A Phenomenological Theory of Correlated Multiple Soft-Breakdown Events in Ultra-thin Gate Dielectrics

Muhammad A. Alam and R. Kent Smith
Agere Systems, 555 Union Blvd., Allentown, PA 18109, alam@agere.com

ABSTRACT

A general theory of the statistics of multiple breakdown events in ultra-thin gate oxides is developed. The standard theory of multiple breakdown events, based on Poisson statistics, allows only qualitative determination of the influence of each breakdown event on the rate of subsequent trap generation. In this paper, we develop a more general theory of multiple breakdown events to quantitatively determine the degree of correlation. This work allows one to calculate, regardless of the degree of correlation among the successive breakdown events, the statistical failure distribution of ICs as well as the time-dependent increase in the total leakage current during their operation. The theoretical description of leakage current not only explains all the experimental data reported in the literature to date, but also suggests a simple, yet fast experimental technique to determine the Weibull slope, the voltage- and the temperature- acceleration factors.

1. INTRODUCTION

There has been growing interest in using gate oxides after a soft-breakdown [SBD] event [1-7] to extend the operating lifetime of silicon ICs, because the first SBD limited safe operating voltage (V^{max}) for sub-2 nm oxide PMOS transistors has now become unacceptably small (Fig. 1, open squares and [3,8]). If SBD events are *uncorrelated* both in space and in time, that is, if the trap generation rate is neither enhanced nor localized after the each SBD event, the operating lifetime of the ICs can improve geometrically [2,3], i.e.

$$(\frac{T_n}{T_1})^{\beta} \approx \frac{C_n}{F_n^{(1-1/n)}} \qquad (1)$$

where $C_n = (n/e)(2\pi n)^{(1/2n)}$, T_n is the lifetime of the IC if n-th SBD in a given transistor causes it to malfunction, and F_n is the acceptable failure fraction (typically 10^{-12}). This increase in lifetime can be traded for higher operating voltage, denoted by the symbol $\times$ in Fig. 1, which would then allow continued scaling of silicon IC with sub-2 nm gate oxides. However, this reliability improvement depends on the degree of correlation, and a significant increase in the trap generation rate after a breakdown event may have a non-negligible impact on the failure statistics of subsequent SBDs, thereby reducing the safe operating voltage below that anticipated by Eq. 1. For example, before the first BD, trap generation (Fig. 2, open squares) is spatially random. However, once the vertical stack of defects form a percolation path, the SBD current I_{SBD} can be orders of magnitude larger than the background current, I_{TUNNEL}. This could raise the background temperature, which may, in principle, increase the trap generation rate (Fig. 2, crossed boxes) either globally (temporal enhancement) or locally in the vicinity of the existing percolation path (spatial enhancement). In either case, successive breakdown events (completion of new vertical stacks) could now occur faster, reducing the time T_n to reach the n-th breakdown in a transistor (i.e. T_n can no longer be described by Eq. 1). This reduction could adversely impact the reliability-speed budget, that is, correlation among breakdown events could decrease the safe operating voltage V^{max} if the IC operating lifetime T_n is held fixed. In this paper, we develop a theoretical framework to quantify the influence of such correlation on the estimates of TDDB lifetime improvement, and in the process, hope to unify the statistical description of the entire spectrum of breakdowns, from soft (weak correlation) to hard (strong correlation).

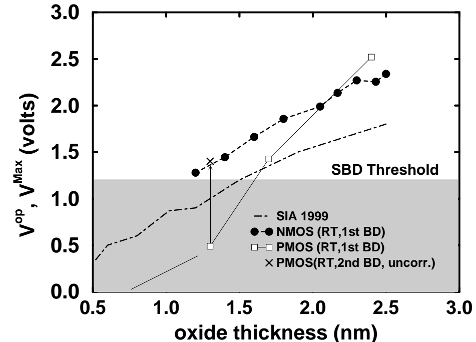

Fig. 1: Industry projections for the safe operating voltage for NMOS and PMOS technologies, based on first BD reliability estimate. According to this projection, while NMOS remain reliable down to 1.1-1.3 nm [10,11] with safe operating voltage in excess of ITRS requirement, PMOS does not [3,8]. This makes the analysis of soft-broken oxides particularly relevant for technologies with sub-2nm oxides.

2. ANALYTICAL THEORY

2.1 Model Equations:

Assume that a set of N samples are being electrically stressed at constant voltage. At any given time t, the samples with precisely n SBD spots, N_n, are assigned to reservoir R_n, so that $\sum_{R_n} N_n(t) = N$ (see Fig. 3), and the populations of the reservoirs are given by

$$\frac{P_n(\chi)}{d\chi} = [k_{n-1}P_{n-1} - k_nP_n] \qquad (2)$$

where $P_n \equiv N_n/N$, $\chi = (t/\eta)^\beta$ is the generalized and normalized time-variable, η is the mean-time to first breakdown, $\beta \propto T_{ox}$ - the oxide thickness, and k_n is the rate at which samples with n-spots develop $(n+1)$ breakdown spots.

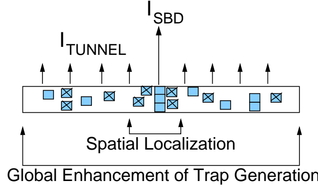

Fig. 2: Before the first-BD, the spatially random trap generation (open squares) can be described by Poisson statistics. However, once the vertical stack of defects complete a percolation path, the trap generation rate (crossed boxes) may increase either globally (temporal enhancement) or locally in the vicinity of the existing percolation path (spatial enhancement). In either case, standard Poisson statistics can no longer be used to describe the failure statistics of subsequent breakdown events.

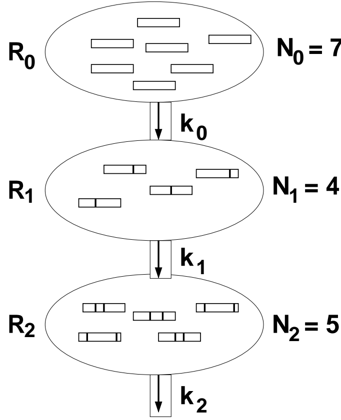

Fig. 3: Since standard Poisson statistics can no longer be used to analyze the correlated breakdown problem, we use a macroscopic Markov technique instead. Each small rectangle represents a stressed oxide (e.g. $N = 16$ in this example). The number of short vertical lines within each rectangle indicates the number of SBD spots in that sample at that instant. Samples are assigned to reservoirs (R_n) according to the number of breakdown spots (n) they have at a given instant, with a rate of transfer k_n characteristic of the trap generation rate of that reservoir. The number of reservoirs (R_n) equals the maximum number of SBD spots before a transistor is destroyed; here we sketch only three for illustration.

At $t = 0$, none of the samples has any breakdown spots, therefore $N = N_o$ and $P_o = e^{-\chi}$. Moreover, since a sample must develop n SBD events before it can have $(n+1)$ (*i.e.* the devices pass through the reservoirs sequentially), Eq. 2 can be solved by integral iteration

$$P_{n+1}(\chi) = k_n \int_o^\chi P_n(\chi')e^{-k_{n+1}(\chi-\chi')}\, d\chi'. \qquad (3)$$

Eq. 3 holds for any enhancement of trap generation rate with $k_n \geq 1$.

An analytical solution of Eq. 3 can be obtained for the case where each soft-breakdown event increases the trap generation rate by a factor of ξ, i.e. $k_n = 1 + n\xi$. With this linear correlation model, the probability distribution functions at any time χ are given by:

$$P_n(\chi,\xi) = \left[\prod_{m=0}^{n-1}(1+m\xi)\right]\left(\frac{1-e^{-\xi\chi}}{\xi\chi}\right)^n\left(\frac{\chi^n e^{-\chi}}{n!}\right). \qquad (4\text{-a})$$

The first two correlated distributions are

$$P_1(\chi,\xi) = \frac{(1-e^{-\xi\chi})e^{-\chi}}{\xi}, \qquad (4\text{-b})$$

$$P_2(\chi,\xi) = \frac{(1+\xi)}{2\xi^2}(1-e^{-\xi\chi})^2 e^{-\chi}. \qquad (4\text{-c})$$

2.2. SBD and HBD Limits and Conservation Rules:

Since Eq. 2 was derived from simple phenomenological considerations, it is important to check if its solutions have well-known asymptotic limits. For example, if the enhancement of trap generation after a SBD event is so small that $k_n \to 1$ for all n, then from Eq. 3 (for $n > 0$)

$$\lim_{k\to 1} P_n(\chi) = \int_o^\chi P_{n-1}(\chi')e^{-(\chi-\chi')}\, d\chi' = (\chi^n/n!)\, e^{-\chi}, \qquad (5\text{-a})$$

which is the Poisson distribution for samples with *exactly* n uncorrelated SBD events [2,3] (the same result could be obtained by setting $\xi \to 0$ in Eq. 4a). To convert this distribution to Weibull distribution, $W_n = ln\,(-ln(1-F_n))$, where F_n is the cumulative failure probability of samples with *at least* n breakdown events, we use the identity $\sum_{i=0}^{n-1} P_i(k) \equiv 1 - F_n$. For small values of χ, it is easy to show that

$$W_n = n\beta\; ln\,(t/\eta) + C(n) \qquad (5\text{-b})$$

where $C(n)$ is a n-dependent constant. Clearly, these Weibull distributions represent the statistics of uncorrelated breakdown, as has been shown in Ref. [1-3] by using binomial statistics.

On the other hand, if enhancement of trap generation after the first BD is so strong that $k \to \infty$, then

$$\lim_{k\to\infty} P_{n>0}(\chi) = \lim_{k\to\infty}\int_o^\chi P_{n-1}(\chi')e^{-k(\chi-\chi')}\, d\chi' \quad \to 0. \qquad (5\text{-c})$$

Once again, one may compute the Weibull distributions by using the identity $\sum_{i=0}^{n-1} P_i(k) \equiv 1 - F_n$. However, since now only $P_o = e^{-\chi}$ and $P_{n>0} \to 0$ otherwise, therefore for HBD, $F_n \to F_1$: in other words, during the highly correlated trap generation due to temperature-induced feedback, all Weibull distributions collapse onto the first Weibull line in a runaway process which is characteristic of HBD.

Also, since all the samples must have n-spots at least once, Eq. 3 allows us to prove the following conservation rules

$$\sum_{time} dP_n = \int_0^\infty k_{n-1} P_{n-1} \, d\chi = 1. \quad \text{(5-d)}$$

Moreover, since the devices must be in one or the other reservoir at any given instant (Fig. 3), that is, since each of the sample must have one or more percolation paths present at any given time, therefore

$$\sum_n P_n = 1 \quad \text{(5-e)}$$

It is easy to show that the iterated probability distributions described in Eq. 3 satisfy this conservation rule as well.

2.3 Numerical Verification

To prove that this analytical theory is not only correct at the asymptotic limits, but also correct in the details, we verify Eq. 4 against the results of a numerical percolation model [12,13]. We shall only compare the analytical and numerical values of the first non-trivial distribution, $P_1(k)$. If they coincide, the conservation rules and the asymptotic limits established in Sec. 2.2 guarantee that other distributions, $P_{n>1}$, should coincide as well.

In the numerical percolation model, we divide the area of the oxide into N_{xy} units and its thickness into N_z units. Cells of this cubic lattice are randomly marked defective until the formation of the first vertical defective column, establishing a low resistive path between substrate and gate. After the formation of this first percolation path, two different processes can occur. In the first scenario, the rate of trap generation may increase "globally" by a factor of p (temporal correlation). The impact of such global increase in trap generation rate on the failure statistics is shown in Figs. 4 and 5, in which we plot the numerical results from the percolation simulation for P_o and P_1 as a function of $\chi(\equiv N_{xy}\left(\frac{N}{N_{xy}N_z}\right)^{N_z})$, for various values of p. Remarkably, Eq. 4b fits the numerical results quantitatively, including the positions of the peaks and the shapes of the distributions (Fig. 4) as well as their exponentially decaying tails at longer times (Fig. 5). Excellent agreement is obtained among these results and the analytical model by using the relationship $k \equiv 1 + \xi = 1 + 2p$ (Fig. 6).

In the second scenario, the increase in the current through the percolation path could, in principle, localize trap generation rate (spatial correlation) within a areal diameter L_c around the first percolation path. Such spatial localization could allow successive breakdown paths to form quickly, leading to a rapid increase in the total leakage current. To numerically simulate such cases, we again begin by randomly marking the cells of the cubic lattice percolation model defective until the first defective column forms. Then, instead of increasing the trap generation rate by a factor p globally, we localize the subsequent trap generation within a distance $L_c/2$ on either sides of the first breakdown spot. Clearly, L_c would become smaller with increasing spatial localization. Fig. 7 shows that even in such cases of increasing spatial localization, the analytical

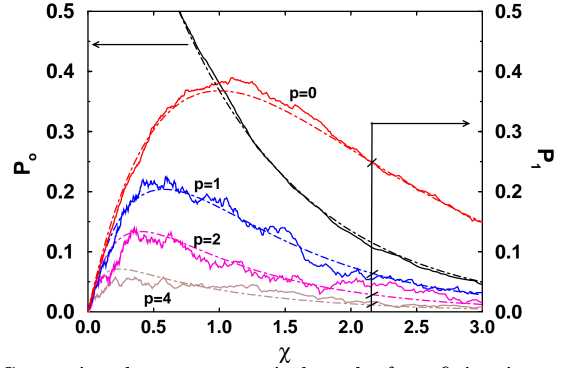

Fig. 4: Comparison between numerical results from finite-size percolation model (solid lines; $N_{xy} = 500, N_z = 5$) and Eq. 4b (dashed lines) for various values of enhancement of trap-generation rates. The excellent agreement between the numerical results and the analytical formula establishes the validity of the theoretical framework.

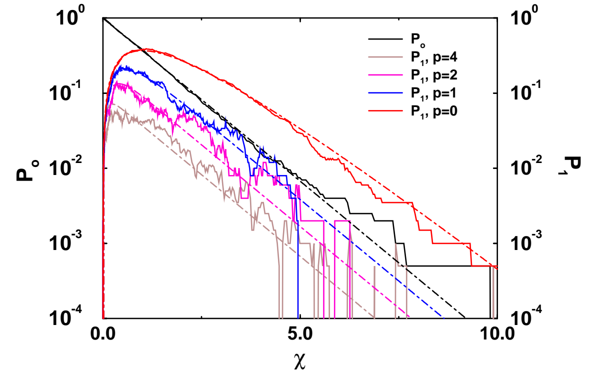

Fig. 5: The data of Fig. 4 is replotted in a log plot to show that even the exponentially decaying probability functions from the percolation calculations (solid lines) are well represented by the analytical probability distributions (dashed lines). The noise in the numerical data reflects the difficulty of computing rare events by Monte Carlo process - highlighting the need for the analytical theory described in this paper.

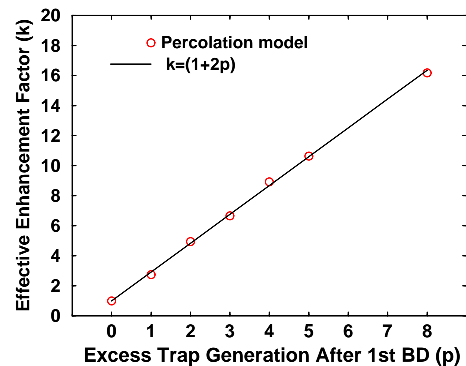

Fig. 6: The phenomenological enhancement constant k is linearly related to the microscopic enhancement of trap generation rate, p.

theory developed in Sec. 2 still provides remarkably good

description of the trap generation process, provided we interpret the enhancement factor k as a spatial average of the trap generation process, i.e. $k \equiv 1+\xi = (1+2p)f(N_{xy}/L_c)$ (see Fig. 8).

It is important to understand that the analytical theory (e.g. Eq. 4a) only allows determination of the correlation parameter $k(\equiv 1+\xi)$ from the measured breakdown statistics, but can not by itself distinguish between spatial-correlation (L_c) or temporal- correlation (p). To make such distinction, one need to use the position measurement techniques described in Refs. [3,14]. Once the spatial localization parameter L_c is known [3], then the experimentally measured ξ will allow the determination of the time-correlation parameter p, thereby providing a complete description of the trap generation process.

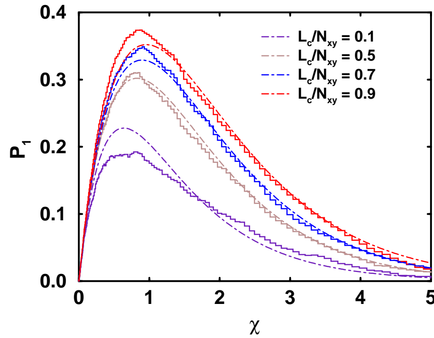

Fig. 7: A cubic lattice of dimension, N_{xy}=500, N_z=5, was used to test the validity of the analytical theory in the presence of spatial correlation. (dot-dashed line: analytical model, solid line: numerical percolation model). The mean-field analytical theory matches the numerical results very well even when spatial localization is strong. Note that the ratio of L_c to N_{xy} indicates the strength of spatial localization.

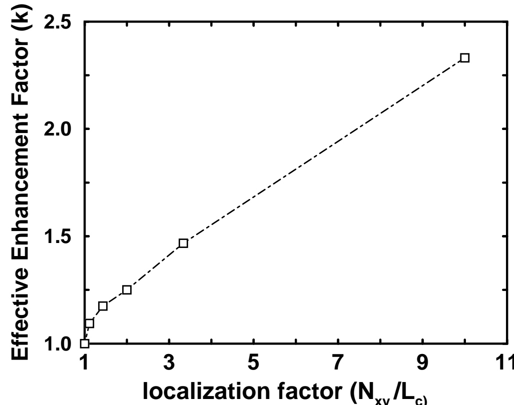

Fig. 8: The relationship between correlation parameter $k(\equiv 1+\xi)$ and the spatial localization parameter N_{xy}/L_c is no longer as simple as the relationship between k and the temporal localization parameter p (Fig. 6), although the relationship is still monotonic. Such deviation from linearity is expected, because the analytical theory developed in this paper treats the trap generation process in an average manner, without regard to its spatial features.

3. APPLICATION OF THE MODEL

3.1 Reliability Projection:

In order to use the analytical model to quantify the improvement of reliability due to multiple SBD with various degrees of correlations, we use the first two cumulative failure distributions to compute $(F_1 - F_2) = P_1(k)$ [3], then determine $k(\equiv 1+\xi)$ by using Eq. 4b. Since $\sum_{i=0}^{n-1} P_i(k) \equiv 1 - F_n$, Eq. 4a can then be used to construct all the successive Weibull distributions, $W_n = ln(-ln(1-F_n))$. For example, as shown in Fig. 9 (based on the data from Ref. [15]), that the increased trap generation rate after the first SBD can reduce the 3rd SBD lifetime by approximately 45 % ($A \to B$), or equivalently, can increase the number of soft-broken transistors by 37× ($A \to C$). In general, it is easy to show that the correlation effects reduce the time to n-th SBD failure time as

$$\left(\frac{T_n(\xi)}{T_1}\right)^{\beta} \approx \frac{1}{(1+(n\xi/2))^{(1-1/n)}} \left(\frac{C_n}{F_n^{(1-1/n)}}\right) \qquad (6)$$

which is a simple generalization of Eq. 1 in the presence of linear correlation. Surprisingly, Eq. 6 predicts that the correction to the previously anticipated enhancement of lifetime due to soft-breakdown events (Eq. 1) would be small even with relatively large increase in the trap generation rate (i.e. $\xi \sim 1$) after the first breakdown event. This reinforces the general validity of the predicted improvement in lifetime (and the safe operating voltage) discussed in Ref. [3]. Indeed, as oxides get thinner than those discussed in Fig. 9 and the operating voltages are reduced following the ITRS roadmap (see Fig. 1), the corresponding reduction in power dissipation through the percolation path [3,4] allows the correlation coefficient ξ to approach 0 quickly. In that case, the breakdown will be essentially uncorrelated as discussed in Ref. [1,3], with a corresponding increase in the lifetime for oxides with multiple breakdown events given by Eq. 1.

On the other hand, if the correlation is so strong that $\xi \to \infty$, then Eq. 6 predicts that $T_n \to T_1$. In this case, no improvement in lifetime due to multiple breakdown events is expected. This is precisely the case for rapid succession of breakdown events during hard breakdown.

3.2 Analysis of Leakage Current:

In addition to making precise determination of the failure statistics of multiple soft-breakdown events in the presence of an arbitrary degree of spatial and temporal correlation, as discussed above, the Markov formulation of multiple breakdown events also allows us to determine the total increase in the leakage current of an IC under such conditions [16-18]. The total leakage current due to soft-breakdown in gate dielectrics is given by the the first moment of the probability distribution function, i.e.

$$L(\xi,t) \equiv \frac{I_{leak}(\xi,t)}{I_o N_T} = \sum_{m=1}^{\infty} m P_m(\xi,t), \qquad (7)$$

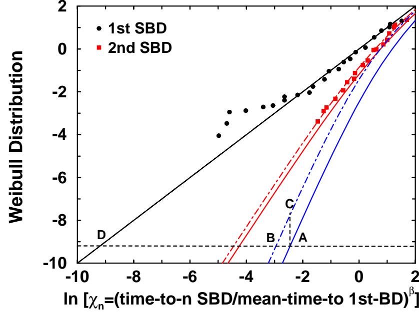

Fig. 9: An example of how correlated breakdown may reduce the estimated enhancement of lifetime due to multiple SBD events. The normalized data from Ref. [15] (collected from a graph, so the values may not be precise) shows that the 2nd SBD is best fitted with $p = 0.25$ or 25% enhancement of trap generation after the first SBD. (Solid lines: uncorrelated, dot-dashed line: weakly correlated). This translates to a 45% correction in extrapolated lifetime for the third SBD ($A \rightarrow B$). However, this correction is small compared to the orders of magnitude improvement of the lifetime of an IC with transistors tolerant to three SBD events ($D \rightarrow A$).

where N_T is the number of transistors in an IC and I_o in the average increase in the leakage current per breakdown event. Assuming a linear correlation model, we can insert the correlation parameters $k_n = 1 + n\xi$ in Eq. 2 to obtain

$$\frac{dP_m(\xi,\chi)}{d\chi} = (1+(m-1)\xi)\,P_{m-1}(\chi) - (1+m\xi)\,P_m(\chi).$$

Multiplying the both sides of the equation by m and summing over all m from 1 to ∞, we derive the following simple differential equation for the leakage current of an IC

$$\frac{dL(\xi,\chi)}{d\chi} = \xi L(\xi,\chi) + 1 \quad (8)$$

which can be solved to obtain

$$L \equiv \frac{I_{leak}}{N_T I_o} = \chi \left[\frac{e^{\xi\chi} - 1}{\xi\chi}\right] \quad (9)$$

Eq. 9 is another key result of this paper, with many implications. Indeed, as we shall see below that Eq. 9 can explain all aspects of leakage current data reported in the literature to date.

Leakage current for uncorrelated breakdown: Let us begin by considering the limiting case when the correlation among breakdown events are weak, so that $k_n \rightarrow 1$ or $\xi \rightarrow 0$. Under this condition, Eq. 9 reduces to

$$\frac{I_{leak}}{N_T I_o} = \chi = \left(\frac{t}{\eta}\right)^{\beta} \quad (10)$$

as discussed in Refs. [3,19]. Inserting the voltage- and temperature- dependence of the mean failure time $\eta = \eta_o e^{-\gamma_V (V-V_o)} e^{-\gamma_T (T-T_o)}$ in Eq. 10, we obtain the following theoretical result for the leakage current

$$ln(I_{leak}) = \beta ln(t) + \gamma_V (V - V_o) + \gamma_T (T - T_o) - ln(\eta_o) \quad (11)$$

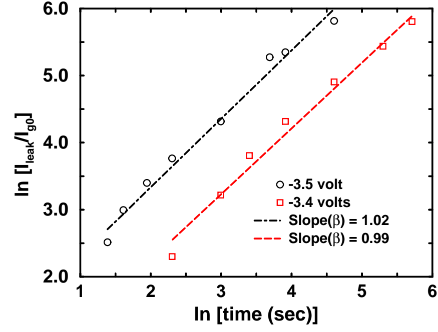

Fig. 10: An example of the increase in the leakage current that follows Eq. 11. The oxide thickness is 1.3 nm (hence the slope, $\beta \sim 1$), the area of the oxide is 100 μm^2, and the voltage acceleration factor is $\gamma_V = 5.4$ dec/volt, close to the expected value of 5.5 dec/volt determined from test of small transistors [8]. We should also compare this number to 5.2 dec/volt reported in Fig. 16, Ref. [18]

where η_o is the mean lifetime at voltage V_o and temperature T_o. According to this formula, the leakage current curves for different voltages and/or temperatures are anticipated to be parallel to each other, with the separation constants given by the voltage- (γ_V) and temperature- (γ_T) acceleration factors. For a given voltage and temperature, the leakage current is expect to increase with time as a power law with slope identical to the Weibull slope (β) for the oxide. The leakage current data shown in Fig. 10 supports this conclusion. Indeed, this theoretically derived formula is identical to the one derived empirically in Ref. [17, Eq. 1-3; 18, Eq. 5] to describe leakage current data, and provides theoretical validation of these empirical results.

The above result also resolves the confusion regarding the linear vs. nonlinear time-dependence of the leakage data reported in the literature. Indeed, for thiner oxides, with $\beta \rightarrow 1$, the leakage current may be well characterized by a linear function of time within a short experimental window. However, if $\beta \neq 1$ exactly, then over a longer measurement window or for thicker oxides, nonlinear power-law time dependence is expected. As seen in Fig. 5 of Ref. 18, the slope of the post-breakdown $ln(I) - ln(t)$ curve is given by approximately 1.2-1.4, consistent with the Weibull slope of the 2 nm oxide described.

A measurement technique for β, γ_V, and γ_T: Remarkably, therefore, the Weibull like formula described in Eq. 11 can be used to determine Weibull slope β, voltage acceleration factor γ_V, and temperature acceleration factor γ_T from a few simple measurements of the leakage currents of *large* area transistors stressed at different voltages and temperatures. This would dramatically reduce the testing time to determine these parameters [19]. Indeed, the failure to detect the first breakdown in large area testers, traditionally a strong argument against using such structures to construct the Weibull distribution, does not apply to this technique at all. The technique simply depends on the slopes and shifts of parallel leakage current lines to compute Weibull slopes and voltage- and temperature acceleration factors.

Leakage current for correlated breakdown: Finally, if the correlation among the breakdown events is strong *i.e.* $\xi \gg 1$, then the leakage current will increase exponentially with time, i.e.

$$ln\left(\frac{I_{leak}}{I_o N_T}\right) \sim \xi\chi = \xi(t/\eta)^{\beta}. \quad (12)$$

Such leakage current is characteristic of increasingly harder breakdown events induced by higher voltages. If β and η are known from standard measurements, Eq. 12 provides an experimental technique to determine the correlation parameter ξ by a simple measurement of the time dependent leakage current. The parameter ξ can then be used to construct the general failure probability distributions using Eq. 4a. Note that while this formula (Eq. 12) applies to the simulation up to the hard breakdown limit, this does not apply to hard breakdown current itself. This is because the trap generation formula assumes a simple linear increase in trap generation rate with each SBD event, whereas a hard-breakdown event includes a melting of the local spot with discontinuous increase in k_n beyond certain critical value of n. Such discontinuous changes are not reflected in the formula for the leakage current just derived. However, the formula should be applicable to computing the leakage current until the initiation of the hard breakdown.

4. SUMMARY

We have developed a phenomenological model for correlated breakdown that correctly reproduces the limiting cases of uncorrelated breakdown, follows all the conservation rules, and reproduces, in detail, the exact results from a finite-size percolation model allowing us to explore the statistics of breakdown events regardless of the degree of correlation present. We have also shown how such a model may revise downward the reliability improvement predicted by simpler models. However, for thinner oxides, we find that any such correction due to correlation effects is negligible because at lower operating voltage, the breakdown is universally soft with very small degree of correlation, confirming the validity of the results discussed in Ref. [1-3]. Since the "numerical noise" in a percolation model makes it difficult to compute extraordinarily rare events such as those used in reliability extrapolation, the phenomenological theory described here is the only way to make appropriate reliability projections in the presence of correlation, and to quantitatively consider the reliability vs. speed (or power dissipation) tradeoff during IC design. Finally, we have provided a detailed analysis of post-breakdown leakage current within the theoretical framework developed in this paper. The formula developed not only provides circuit designers specific design targets for increased leakage current, but also explains a wide variety of experimental data reported in the literature. This analysis of the leakage current also allows a new, simple, and direct measurement technique to determine Weibull slopes and voltage acceleration factors for ultra-thin gate dielectrics.

Acknowledgement

The authors would like to thank B. Weir, P. Silverman, P. Mason, and D. Monroe for many stimulating discussions. We are also grateful to J. Suñé for pointing out that the authors of Ref. [19] had also arrived at the same expression for the leakage current for uncorrelated breakdown (Eq. 10) using the properties of extreme-value distribution.

REFERENCES

1. M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Uncorrelated breakdown of silicon integrated circuits," Nature, 6914, 378, 2002.

2. J. Suñé and E. Wu, "Statistics of successive breakdown events for ultrathin gate oxides," IEDM Tech. Digest, pp. 147-150, 2002.

3. M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Statistically independent soft-breakdowns redefine oxide reliability specifications," IEDM Tech. Digest, pp. 151-154, (2002).

4. M. Alam, B. Weir, and P. Silverman, "A study of soft and hard breakdown," IEEE Trans. Elec. Dev., **49**, pp. 239-246, 2002.

5. M. Alam, B. Weir, and P. Silverman, "A Future of function or failure," IEEE Electronics and Photonics Magazine, **18**(2), pp. 42-48, 2002.

6. B. Linder *et al.*, "Gate oxide breakdown under current limited constant voltage stress," VLSI Digest, 214, 2000,

7. B. Kaczer *et al.*, "Impact of MOSFET oxide breakdown on digital circuit operation and reliability," IEDM Tech. Digest, pp. 553-556, 2000.

8. B. Weir, M. Alam, P. Silverman, and Y. Ma, "Low voltage gate dielectric reliability," ECS Proc., 2002-2, pp. 465-474.

9. J. Bude, B. Weir, and P. Silverman, "Explanation of stress induced damage in thin oxides", IEDM Tech. Digest, pp. 179-182, 1998.

10. M. Alam, B. Weir, and P. Silverman, "Physics and Prospects of sub-2 nm Oxides," ECS Proc. 2000-2, 365.

11. E. Wu, J. Stathis, and L.-K. Han, "Ultrathin oxide reliability for ULSI applications," Semi. Sci. Tech., **15**, 425, 2000.

12. R. Degraeve *et al.*, "New insights in the relation between electron defect generation and the statistical properties of oxide breakdown," *IEEE Tran. Elec. Dev.*, **45**, 904-911, (1998).

13. J. Stathis, "Percolation Models for Gate Oxide Breakdown," *J. Appl. Phys.*, **86**, 5757-5766, (1999).

14. R. Degraeve, B. Kaczer, An De Keersgieter, and Guido Groeseneken, "Relation between breakdown mode and breakdown location in short channel NMOSFETs and its impact on reliability specifications," IRPS Proc., pp. 360-366, 2001.

15. J. Suñé, G. Mura, and E. Miranda, "Are soft breakdown and hard breakdown of ultrathin gate oxides actually different failure mechanisms?," EDL, **21**, 4, p. 167, 2001.

16. K. Okada, H. Kubo, A. Ishinaga, and K. Yoneda, "A concept of gate oxide lifetime limited by 'B-mode' stress induced leakage currents in direct tunneling regime," VLSI Digest, pp. 57-58, 1999.

17. T. Hosoi, P. Lo Re, Y. Kamakura, and K. Taniguchi, "A new model of time evolution of gate leakage current after soft breakdown in ultra-thin gate oxides, IEDM Tech. Digest, pp. 155-158, 2002.

18. F. Monsieur *el al.*, "A thorough investigation of progressive breakdown in ultra-thin oxides. Physical understanding and Application for industrial reliability assessment." IRPS Proceedings, pp. 45-54, 2002.

19. E. Farrés, M. Nafria, J. Suñé, and X. Aymerich, "The statistical distribution of breakdown from multiple breakdown events in one sample," J. Phys. D: Appl. Phys. 24, 407-414, (1991).