

A Study of Soft and Hard Breakdown—Part I: Analysis of Statistical Percolation Conductance

Muhammad Ashraful Alam, *Senior Member, IEEE*, Bonnie E. Weir, *Member, IEEE*, and Paul J. Silverman

Abstract—A theory of the statistical origin of soft and hard breakdown, that can explain a wide range of experimental data, is proposed. The theory is based on the simple premise that the severity of breakdown depends on the magnitude of the power dissipation through the sample-specific, statistically distributed percolation conductance, rather than on any physical difference between the traps involved. This model (a) establishes the connection between the statistical distribution of the theoretically predicted percolation conductance and the distribution of experimentally measured conductances after soft breakdown (Part I), and (b) explains the thickness, voltage, stress, and circuit configuration dependence of soft and hard breakdown (Part II). Connections to previous theories are made explicit, and contradictions to alternate models are resolved.

Index Terms—Hard breakdown, MOS devices, reliability, semiconductor device modeling, soft breakdown.

I. INTRODUCTION

WHEN stressed by an applied voltage, an oxide film loses its insulating properties in two stages [1]–[4]. First, in the wearout phase, traps are generated within the oxide that increase the leakage current through the film. Eventually, these traps complete a percolation path that bridges the two electrodes across the oxide. Various models, e.g., anode hole injection (AHI) theory [5]–[7], hydrogen release model [8], [9], [35], and electrochemical model [10], etc., have been proposed to explain this wearout phase of oxide degradation. Then, once the percolation path is completed, the power dissipation through it controls the second stage of the breakdown transient (characterized by submicrosecond time-constants, segment I-II or segment I-III, see Fig. 1) which determines the postbreakdown conduction property of the oxide.

This second stage of oxide degradation for thinner oxides tested at lower voltages (soft) appears to be fundamentally different from breakdown for thicker oxides tested at higher voltages (hard) [2], [12]–[17]. Hard breakdown is characterized by a large change in voltage or current during the breakdown transient and a postbreakdown current–voltage (I – V) characteristic that is essentially ohmic (see Fig. 1, long dashed line). Soft breakdown is detected by a much smaller change of voltage or current after breakdown and by postbreakdown I – V characteristics which can be described by a power law $I = G_o V^\delta$ (see Fig. 1, short-dashed line). Although an impressive set of

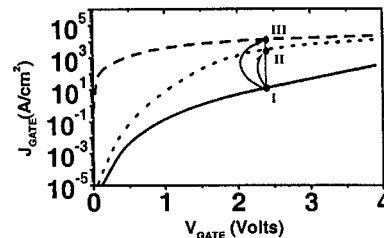


Fig. 1. One monitors the pre- and postbreakdown I – V characteristics to determine soft (nonlinear, power-law conduction: short dashed line) versus hard (ohmic conduction: long dashed line) breakdown.

data has been collected for soft-breakdown over the years [11], [16], [20], and the possibility of using soft-broken oxides in circuits has been discussed [11]–[13], a consistent theory of soft and hard breakdown that explains available experimental results is still lacking. To be useful for reliability projections and future oxide scaling, we must be able to translate the available data—collected from large area capacitors stressed at high voltages—to ultrasmall IC transistors operating at low voltages for oxides which are yet to be fabricated. Moreover, this scaling must be performed with an eye to the actual circuit conditions which may stress these oxides in some combination of constant voltage and constant current stress. The theory proposed in this paper is an effort to achieve these objectives.

In this paper, after the introduction, we propose a theory of breakdown transient (segment I-II or I-III, see Fig. 1) in Section II. This theory is based on two simple concepts: statistical distribution of the percolation conductance (G_o) and the existence of a threshold power density (P_{crit}). Section III discusses the concept of percolation conductance G_o and its time evolution. Our conclusions regarding the statistical properties of G_o are summarized in Section IV. The companion paper [34] will then analyze the consequences of the statistically-distributed percolation conductance (G_o) and critical threshold power density (P_{crit}) in determining hard and soft breakdown as a function of capacitor area, oxide thickness, and stress conditions, establishing the scaling principles for reliability projections.

II. CIRCUIT MODEL FOR BREAKDOWN TRANSIENT

Consider an oxide capacitor stressed with a constant current stress (CCS) [or a constant voltage stress (CVS)], as shown in Fig. 2. During the wearout phase, the current-flow from the electrodes is balanced by spatially uniform tunneling current through the oxide [which may include both direct or Fowler-Nordheim tunneling current as well as the stress-induced leakage current (SILC)]. Therefore, at time $t < 0$ (the

Manuscript received March 20, 2001; revised September 17, 2001. The review of this paper was arranged by Editor G. Groeseneken.

The authors are with Agere Systems, Murray Hill, NJ 07974 (e-mail: alam@agere.com).

Publisher Item Identifier S 0018-9383(02)00821-3.

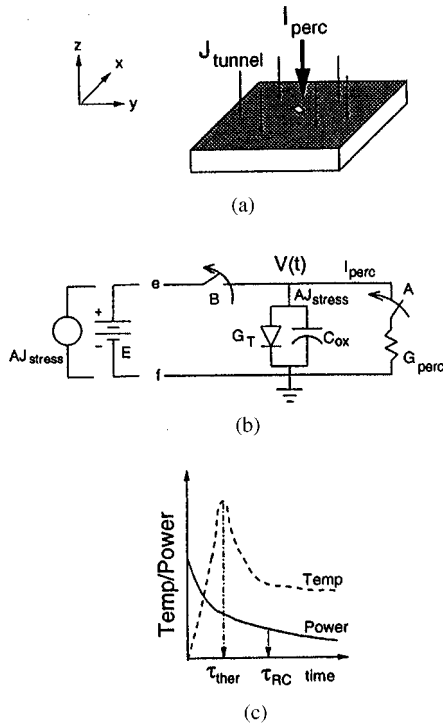


Fig. 2. (a) Oxide sample is stressed uniformly before breakdown; after breakdown a localized percolation path provides a less resistive path between the two electrodes. (b) The conduction through the oxide can be represented by an equivalent circuit. For constant current stress, current source J_{stress} is attached to points e - f , for constant voltage stress, voltage source E is attached to terminals e - f . $C_{\text{ox}} [= (\epsilon A/T_{\text{ox}})]$ and $G_T [= A(dJ_{\text{tunnel}}/dV)]$ represent the capacitance and tunneling conductance of the capacitor. At breakdown, the percolation path with conductance G_p is switched in (switch A), and power dissipation begins through the localized percolation spot. The power dissipation can be stopped by adjusting the compliance limit (which act as variable time delay switch B). (c) The power dissipation through the percolation path (solid line), which is time dependent, raises the temperature of the localized spot (dashed line) which may cause local melting and lead to hard breakdown.

instant when the percolation path is completed defines the reference time $t = 0$

$$AJ_{\text{stress}} = AJ_{\text{tunnel}} \quad (1)$$

$$J_{\text{tunnel}} = \alpha(T_{\text{ox}})e^{-\beta(T_{\text{ox}})/V(t < 0)} \quad (2)$$

where A is the area of the capacitor, J_{stress} is the stress current density, and the J_{tunnel} is the tunneling current density through the oxide.

After the localized percolation path is completed at $t > 0$ [schematically shown as a localized spot in Fig. 2(a) and is represented by a sudden closure of switch A in Fig. 2(b)], significant current begins to flow through this high-conductance percolation path [11], [17]. For constant current stress, the total current from the current source must remain constant at AJ_{stress} , therefore the applied voltage across the capacitor, $V(t > 0)$, has to decrease as a function of time. For $t > 0$

$$AJ_{\text{stress}} = AJ_{\text{tunnel}} + AJ_{\text{disp}} + \{I_{\text{perc}}(t)\} \quad (3)$$

$$J_{\text{disp}} \equiv \frac{\epsilon}{T_{\text{ox}}} \frac{dV(t)}{dt} \quad (4)$$

$$\{I_{\text{perc}}\} \equiv \{G_o(t)\}V(t)^{\{\delta\}}. \quad (5)$$

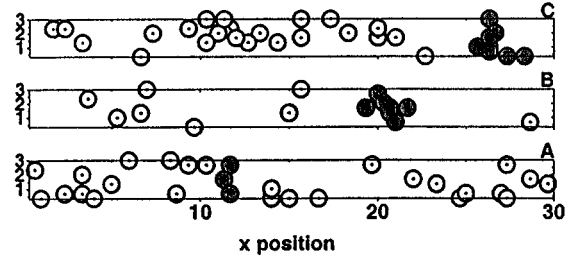


Fig. 3. Three examples of the percolation path formed in an oxide whose thickness is $\sim 3X$ the trap diameter D . The number of traps in the percolation path (shaded circles), N_{perc} , can be as small as 3 (sample A) or as large as 8 (sample C), with different conductance. Note that these percolation paths were schematically represented by the white spot near the center of the oxide in Fig. 2(a). The total number of traps at breakdown (the sum of the shaded and unshaded ones) for each sample is called the critical trap density N_{BD} which is obviously sample-specific.

The displacement current through the oxide J_{disp} depends on the rate of change in voltage, and on the oxide thickness T_{ox} , which includes the quantization and poly-depletion effects. For constant-voltage stress, the voltage source holds the voltage across the oxide constant at the point e (see Fig. 2) so that $dV/dt = 0$. Therefore, J_{disp} need not be considered for constant voltage stress.

The third term on the right-hand side of (3), which describes the current through the percolation path is an area-independent, sample-specific [see Fig. 3—each sample has different, statistically distributed conductance, hence the curly braces in (3)–(5)], time-dependent (G_o may change as power through the path changes with time) current. The exact analytical formula for I_{perc} —we use a power law form—is not important for our discussion, because we shall not analyze the physical mechanism underlying the nonlinear voltage dependence [18]–[21], [32] of the percolation current. We shall adopt a simple nonlinear network of conductors (more on this in Section III) to parameterize this voltage dependence. Since I_{perc} in (3) is sample-specific, the power dissipation through the percolation path

$$P_{\text{perc}}(t > 0) = V(t > 0)I_{\text{perc}}(t > 0) \quad (6)$$

will also vary from sample to sample. For constant voltage stress, $V(t) = E$, so that this power dissipation takes a simple form, i.e.,

$$P_{\text{perc}} = EI_{\text{perc}} = G_o E^{\delta+1} = G_p E^2 / \delta \quad (7)$$

$$G_p \equiv dI_{\text{perc}}/dV. \quad (8)$$

The local temperature T_{perc} due to this power dissipation through the percolation path is obtained by solving the heat diffusion equation [22]

$$c_p \rho \frac{dT_{\text{perc}}}{dt} + \kappa \nabla^2 T_{\text{perc}} = P_{\text{perc}}(t). \quad (9)$$

$T_{\text{perc}}(t)$ depends on heat capacity c_p , density ρ , specific thermal conductivity κ , of the oxide and of silicon surrounding the percolation spot, i.e.,

$$T_{\text{perc}}(t) = T_A + f(\kappa_{\text{ox}}, \kappa_{\text{si}}, T_{\text{ox}}, P_{\text{perc}}(t)) \quad (10)$$

where T_A is the ambient temperature. If the thermal conductivities are relatively constant (they may depend weakly on doping), then for an oxide of a given thickness

$$T_{\text{perc}}(t) \approx T_A + f(P_{\text{perc}}(t)). \quad (11)$$

Therefore, if the P_{perc} exceeds a certain critical power density (P_{crit}) for at least the duration of the thermal response time (τ_{ther}), T_{perc} will be high enough to melt silicon near the percolation spot ($>T_{\text{crit}}$) and allow it to flow through the oxide. This will cause an ohmic short-circuit across the oxide [23], [25], [30]. This, in our definition, is hard breakdown. However, if $P_{\text{perc}} < P_{\text{crit}}$, then the percolation path will be preserved even after breakdown (although leakage current through the oxide will increase). This is soft breakdown. Interestingly, since thermal resistances ($\kappa_{\text{ox}}, \kappa_{\text{si}}$) depend almost linearly on oxide thickness (see [22], [31], and [33]), thinner oxides require higher P_{crit} to reach a given T_{crit} (see [34]).

III. DISTRIBUTION OF PERCOLATION CONDUCTANCE

Since the assumption of the statistical distribution of the percolation current I_{perc} (or equivalently that of percolation conductance G_p [see Fig. 3]) is central to the analysis of soft and hard breakdown [see (3)], we explore this concept in detail in this section. In Section III-A, we describe the techniques we used to determine the percolation conductance from theory and from experiment. We then compare (in Section III-B), as a function of thickness and of power dissipation, the theoretical predictions of percolation conductance with the statistical distribution of the postbreakdown conductance (G_p) determined from measurements. We also see, in Section III-C, how percolation conductance evolves during the breakdown transient, and how it relates to the critical trap density N_{BD} (and indirectly to time-to-breakdown T_{BD}) providing microscopic insight into the breakdown process. Our analysis also helps determine (see Section III-D) the upper limit for trap-diameter D which is a critical, yet controversial parameter for predictive percolation models [24], [26], [27]. We find that $D < 1.5$ nm.

A. Measurement Technique and Theoretical Model for Percolation Conductance

To determine the percolation conductance, we measure the postbreakdown I - V characteristics for a large number of samples (~ 100). The area of the capacitors is kept small so that the first breakdown, be it soft or hard, is easily detected by a large change in current ($I_{\text{perc}} \gg AJ_{\text{stress}}$). Moreover, we use a constant voltage source with a low compliance limit so that power dissipation through the percolation path is minimized, and the conductance of the virgin percolation path is preserved. We then obtain G_o and δ for each of the samples by fitting the measured sample-specific postbreakdown I_{perc} - V characteristics (see Fig. 1, short dashed line,) with a power-law (5). The histogram of these conductivities gives us the experimentally determined probability distribution function (pdf). In the pdf, hard breakdown is characterized by $\delta < 2$ (essentially ohmic conduction though the percolation path), and soft breakdown by $\delta > 3$. We must detect the first breakdown event, be it hard

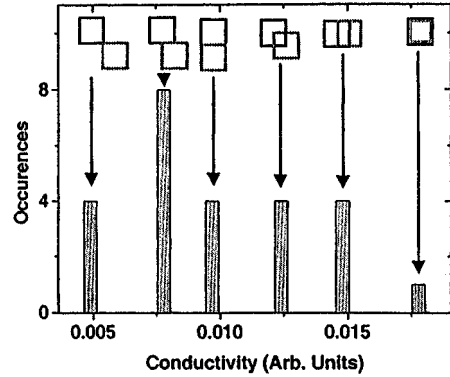


Fig. 4. Simple model explains the various configurations (top view) and conductivities for digitized overlap of two traps (solid square, at bottom; gray squares, at top) and predicts the salient features of the conductance distribution.

or soft, in analyzing the theory described above. If we continue stressing the oxide sufficiently long in constant voltage stress after the first breakdown, all oxides will eventually experience hard breakdown. These subsequent hard breakdowns will not be related to the initial distribution of percolation conductance and, therefore, cannot be described by the theory in this paper.

To determine the percolation conductance theoretically, we use a lattice-percolation algorithm [27]. A three-dimensional grid (x - y representing the area, and z representing the thickness direction) divides the oxides into many small cubes (see Fig. 4). Each trap is represented by a larger cube made of 6^3 smaller units. This discretization of the larger trap into smaller cubes allows for partial overlap of the traps which is important for thin oxides. The traps are randomly generated until a connected path joins the top and bottom electrodes. The total number of traps created in the sample before breakdown (N_{BD}) (see Fig. 3) and a subset of these traps (N_{perc}) that contribute to the percolation path, are both determined from this percolation model. To compute the conductance of this path, we replace each side of the smaller cubes which make up the percolation path, with a nonlinear conductor $g = iv^{-\delta}$, where v and i are the voltage across and current through the side of the smaller cube [21]. With $V = V_{\text{app}}$ across the oxide, we can now compute the net conductance of these randomly configured percolation paths with g as a (“hopping”) parameter of these traps, i.e., $G_o \equiv G_o(g, N_{\text{perc}}, \text{config})$. Both N_{perc} and the relevant overlap configuration comes from the percolation theory, but the “hopping parameter” g must be determined empirically from measurement by comparing a predicted distribution of $G_o^{(\text{expt})}$ (last paragraph) with that of $G_o^{(\text{theory})}(g)$. Once g is determined, the value is kept constant for all subsequent calculations involving different oxide thicknesses or stress conditions to make relative comparisons meaningful (of course, N_{perc} and P_{perc} will depend on oxide thickness and stress conditions). Obviously, this model does not explain the physical origin of the parameter g and the reason power-law conductance exists in the first place—by definition the model presupposes the power-law form of the net conductance.

Before making a detailed comparison between the theory and measurement, it is important to understand how the trap configurations affect G_o . Consider a model problem of computing the

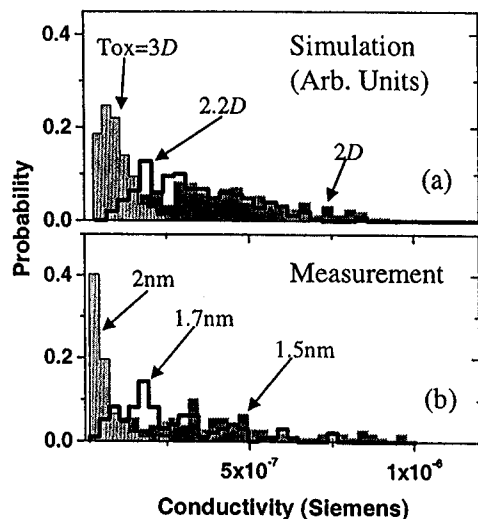


Fig. 5. (a) Probability distribution of percolation conductance for oxides with different thicknesses from percolation calculation. (b) Experimental data exhibiting the same trend. NFETs (Area = 2.5×10^{-8} cm²) were tested in inversion at 4 V. After soft breakdown, detected by at least 30% increase in current, a I - V measurement was performed. During stress the compliance was set to 100 μ A.

configuration phase-space of the traps in an oxide twice as thick as the trap diameter so that two traps can form a path across the oxide (see Fig. 4). Percolation paths for this oxide are dominated by vertical stacks (z -axis) involving just two traps—one on the top of the other ($N_{\text{perc}} = 2$), but having different degrees of overlap in the x - y plane (see inset of Fig. 4). A reduction in overlap between the traps decreases conductance, but increases the number of possible instances of the particular geometrical configuration [28]. If the traps are perfectly aligned, the conductance is maximum, but only one such configuration exists (see Fig. 4). Four other configurations all have just four instances, while one other configuration, with traps touching each other halfway along the edge, has eight instances. While this analogy is simple—since traps are not cubes and may not have such specific discreteness—it does predict some general features of the expected distribution for percolation conductivity. Indeed, this simple model predicts (a) the ratio of G_o^{max} to $G_o^{\text{min}} \sim 4$, (b) the asymmetry of the distribution profile, and (c) a distribution that can have discrete peaks. The peaks in this idealized distribution may be considerably broadened in practice (see Fig. 5), because the bonding configurations between traps may allow an increased number of discrete overlaps.

B. Distribution of Percolation Conductance: Theory Versus Experiment

Fig. 5(a) shows several key theoretical predictions based on detailed numerical calculation for the distribution of percolation conductivity as discussed in Section III-A. These predictions are then verified by the experimental data in Fig. 5(b) [14]. To construct each of these histograms, about 100 devices were tested, and their postbreakdown conductivities were noted and binned. As expected from the simple analysis discussed in the last paragraph of Section III-A, these distributions are fairly localized

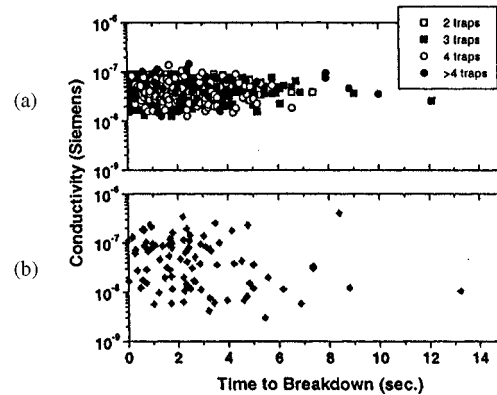


Fig. 6. (a) Conductivity after breakdown versus T_{BD} from a percolation calculation for $T_{\text{ox}} = 2D$. (b) Data from a 1.7-nm oxide (25 μ A compliance) shows the same general shape. Note that a given conductivity may arise from paths involving different numbers of traps.

and asymmetric, the ratio of maximum to minimum conductivities (defined with $\text{Prob}(G_p) > .04$) is approximately 5–10 [2], and the histograms for thinner oxides indeed show discreteness that disappears with increasing thickness (the sample is small, so the existence of discreteness at lower thickness is not conclusive). It is also clear from the experimental data that post-breakdown conductivity depends weakly on the thickness of the oxide. Such a systematic, albeit weak, dependence of the percolation conductance on thickness may require refinement of the ballistic quantum-point-contact model, which has been proposed, in [19] and [32], as an explanation for the power-law behavior of the I_{perc} - V characteristics.

To further validate the percolation model discussed in Section III-A, we plot G_p as a function of T_{BD} (Fig. 6). To determine G_p and T_{BD} from the percolation theory, note that $G_p = G_o(g, N_{\text{perc}}, \text{config})V^\delta$ and $T_{\text{BD}} \sim N_{\text{BD}}^m$ with ($0.3 < m < 0.7$). Experimentally G_p and T_{BD} are determined directly. The theory and experiment show remarkable similarities (see Fig. 6) regarding relative T_{BD} distributions: the conductivity-spread is about an order of magnitude, only a few breakdowns occur at very long times (and these long-lived oxides have neither the most nor the least conductivities), and rather surprisingly, there is no correlation between N_{perc} and T_{BD} . Fig. 6 also illustrates the expected one-to-many mapping between G_p and N_{perc} ; even though the N_{perc} and the geometrical configurations of two samples may be different, they may still conspire to give identical percolation conductances.

C. Compliance Limit, Reconfiguration of the Percolation Path, and Power Threshold

We have noted in discussing (5) that percolation conductance is expected to change during the breakdown transient. Therefore, to preserve the nature of the percolation path just before the breakdown (which allows comparison with predictions from the percolation theory), we minimized the energy discharge during the breakdown transient (which can produce a temperature-induced reconfiguration) by using a low current compliance for the test setup [29]. The agreement between various experimental results and theory in Section III-B shows that this can indeed be

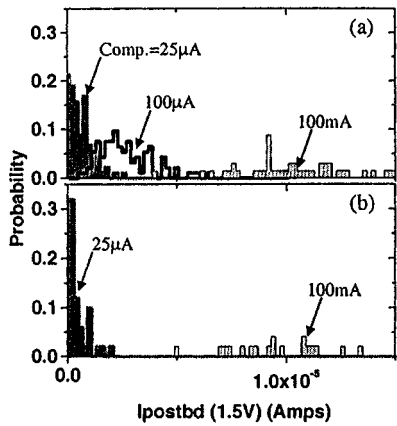


Fig. 7. (a) Distribution of percolation conductance for three sets of 1.7 nm oxides stressed at 4 V. Each set was stressed with a different compliance limit to control the amount of energy discharge through the percolation path. A systematic increase in conductivity is observed. (b) The same trend as in (a) is observed if a set of oxides is stressed first with low compliance and then, after breakdown, the stress is reapplied with higher compliance limit.

achieved. Interestingly, this technique of changing the amount of energy delivered to the percolation path by limiting the compliance [29] also allows us to probe the dynamics of reconfiguration [and, therefore, $G_p(t)$] of the percolation paths during the breakdown transient (see Fig. 7).

We probe this dynamic reconfiguration in two ways—first, we take three sets of samples (each about 100 devices) and stress them with the same voltage but with a different compliance limit. Since setting the compliance limit is similar to setting the postbreakdown time-delay for opening the switch that connects the stress source to the device under test (see Fig. 2, switch B), the subsequent conductance distribution will reflect the dynamic conductance distribution at the moment when the “switch was opened” (see Fig. 7). To ensure that these snap-shots of the conductivity distribution, taken from different sets of samples, actually correspond to the intrinsic time-dependent distribution, we can test another set of samples sequentially in the following way: first use a low compliance setup to interrupt the energy discharge at an early phase of the breakdown transient, note the corresponding conductivity distribution $G(t)$ and then restress the same oxides with a larger compliance limit allowing for more energy discharge (similar to reclosing switch B of Fig. 2(b) and then reopening it), and then re-determine the conductivity profile. Since the distribution determined from both procedures shows similar profiles (see Fig. 7), we are assured that the measured conductivity actually reflects the dynamic conductivity of the samples, rather than any measurement artifacts. The broadening of the conductance distribution with the increase in the compliance limit reflects both the reconfigurations as well as the generation of new traps within the percolation path during the breakdown transient, i.e., $G_o(t) \equiv G_o(N_{\text{perc}}(t), \text{config}(t))$. While we have used the compliance limit as a proxy for a time-delayed switch, one can also probe this transient distribution by changing the energy delivered to G_p by changing the stress voltage E , but keeping the compliance limit constant. In this case, $P_{\text{perc}} \sim G_p E^2$ increases with increasing voltage and

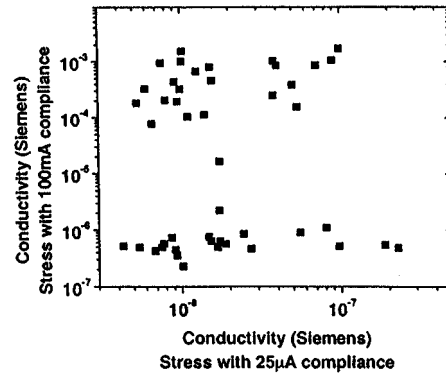


Fig. 8. Same data as in Fig. 7 replotted to determine the correlation between the low compliance conductivity and high compliance conductivity for each sample. No correlation is found.

changes the rate of new trap generation within the percolation path $dN_{\text{perc}}(t)/dt \sim E^2$. Therefore, a given compliance limit, even with constant time delay, can provide snap-shots of different stages of the conductivity evolution.

Given this reconfiguration of the percolation conductivity distribution, it is interesting to consider whether or not this dynamic evolution preserves the sequence of conductivities [30]. In other words, will the oxide with the largest initial postbreakdown conductance have the largest conductance after further conductivity evolution? The sequential stressing method for determining conductance evolution [see Fig. 7(b)], as described in the previous paragraph, does allow us to answer this question directly. In this measurement, oxides were biased with a constant voltage stress. For all these samples, we interrupt the energy discharge at the beginning of the breakdown transient so that the initial percolation paths are preserved (dark gray, Fig. 7(b); x -axis, Fig. 8). Subsequently, the stress is reapplied with higher compliance so that the percolation paths experience full power dissipation (light gray, Fig. 7(b); y -axis, Fig. 8). Remarkably, the initial G_p distribution early in the breakdown transient is not correlated to the final G_p distribution at the end of the breakdown transient. For a correlated distribution, the data points would have been clustered diagonally, not as randomly as shown in Fig. 8. This means that as the conductance distribution begins to evolve within the breakdown transient, this evolution is determined not only by the initial G_p itself, but also by the configuration of the defects that gave rise to that conductance. Since the conductance distribution and the trap configuration do not have a one-to-one relationship (see Fig. 6), the evolution of $G_p(t)$ need not preserve the sequence. Predicting such a configuration-dependent change in conductivity (similar to electro-migration in metal lines?) is beyond the scope of the breakdown model presented in Section II. Therefore, we conclude that while we can safely describe the statistical evolution of the conductivity in probabilistic terms, we cannot trace the conductivity evolution deterministically for individual samples.

D. Implication for Percolation Trap Diameter

Before we conclude this section on the analysis of the percolation conductance, let us emphasize that the existence of a

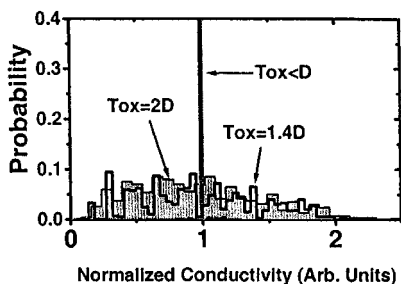


Fig. 9. Distribution of normalized conductivities (G/G_{avg}) obtained from the percolation calculation for several ratios of T_{ox} to the defect size, D . The width of the distribution decreases as the defect size becomes comparable to the oxide thickness. For defect sizes larger than twice the oxide thickness, a single defect can short-circuit the electrodes making the distribution nearly delta-distributed.

broad conductance distribution for 1.5-nm oxides requires that the trap diameter (D) be smaller than 1.5 nm (for our simulation, $D = 0.8$ nm). Indeed, if D were greater than T_{ox} , especially if $D \approx 2T_{ox}$ [27], the percolation conductances would have been nearly delta-distributed with the conductance range given by (see Fig. 9), $(G_{(D=2T_{ox})}^{max}/G_{(D=2T_{ox})}^{min}) = (1/2\sqrt{e})/(1/(1+e)) \approx 1.12$. Such a localized conductivity distribution is not observed in Figs. 5–7, establishing that the trap diameter D cannot exceed the oxide thickness T_{ox} . Contradicting this conclusion, however, one experimental paper (see [31] and [33]) is often cited to support the claim that $D \sim 3$ nm even for oxides with $T_{ox} = 1.5$ nm (i.e., $D \sim 2T_{ox}$).¹ However, the data in [31] and [33] only requires that the number of weak spots with “almost-completed” percolation paths be about two orders of magnitude smaller than the areal density of traps which in the percolation model is defined as the average number of traps, per square cm, in a given layer of the oxide. The total number of almost-completed percolation paths (defined as those paths that require only one more trap to complete the bridge across the two electrodes), determines the number of weak spots at the point of breakdown. The ratio of weak spots to the areal density of traps, determined theoretically from the square lattice percolation model described in Section II and based on trap-diameter of $D = 0.8$ nm, matches the results in [31] and [33] well (see Fig. 10). This agreement between theory and experiment, based on $D = 0.8$ nm, makes it unnecessary to assume $D \sim 3$ nm to interpret this experiment. Although only one experimental data point is shown, the main point here

¹In percolation models, the “size” of the traps is determined by requiring that the experimental Weibull slopes match the theoretical ones as a function of oxide thickness. If the trap generation is assumed linear, the predicted trap size is larger [27]; and if the trap generation is assumed nonlinear, the predicted trap size is smaller [24], [26]. The critical density of traps computed from the “linear” percolation model is generally smaller (due to the larger trap size) than the measured critical trap density. Therefore, the author in [27] assumes that two types of percolation paths exist: active ones which cause breakdown, and inactive ones which contribute only to SILC and causes local weak spots, but do not help oxide breakdown. In contrast, the “nonlinear” percolation model [26], with smaller trap sizes, predicts critical trap densities comparable to experimentally measured ones and, therefore, does not distinguish between electrically active and electrically inactive percolation paths. Since both models can predict the N_{BD} with their respective trap sizes and scaling factors, the distribution of percolation conductances and an alternate explanation of the “weak spots” provide additional experimental information to distinguish between them.

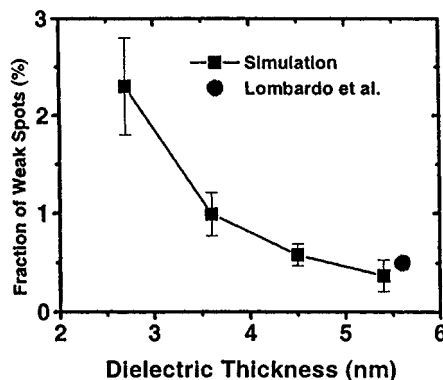


Fig. 10. Calculation of the ratio of “almost completed” path (just lacking one more trap to complete the percolation path) to the areal trap density plotted as a function of dielectric thickness. The calculation is in agreement with [31] even though 0.8 nm is used as the trap diameter.

is that the theory and experiment both agree that the ratio must be within a few percent without having to assume that the trap diameter exceeds 3 nm.

IV. CONCLUSION

To summarize, we have shown through measurement and percolation calculation that the postbreakdown conductivity is area-independent, sample-specific, statistically distributed, and asymmetric, and that this conductivity changes systematically during breakdown transient. We have noted that while the distribution broadens as a function of time during the transient, indicating new trap generation, it does not do so sequentially. In other words, the number of new traps generated during the breakdown transient, depends on other factors (configuration, for example) in addition to the total number of traps in the sample initially. In addition, this systematic agreement between theory and experiment convinces us that the trap diameter must be less than 1.5 nm for the distribution to be as broad as seen experimentally. In the companion paper [34], we shall use these properties of percolation conductance to study the scaling principles of soft and hard breakdown.

REFERENCES

- [1] S.-H. Lee, H.-J. Cho, J.-C. Kim, and S.-H. Choi, “Quasi-breakdown in ultrathin gate oxides under high field stress,” in *IEDM Tech. Dig.*, 1994, pp. 605–608.
- [2] K. Okada, S. Kawasaki, and Y. Hirofuji, “New experimental findings on stress induced leakage current of ultrathin silicon dioxides,” in *Int. Conf. Solid State Devices Material*, 1994, pp. 565–567.
- [3] M. Depas, T. Nigam, and M. H. Heynes, “Soft breakdown in ultrathin gate oxide layers,” *IEEE Trans. Electron Devices*, vol. 43, pp. 1499–1507, Sept. 1996.
- [4] M. Depas, R. Degraeve, G. Groeseneken, and M. Heynes, “Reliability of ultrathin gate oxide below 3 nm in the direct tunneling regime,” *Jpn. J. Appl. Phys.*, vol. 36, no. 38, pp. 1602–1608, 1997.
- [5] K. F. Schuegraf *et al.*, “Hole injection in SiO₂ breakdown model for very low voltage lifetime extrapolation,” *IEEE Trans. Electron Devices*, vol. 41, pp. 761–766, May 1994.
- [6] J. Bude, B. Weir, and P. Silverman, “Explanation of stress induced damage in thin oxides,” in *IEDM Tech. Dig.*, 1998, pp. 179–181.
- [7] M. Alam, J. Bude, and A. Ghetti, “Field acceleration for oxide breakdown—Can an accurate anode hole injection model resolve the E versus $1/E$ controversy,” in *Proc. 38th IRPS*, 2000, pp. 21–26.
- [8] J. H. Stathis and D. J. DiMaria, “Reliability projection for ultrathin oxides at low voltage,” in *IEDM Tech. Dig.*, 1998, pp. 167–170.

- [9] —, "Defect generation and reliability of ultrathin SiO₂ at low voltage," in *Proc. ECS*, H. Massoud, Ed., 2000, pp. 33–44.
- [10] J. W. McPherson and H. C. Mogul, "Underlying physics of the thermochemical *E* model in describing low field time dependent dielectric breakdown in SiO₂ thin films," *J. Appl. Phys.*, vol. 84, no. 3, pp. 1513–1523, 1998.
- [11] B. Weir *et al.*, "Ultrathin gate dielectrics: They break down, but do they fail?," in *IEDM Tech. Dig.*, 1997, pp. 73–76.
- [12] E. Wu, K. Nowak, J. Aitken, W. Abadeer, L. K. Han, and S. Lo, "Structural dependence of dielectric breakdown in ultrathin gate oxides and its relationship to soft broken modes and device failure," in *IEDM Tech. Dig.*, 1998, pp. 187–190.
- [13] T. Pompl, H. Wurzer, M. Kerber, W. Wilkins, and I. Eisele, "Influence of soft breakdown on NMOSFET device characteristics," in *Proc. 37th IRPS Conf.*, 1999, pp. 82–87.
- [14] F. Crupi, R. Degraeve, G. Groeseneken, T. Nigam, and H. Maes, "On the properties of the gate and substrate current after soft breakdown in ultrathin oxides layers," *IEEE Trans. Electron Devices*, vol. 45, pp. 2329–2334, Dec. 1998.
- [15] J. Suehle, E. Vogel, B. Wang, and J. Bernstein, "Temperature dependence of soft breakdown and wear-out in sub-3-nm SiO₂ films," in *Proc. IRPS*, 2000, pp. 33–39.
- [16] K. Okada, "The gate oxide lifetime limited by B-mode stressed induced leakage current and the scaling limit of silicon dioxide in the direct tunnelling region," *Semicond. Sci. Technol.*, vol. 15, no. 5, pp. 478–484, 2000.
- [17] K. Okada and S. Kawasaki, "New dielectric breakdown model of local wearout in ultrathin silicon dioxides," in *Int. Conf. Solid State Devices Material*, 1995, pp. 473–475.
- [18] K. Okada and K. Taniguchi, "Electrical stress induced variable range hopping conduction in ultrathin silicon dioxides," *Appl. Phys. Lett.*, vol. 70, no. 3, pp. 351–353, 1997.
- [19] J. Suñé, E. Miranda, M. Nafria, and X. Aymerich, "Failure physics of ultrathin SiO₂ gate oxides near their scaling limit," *Semicond. Sci. Technol.*, vol. 15, pp. 445–454, 2000.
- [20] T. Nigam, "Growth kinetics, electrical characterization, and reliability study of sub 5-nm gate dielectrics," Ph.D. dissertation, Katholieke Univ., Lueven, Belgium, May 1999.
- [21] M. Houssa, T. Nigam, P. Mertens, and M. Heynes, "Soft breakdown in ultrathin gate oxides: Correlation with the percolation theory of non-linear conductors," *Appl. Phys. Lett.*, vol. 73, no. 4, pp. 514–516, 1998.
- [22] B. K. Ridley, "Mechanism of electrical breakdown in SiO₂ films," *J. Appl. Phys.*, vol. 46, no. 3, pp. 998–1007, 1975.
- [23] N. Klein, "Switching and breakdown in films," *Thin Solid Films*, p. 149, 1971.
- [24] M. Alam, J. Bude, B. Weir, P. Silverman, A. Ghetti, D. Monroe, K. P. Cheung, and S. Moccio, "An anode hole injection percolation model for oxide breakdown—The Doom's day scenario revisited," in *IEDM Tech. Dig.*, 1999, pp. 715–717.
- [25] M. Alam, B. Weir, J. Bude, P. Silverman, and D. Monroe, "Explanation of soft and hard breakdown and its consequences for area scaling," in *IEDM Tech. Dig.*, 1999, pp. 449–452.
- [26] R. Degraeve *et al.*, "New Insights in the relation between electron trap generation and the statistical properties of oxide breakdown," *IEEE Trans. Electron Devices*, vol. 45, pp. 904–911, May 1998.
- [27] J. Stathis, "Percolation models for gate oxide breakdown," *J. Appl. Phys.*, vol. 86, pp. 5757–5766, 1999.
- [28] J. Krans *et al.*, "The signature conductance quantization in metallic point contacts," *Nature*, vol. 375, p. 767, 1995.
- [29] B. Linder, J. Stathis, R. Wachnik, E. Wu, S. Cohen, A. Ray, and A. Vayshenker, "Gate oxide breakdown under current limited constant voltage stress," in *VLSI Tech. Dig.*, 2000, pp. 214–215.
- [30] H. Satake and A. Toriumi, "Dielectric breakdown mechanism of thin-SiO₂ studied by post-breakdown resistance statistics," in *VLSI Tech. Dig.*, 1999, pp. 61–62.
- [31] S. Lombardo, F. Crupi, A. La Magna, C. Spinella, A. Terrasi, and A. La Mantia, "Electrical and thermal transient during dielectric breakdown of thin oxides in metal SiO₂-Si capacitors," *J. Appl. Phys.*, vol. 84, no. 1, pp. 472–479, 1998.
- [32] —, "Modeling the breakdown spots in silicon dioxide films as point contacts," *Appl. Phys. Lett.*, vol. 75, no. 7, pp. 959–961, 1999.
- [33] S. Lombardo *et al.*, "Degradation and hard breakdown spots in silicon dioxide films as point contacts," *J. Appl. Phys.*, vol. 86, no. 11, pp. 6382–6391, 1999.
- [34] M. A. Alam, B. E. Weir, and P. J. Silverman, "A study of soft and hard breakdown—Part II: Principles of area, thickness, and voltage scaling," *Trans. Electron Devices*, vol. 49, pp. 239–246, Feb. 2002.
- [35] D. J. DiMaria and J. W. Stasiak, "Trap creation in silicon dioxide produced by hot electrons," *J. Appl. Phys.*, vol. 65, no. 6, pp. 2343–2355, 1989.



Muhammad Ashraf Alam (M'97–SM'01) was born in Dhaka, Bangladesh, in 1964. He received the B.S.E.E. degree from Bangladesh University of Engineering and Technology in 1988, the M.S. degree from Clarkson College of Technology, Potsdam, NY, in 1991, and the Ph.D. degree from Purdue University, West Lafayette, IN, in 1994, all in electrical engineering.

From 1995 to 2000, he was with Bell Laboratories, Lucent Technologies, as Member of Technical Staff in the Silicon ULSI Research Department. Currently,

he is a Distinguished Member of Technical Staff of Agere Systems, Murray Hill, NJ, a spin-off of Lucent Technologies which specializes in the manufacture of electronic and optoelectronic devices. His research interest involves the physics of carrier transport in semiconductor devices, and he has worked on theoretical aspects of transport models, quasi-ballistic transport in bipolar transistors, MOCVD crystal growth, laser dynamics, and most recently, on the theory of oxide reliability.

Bonnie E. Weir (M'95) was born in Osaka, Japan, in 1965. She received the B.A. degree (with honors) in physics from Swarthmore College, Swarthmore, PA, in 1988, and the Ph.D. degree in physics from Stevens Institute of Technology, Hoboken, NJ, in 1994.

From 1988 to 2001, she was Member of Technical Staff at Bell Laboratories, working for AT&T and then Lucent Technologies. She is currently with Agere Systems, Murray Hill, NJ. Her former research areas include Rutherford backscattering and boron δ -doping. She is currently active in experimental research on ultrathin oxide reliability. She holds a patent on dielectric breakdown detection and has authored or coauthored more than 50 technical papers.

Paul J. Silverman was born in Philadelphia, PA, in 1948. He received the B.S. degree in physics from Drexel University, Philadelphia, and the M.S. degree in physics from Purdue University, West Lafayette, IN.

He joined Bell Laboratories in 1969, where he worked on loudspeaker design, Rutherford backscattering spectrometry, scanning tunneling microscopy, and molecular beam epitaxy. He is now working on electrical device characterization and dielectric reliability testing at Agere Systems (formerly Lucent Technologies' Microelectronics division), Murray Hill, NJ.