasing SiO₂ thickness, but then falls off; gate leakage current continuously increases, as expected, with decreasing SiO₂ thickness. Thus for SiO₂ layers thinner than about 1.3 nm there is no advan-

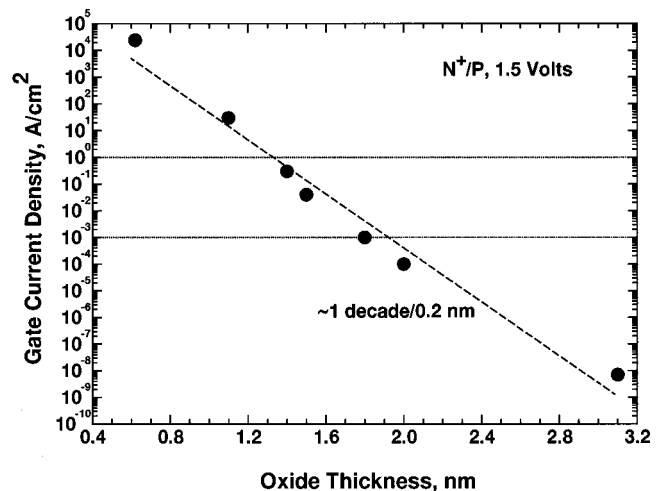


FIG. 5. Gate leakage current measured at 1.5 V as a function of oxide thickness (measured by TEM) for 35 nm NMOSFETs. Leakage current increases one order of magnitude for every 0.2 nm decrease in SiO₂ thickness.

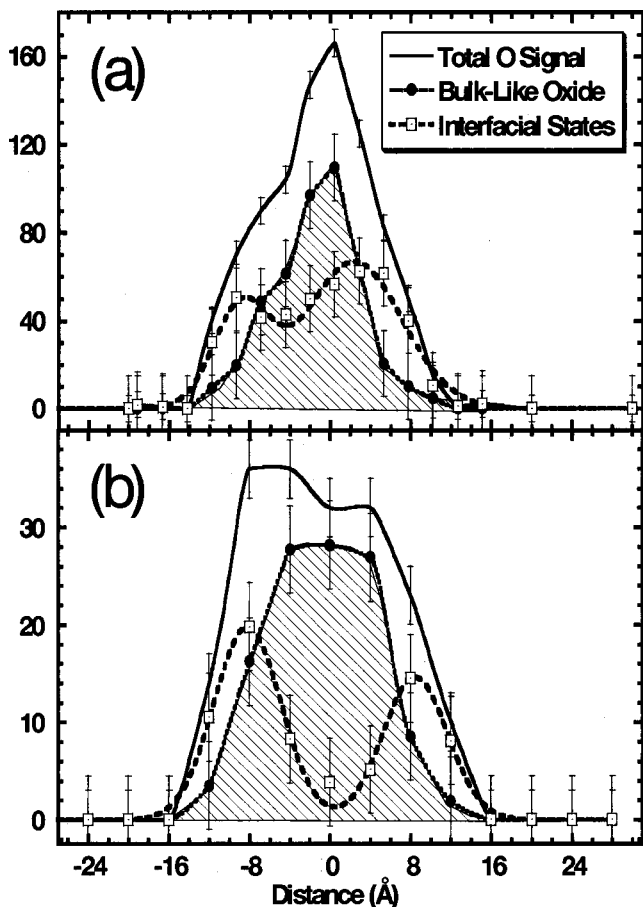


FIG. 6. Oxygen bonding profiles measured by STEM-EELS. The Si substrate is on the left and the gate polycrystalline Si is on the right. (a) 1.0 nm (ellipsometric) oxide, annealed at 1050 °C/10 s. The bulk-like O signal (*y* axis, arbitrary scale) yields a FWHM of 0.85 nm, whereas the total O signal yields a FWHM of 1.3 nm. The overlap of the two interfacial regions has been correlated with the observation of a very high gate leakage current, 10^2 A/cm². (b) a thicker (1.8 nm ellipsometric) oxide, also annealed. The interfacial regions no longer overlap and the gate leakage current is 10^{-5} A/cm² [from Muller *et al.* (Ref. 17)].

tage in performance (drive current) for incurring the burden of an ever-increasing gate leakage current. This would suggest that SiO₂ layers thinner than 1.3 nm no longer deliver any performance advantage. The cause of the decreased

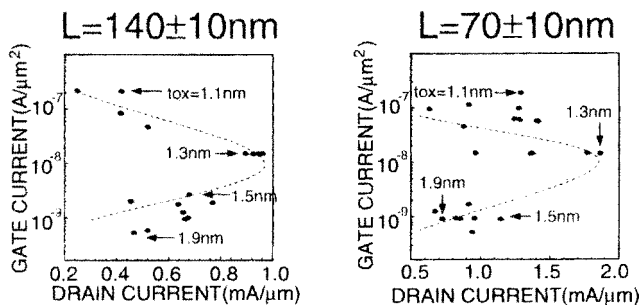


FIG. 7. Drive current vs leakage current for two ultrasmall (gate lengths of 70 and 140 nm) NMOSFETs. In both cases it can be seen that gate current increases, as is expected, with decreasing SiO₂ thickness. However, drain

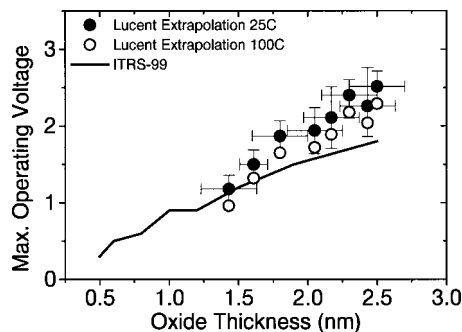


FIG. 8. Predicted maximum allowable operating voltage as a function of oxide thickness (Ref. 16) (solid line), modeled so that requirements of power dissipation and circuit speed for successive technology generations are met. The Lucent data, at either 25 or 100 °C, estimates the safe operating voltages (for 10 year lifetime reliability) for sub-2.5 nm oxides, based on measured results and extrapolation principles (Ref. 62). The Lucent data clearly show that the 10 year reliability of sub-2.5 nm oxides meets the ITRS specifications for SiO₂ layers as thin as 1.4 nm.

drive current is not fully understood. One possibility is an additional scattering component from the upper (SiO₂/polycrystalline Si gate) interface. Some experimental evidence exists for this case⁴⁹ but the observed effect is not enough to explain the data in Fig. 7. Another cause could be a universal mobility curve effect, i.e., lowered mobility due to enhanced scattering because of extreme carrier confinement in the inversion layer of the ultrathin oxide. The issue of long-range electrostatic interactions between charges in very heavily doped gate, source, and drain regions, and electrons traveling in the channel, was recently theoretically studied⁵⁰ using both semiclassical two-dimensional self-consistent Monte Carlo–Poisson simulations and a quantum mechanical model based on electron scattering from gate–oxide interface phonons. It was shown that excitation and absorption of plasma modes in the gate region may result in a net momentum loss of carriers in the channel, thus decreasing their velocity, and leading to reduced drain current.

Reliability (lifetime to breakdown) of ultrathin SiO₂ is a major concern for oxide scaling into the sub-2 nm range and currently a contentious issue.^{35,40,51–62} Electrons traveling through the SiO₂ layer may create defects such as electron traps and interface states^{63–65} that in turn, upon accumulation to some critical density, degrade the insulating properties of the oxide. The accumulated charge the film can withstand before its breakdown (Q_{bd}) decreases with oxide thickness.⁵⁴ Recently, it was predicted that oxide films thinner than about 2.2 nm would not have the reliability required by the industry roadmap.⁵⁴ Data from another research group,⁶² shown in Fig. 8, indicates that acceptable reliability will be achievable for SiO₂ thicknesses as low as 1.4 nm. At about 1.0 nm thickness, the statistical probability of a percolation path may reduce reliability to an unacceptable level.⁶⁶ One should mention that all reliability data is model dependent. Unlike directly measured parameters such as the gate leakage and drive currents, reliability studies always involve extrapolations from relatively high (~2.5–4 V) stress voltages to real device operating voltages (~1.0–1.2 V). The extrapolations

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.