

Key measurements of ultrathin gate dielectric reliability and in-line monitoring

by W. W. Abadeer
A. Bagramian
D. W. Conkle
C. W. Griffin
E. Langlois
B. F. Lloyd
R. P. Mallette
J. E. Massucco
J. M. McKenna
S. W. Mittl
P. H. Noel

High-performance CMOS products depend upon the reliability of ultrathin gate dielectrics. In this paper a methodology for measuring thin gate dielectric reliability is discussed in which the focus is upon the elements of those test structures used in the evaluation, the design of the reliability stress matrix, and the generation of engineering design models. Experimental results are presented which demonstrate the reliability of ultrathin gate dielectrics measured on a wide variety of test structures with dielectric thicknesses ranging from 7 to 3.5 nm. An overview is provided for thin gate oxide reliability that was measured on integrated functional chips—high-performance microprocessors and static random-access memory (SRAM) chips. The data from these measurements spanned the period from early process and device development to full production. Manufacturing in-line monitoring for thin gate dielectric yield and reliability is also discussed, with several case histories presented which show the effectiveness of monitors in detecting process-induced dielectric failures. Finally, causes of oxide fails are discussed, leading to the process actions necessary for controlling thin gate dielectric defects.

Introduction

Thin gate dielectric integrity and reliability are requirements for the development and manufacture of VLSI and ULSI semiconductor devices. Achieving reliable and high-quality thin gate dielectrics requires research and development efforts to meet the demands for smaller device geometry and better device performance. Because the reduction of a chip's power supply voltage has not scaled with device geometry in successive generations, operating electric fields for the dielectric have increased [1]. For many years, thermally grown silicon dioxide has been the primary gate dielectric and has demonstrated robustness and effectiveness. For ultrathin (<3.0-nm) gate dielectrics, oxynitrides (N incorporated in SiO₂), are used because of their greater immunity to electrical stress and suppression of boron penetration [2–4].

There are two mechanisms by which dielectric failure (or breakdown) occurs: extrinsic or intrinsic failure. Extrinsic failures occur with a decreasing failure rate over time and are caused by process defects such as metallic, organic, or other contaminants on the crystalline silicon surface, and by surface roughness [5–7]. Most extrinsic failures occur early in the lifetime of a device and can cause significant reliability problems as the number of dielectric fails exceed a predefined specification. Failure rates generated extrinsically can be reduced by applying a burn-in or voltage screening process to the devices. The failures are accelerated at elevated

©Copyright 1999 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/99/\$5.00 © 1999 IBM

voltages and temperatures. These procedures screen out dielectric defects before the product is shipped, thereby containing early life failure rates within product specifications. Minimizing extrinsic defects in a process and reducing dielectric fails early in the life cycle of a product is important for processing ultrathin gate dielectrics. It is also critical that reliability testing detect systematic and/or random types of process-induced defects that may contribute to an increase in number of dielectric failures. In full-scale product manufacturing, in-line monitoring of both yield and reliability ensures the tight process control of defects to minimize the device and/or chip failures.

The second type of failure for a thin gate dielectric, sometimes referred to as wear-out, is intrinsic to the material. Intrinsic failures occur with an increasing failure rate over time and are usually caused by an inherent imperfection in the dielectric material. It is essential that these fails do not occur during the intended useful lifetime of the device when it is operating under specified conditions. The intrinsic failure of ultrathin gate dielectric has been the subject of intensive research [8–18]. The process of intrinsic degradation in gate dielectric begins with trap creation and formation of interface states as the device gate dielectric is stressed at elevated voltages and temperatures. The creation of these defects continues with the injected current (and therefore time) until the defect density reaches a critical value, after which dielectric breakdown occurs. Parameters that affect when and how the dielectric breakdown occurs include the applied stress voltage, temperature, dielectric thickness, device dielectric area, and intrinsic dielectric lifetime. The procedures and methodologies for characterizing the ultrathin gate dielectrics used in current metal–oxide–semiconductor (MOS) devices are outlined in this paper. Examples of the intrinsic behavior of SiO_2 -based dielectrics in the thickness range of 7 nm to 3.5 nm are also discussed.

During product manufacturing, routine in-line monitoring of thin gate dielectric yield and reliability is essential. The monitoring is intended to detect dielectric problems early in the manufacturing process and to measure the dielectric performance of the final product. In-line monitoring is usually performed at the wafer level (i.e., prior to dicing and bonding of individual chips), using specially designed test structures manufactured on product wafers. It is important that the test structures and the procedures for in-line monitoring be able to detect both intrinsic and extrinsic dielectric defects so that corrective action can be quickly taken.

Experimental procedures

From early device development to qualification and early production, specific test structures for ultrathin gate dielectric characterization are required in order to

understand and achieve the consistency required for manufacturing. A minimum set of test structures is developed so that manufacturing processes can be consistently controlled. Our experimental results are obtained from both short-loop test structures and fully integrated chips in a complementary metal–oxide–semiconductor (CMOS) dual-work-function technology (i.e., n^+ -doped polysilicon gates for n-FETs and p^+ -doped polysilicon gates for p-FETs). The polysilicon minimum-linewidth-lithography-dimension technologies used in this study range from 0.33 μm to 0.2 μm . The nominal thin gate dielectric thickness ranged from 7 nm to 3.5 nm, with the respective power-supply voltages (V_{DD}) ranging from 3.3 V to 1.8 V. The gate dielectrics used for these devices were silicon dioxide (SiO_2) thermally grown in O_2 , although some films were also annealed in N_2O . All technologies investigated used shallow-trench isolation (STI). A more detailed description of the process details can be found in [19].

The following test structures were used to study the extrinsic and intrinsic properties of thin gate dielectrics and to establish engineering models for dielectric intrinsic breakdown:

- Large-area capacitors with both diffusion and STI-bounded thin oxide. Some devices had floating gates (unprotected), while others had the diffusion tied to the gate at first-level metal (protected). Both n^+ and p^+ polysilicon gates were included in these test structures. Some of these gates included polysilicon geometry with 1) large square areas, 2) a number of small rectangles, or 3) narrow finger lines. Each structure contained the same oxide area, with perimeter-to-area ratios (of the thin gate dielectric) varying from approximately 0.01 to 1.0 (and higher).
- Product-like arrays of n-FETs and p-FETs, with each array consisting of many individual FETs, each having a small channel length and large channel width. Several arrays are contained on each chip to allow for the testing of different total thin gate dielectric areas.
- Large-area decoupling capacitors.
- n-FETs and p-FETs with different types of antennas formed by either a single plate or fingers of polysilicon and metal lines. (An antenna structure is a device in which the polysilicon area is significantly larger than the active device gate dielectric area.)

These test structures were designed so that the effect of the oxide area upon the device lifetime could be determined. Test structures with similar layouts but different thin gate dielectric areas were included in these measurements. The largest thin gate dielectric area on a single structure was of the order of several square millimeters. With the use of these different test structures,

the sensitivity of the gate dielectric yield and/or reliability to process layout could be determined. Device problems such as process-induced defects, oxide thinning at the STI edge, oxide perimeter fails resulting from polysilicon etch, or issues related to the spacer could be measured using the aforementioned devices. These structures were subjected to time-dependent dielectric breakdown tests (TDDB) at elevated voltage and temperature conditions. Stress temperatures varied from 85°C to 180°C, while stress voltages were adjusted so that the dielectric electric field was well above 8 MV/cm. In this high-field regime, the thin gate dielectric current results from Fowler–Nordheim tunneling [9, 10], and the intrinsic-mode behavior for the gate dielectric breakdown is observed. Dielectric breakdown was identified by any instantaneous increase (usually more than 2×) in gate current.

Accelerated lifetime stress data were collected from various individual devices, integrated SRAM and microprocessor chips using the test structures described earlier in which the thin gate dielectric thickness ranged from 7 nm to 3.5 nm. Stress temperatures varied from 100°C to 140°C, with stress duration ranging up to 1000 hours or more. Stress values for power-supply voltage (V_{DD}) ranged from 2.5 V to 5.5 V. Such stress conditions were suitable for studying the extrinsic-mode behavior, in which thin gate dielectric failures occur because of process defects. Functional stress failures were subjected to physical failure analysis to determine their cause.

Routine in-line monitoring of thin gate dielectric yield and reliability during manufacturing is necessary to detect problems early and to define corrective actions. The test structures and procedure for in-line monitoring of thin gate dielectrics were designed to detect problems with either intrinsic or extrinsic thin gate dielectric failures. The test structures include components for both STI and diffusion-bounded thin gate dielectric regions in n^+ and p^+ polysilicon-gated devices. In-line testing of thin gate dielectrics was performed at the wafer level at room temperature. The yield was initially determined by measuring gate dielectric leakage at nominal operating voltage conditions and then performing a voltage ramp test. The ramp test was performed at voltage ramp rates of 10 V/s or less, with the gate voltage ramped to exceed an electric field of ~ 12 MV/cm, thus inducing an intrinsic breakdown on all chips. The cumulative percentage of gate dielectric failures as a function of breakdown voltage was determined from this test. In-line testing was also performed on antenna devices to detect problems caused by in-line charging effects.

Results and discussion

Figures 1–3 show examples of Weibull (statistical) distributions of the cumulative fraction (F) of ultrathin

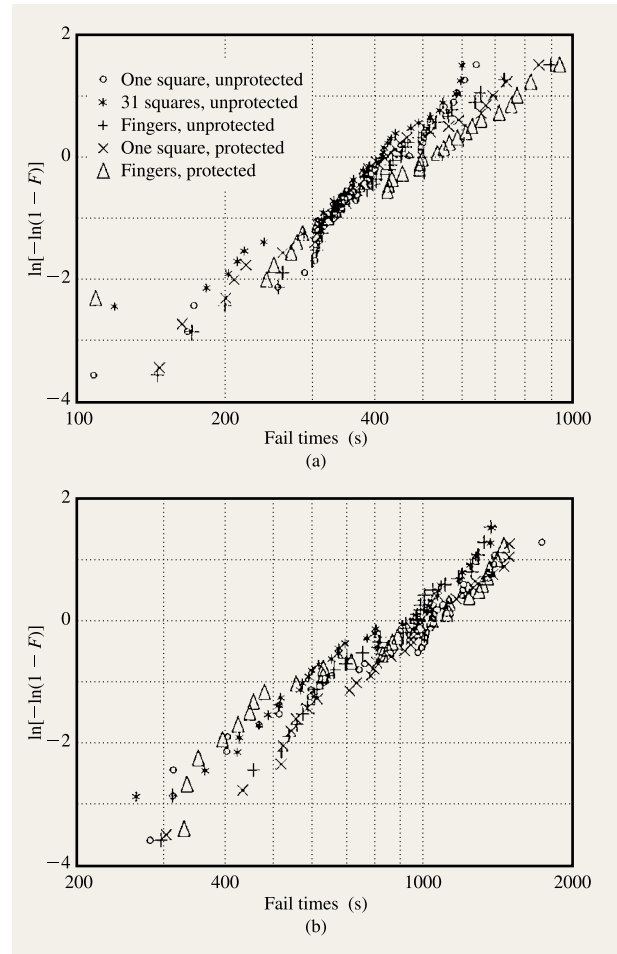


Figure 1

Weibull plot for cumulative fraction fails vs. stress time for (a) n-FET and (b) p-FET for oxide thickness of 6.2 nm. Gate stress voltages are 7 V for n-FET and -7 V for p-FET, stress temperature is 140°C, and oxide area is 6.2×10^{-4} cm².

gate dielectric breakdown (failures) as a function of stress time. The accelerated life stress data were obtained by wafer-level stressing at a constant voltage. Details concerning test structures, gate stress voltages, stress temperatures, and gate dielectric thickness are given in **Table 1**. For all stress conditions, the diffusions, substrate, and n-well terminals are grounded, while positive and negative gate stress voltages were respectively used for n^+ and p^+ diffusion-bounded capacitors. Under these bias configurations, the silicon surface is inverted; therefore, a uniform electric field is present across the thin gate dielectric. The gate dielectric thickness indicated in Table 1 for the various cases was determined by measuring STI-bounded thin dielectric capacitors with the gate biased to

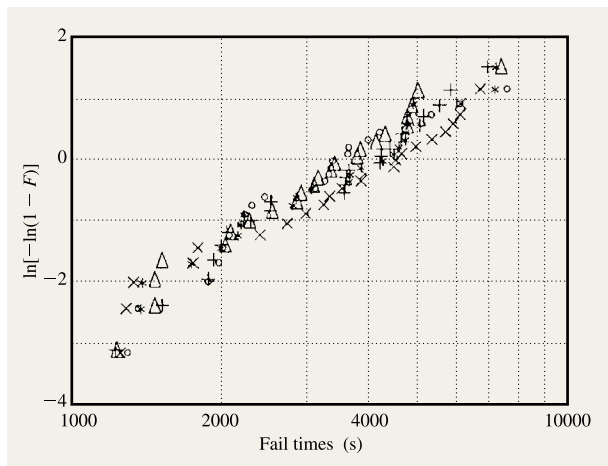


Figure 2

Weibull plot for cumulative fraction fails vs. stress time for n-FET for oxide thickness of 3.63 nm. Gate stress voltage is 4.1 V, stress temperature is 180°C, and oxide area is $6.2 \times 10^{-4} \text{ cm}^2$.

accumulate the silicon surface (negative for p-substrate and positive for n-well). This thickness was determined from the gate capacitance at low gate bias values.

The reader should observe that in Figures 1 and 2, all five test structures have very similar Weibull distributions for gate breakdown. Specifically, no significant difference (<32%) exists between any two of the five structures as far as characteristic lifetime (time to reach 63.2% cumulative percent fails) is concerned. The characteristic lifetime (63.2%) is the time at which the Weibull function $\ln[-\ln(1 - F)]$ is equal to zero. Also, there are no statistically significant differences in the values of the shape factor (i.e., the slope of the Weibull distribution) for all five structures. Table 1 demonstrates that there were no significant differences between the breakdown distributions of various structures where the perimeter-to-area ratio was varied from 0.01 to 1.0. Gate dielectric breakdown distributions were a function only of the oxide area and not the perimeter-to-area ratio. (All five structures have the same total dielectric area.) In addition, there were no significant differences between structures in which the gate was or was not connected to a diffusion to substrate or an n-well (protected versus unprotected). This indicates that for the structures listed in Table 1, in-line charging (which usually arises during the polysilicon etch) did not affect breakdown distributions. Another significant observation can be made about the results shown in Figures 1 and 2: For all five structures, the Weibull breakdown distribution can be approximated by a straight-line fit, which is indicative of the intrinsic wear-out of the thin gate dielectric. With a straight-line fit for the Weibull

distributions in Figures 1 and 2, the shape factor values were 2.71 for the n-FET at 6.2 nm [Figure 1(a)], 2.70 for the p-FET at 6.2 nm [Figure 1(b)], and 2.22 for the n-FET at 3.63 nm (Figure 2). These results show a reduction in the shape factor as dielectric thickness is reduced. The values are consistent with previously published results [20], where the shape factor is presented as a function of dielectric thickness for the intrinsic wear-out. There is no evidence of a significant extrinsic defect failure mode, which is usually associated with a much lower value of the shape factor. These results show that a single failure mode is evident which is consistent with intrinsic breakdown.

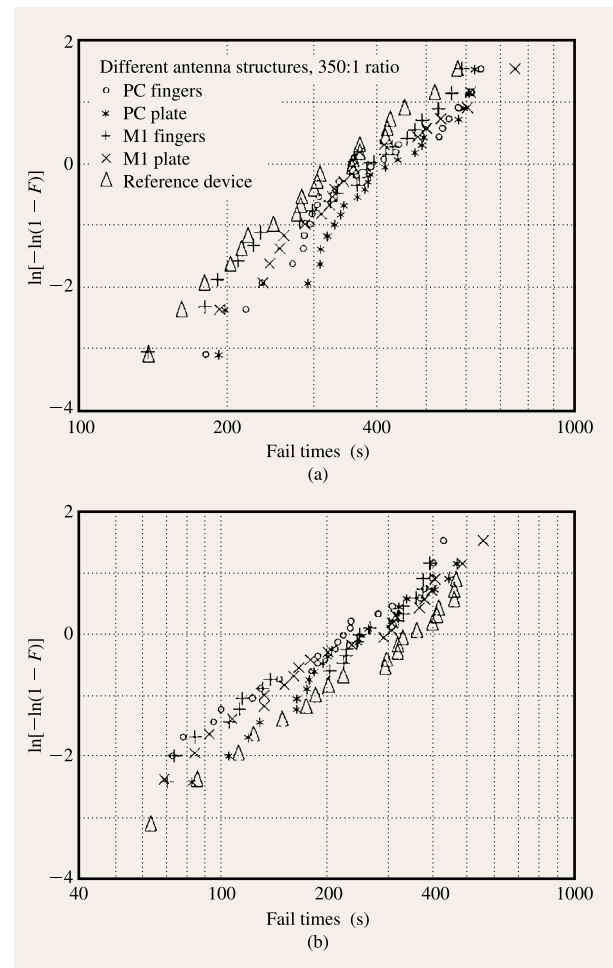


Figure 3

Weibull plot for cumulative fraction fails vs. stress time for (a) n-FET and (b) p-FET for oxide thickness of 5 nm with different antenna structures. Gate stress voltages are 6.7 V for n-FET and -7 V for p-FET, stress temperature is 140°C, and oxide area is $5 \times 10^{-8} \text{ cm}^2$.

Figure 3 shows examples for Weibull distributions of the cumulative fraction fails versus stress time for (a) n-FETs and (b) p-FETs, respectively, each with a different shape of antenna structure but with a constant antenna ratio of 350:1. Antenna device structures include polysilicon fingers, polysilicon plates, metal fingers, and metal plates. Figure 3 also shows data from a reference device with no antenna. It should be observed that for all device structures (reference and antenna) the Weibull distribution for cumulative fraction fails can be approximated by a single straight-line fit which is representative of the intrinsic mode of dielectric breakdown. Therefore, no extrinsic defect failures were observed for either the reference FETs or the antenna devices. In addition, <30% difference was found between the characteristic lifetime of the antenna devices and the reference devices, which suggests there is no significant in-line charging effect on thin gate dielectric breakdown.

A key parameter for intrinsic dielectric breakdown is the thin gate dielectric area. As demonstrated earlier, the breakdown of these devices does not depend on the perimeter-to-area ratio, indicating no particular process weakness or sensitivity at the dielectric perimeter. **Figure 4** shows a logarithmic plot for the normalized lifetime at 50% cumulative fails versus dielectric area for n^+ diffusion-bounded 5-nm thin gate dielectric. All structures with different areas were stressed at 140°C and 5.75 V. Lifetime normalization is computed by dividing each device lifetime by the lifetime of the 10 000- μm^2 device. Six different dielectric areas ranging from 2.5 μm^2 to 9.0 mm^2 were used to generate these data. A straight line on a log-log scale provides an excellent fit to the data in Figure 4, which suggests a power-law relationship between the intrinsic lifetime and the thin gate dielectric area. This area dependence of the cumulative probability of failure (F) follows a Poisson model in which $F = \exp(-DA)$, where A is the dielectric area and D is a defect density.

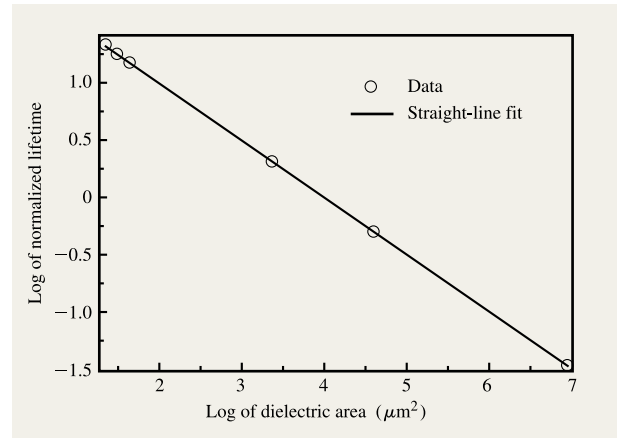


Figure 4

Logarithmic plot of cumulative fraction fails vs. oxide area for an n-FET with an oxide thickness of 5 nm.

The Poisson-like nature of this area dependence demonstrates that the intrinsic dielectric failures (for the devices discussed above) are randomly distributed over a given dielectric area.

The value of the applied stress voltage has a significant effect on the intrinsic dielectric breakdown. The dependence of the dielectric lifetime on gate voltage was obtained by stressing many groups of dielectric samples in which the stress voltage was the only parameter varied. All other device and measurement parameters (i.e., dielectric thickness, stress temperature, and dielectric area) were held constant. In **Figure 5** the normalized lifetime (at 50% cumulative failure) is shown as a function of the gate stress voltage. The normalized lifetime is obtained by dividing the lifetime of each sample by the lifetimes at a

Table 1 Details of test structures, gate dielectric thickness, and stress conditions for Figures 1–3.

Figure	Polysilicon gate	Gate edge	Type of layout	Oxide area (μm^2)	Perimeter/area	Thickness (nm)	Gate voltage (V)	Temperature (°C)
1	n^+ & p^+	n^+ & p^+ diffusion	One square, squares, fingers	62,000 62,000 62,000	0.01 0.12 1.00	6.2	7.0	140
2	n^+	n^+ diffusion	One square, squares, fingers	62,000 62,000 62,000	0.01 0.12 1.00	3.63	4.1	180
3	n^+ & p^+	n^+ & p^+ diffusion	n-FET 350:1 antennas; polysilicon, metal	5.0	8.1	5.0	6.7	140

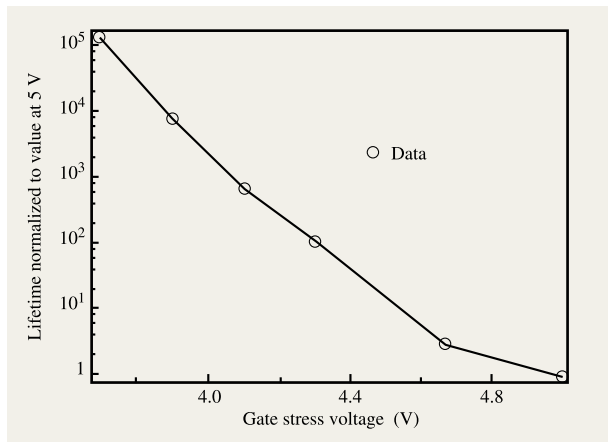


Figure 5

Normalized lifetime vs. gate stress voltage for an n-FET with a dielectric thickness of 3.6 nm.

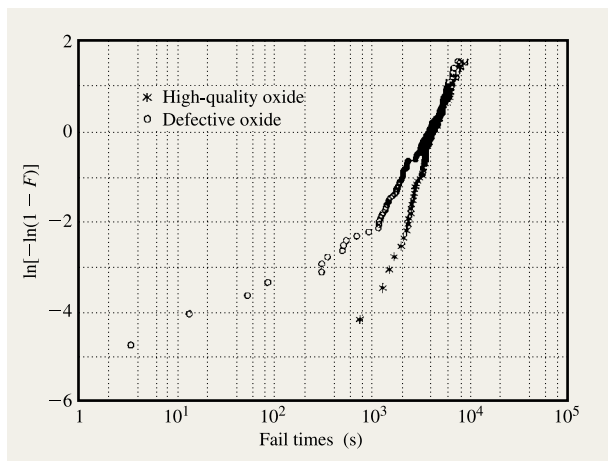


Figure 6

Weibull plot for cumulative fraction fails vs. stress time for both high-quality and defective thin-oxide n-FET arrays.

stress voltage of 5 V. The data shown in Figure 5 are for an n-FET with a dielectric thickness of 3.6 nm, an oxide area of 62 000 μm^2 , and a stress temperature of 180°C. These data indicate that the relationship between the logarithm of device lifetime and the applied gate voltage is approximately linear at voltages equal to or less than 4 V. Above 4 V this relationship becomes highly nonlinear. For smaller dielectric thicknesses, the relationship between the logarithm of lifetime and stress voltage approaches linearity at low voltages. These data support the use of a

model in which an exponential dependence of the device lifetime on the applied stress voltage is valid at low applied bias conditions. This model is commonly referred to as the E model [5, 9, 20]. The dependence of dielectric lifetime on stress temperature was obtained in a manner similar to that used to obtain the voltage dependence (i.e., only the stress temperature was varied and all other parameters held constant). An activation energy in the temperature range of interest was determined from an Arrhenius plot of the data. The dependence of the lifetime on the dielectric thickness was found to be nonlinear. This dependence shows a decreasing lifetime with decreasing dielectric thickness [20]. Using this dielectric wear-out model, the maximum operating conditions for voltage and temperature for any desired oxide area, oxide thickness, and failure criterion (e.g., parts-per-million failures at the end of a desired useful lifetime) can be determined.

In-line monitoring for both yield and reliability was performed on n^+ and p^+ diffusion-bounded test structures which have a sufficient STI perimeter with the thin oxide. The voltage was ramped at a rate of 10 V/s or lower for test structures held at 30°C. The total oxide area was in excess of 0.1 mm^2 . The polarity of gate stress voltage produced an inverted silicon surface. A good correlation between the results of accelerated time-dependent dielectric breakdown (TDDB) testing and in-line voltage ramp test was obtained. In **Figure 6**, accelerated TDDB stress results are shown for both high-quality and defective 5.5-nm thin gate oxides. Test structures for both types of oxide were stressed at a gate voltage of 5.75 V at 140°C. The high-quality oxide data from Figure 6 show a steep intrinsic wear-out behavior with a shape factor for the Weibull distribution in excess of 3.0. On the other hand, the defective oxide data points in Figure 6 show degraded performance, with a significant extrinsic defect contribution where the shape factor is only 0.43. Voltage ramp test results for the defective oxides are shown in **Figure 7**. These data show a linear relationship between the logarithm of lifetime (i.e., under TDDB testing conditions) and the ramp breakdown voltage [21–24]. Correspondingly, a linear relationship exists between the Weibull function of cumulative failures and the ramp breakdown voltage. Figure 7 shows that the high-quality oxides exhibit steep wear-out, while defective oxides have a significantly shallower distribution that is indicative of a degraded reliability. Results in Figures 6 and 7 show that a voltage ramp test can be very effective in detecting dielectric reliability problems.

Test structures for in-line monitors of thin gate dielectrics were designed to detect dielectric problems caused by extrinsic defects at either the STI or diffusion perimeter with the thin gate dielectric. A detailed study of the manner in which the dielectric thickness affects yield

and reliability revealed that a reduction of thin gate dielectric thickness at the STI edge caused above-average levels of dielectric fails on chips at the perimeter of a wafer. This problem, detected with in-line dielectric monitors, was corrected. Further studies of this defect used samples of large-area capacitors with dielectric thicknesses ranging from 7 nm to 4.5 nm. The yield of these devices was first tested at normal operating voltages (3.3 V–2.5 V) and then stressed at elevated voltage and temperature conditions. The percentage of cumulative failure for ten years of continuous use at normal operating conditions of voltage and temperature was calculated from the stress results. **Table 2** shows the results for three different groups of capacitors with dielectric thicknesses of 7 nm, 5.5 nm, and 4.5 nm. The total thin gate dielectric area for each group of capacitors exceeded 50 cm². The initial value of gate dielectric leakage was used as the measure of yield, where a gate current in excess of 25 nA (at a predetermined voltage) was considered a failure. For a good device the normal gate current was $\ll 25$ nA. The normalized yield values (as determined using the gate leakage criterion mentioned above) shown in Table 2 were obtained by dividing the yield for each group of capacitors by the yield for the capacitor with a 7-nm dielectric thickness. The number of reliability failures for ten years of use at operating conditions was normalized for each group to the value that corresponds to 7-nm dielectric thickness. Table 2 shows that the effect of the process-induced extrinsic defects on both yield and reliability increases as gate dielectric thickness is reduced. Since these results are only for the STI edge dielectric problem, no attempt has been made to extend them to a more general relationship between dielectric yield and reliability. As shown in **Figure 8** for these types of defects, gate dielectric problems can be observed in both yield and reliability measurements.

The assertion that the STI defect mechanism mentioned above is an extrinsic defect mechanism is observed in the burn-in results given in **Figures 9(a)** and **9(b)**. Pre-burn-in test failures on a high-performance microprocessor with a thin gate dielectric problem are shown in Figure 9(a). The contribution of decoupling capacitor failures relative to all failures other than dielectric on the chip, prior to burn-in, is also shown. Decoupling capacitors on microprocessor chips contribute significantly to thin gate dielectric perimeter with STI and therefore are expected to be sensitive to this particular defect. Accelerated life test results for a microprocessor, following a conventional burn-in at 140°C, are shown in Figure 9(b). No contributions from the decoupling capacitors were observed after burn-in, indicating that burn-in is effective in screening out extrinsic dielectric defects.

The data shown in **Figure 10** demonstrate the effectiveness of in-line thin gate dielectric monitors in

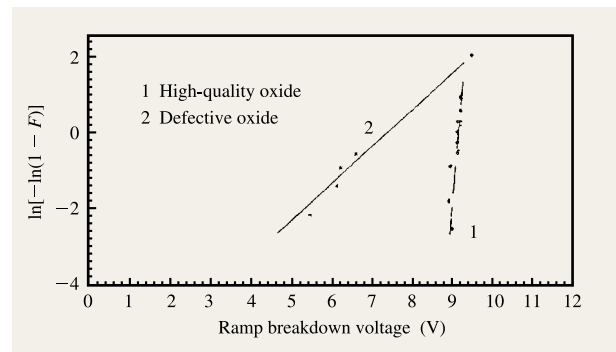


Figure 7

Weibull plot for cumulative fraction fails vs. ramp breakdown voltage for both high-quality and defective thin-oxide n-FET arrays.

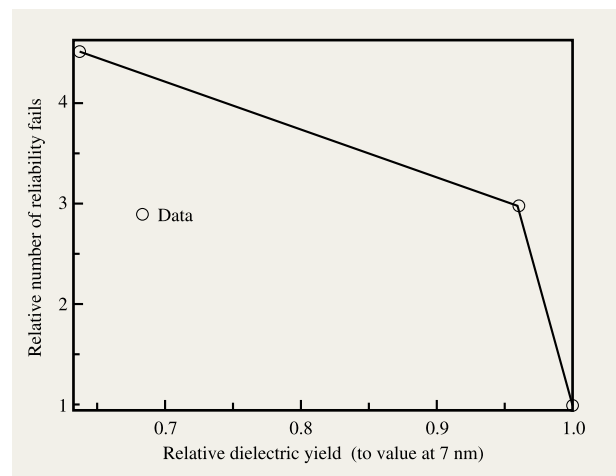


Figure 8

Relationship between dielectric yield and reliability for a dielectric problem at STI edge.

Table 2 Normalized (to results at 7 nm) yield and number of reliability fails for large-area capacitors with different oxide thicknesses for hardware with known dielectric problem at STI edge.

Oxide thickness (nm)	Normalized yield	Normalized number of reliability fails
7.0	1.00	1.00
5.5	0.96	3.00
4.5	0.64	4.50

identifying reduced dielectric yield during a specific period on one of three etch tools. Here, dielectric monitor results

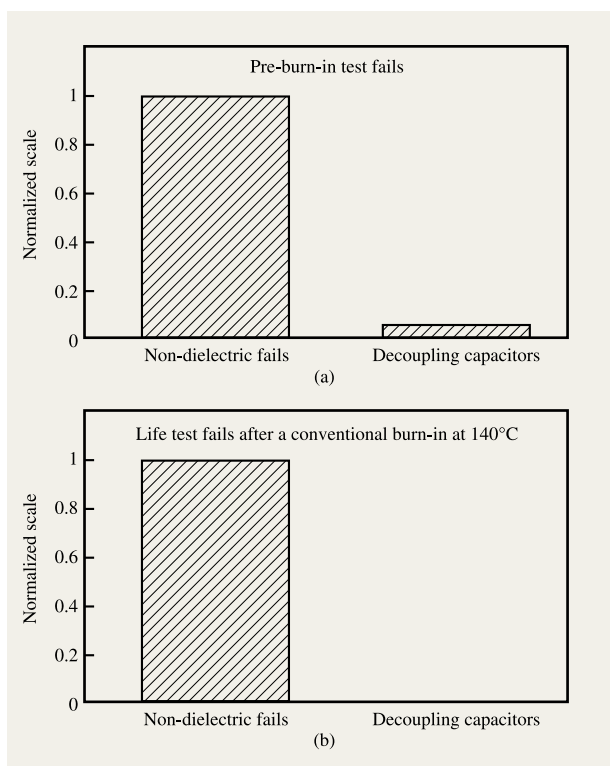


Figure 9

(a) Pre-burn-in test relative yield results of high-performance microprocessor, including contribution of decoupling capacitors with an STI edge dielectric problem. (b) Post-burn-in accelerated life test results showing the effectiveness of burn-in, with no fails from decoupling capacitors.

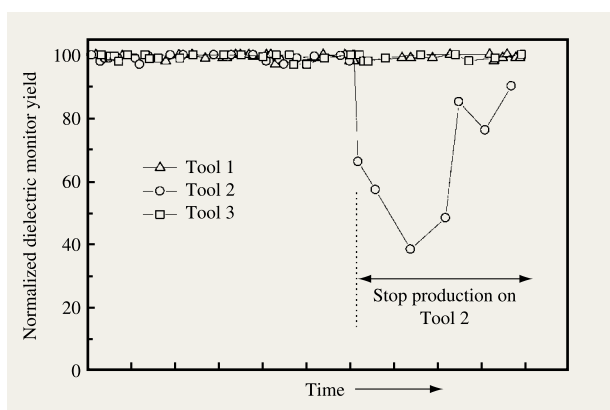


Figure 10

Effect of a process tool problem on dielectric monitor yield.

and physical failure analyses were used to determine the cause of a process problem that resulted in a significant yield loss (Tool 2 in Figure 10) on these tools. This problem was related to a spacer etch tool and dielectric fails that occurred at the diffusion perimeter. When this problem was corrected, dielectric yield returned to normal.

Physical failure analysis is important for thin gate dielectric integrity in both development and manufacturing, since it tracks and identifies gate dielectric fails on chips for engineering, qualification, or production samples. It can also identify a failure, i.e., whether it occurred at the perimeter or in some particular device area, as well as the type of device or circuit involved. This analysis can then be correlated with chemical, optical, or electrical analyses performed. In this and other studies, it has been found that for chips manufactured with thin gate dielectric thicknesses of 7 nm to 3.5 nm, thin gate dielectric failures account for fewer than 10% of all chip failures. This low percentage includes both yield and reliability failures for SRAMs, microprocessors, and custom logic chips. The oxide failures are randomly distributed across the gate dielectric area, with metallic particle contaminants identified at the site of failure.

The removal of metallic contaminants, most commonly iron, is a major challenge for reducing defect density in ULSI MOSFET chip manufacturing. Three major components must be addressed in order to control this type of contamination: 1) substrate design, specification, and qualification; 2) control of pre-gate wet cleaning; and 3) qualification and control of device hot-process tools. For substrate design and qualification, the objective is to minimize the effects of metal contaminants in device active areas. Back-side surface roughness and p^+ substrate regions with high boron and dissolved oxygen concentrations are gettering sites for impurity metals that are dissolved in the silicon. All of these sites are located well below the active device region. To minimize back-side impurity gettering effects, a pad oxide is grown and a pad nitride deposited on all wafers. The pad nitride is one of many barrier films on the back side of the wafer that lowers the sensitivity to gate yield loss caused by metal contamination from robot handlers and/or from measurement tools. The addition of silicon nitride as a diffusion barrier prevents back-side metal contamination from diffusing into the silicon and possibly into the active areas of a device. A more detailed discussion of metallic contamination and its control can be found in [25].

Summary

Reliability data for dielectrics with thicknesses ranging from 7 nm to 3.5 nm have demonstrated that thin gate dielectric in high-performance CMOS products can be fabricated without significant contribution arising

from extrinsic defects or in-line charging. The Weibull distribution characteristics of the cumulative fraction of failures for different device structures are described for the intrinsic failure mode using the distribution shape factor and its dependence on oxide thickness. Also, the characteristic life of the distribution has a dependence on oxide area (not perimeter-to-area ratio). The methodology for determining the dependence of dielectric lifetime on stress voltage, oxide area, stress temperature, and dielectric thickness is discussed. For chips manufactured with 7-nm and 3.5-nm gate dielectrics, failures associated solely with gate dielectric account for fewer than 10% of all chip failures. The use of dielectric in-line monitors for detection of dielectric degradation and control of the quality of manufacturing production was described. Data from product chips demonstrated that in-line monitors were effective in detecting dielectric problems in the yield and reliability for chips and/or devices.

Acknowledgments

Many individuals in a variety of development and manufacturing areas in the IBM Microelectronics Division at the Essex Junction, Vermont, and East Fishkill, New York, facilities contributed to this work. The authors would like to give special thanks to Ann Swift, Roger Dufresne, David Brochu, Benjamin Morse, Asit Ray, Ernest Wu, Paul Agnello, Mark Johnson, Sue Bowers, Mark Burns, Keith Stevens, Gary Hueckel, and Julie Lee. The authors are also grateful to Peter Cottrell of the IBM Microelectronics Division, Essex Junction, Vermont, and to Douglas Buchanan of the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, for helpful discussions.

References

1. *The National Technology Roadmap for Semiconductors: Technology Needs*, Semiconductor Industry Association, San Jose, CA, 1997.
2. T. Kuroi, S. Shimizu, S. Ogino, A. Teramoto, M. Shirahata, Y. Okumura, and M. Inuishi, "Sub-Quarter-Micron Dual Gate CMOSFETs with Ultrathin Gate Oxide of 2nm," *Digest of Technical Papers, Symposium on VLSI Technology*, 1996, p. 210.
3. A. I. Chou, C. Lin, K. Kumar, P. Chowdhury, M. Gardner, M. Gilmer, J. Fulford, and J. C. Lee, "The Effects of Nitrogen Implant into Gate Electrode on the Characterization of Dual-gate MOSFETs with Ultrathin Oxide and Oxynitrides," *Proceedings of the International Reliability Physics Symposium*, 1997, p. 174.
4. B. Maiti, "Nitrided Oxides in CMOS Technology," *International Reliability Physics Symposium (IRPS) Tutorial Proceedings*, 1998, Section 3, p. 3.1.
5. R. Degraeve, J. L. Ogier, R. Bellens, Ph. Roussel, G. Groeseneken, and H. E. Maes, "On the Field Dependence of Intrinsic and Extrinsic Time-Dependent Dielectric Breakdown," *Proceedings of the International Reliability Physics Symposium*, 1996, p. 44.
6. J. L. Ogier, R. Degraeve, P. Roussel, G. Groeseneken, and H. E. Maes, "Analysis of the Early Failure Rate Prediction of Time-Dependent Dielectric Breakdown in Thin Oxides," *Proceedings of the European Solid State Device Research Conference (ESSDERC)*, 1995, p. 299.
7. W. R. Hunter, "A Failure Rate Based Methodology for Determining the Maximum Operating Gate Electric Field, Comprehending Defect Density and Burn-In," *Proceedings of the International Reliability Physics Symposium*, 1996, p. 37.
8. I. C. Chen, S. Holland, and C. Hu, "Hole Trapping and Breakdown in SiO₂," *IEEE Electron Device Lett.* **ED-7**, 164 (1986).
9. David J. Dumin, Sai K. Mopuri, S. Vanchinathan, R. S. Scott, R. Subramoniam, and Terrill C. Lewis, "High-Field Related Thin Oxide Wearout and Breakdown," *IEEE Trans. Electron Devices* **42**, 760 (1995).
10. E. Harari, "Conduction and Trapping of Electrons in Highly Stressed Ultrathin Films of Thermal SiO₂," *Appl. Phys. Lett.* **30**, 601 (1977).
11. J. S. Suehle, P. Chaparala, C. Messick, W. M. Miller, and K. C. Boyko, "Field and Temperature Acceleration of Time-Dependent Dielectric Breakdown in Intrinsic Thin SiO₂," *Proceedings of the International Reliability Physics Symposium*, 1994, p. 120.
12. Hideki Satake, Shin-ichi Takagi, and Akira Toriumi, "Evidence of Electron-Hole Cooperation in SiO₂ Dielectric Breakdown," *Proceedings of the International Reliability Physics Symposium*, 1997, p. 156.
13. C.-F. Chen, C.-Y. Wu, M.-K. Lee, and C.-N. Chen, "The Dielectric Reliability of Intrinsic Thin SiO₂ Films Thermally Grown on a Heavily Doped Si Substrate—Characterization and Modeling," *IEEE Trans. Electron Devices* **ED-34**, 1540 (1987).
14. R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A Consistent Model for the Thickness Dependence of Intrinsic Breakdown in Ultrathin Oxides," *Proceedings of the International Electron Devices Meeting*, 1995, p. 863.
15. R. Degraeve, J. De Blauwe, J. L. Ogier, Ph. Roussel, G. Groeseneken, and H. E. Maes, "A New Polarity Dependence of the Reduced Trap Generation During High-Field Degradation of Nitrided Oxides," *Proceedings of the International Electron Devices Meeting*, 1996, p. 327.
16. J. H. Stathis, "Quantitative Model of the Thickness Dependence of Breakdown in Ultra Thin Oxides," *Microelectron. Eng.* **36**, 325 (1997).
17. D. J. Dimaria and J. H. Stathis, "Ultimate Limit for Defect Generation in Ultrathin Silicon Dioxides," *Appl. Phys. Lett.* **71**, 3230 (1997).
18. R.-P. Vollertsen and W. W. Abadeer, "Comprehensive Gate-Oxide Reliability Evaluation for DRAM Processes," *Microelectron. & Reliabil.* **36**, 1631 (1996).
19. L. Su, S. Subbanna, E. Crabbe, P. Agnello, E. Nowak, R. Schulz, S. Rauch, H. Ng, T. Newmann, A. Ray, M. Hargrove, A. Acovic, J. Snare, S. Crowder, B. Chen, J. Sun, and B. Davari, "A High-Performance 0.8 μ m CMOS," *Proceedings of the 1996 Symposium on VLSI Technology*, 1996, p. 12.
20. J. H. Stathis and D. J. DiMaria, "Reliability Projection for Ultrathin Oxides at Low Voltage," *Proceedings of the International Electron Devices Meeting*, 1998, p. 167.
21. M. Shatzkes, M. Av-Ron, and K. V. Srikrishnan, "Determination of Reliability from Ramped Voltage Breakdown Experiments; Application to Dual Dielectric MIM Capacitors," *Proceedings of the 2nd International Reliability Physics Symposium*, 1984, pp. 138–139.
22. M. Shatzkes and M. Av-Ron, "Determination of Breakdown Rates and Defect Densities in Si-SiO₂," *Thin Solid Films* **91**, 217 (1982).
23. E. S. Anolick and Li Yu Chen, "Application of Step Stress to Time-Dependent Breakdown," *Proceedings of the 19th International Reliability Physics Symposium*, 1981, pp. 23–27.

24. A. Berman, "Time-Zero Dielectric Reliability Test by a Ramp Method," *Proceedings of the 19th International Reliability Physics Symposium*, 1981, pp. 204–209.
25. W. Abadeer, W. Cote, A. LeBlanc, M. Levy, R. Vollertsen, H. Akatsu, V. Penka, R. Ludeke, A. Kluwe, P. Schulz, S. Nadahara, A. Antreasyan, and D. Cote, "Advanced Technology Processing and State-of-the-Art Solutions to Cleaning Contamination Control, and Integration Problems," presented at the Spring Meeting of the Materials Research Society, San Francisco, April 17–21, 1995.

Received March 11, 1998; accepted for publication March 24, 1999

W. W. (Bill) Abadeer *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (ABADEER at BTVLABVM, abadeer@us.ibm.com).* Dr. Abadeer is a Senior Engineer with the SRDC Quality and Reliability group at the Burlington facility. He joined IBM at the Burlington facility in 1976, and has since worked on the reliability of semiconductor devices. Dr. Abadeer received a B.S. in electrical engineering from Assiut University, Egypt, in 1966. From 1966 to 1968 he worked as a teaching assistant at Assiut University. He received his M.S. and Ph.D. degrees in electrical engineering in 1970 and 1976, respectively, from the University of Vermont. Dr. Abadeer is a member of the IEEE and the Electrochemical Society.

Asmik Bagramian *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (ABAGRAMI at BTVMANVM, abagrami@us.ibm.com).* Ms. Bagramian is a Staff Engineer in the ASIC/Custom Qualification Department at the IBM Burlington facility, working on RS/6000 and AS/400 microprocessor product qualifications. She joined IBM in 1984 and was involved with the manufacturing line monitoring strategy for several years. After taking a two-year leave from IBM, Ms. Bagramian received a B.S. degree in electrical engineering from the University of Vermont in 1990. She has received an IBM Outstanding Contribution Award for her effort in the success of the Power-2 Super Chip Products and continues to support the qualification of the new generation of Power Processors.

David W. Conkle *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (CONKLED at BTVVMOFS).* Mr. Conkle is a Staff-level Manager for the Kerf Characterization Department at the IBM Burlington facility. He received a B.S. degree in physics from Sewanee University in 1990 and subsequently joined the United States Navy, where he studied naval nuclear engineering. He left the Navy in 1995 and joined IBM at the Burlington facility in Vermont.

Charles W. Griffin *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (GRIFFINC at BTVLABVM, griffinc@us.ibm.com).* Mr. Griffin is currently with the SRDC Quality and Reliability group at the IBM Burlington facility, focusing on gate dielectric reliability and test structure design. He received his M.S. and B.S. degrees in electrical engineering and a B.A. degree in psychology. Mr. Griffin joined IBM in 1979 at the Burlington facility.

Eric Langlois *Department of Electrical Engineering, Arizona State University, University Drive and Mill Avenue, Tempe, Arizona 85287 (elanglo@asu.edu).* Mr. Langlois is currently a Graduate Teaching Associate at the University of Arizona, Tempe, where he is studying for a Ph.D. in solid-state electronics. He received a B.S.E. in engineering physics and

an M.S. in electrical engineering from RPI. Mr. Langlois was an IBM employee for 2½ years, beginning in 1996. He is a member of Tau Beta Pi.

Brian F. Lloyd *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (blloyd@us.ibm.com).* Mr. Lloyd is a physicist and an Advisory Engineer at the IBM Burlington facility; he is primarily responsible for metal contamination detection and control in semiconductor manufacturing. He is the technical leader of the Electronic Test Center, which performs the metal monitoring measurements for the manufacturing lines. Mr. Lloyd has published on variable retention time characterization and control. He joined IBM in 1980 after teaching and administration at the University of Vermont.

Raymond P. Mallette *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (mallette@us.ibm.com).* Mr. Mallette is currently an Electrical Test Yield Characterization Engineer for microprocessors at the IBM Burlington facility. He received his B.S. degree in physics from San Jose State University in 1985 and his M.S. in electrical engineering from the University of Santa Clara in 1990. Since joining IBM in 1977, he has worked in the Storage Systems Division and, since 1992, in the Microelectronics Division.

James E. Massucco *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (jmassucco@us.ibm.com).* Mr. Massucco is a Staff Engineer in the ASICS Failure Analysis group within the Technology Analysis Project at the IBM Burlington facility. He joined IBM in 1974 after receiving an A.S.E.E. degree from Vermont Technical College. Mr. Massucco has spent his entire career in the failure analysis arena and has worked on various technologies including DRAM, SRAM, ASICS, SiGe, and SOI.

Jonathan M. McKenna *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (jmmckenn@us.ibm.com).* Mr. McKenna is currently an Engineer with the SRDC Quality and Reliability group at the IBM Burlington facility, working on gate dielectric reliability. He joined IBM in 1966 and has been working on gate dielectric breakdown mechanisms and reliability since then. He received his B.S. in electrical engineering from the University of Notre Dame in 1996.

Steven W. Mittl *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (mittl@us.ibm.com).* Mr. Mittl is currently an Advisory Engineer with the SRDC Quality and Reliability group at the IBM Burlington facility, with special interest in NVRAM reliability. He received the B.S. and M.S. degrees in electrical engineering from Lehigh University in 1983 and the University of Vermont in 1986, respectively. He joined IBM at the Burlington facility in 1983 and, since 1986, has been involved in hot-carrier reliability on DRAM and logic technologies, and hot-carrier device and circuit modeling. Mr. Mittl holds a patent on product hot-carrier-reliability test methods; he has authored/coauthored numerous papers on hot-carrier effects.

Philip H. Noel *IBM Microelectronics Division, Burlington facility, Essex Junction, Vermont 05452 (pnoel@us.ibm.com).* Mr. Noel is currently a Staff Engineer in the Technology Analysis Project at the IBM Burlington facility. He joined IBM in 1966 and has worked in failure analysis since 1970. Mr. Noel has been a Group Leader for the DRAM Failure Analysis Team; he is currently the Group Leader for SRAM and CMOS F/A Development.